1, TITLE: Understanding Embodied Reference with Touch-Line Transformer

https://openreview.net/forum?id=ugA1HX69sf

AUTHORS: Yang Li, Xiaoxue Chen, Hao Zhao, Jiangtao Gong, Guyue Zhou, Federico Rossano, Yixin Zhu HIGHLIGHT: We study embodied reference understanding: locating referents using embodied gestural cues and language references.

2, TITLE: ISS: Image as Stepping Stone for Text-Guided 3D Shape Generation

https://openreview.net/forum?id=GMRodZ8OlVr

AUTHORS: Zhengzhe Liu, Peng Dai, Ruihui Li, XIAOJUAN QI, Chi-Wing Fu

HIGHLIGHT: This paper presents a new framework called Image as Stepping Stone (ISS) for the task by introducing 2D image as a stepping stone to connect the two modalities and to eliminate the need for paired text-shape data.

3, TITLE: Structured World Representations via Block-Slot Attention

https://openreview.net/forum?id=ZPHE4fht19t

AUTHORS: Gautam Singh, Yeongbin Kim, Sungjin Ahn

HIGHLIGHT: In this paper, we propose a novel object-centric representation, called Block-Slot Representation which, unlike the conventional slot representation, provides concept-level disentanglement within a slot.

4, TITLE: Learning to Estimate Single-View Volumetric Flow Motions without 3D Supervision https://openreview.net/forum?id=2vmGv5wPDBZ

AUTHORS: Erik Franz, Barbara Solenthaler, Nils Thuerey

HIGHLIGHT: We address the challenging problem of jointly inferring the 3D flow and volumetric densities moving in a fluid from a monocular input video with a deep neural network.

5, TITLE: DBQ-SSD: Dynamic Ball Query for Efficient 3D Object Detection

https://openreview.net/forum?id=ZccFLU-Yk65

Jinrong Yang, Lin Song, Songtao Liu, Weixin Mao, Zeming Li, Xiaoping Li, Hongbin Sun, Jian Sun, Nanning AUTHORS: Zheng

HIGHLIGHT: Different from them, we propose a Dynamic Ball Query (DBQ) network to adaptively select a subset of input points according to the input features, and assign the feature transform with a suitable receptive field for each selected point.

6, TITLE: Edgeformers: Graph-Empowered Transformers for Representation Learning on Textual-Edge Networks

https://openreview.net/forum?id=2YQrge4RNv AUTHORS: Bowen Jin, Yu Zhang, Yu Meng, Jiawei Han

HIGHLIGHT: In this paper, we propose Edgeformers, a framework built upon graph-enhanced Transformers, to perform edge and node representation learning by modeling texts on edges in a contextualized way.

7, TITLE: Protein Representation Learning by Geometric Structure Pretraining

https://openreview.net/forum?id=to3qCB3tOh9

AUTHORS: Zuobai Zhang, Minghao Xu, Arian Rokkum Jamasb, Vijil Chenthamarakshan, Aurelie Lozano, Payel Das, Jian Tang

HIGHLIGHT: In this paper, we propose to pretrain protein representations according to their 3D structures.

8, TITLE: Any-scale Balanced Samplers for Discrete Space

https://openreview.net/forum?id=lEk10jdSb7B

AUTHORS: Haoran Sun, Bo Dai, Charles Sutton, Dale Schuurmans, Hanjun Dai

HIGHLIGHT: Instead, we propose any-scale balanced samplers to repair the gap in non-local proposals.

9, TITLE: Truthful Self-Play

https://openreview.net/forum?id=WVRb98rwbv9

AUTHORS: Shohei Ohsawa

HIGHLIGHT: We present a general framework for evolutionary learning to emergent unbiased state representation without any supervision.

10, TITLE: Rethinking Symbolic Regression: Morphology and Adaptability in the Context of Evolutionary Algorithms https://openreview.net/forum?id=OPGy07PojsZ

AUTHORS: Kei Sen Fong, Shelvia Wongso, Mehul Motani

In this paper, we rethink SR from 2 perspectives: morphology and adaptability.For researchers interested in HIGHLIGHT: equation-recovery problems, we also propose a set of conventions that can be used to promote fairness in comparison across SR methods and to reduce unintentional bias.

11, TITLE: Equivariant Shape-Conditioned Generation of 3D Molecules for Ligand-Based Drug Design https://openreview.net/forum?id=4MbGnp4iPQ AUTHORS: Keir Adams, Connor W. Coley

HIGHLIGHT: We introduce a new multimodal 3D generative model that enables shape-conditioned 3D molecular design by equivariantly encoding molecular shape and variationally encoding chemical identity.

12, TITLE: TranSpeech: Speech-to-Speech Translation With Bilateral Perturbation https://openreview.net/forum?id=UVAmFAtC5ye AUTHORS: Rongjie Huang, Jinglin Liu, Huadai Liu, Yi Ren, Lichao Zhang, Jinzheng He, Zhou Zhao HIGHLIGHT: In this work, we propose TranSpeech, a speech-to-speech translation model with bilateral perturbation. 13, TITLE: A law of adversarial risk, interpolation, and label noise https://openreview.net/forum?id=0_TxFpAsEI AUTHORS: Daniel Paleka, Amartya Sanyal HIGHLIGHT: We show that interpolating label noise induces adversarial vulnerability, and prove the first theorem showing the relationship between label noise and adversarial risk for any data distribution. 14, TITLE: Short-Term Memory Convolutions https://openreview.net/forum?id=4DU HCijfJp AUTHORS: Grzegorz Stefa?ski, Krzysztof Arendt, Pawe? Daniluk, Bart?omiej Jasik, Artur Szumaczuk HIGHLIGHT: Although they often have superior quality compared to standard DSP methods, this advantage is diminished by higher latency. In this work we propose a method for minimization of latency and memory consumption, called Short-Term Memory Convolution (STMC) and its transposed counterpart. 15. TITLE: StyleMorph: Disentangling Shape, Pose and Appearance through 3D Morphable Image and Geometry Generation https://openreview.net/forum?id=Ojpb1y8jflw AUTHORS: Eric-Tuan Le, Edward Bartrum, Iasonas Kokkinos We introduce StyleMorph, a 3D generative model that relies on the 3D morphable model paradigm to HIGHLIGHT: disentangle shape, pose, object and scene texture for high quality image synthesis. 16, TITLE: SlotFormer: Unsupervised Visual Dynamics Simulation with Object-Centric Models https://openreview.net/forum?id=TFbwV6I0VLg Ziyi Wu, Nikita Dvornik, Klaus Greff, Thomas Kipf, Animesh Garg AUTHORS: HIGHLIGHT: In this paper, we successfully apply SlotFormer to perform video prediction on datasets with complex object interactions. Deconstructing Distributions: A Pointwise Framework of Learning 17, TITLE: https://openreview.net/forum?id=9IaN4FkVSR1 AUTHORS: Gal Kaplun, Nikhil Ghosh, Saurabh Garg, Boaz Barak, Preetum Nakkiran HIGHLIGHT: In this work, we propose a new approach: we measure the performance of a collection of models when evaluated at *single input point*. 18. TITLE: Trading Information between Latents in Hierarchical Variational Autoencoders https://openreview.net/forum?id=eWtMdr6yCmL AUTHORS: Tim Z. Xiao, Robert Bamler HIGHLIGHT: The proposal of \$\beta\$-VAEs breaks this interpretation and generalizes VAEs to application domains beyond generative modeling (e.g., representation learning, clustering, or lossy data compression) by introducing an objective function that allows practitioners to trade off between the information content (``bit rate") of the latent representation and the distortion of reconstructed data. In this paper, we reconsider this rate/distortion trade-off in the context of hierarchical VAEs, i.e., VAEs with more than one layer of latent variables. FairGBM: Gradient Boosting with Fairness Constraints **19. TITLE:** https://openreview.net/forum?id=x-mXzBgCX3a AUTHORS: André Cruz, Catarina G Belém, João Bravo, Pedro Saleiro, Pedro Bizarro HIGHLIGHT: We present FairGBM, a dual ascent learning framework for training GBDT under fairness constraints, with little to no impact on predictive performance when compared to unconstrained GBDT. 20, TITLE: DySR: Adaptive Super-Resolution via Algorithm and System Co-design https://openreview.net/forum?id=Pgtn4l6eKjv AUTHORS: Syed Zawad, Cheng Li, Zhewei Yao, Elton Zheng, Yuxiong He, Feng Yan HIGHLIGHT: Nevertheless, there is no SR model or machine learning system that supports adaptive SR, and enabling adaptive SR model on mobile devices is challenging because adapting model can cause significant framerate drop or even service interruption. To address this challenge, we take an algorithm and system co-design approach and propose DySR that maintains QoS while maximizing the model performance. A CMDP-within-online framework for Meta-Safe Reinforcement Learning 21. TITLE: https://openreview.net/forum?id=mbxz9Cjehr AUTHORS: Vanshaj Khattar, Yuhao Ding, Bilgehan Sel, Javad Lavaei, Ming Jin HIGHLIGHT: In this paper, we study the problem of meta-safe reinforcement learning (meta-SRL) through the CMDP-withinonline framework. 22, TITLE: Mastering the Game of No-Press Diplomacy via Human-Regularized Reinforcement Learning and Planning https://openreview.net/forum?id=F61FwJTZhb Anton Bakhtin, David J Wu, Adam Lerer, Jonathan Gray, Athul Paul Jacob, Gabriele Farina, Alexander H AUTHORS: Miller, Noam Brown HIGHLIGHT: We address this shortcoming by first introducing a planning algorithm we call DiL-piKL that regularizes a reward-maximizing policy toward a human imitation-learned policy. We prove that this is a no-regret learning algorithm under a modified utility function.

23. TITLE: Pruning Deep Neural Networks from a Sparsity Perspective https://openreview.net/forum?id=i-DleYh34BM AUTHORS: Enmao Diao, Ganghua Wang, Jiawei Zhang, Yuhong Yang, Jie Ding, Vahid Tarokh HIGHLIGHT: However, existing approaches lack a quantifiable measure to estimate the compressibility of a sub-network during each pruning iteration and thus may under-prune or over-prune the model. In this work, we propose PQ Index (PQI) to measure the potential compressibility of deep neural networks and use this to develop a Sparsity-informed Adaptive Pruning (SAP) algorithm. 24, TITLE: Everybody Needs Good Neighbours: An Unsupervised Locality-based Method for Bias Mitigation

https://openreview.net/forum?id=pOnhudsvzR

AUTHORS: Xudong Han, Timothy Baldwin, Trevor Cohn

To this end, we propose a new meta-algorithm for debiasing representation learning models, which combines HIGHLIGHT: the notions of data locality and accuracy of model fit, such that a supervised debiasing method can optimise fairness between neighbourhoods of poorly vs. well modelled instances as identified by our method.

25. TITLE: Spacetime Representation Learning

https://openreview.net/forum?id=qV_M_rhYajc

AUTHORS: Marc T. Law, James Lucas

HIGHLIGHT: In this work, we introduce a general family of representations for directed graphs through connected timeoriented Lorentz manifolds, called "spacetimes" in general relativity.

Quasi-optimal Learning with Continuous Treatments 26. TITLE:

https://openreview.net/forum?id=O8Vc52xFSUR

AUTHORS: Yuhan Li, Wenzhuo Zhou, Ruoging Zhu

HIGHLIGHT: Hence, it is important to induce a policy class whose support only contains near-optimal actions, and shrink the action-searching area for effectiveness and reliability. To achieve this, we develop a novel \emph{quasi-optimal learning algorithm}, which can be easily optimized in off-policy settings with guaranteed convergence under general function approximations.

Learning to Extrapolate: A Transductive Approach 27. TITLE: https://openreview.net/forum?id=lid14UkLPd4 Aviv Netanyahu, Abhishek Gupta, Max Simchowitz, Kaiqing Zhang, Pulkit Agrawal AUTHORS: HIGHLIGHT: In this work, we tackle the problem of developing machine learning systems that retain the power of overparametrized function approximators, while enabling extrapolation to out-of-support testing points when possible. 28, TITLE: Label-free Concept Bottleneck Models https://openreview.net/forum?id=FlCg47MNvBA

AUTHORS: Tuomas Oikarinen, Subhro Das, Lam M. Nguyen, Tsui-Wei Weng

HIGHLIGHT: This poor performance creates a barrier for adoption in practical real world applications. Motivated by these challenges, we propose \textit{Label-free} CBM which is a framework to transform any neural network into an interpretable CBM without labeled concept data, while retaining a high accuracy.

CLIP-Dissect: Automatic Description of Neuron Representations in Deep Vision Networks 29. TITLE: https://openreview.net/forum?id=iPWiwWHc1V AUTHORS: Tuomas Oikarinen, Tsui-Wei Weng HIGHLIGHT: In this paper, we propose CLIP-Dissect, a new technique to automatically describe the function of individual

hidden neurons inside vision networks.

30, TITLE: Predicting Cellular Responses with Variational Causal Inference and Refined Relational Information https://openreview.net/forum?id=ICYasJBlZNs

AUTHORS: Yulun Wu, Rob Barton, Zichen Wang, Vassilis N. Ioannidis, Carlo De Donno, Layne C Price, Luis F. Voloch, George Karypis

HIGHLIGHT: Predicting the responses of a cell under perturbations may bring important benefits to drug discovery and personalized therapeutics. In this work, we propose a novel graph variational Bayesian causal inference framework to predict a cell's gene expressions under counterfactual perturbations (perturbations that this cell did not factually receive), leveraging information representing biological knowledge in the form of gene regulatory networks (GRNs) to aid individualized cellular response predictions.

31. TITLE: Hard-Meta-Dataset++: Towards Understanding Few-Shot Performance on Difficult Tasks https://openreview.net/forum?id=wq0luyH3m4

AUTHORS: Samyadeep Basu, Megan Stanley, John F Bronskill, Soheil Feizi, Daniela Massiceti

HIGHLIGHT: This phenomenon has real-world consequences for deployed few-shot systems where safety and reliability are paramount, yet little has been done to understand these failure cases. In this paper, we study these difficult tasks to gain a more nuanced understanding of the limitations of current methods.

32, TITLE: Data Continuity Matters: Improving Sequence Modeling with Lipschitz Regularizer https://openreview.net/forum?id=27uBgHuoSQ

AUTHORS: Eric Qu, Xufang Luo, Dongsheng Li

However, few attempts have been made to understand the inherent data property of sequence data, neglecting HIGHLIGHT: the critical factor that may significantly affect the performance of sequence modeling. In this paper, we theoretically and empirically analyze a generic property of sequence data, i.e., continuity, and connect this property with the performance of deep models.

33, TITLE: Symbolic Physics Learner: Discovering governing equations via Monte Carlo tree search https://openreview.net/forum?id=ZTK3SefE8 Z

AUTHORS: Fangzheng Sun, Yang Liu, Jian-Xun Wang, Hao Sun

HIGHLIGHT: Distilling analytical expressions that govern nonlinear dynamics from limited data remains vital but challenging. To tackle this fundamental issue, we propose a novel Symbolic Physics Learner (SPL) machine to discover the mathematical structure of nonlinear dynamics.

34, TITLE: Neural Implicit Shape Editing using Boundary Sensitivity

https://openreview.net/forum?id=CMPIBjmhpo

AUTHORS: Arturs Berzins, Moritz Ibing, Leif Kobbelt

HIGHLIGHT: Compared to classic geometry representations, however, neural representations do not allow the user to exert intuitive control over the shape. Motivated by this, we leverage \emph{boundary sensitivity} to express how perturbations in parameters move the shape boundary.

35, TITLE: Understanding the Role of Nonlinearity in Training Dynamics of Contrastive Learning https://openreview.net/forum?id=s130rTE3U_X AUTHORS: Yuandong Tian

HIGHLIGHT: In this paper, we study the role of nonlinearity in the training dynamics of contrastive learning (CL) on one and two-layer nonlinear networks with homogeneous activation h(x) = h'(x)x.

36, TITLE: Towards Understanding and Mitigating Dimensional Collapse in Heterogeneous Federated Learning https://openreview.net/forum?id=EXnIyMVTL8s

AUTHORS: Yujun Shi, Jian Liang, Wenqing Zhang, Vincent Tan, Song Bai

HIGHLIGHT: In addition, we theoretically analyze the gradient flow dynamics to shed light on how data heterogeneity result in dimensional collapse for local models. To remedy this problem caused by the data heterogeneity, we propose FedDecorr, a novel method that can effectively mitigate dimensional collapse in federated learning.

37, TITLE: CodeGen: An Open Large Language Model for Code with Multi-Turn Program Synthesis https://openreview.net/forum?id=iaYcJKpY2B

AUTHORS: Erik Nijkamp, Bo Pang, Hiroaki Hayashi, Lifu Tu, Huan Wang, Yingbo Zhou, Silvio Savarese, Caiming Xiong HIGHLIGHT: The prevalence of large language models advances the state-of-the-art for program synthesis, though limited training resources and data impede open access to such models. To democratize this, we train and release a family of large language models up to 16.1B parameters, called CODEGEN, on natural language and programming language data, and open source the training library JAXFORMER.

38, TITLE: On the complexity of nonsmooth automatic differentiation

https://openreview.net/forum?id=uqg3FhRZaq

AUTHORS: Jerome Bolte, Ryan Boustany, Edouard Pauwels, Béatrice Pesquet-Popescu

HIGHLIGHT: Using the notion of conservative gradient, we provide a simple model to estimate the computational costs of the backward and forward modes of algorithmic differentiation for a wide class of nonsmooth programs.

 39, TITLE:
 M-L2O: Towards Generalizable Learning-to-Optimize by Test-Time Fast Self-Adaptation

 https://openreview.net/forum?id=s7oOe6cNRT8

 AUTHORS:
 Junjie Yang, Xuxi Chen, Tianlong Chen, Zhangyang Wang, Yingbin Liang

 HIGHLIGHT:
 However, such learned optimizers often struggle when new test problems come with a substantially deviation

 from the training task
 distribution. This paper investigates a potential solution to this open challenge, by meta-training an L2O

optimizer that can perform fast test-time self-adaptation to a out-of-distribution task, in only a few steps.

 40, TITLE:
 Benchmarking Deformable Object Manipulation with Differentiable Physics

 https://openreview.net/forum?id=1NAzMofMnWl

 AUTHORS:
 Siwei Chen, Cunjun Yu, Yiqing Xu, Linfeng Li, Xiao Ma, Zhongwen Xu, David Hsu

 HIGHLIGHT:
 In this work, we present DaXBench, a differentiable DOM benchmark with a wide object and task coverage.

41, TITLE: DiffMimic: Efficient Motion Mimicking with Differentiable Physics

https://openreview.net/forum?id=06mk-epSwZ

AUTHORS: Jiawei Ren, Cunjun Yu, Siwei Chen, Xiao Ma, Liang Pan, Ziwei Liu

HIGHLIGHT: In this work, we leverage differentiable physics simulators (DPS) and propose an efficient motion mimicking method dubbed \$\textbf{DiffMimic}\$.

 42, TITLE:
 Thalamus: a brain-inspired algorithm for biologically-plausible continual learning and disentangled

 representations
 https://openreview.net/forum?id=6orC5MvgPBK

 AUTHORS:
 Ali Hummos

HIGHLIGHT: Inspired by the brain thalamocortical circuit, we introduce a simple algorithm that uses optimization at inference time to generate internal representations of the current task dynamically.

43, TITLE: Adversarial Diversity in Hanabi

https://openreview.net/forum?id=uLE3WF3-H_5

AUTHORS: Brandon Cui, Andrei Lupu, Samuel Sokota, Hengyuan Hu, David J Wu, Jakob Nicolaus Foerster

HIGHLIGHT: In this work, we propose a novel approach to diverse policy generation for turn-based Dec-POMDPs with public actions, which relies on off-belief learning to encourage reasonableness and skill, and on ``repulsive" fictitious transitions to encourage diversity.

44, TITLE:	CogVideo: Large-scale Pretraining for Text-to-Video Generation via Transformers
https://openreview.ne	t/forum?id=rB6TpjAuSRy
AUTHORS:	Wenyi Hong, Ming Ding, Wendi Zheng, Xinghan Liu, Jie Tang
HIGHLIGHT:	In this work, we present CogVideo, a 9B-parameter transformer for text-to-video generation.
45, TITLE:	Interpretability in the Wild: a Circuit for Indirect Object Identification in GPT-2 Small
https://openreview.ne	t/forum?id=NpsVSN604ul
AUTHORS:	Kevin Ro Wang, Alexandre Variengien, Arthur Conmy, Buck Shlegeris, Jacob Steinhardt
HIGHLIGHT:	However, most previous work either focuses on simple behaviors in small models, or describes complicated
behaviors in larger m performs a natural lar	odels with broad strokes. In this work, we bridge this gap by presenting an explanation for how GPT-2 small nguage task that requires logical reasoning: indirect object identification (IOI).
46, TITLE:	Causal Reasoning in the Presence of Latent Confounders via Neural ADMG Learning
https://openreview.ne	tt/forum?id=dcN0CaXQhT
AUTHORS:	Matthew Ashman, Chao Ma, Agrin Hilmkil, Joel Jennings, Cheng Zhang
HIGHLIGHT:	We first show that the presence of latent confounding is identifiable under the assumptions of bow-free ADMGs
with nonlinear additiv	ve noise models. With this insight, we propose a novel neural causal model based on autoregressive flows.
47, TITLE:	Offline RL for Natural Language Generation with Implicit Language Q Learning
https://openreview.ne	t/forum?id=aBH_DydEvoH
AUTHORS:	Charlie Victor Snell, Ilya Kostrikov, Yi Su, Sherry Yang, Sergey Levine
HIGHLIGHT:	This issue can be addressed by finetuning such models via supervised learning on curated datasets, or via
reinforcement learnin	g. In this work, we propose a novel offline RL method, implicit language Q-learning (ILQL), designed for use on
language models, that	t combines both the flexible utility maximization framework of RL algorithms with the ability of supervised
learning to leverage p	previously collected data, as well as its simplicity and stability.
48, TITLE:	Mid-Vision Feedback for Convolutional Neural Networks
https://openreview.ne	t/forum?id=4oLK1_k71Tz
AUTHORS:	Michael Maynord, Eadom T Dessalene, Cornelia Fermuller, Yiannis Aloimonos
HIGHLIGHT:	We introduce a novel mechanism which modulates perception in Convolutional Neural Networks (CNNs) based
on high level categor	ical expectations: Mid-Vision Feedback (MVF).
49, TITLE:	HiViT: A Simpler and More Efficient Design of Hierarchical Vision Transformer
https://openreview.ne	t/forum?id=3F61-0-57SC
AUTHORS:	Xiaosong Zhang, Yunjie Tian, Lingxi Xie, Wei Huang, Qi Dai, Qixiang Ye, Qi Tian
HIGHLIGHT:	In this paper, we delve deep into the comparison between ViT and Swin, revealing that (i) the performance gain
of Swin is mainly bro	ught by a deepened backbone and relative positional encoding, (ii) the hierarchical design of Swin can be
simplified into hierard removed.	chical patch embedding (proposed in this work), and (iii) other designs such as shifted-window attentions can be
50. TITLE:	Generalizing and Decoupling Neural Collapse via Hyperspherical Uniformity Gap
https://openreview.ne	// forum?id=inU2auhGdNU
AUTHORS:	Weivang Liu, Longhui Yu, Adrian Weller, Bernhard Schölkopf
HIGHLIGHT:	Inspired by how NC characterizes the training target of neural networks, we decouple NC into two objectives:
minimal intra-class va	ariability and maximal inter-class separability. We then introduce the concept of hyperspherical uniformity
(which characterizes	the degree of uniformity on the unit hypersphere) as a unified framework to quantify these two objectives.

51, TITLE: Score-based Generative 3D Mesh Modeling

https://openreview.net/forum?id=0cpM2ApF9p6

AUTHORS: Zhen Liu, Yao Feng, Michael J. Black, Derek Nowrouzezahrai, Liam Paull, Weiyang Liu

HIGHLIGHT: Existing scalable methods for generating meshes typically rely on sub-optimal post-processing, and they tend to produce overly-smooth or noisy surfaces without fine-grained geometric details. To overcome these shortcomings, we take advantage of the regular graph structure of meshes and use a simple yet very effective generative modeling method to generate 3D meshes.

52, TITLE: ISAAC Newton: Input-based Approximate Curvature for Newton's Method

https://openreview.net/forum?id=0paCJSFW7j

AUTHORS: Felix Petersen, Tobias Sutter, Christian Borgelt, Dongsung Huh, Hilde Kuehne, Yuekai Sun, Oliver Deussen HIGHLIGHT: We present ISAAC (Input-baSed ApproximAte Curvature), a novel method that conditions the gradient using selected second-order information and has an asymptotically vanishing computational overhead, assuming a batch size smaller than the number of neurons.

53, TITLE: Language Models Can Teach Themselves to Program Better

https://openreview.net/forum?id=SaRj2ka1XZ3

AUTHORS: Patrick Haluptzok, Matthew Bowers, Adam Tauman Kalai

HIGHLIGHT: We show that it is possible for an LM to synthesize programming problems and solutions, which are filtered for correctness by a Python interpreter.

54, TITLE: Latent Bottlenecked Attentive Neural Processes https://openreview.net/forum?id=yIxtevizEA AUTHORS: Leo Feng, Hossein Hajimirsadeghi, Yoshua Bengio, Mohamed Osama Ahmed HIGHLIGHT: Conversely, existing sub-quadratic NP variants perform significantly worse than that of TNPs. Tackling this issue, we propose Latent Bottlenecked Attentive Neural Processes (LBANPs), a new computationally efficient sub-quadratic NP variant, that has a querying computational complexity independent of the number of context datapoints. 55, TITLE: Embed to Control Partially Observed Systems: Representation Learning with Provable Sample Efficiency https://openreview.net/forum?id=8oJHwb3Sgp Lingxiao Wang, Qi Cai, Zhuoran Yang, Zhaoran Wang AUTHORS: HIGHLIGHT: To this end, we propose a reinforcement learning algorithm named Embed to Control (ETC), which learns the representation at two levels while optimizing the policy. Learning Kernelized Contextual Bandits in a Distributed and Asynchronous Environment 56. TITLE: https://openreview.net/forum?id=-G1kjTFsSs AUTHORS: Chuanhao Li, Huazheng Wang, Mengdi Wang, Hongning Wang HIGHLIGHT: In this paper, in order to improve the robustness against delays and unavailability of clients that are common in practice, we propose the first asynchronous solution based on approximated kernel regression for distributed kernel bandit learning. 57, TITLE: GReTo: Remedying dynamic graph topology-task discordance via target homophily https://openreview.net/forum?id=8duT3mi 5n AUTHORS: Zhengyang Zhou, qihe huang, Gengyu Lin, Kuo Yang, LEI BAI, Yang Wang HIGHLIGHT: In this work, we revisit node-wise relationships and explore novel homophily measurements on dynamic graphs with both signs and distances, capturing multiple node-level spatial relations and temporal evolutions. 58. TITLE: DocPrompting: Generating Code by Retrieving the Docs https://openreview.net/forum?id=ZTCxT2t2Ru AUTHORS: Shuyan Zhou, Uri Alon, Frank F. Xu, Zhengbao Jiang, Graham Neubig HIGHLIGHT: In contrast, when human programmers use functions and libraries for the first time, they frequently refer to textual resources such as code manuals and documentation, to explore and understand the available functionality. Inspired by this observation, we introduce DocPrompting: a natural-language-to-code generation approach that explicitly leverages documentation by (1) retrieving the relevant documentation pieces given an NL intent, and (2) generating code based on the NL intent and the retrieved documentation. 59, TITLE: SWIFT: Rapid Decentralized Federated Learning via Wait-Free Model Communication https://openreview.net/forum?id=jh1nCir1R3d AUTHORS: Marco Bornstein, Tahseen Rabbani, Evan Z Wang, Amrit Bedi, Furong Huang HIGHLIGHT: In this work, we propose SWIFT: a novel wait-free decentralized FL algorithm that allows clients to conduct training at their own speed. RoPAWS: Robust Semi-supervised Representation Learning from Uncurated Data 60. TITLE: https://openreview.net/forum?id=G1H4NSATlr AUTHORS: Sangwoo Mo, Jong-Chyi Su, Chih-Yao Ma, Mido Assran, Ishan Misra, Licheng Yu, Sean Bell HIGHLIGHT: We propose RoPAWS, a robust extension of PAWS that can work with real-world unlabeled data. 61, TITLE: Interpretable Geometric Deep Learning via Learnable Randomness Injection https://openreview.net/forum?id=6u7mf9s2A9 AUTHORS: Siqi Miao, Yunan Luo, Mia Liu, Pan Li HIGHLIGHT: This work proposes a general mechanism based on \learnable randomness injection (LRI) that allows building inherently interpretable models with general GDL backbones. We also propose four scientific datasets in the domains of high energy physics and biochemistry to evaluate LRI. 62, TITLE: Machine Unlearning of Federated Clusters https://openreview.net/forum?id=VzwfoFyYDga Chao Pan, Jin Sima, Saurav Prakash, Vishal Rana, Olgica Milenkovic AUTHORS: HIGHLIGHT: This work proposes the first known unlearning mechanism for federated clustering with privacy criteria that support simple, provable, and efficient data removal at the client and server level. 63, TITLE: PerFedMask: Personalized Federated Learning with Optimized Masking Vectors https://openreview.net/forum?id=hxEIgUXLFF AUTHORS: Mehdi Setayesh, Xiaoxiao Li, Vincent W.S. Wong HIGHLIGHT: To this end, we propose a personalized FL algorithm with optimized masking vectors called PerFedMask. 64, TITLE: A Minimalist Dataset for Systematic Generalization of Perception, Syntax, and Semantics https://openreview.net/forum?id=kIPyTuEZuAK AUTHORS: Qing Li, Siyuan Huang, Yining Hong, Yixin Zhu, Ying Nian Wu, Song-Chun Zhu HIGHLIGHT: Inspired by humans' remarkable ability to master arithmetic and generalize to unseen problems, we present a new dataset, HINT, to study machines' capability of learning generalizable concepts at three levels: perception, syntax, and semantics.

65, TITLE: Conservative Bayesian Model-Based Value Expansion for Offline Policy Optimization https://openreview.net/forum?id=dNqxZgyjcYA AUTHORS: Jihwan Jeong, Xiaoyu Wang, Michael Gimelfarb, Hyunwoo Kim, Baher abdulhai, Scott Sanner HIGHLIGHT: To this end, we derive an elegant and simple methodology called conservative Bayesian model-based value expansion for offline policy optimization (CBOP), that trades off model-free and model-based estimates during the policy evaluation step according to their epistemic uncertainties, and facilitates conservatism by taking a lower bound on the Bayesian posterior value estimate. 66, TITLE: SAM as an Optimal Relaxation of Bayes https://openreview.net/forum?id=k4fevFqSQcX AUTHORS: Thomas Möllenhoff, Mohammad Emtiyaz Khan HIGHLIGHT: Here, we establish SAM as a relaxation of the Bayes objective where the expected negative-loss is replaced by the optimal convex lower bound, obtained by using the so-called Fenchel biconjugate. 67. TITLE: Masked Vision and Language Modeling for Multi-modal Representation Learning https://openreview.net/forum?id=ZhuXksSJYWn AUTHORS: Gukyeong Kwon, Zhaowei Cai, Avinash Ravichandran, Erhan Bas, Rahul Bhotika, Stefano Soatto HIGHLIGHT: In this paper, we study how to use masked signal modeling in vision and language (V+L) representation learning. 68, TITLE: Extreme Q-Learning: MaxEnt RL without Entropy https://openreview.net/forum?id=SJ0Lde3tRL AUTHORS: Divyansh Garg, Joey Hejna, Matthieu Geist, Stefano Ermon HIGHLIGHT: In this work, we introduce a new update rule for online and offline RL which directly models the maximal value using Extreme Value Theory (EVT) inspired by Economics. Direct Embedding of Temporal Network Edges via Time-Decayed Line Graphs 69, TITLE: https://openreview.net/forum?id=Qamz7Q Ta1k AUTHORS: Sudhanshu Chanpuriya, Ryan A. Rossi, Sungchul Kim, Tong Yu, Jane Hoffswell, Nedim Lipka, Shunan Guo, Cameron N Musco First, time is assumed to be discretized, so if the time data is continuous, the user must determine the HIGHLIGHT: discretization and discard precise time information. Second, edge representations can only be calculated indirectly from the nodes, which may be suboptimal for tasks like edge classification. We present a simple method that avoids both shortcomings: construct the line graph of the network, which includes a node for each interaction, and weigh the edges of this graph based on the difference in time between interactions. Scaling Forward Gradient With Local Losses 70, TITLE: https://openreview.net/forum?id=JxpBP1JM15-Mengye Ren, Simon Kornblith, Renjie Liao, Geoffrey Hinton AUTHORS: HIGHLIGHT: In this paper, we propose to scale forward gradient by adding a large number of local greedy loss functions. 71. TITLE: Latent Variable Representation for Reinforcement Learning https://openreview.net/forum?id=mQpmZVzXK1h AUTHORS: Tongzheng Ren, Chenjun Xiao, Tianjun Zhang, Na Li, Zhaoran Wang, sujay sanghavi, Dale Schuurmans, Bo Dai HIGHLIGHT: In this paper, we provide a representation view of the latent variable models for state-action value functions, which allows both tractable variational learning algorithm and effective implementation of the optimism/pessimism principle in the face of uncertainty for exploration. 72, TITLE: Learning in temporally structured environments https://openreview.net/forum?id=z0 V5O9cmNw AUTHORS: Matt Jones, Tyler R. Scott, Mengye Ren, Gamaleldin Fathy Elsayed, Katherine Hermann, David Mayo, Michael Curtis Mozer HIGHLIGHT: This paper advances a multiscale learning model in which each weight in a neural network is decomposed into a sum of subweights learning independently with different learning and decay rates. 73. TITLE: Learning QUBO Forms in Quantum Annealing https://openreview.net/forum?id=isiQ5KIXbjj AUTHORS: Marcel Seelbach Benkner, Maximilian Krahn, Edith Tretschk, Zorah Lähner, Michael Moeller, Vladislav Golyanik HIGHLIGHT: Moreover, such explicit formulations impose tangible constraints on solution encodings. In stark contrast to prior work, this paper proposes to learn QUBO forms from data through gradient backpropagation instead of deriving them. 74, TITLE: The Generalized Eigenvalue Problem as a Nash Equilibrium https://openreview.net/forum?id=PEgBEB74JjB AUTHORS: Ian Gemp, Charlie Chen, Brian McWilliams In this work, we develop a game-theoretic formulation of the top-\$k\$ SGEP whose Nash equilibrium is the set HIGHLIGHT: of generalized eigenvectors.

75, TITLE: \$O(T^{-1})\$ Convergence of Optimistic-Follow-the-Regularized-Leader in Two-Player Zero-Sum Markov Games https://openreview.net/forum?id=VWqiPBB EM AUTHORS: Yuepeng Yang, Cong Ma HIGHLIGHT: We prove that the optimistic-follow-the-regularized-leader (OFTRL) algorithm, together with smooth value updates, finds an \$O(T^{?1})\$ approximate Nash equilibrium in \$T\$ iterations for two-player zero-sum Markov games with full information. 76, TITLE: Sample Complexity of Nonparametric Off-Policy Evaluation on Low-Dimensional Manifolds using Deep Networks https://openreview.net/forum?id=9x3CO0ZU9LR AUTHORS: Xiang Ji, Minshuo Chen, Mengdi Wang, Tuo Zhao HIGHLIGHT: We consider the off-policy evaluation problem of reinforcement learning using deep convolutional neural networks. 77. TITLE: Critic Sequential Monte Carlo https://openreview.net/forum?id=ObtGcyKmwna AUTHORS: Vasileios Lioutas, Jonathan Wilder Lavington, Justice Sefas, Matthew Niedoba, Yunpeng Liu, Berend Zwartsenberg, Setareh Dabiri, Frank Wood, Adam Scibior HIGHLIGHT: We introduce CriticSMC, a new algorithm for planning as inference built from a composition of sequential Monte Carlo with learned Soft-Q function heuristic factors. 78, TITLE: Basic Binary Convolution Unit for Binarized Image Restoration Network https://openreview.net/forum?id=h8T5dZWTZ-Z AUTHORS: Bin Xia, Yulun Zhang, Yitong Wang, Yapeng Tian, Wenming Yang, Radu Timofte, Luc Van Gool HIGHLIGHT: In this study, we reconsider components in binary convolution, such as residual connection, BatchNorm, activation function, and structure, for IR tasks. Knowledge Distillation based Degradation Estimation for Blind Super-Resolution 79, TITLE: https://openreview.net/forum?id=Fg3mYW8owg AUTHORS: Bin Xia, Yulun Zhang, Yitong Wang, Yapeng Tian, Wenming Yang, Radu Timofte, Luc Van Gool HIGHLIGHT: In this paper, we propose a Knowledge Distillation based Blind-SR network (KDSR). Spectral Decomposition Representation for Reinforcement Learning 80, TITLE: https://openreview.net/forum?id=FBMLeaXpZN Tongzheng Ren, Tianjun Zhang, Lisa Lee, Joseph E. Gonzalez, Dale Schuurmans, Bo Dai AUTHORS: HIGHLIGHT: However, current spectral methods suffer from limited applicability because they are constructed for state-only aggregation and are derived from a policy-dependent transition kernel, without considering the issue of exploration. To address these issues, we propose an alternative spectral method, Spectral Decomposition Representation (SPEDER), that extracts a state-action abstraction from the dynamics without inducing spurious dependence on the data collection policy, while also balancing the exploration-versus-exploitation trade-off during learning. 81, TITLE: Fake It Until You Make It : Towards Accurate Near-Distribution Novelty Detection https://openreview.net/forum?id=QWQM0ZwZdRS AUTHORS: Hossein Mirzaei, Mohammadreza Salehi, Sajjad Shahabi, Efstratios Gavves, Cees G. M. Snoek, Mohammad Sabokrou, Mohammad Hossein Rohban HIGHLIGHT: We first demonstrate existing methods could experience up to 20\% decrease in their AUCs in the neardistribution setting. Next, we propose to exploit a score-based generative model to produce synthetic near-distribution anomalous data. Our model is then fine-tuned to distinguish such data from the normal samples. We make quantitative as well as qualitative evaluation of this strategy, and compare the results with a variety of GAN-based models. 82, TITLE: Contextual Image Masking Modeling via Synergized Contrasting without View Augmentation for Faster and Better Visual Pretraining https://openreview.net/forum?id=A3sgyt4HWp AUTHORS: Shaofeng Zhang, Feng Zhu, Rui Zhao, Junchi Yan HIGHLIGHT: We propose a new contextual masking image modeling (MIM) approach called contrasting-aided contextual MIM (ccMIM), under the MIM paradigm for visual pretraining. 83, TITLE: Patch-Level Contrasting without Patch Correspondence for Accurate and Dense Contrastive Representation Learning https://openreview.net/forum?id=10R bcjFwJ AUTHORS: Shaofeng Zhang, Feng Zhu, Rui Zhao, Junchi Yan HIGHLIGHT: We propose ADCLR: $\$ and $\$ and $\$ and $\$ best of the set of \underline {R}epresentation \underline {L}earning, a novel self-supervised learning framework for learning accurate and dense vision representation. 84, TITLE: A Learning Based Hypothesis Test for Harmful Covariate Shift https://openreview.net/forum?id=rdfgqiwz71Z

AUTHORS: Tom Ginsberg, Zhongyuan Liang, Rahul G Krishnan

HIGHLIGHT: In this work, we define harmful covariate shift (HCS) as a change in distribution that may weaken the generalization of a predictive model.

85, TITLE: Backpropagation at the Infinitesimal Inference Limit of Energy-Based Models: Unifying Predictive Coding, Equilibrium Propagation, and Contrastive Hebbian Learning

https://openreview.net/forum?id=nIMifqu2EO

AUTHORS: Beren Millidge, Yuhang Song, Tommaso Salvatori, Thomas Lukasiewicz, Rafal Bogacz

HIGHLIGHT: Here, we provide a comprehensive theory of the conditions under which EBMs can approximate BP, which lets us unify many of the BP approximation results in the literature (namely, predictive coding, equilibrium propagation, and contrastive Hebbian learning) and demonstrate that their approximation to BP arises from a simple and general mathematical property of EBMs at free-phase equilibrium.

86, TITLE: Re-Imagen: Retrieval-Augmented Text-to-Image Generator

https://openreview.net/forum?id=XSEBx0iSjFQ

AUTHORS: Wenhu Chen, Hexiang Hu, Chitwan Saharia, William W. Cohen

HIGHLIGHT: Though state-of-the-art models can generate high-quality images of common entities, they often have difficulty generating images of uncommon entities, such as 'Chortai (dog)' or 'Picarones (food)'. To tackle this issue, we present the Retrieval-Augmented Text-to-Image Generator (Re-Imagen), a generative model that uses retrieved information to produce high-fidelity and faithful images, even for rare or unseen entities.

87, TITLE: Task-customized Masked Autoencoder via Mixture of Cluster-conditional Experts https://openreview.net/forum?id=j8IiQUM33s

AUTHORS: Zhili LIU, Kai Chen, Jianhua Han, Lanqing HONG, Hang Xu, Zhenguo Li, James Kwok HIGHLIGHT: However, when the various downstream tasks have data distributions different from the pre-training data, the semantically irrelevant pre-training information might result in negative transfer, impeding MAE's scalability. To address this issue, we propose a novel MAE based pre-training paradigm, named Mixture of Cluster-conditional Experts (MoCE), which can be trained once but provide customized pre-training models for diverse downstream tasks.

 88, TITLE:
 A Theoretical Framework for Inference and Learning in Predictive Coding Networks

 https://openreview.net/forum?id=ZCTvSF_uVM4

 AUTHORS:
 Beren Millidge, Yuhang Song, Tommaso Salvatori, Thomas Lukasiewicz, Rafal Bogacz

HIGHLIGHT: In this paper, we provide a comprehensive theoretical analysis of the properties of PCNs trained with prospective configuration.

89, TITLE: Learning to Grow Pretrained Models for Efficient Transformer Training https://openreview.net/forum?id=cDYRS5iZ16f

AUTHORS: Peihao Wang, Rameswar Panda, Lucas Torroba Hennigen, Philip Greengard, Leonid Karlinsky, Rogerio Feris, David Daniel Cox, Zhangyang Wang, Yoon Kim

HIGHLIGHT: This paper describes an approach for accelerating transformer training by learning to grow pretrained transformers, where we learn to linearly map the parameters of the smaller model to initialize the larger model.

90, TITLE: A Call to Reflect on Evaluation Practices for Failure Detection in Image Classification https://openreview.net/forum?id=YnkGMIh0gvX

AUTHORS: Paul F Jaeger, Carsten Tim Lüth, Lukas Klein, Till J. Bungert

HIGHLIGHT: Curiously, while these approaches all state to address the same eventual goal of detecting failures of a classifier upon real-life application, they currently constitute largely separated research fields with individual evaluation protocols, which either exclude a substantial part of relevant methods or ignore large parts of relevant failure sources. In this work, we systematically reveal current pitfalls caused by these inconsistencies and derive requirements for a holistic and realistic evaluation of failure detection.

91, TITLE: Generating Intuitive Fairness Specifications for Natural Language Processing

https://openreview.net/forum?id=N_g8TT9Cy7f

AUTHORS: Florian E. Dorner, Momchil Peychev, Nikola Konstantinov, Naman Goel, Elliott Ash, Martin Vechev HIGHLIGHT: While existing research has started to address this gap, current methods are based on hardcoded word replacements, resulting in specifications with limited expressivity or ones that fail to fully align with human intuition (e.g., in cases of asymmetric counterfactuals). This work proposes novel methods for bridging this gap by discovering expressive and intuitive individual fairness specifications.

92, TITLE: PiFold: Toward effective and efficient protein inverse folding

https://openreview.net/forum?id=oMsN9TYwJ0j

AUTHORS: Zhangyang Gao, Cheng Tan, Stan Z. Li

HIGHLIGHT: Structure-based protein design has attracted increasing attention in recent years; however, few methods can simultaneously improve the accuracy and efficiency due to the lack of expressive features and autoregressive sequence decoder. To address these issues, we propose PiFold, which contains a novel residue featurizer and PiGNN layers to generate protein sequences in a one-shot way with improved recovery.

93, TITLE: Contrastive Learning Can Find An Optimal Basis For Approximately View-Invariant Functions https://openreview.net/forum?id=AjC0KBjiMu

AUTHORS: Daniel D. Johnson, Ayoub El Hanchi, Chris J. Maddison

HIGHLIGHT: We give generalization bounds for downstream linear prediction using our kernel PCA representation, and show empirically on a set of synthetic tasks that applying kernel PCA to contrastive learning models can indeed approximately recover the Markov chain eigenfunctions, although the accuracy depends on the kernel parameterization as well as on the augmentation strength.

94. TITLE: Provably Auditing Ordinary Least Squares in Low Dimensions https://openreview.net/forum?id=DlpCotqdTy AUTHORS: Ankur Moitra, Dhruv Rohatgi HIGHLIGHT: We develop provable and efficient algorithms for estimating stability of OLS to dropping samples in the lowdimensional regime. 95, TITLE: Learning Sparse Group Models Through Boolean Relaxation https://openreview.net/forum?id=Do9MOlwWHu0 AUTHORS: Yijie Wang, Yuan Zhou, Xiaoqing Huang, Kun Huang, Jie Zhang, Jianzhu Ma HIGHLIGHT: We introduce an efficient algorithmic framework for learning sparse group models formulated as the natural convex relaxation of a cardinality-constrained program with Boolean variables. 96, TITLE: QAID: Question Answering Inspired Few-shot Intent Detection https://openreview.net/forum?id=gNI4 85Cyve AUTHORS: Asaf Yehudai, Matan Vetzler, Yosi Mass, Koren Lazar, Doron Cohen, Boaz Carmeli Our method achieve SOTA results on few-shot intent detection by combining Question-Answering architecture, HIGHLIGHT: Contrastive Learning techniques and use of the intent name as answer. 97, TITLE: Out-of-distribution Representation Learning for Time Series Classification https://openreview.net/forum?id=gUZWOE42160 AUTHORS: Wang Lu, Jindong Wang, Xinwei Sun, Yiqiang Chen, Xing Xie HIGHLIGHT: In this paper, we propose to view time series classification from the distribution perspective. 98, TITLE: Neural DAG Scheduling via One-Shot Priority Sampling https://openreview.net/forum?id=WL8FlAugqQ AUTHORS: Wonseok Jeon, Mukul Gagrani, Burak Bartan, Weiliang Will Zeng, Harris Teague, Piero Zappi, Christopher Lott HIGHLIGHT: We consider the problem of scheduling operations/nodes, the dependency among which is characterized by a Directed Acyclic Graph (DAG). **99. TITLE:** Efficiently Computing Nash Equilibria in Adversarial Team Markov Games https://openreview.net/forum?id=mjzm6btqgV AUTHORS: Fivos Kalogiannis, Ioannis Anagnostides, Ioannis Panageas, Emmanouil-Vasileios Vlatakis-Gkaragkounis, Vaggos Chatziafratis, Stelios Andrew Stavroulakis HIGHLIGHT: In this work, we depart from those prior results by investigating infinite-horizon \emph{adversarial team Markov games}, a natural and well-motivated class of games in which a team of identically-interested players---in the absence of any explicit coordination or communication --- is competing against an adversarial player. 100. TITLE: Graph Neural Network-Inspired Kernels for Gaussian Processes in Semi-Supervised Learning https://openreview.net/forum?id=flap0Bo6TK AUTHORS: Zehao Niu, Mihai Anitescu, Jie Chen HIGHLIGHT: Their competitive performance is often attributed to a proper capturing of the graph inductive bias. In this work, we introduce this inductive bias into GPs to improve their predictive performance for graph-structured data. SeaFormer: Squeeze-enhanced Axial Transformer for Mobile Semantic Segmentation 101. TITLE: https://openreview.net/forum?id=-qg8MQNrxZw Qiang Wan, Jiachen Lu, Zilong Huang, Gang YU, Li Zhang AUTHORS: HIGHLIGHT: In this paper, we introduce a new method squeeze-enhanced Axial Transformer (SeaFormer) for mobile semantic segmentation. 102, TITLE: Differentiable Gaussianization Layers for Inverse Problems Regularized by Deep Generative Models https://openreview.net/forum?id=OXP9Ns0gnIq AUTHORS: Dongzhuo Li In such cases, deep generative models are ineffective in attaining high-fidelity solutions. To address this issue, HIGHLIGHT: we propose to reparameterize and Gaussianize the latent tensors using novel differentiable data-dependent layers wherein custom operators are defined by solving optimization problems. 103, TITLE: Approximate Vanishing Ideal Computations at Scale https://openreview.net/forum?id=3ZPESALKXO AUTHORS: Elias Samuel Wirth, Hiroshi Kera, Sebastian Pokutta HIGHLIGHT: In this paper, we scale up the Oracle Approximate Vanishing Ideal algorithm (OAVI), the only generatorconstructing algorithm with known learning guarantees. 104, TITLE: SoftMatch: Addressing the Quantity-Quality Tradeoff in Semi-supervised Learning https://openreview.net/forum?id=ymt1zQXBDiF

AUTHORS: Hao Chen, Ran Tao, Yue Fan, Yidong Wang, Marios Savvides, Jindong Wang, Bhiksha Raj, Xing Xie, Bernt Schiele HIGHLIGHT: To this end, we propose SoftMatch to overcome the trade-off by maintaining both high quantity and high quality of pseudo-labels during training, effectively exploiting the unlabeled data. 105. TITLE: Learning Uncertainty for Unknown Domains with Zero-Target-Assumption https://openreview.net/forum?id=pWVASryOyFw AUTHORS: Yu Yu, Hassan Sajjad, Jia Xu HIGHLIGHT: We introduce our Maximum-Entropy Rewarded Reinforcement Learning (MERRL) framework that selects training data for more accurate Natural Language Processing (NLP). Scalable and Equivariant Spherical CNNs by Discrete-Continuous (DISCO) Convolutions 106, TITLE: https://openreview.net/forum?id=eb cpjZZ3GH Jeremy Ocampo, Matthew Alexander Price, Jason McEwen AUTHORS: HIGHLIGHT: We develop a hybrid discrete-continuous (DISCO) group convolution that is simultaneously equivariant and computationally scalable to high-resolution. 107, TITLE: FreeMatch: Self-adaptive Thresholding for Semi-supervised Learning https://openreview.net/forum?id=PDrUPTXJI_A Yidong Wang, Hao Chen, Qiang Heng, Wenxin Hou, Yue Fan, Zhen Wu, Jindong Wang, Marios Savvides, AUTHORS: Takahiro Shinozaki, Bhiksha Raj, Bernt Schiele, Xing Xie HIGHLIGHT: Based on the analysis, we hence propose FreeMatch to define and adjust the confidence threshold in a selfadaptive manner according to the model's learning status. 108, TITLE: Can discrete information extraction prompts generalize across language models? https://openreview.net/forum?id=sbWVtxq8-zE Nathanaël Carraz Rakotonirina, Roberto Dessi, Fabio Petroni, Sebastian Riedel, Marco Baroni AUTHORS: HIGHLIGHT: We introduce a way to induce prompts by mixing language models at training time that results in prompts that generalize well across models. 109, TITLE: Disentangling the Mechanisms Behind Implicit Regularization in SGD https://openreview.net/forum?id=LE5LxBgjB4V AUTHORS: Zachary Novack, Simran Kaur, Tanya Marwah, Saurabh Garg, Zachary Chase Lipton HIGHLIGHT: In this paper, we conduct an extensive empirical evaluation, focusing on the ability of various theorized mechanisms to close the small-to-large batch generalization gap. Transformer-based World Models Are Happy With 100k Interactions 110, TITLE: https://openreview.net/forum?id=TdBaDGCpjly AUTHORS: Jan Robine, Marc Höftmann, Tobias Uelwer, Stefan Harmeling HIGHLIGHT: To build a sample-efficient world model, we apply a transformer to real-world episodes in an autoregressive manner: not only the compact latent states and the taken actions but also the experienced or predicted rewards are fed into the transformer, so that it can attend flexibly to all three modalities at different time steps. Discrete Contrastive Diffusion for Cross-Modal Music and Image Generation 111, TITLE: https://openreview.net/forum?id=1-MBdJssZ-S AUTHORS: Ye Zhu, Yu Wu, Kyle Olszewski, Jian Ren, Sergey Tulyakov, Yan Yan HIGHLIGHT: In this work, we take a different route---we explicitly enhance input-output connections by maximizing their mutual information. 112, TITLE: Confidential-PROFITT: Confidential PROof of FaIr Training of Trees https://openreview.net/forum?id=iIfDQVyuFD AUTHORS: Ali Shahin Shamsabadi, Sierra Calanda Wyllie, Nicholas Franzese, Natalie Dullerud, Sébastien Gambs, Nicolas Papernot, Xiao Wang, Adrian Weller HIGHLIGHT: We introduce a method to provide a confidential proof of fairness for training, in the context of widely used decision trees, which we term Confidential-PROFITT. 113. TITLE: DCI-ES: An Extended Disentanglement Framework with Connections to Identifiability https://openreview.net/forum?id=462z-gLgSht AUTHORS: Cian Eastwood, Andrei Liviu Nicolicioiu, Julius Von Kügelgen, Armin Keki?, Frederik Träuble, Andrea Dittadi, Bernhard Schölkopf HIGHLIGHT: Our main idea is that the functional capacity required to use a representation is an important but thus-far neglected aspect of representation quality, which we quantify using explicitness or ease-of-use (E). 114, TITLE: Bort: Towards Explainable Neural Networks with Bounded Orthogonal Constraint https://openreview.net/forum?id=My57qBufZWs AUTHORS: Borui Zhang, Wenzhao Zheng, Jie Zhou, Jiwen Lu However, existing methods rely on intuitive assumptions and lack mathematical guarantees. To bridge this gap, HIGHLIGHT: we introduce Bort, an optimizer for improving model explainability with boundedness and orthogonality constraints on model parameters, derived from the sufficient conditions of model comprehensibility and transparency.

115, TITLE: Faster federated optimization under second-order similarity

https://openreview.net/forum?id=ElC6LYO4MfD

AUTHORS: Ahmed Khaled, Chi Jin

HIGHLIGHT: We consider finite-sum federated optimization under a second-order function similarity condition and strong convexity, and propose two new algorithms: SVRP and Catalyzed SVRP.

The Augmented Image Prior: Distilling 1000 Classes by Extrapolating from a Single Image 116, TITLE:

https://openreview.net/forum?id=6kxApT2r2i

AUTHORS: Yuki M Asano, Aaqib Saeed

While any image obviously cannot contain the multitudes of all existing objects, scenes and lighting conditions HIGHLIGHT: -- within the space of all \$256^{3\cdot224\cdot224}\$ possible \$224\$-sized square images, it might still provide a strong prior for natural images. To analyze this ``augmented image prior" hypothesis, we develop a simple framework for training neural networks from scratch using a single image and augmentations using knowledge distillation from a supervised pretrained teacher.

117. TITLE: Self-Supervised Category-Level Articulated Object Pose Estimation with Part-Level SE(3) Equivariance https://openreview.net/forum?id=20GtJ6hIaPA

AUTHORS: Xueyi Liu, Ji Zhang, Ruizhen Hu, Haibin Huang, He Wang, Li Yi

HIGHLIGHT: Our key idea is to factorize canonical shapes and articulated object poses from input articulated shapes through part-level equivariant shape analysis.

118, TITLE: Schema Inference for Interpretable Image Classification

https://openreview.net/forum?id=VGI9dSmTgPF

AUTHORS: Haofei Zhang, Xiaokang Liu, Mengqi Xue, Kaixuan Chen, Jie Song, Mingli Song HIGHLIGHT: In this paper, we study a novel inference paradigm, termed as schema inference, that learns to deductively infer the explainable predictions by rebuilding the prior deep neural network (DNN) forwarding scheme, guided by the prevalent philosophical cognitive concept of schema.

119, TITLE: Autoencoders as Cross-Modal Teachers: Can Pretrained 2D Image Transformers Help 3D Representation Learning?

https://openreview.net/forum?id=8Oun8ZUVe8N

AUTHORS: Runpei Dong, Zekun Qi, Linfeng Zhang, Junbo Zhang, Jianjian Sun, Zheng Ge, Li Yi, Kaisheng Ma HIGHLIGHT: In this paper, we revisit masked modeling in a unified fashion of knowledge distillation, and we show that foundational Transformers pretrained with 2D images or natural languages can help self-supervised 3D representation learning through training Autoencoders as Cross-Modal Teachers (ACT).

Partially Observable RL with B-Stability: Unified Structural Condition and Sharp Sample-Efficient Algorithms 120, TITLE: https://openreview.net/forum?id=n05upKp02kQ

AUTHORS: Fan Chen, Yu Bai, Song Mei

HIGHLIGHT: However, this line of research is still in its infancy, where (1) unified structural conditions enabling sampleefficient learning are lacking; (2) existing sample complexities for known tractable subclasses are far from sharp; and (3) fewer sample-efficient algorithms are available than in fully observable RL. This paper advances all three aspects above for Partially Observable RL in the general setting of Predictive State Representations (PSRs).

Towards Lightweight, Model-Agnostic and Diversity-Aware Active Anomaly Detection 121, TITLE: https://openreview.net/forum?id=-vKlt84fHs

AUTHORS: Xu Zhang, Yuan Zhao, Ziang Cui, Liqun Li, Shilin He, Qingwei Lin, Yingnong Dang, Saravan Rajmohan, Dongmei Zhang

HIGHLIGHT: Besides, most existing AAD approaches are specially tailored for a certain unsupervised detector, making it difficult to extend to other detection models. To tackle these problems, we propose a lightweight, model-agnostic and diversity-aware AAD method, named LMADA.

122, TITLE: Complexity-Based Prompting for Multi-step Reasoning

https://openreview.net/forum?id=yf1icZHC-19

AUTHORS: Yao Fu, Hao Peng, Ashish Sabharwal, Peter Clark, Tushar Khot HIGHLIGHT: In this work, we propose complexity-based prompting, a simple and effective example selection scheme for multi-step reasoning.

123, TITLE: Gromov-Wasserstein Autoencoders

https://openreview.net/forum?id=sbS10BCtc7

AUTHORS: Nao Nakagawa, Ren Togo, Takahiro Ogawa, Miki Haseyama

HIGHLIGHT: In this paper, we propose a novel representation learning method, Gromov-Wasserstein Autoencoders (GWAE), which directly matches the latent and data distributions using the variational autoencoding scheme.

124, TITLE: Moving Forward by Moving Backward: Embedding Action Impact over Action Semantics https://openreview.net/forum?id=vmjctNUSWI AUTHORS:

Kuo-Hao Zeng, Luca Weihs, Roozbeh Mottaghi, Ali Farhadi

HIGHLIGHT: Instead of relying that the impact of an action stably reflects its pre-defined semantic meaning, we propose to model the impact of actions on-the-fly using latent embeddings.

125, TITLE: Decomposed Prompting: A Modular Approach for Solving Complex Tasks

https://openreview.net/forum?id= nGgzQjzaRy AUTHORS: Tushar Khot, Harsh Trivedi, Matthew Finlayson, Yao Fu, Kyle Richardson, Peter Clark, Ashish Sabharwal HIGHLIGHT: However, this approach struggles as the task complexity increases or when the individual reasoning steps of the task themselves are hard to learn, especially when embedded in more complex tasks. To address this, we propose Decomposed Prompting, a new approach to solve complex tasks by decomposing them (via prompting) into simpler sub-tasks that can be delegated to a library of prompting-based LLMs dedicated to these sub-tasks. 126. TITLE: UNICORN: A Unified Backdoor Trigger Inversion Framework https://openreview.net/forum?id=Mj7K4lglGyj AUTHORS: Zhenting Wang, Kai Mei, Juan Zhai, Shiqing Ma HIGHLIGHT: This work formally defines and analyzes the trigger and the inversion problem. How gradient estimator variance and bias impact learning in neural networks 127, TITLE: https://openreview.net/forum?id=EBC60mxBwyw AUTHORS: Arna Ghosh, Yuhan Helena Liu, Guillaume Lajoie, Konrad Kording, Blake Aaron Richards HIGHLIGHT: Here, we show that variance and bias can impair learning on the training data, but some degree of variance and bias in a gradient estimator can be beneficial for generalization. 128, TITLE: Sampling is as easy as learning the score: theory for diffusion models with minimal data assumptions https://openreview.net/forum?id=zyLVMgsZ0U AUTHORS: Sitan Chen, Sinho Chewi, Jerry Li, Yuanzhi Li, Adil Salim, Anru Zhang HIGHLIGHT: We provide theoretical convergence guarantees for score-based generative models (SGMs) such as denoising diffusion probabilistic models (DDPMs), which constitute the backbone of large-scale real-world generative models such as DALL\$\cdot\$E 2. 129, TITLE: Post-hoc Concept Bottleneck Models https://openreview.net/forum?id=nA5AZ8CEyow Mert Yuksekgonul, Maggie Wang, James Zou AUTHORS: HIGHLIGHT: Moreover, CBMs often do not match the accuracy of an unrestricted neural network, reducing the incentive to deploy them in practice. In this work, we address these limitations of CBMs by introducing Post-hoc Concept Bottleneck models (PCBMs). 130, TITLE: Is a Caption Worth a Thousand Images? A Study on Representation Learning https://openreview.net/forum?id=cYijsVZhb5 AUTHORS: Shibani Santurkar, Yann Dubois, Rohan Taori, Percy Liang, Tatsunori Hashimoto HIGHLIGHT: Motivated by our findings, we devise simple data and algorithmic interventions to improve the transfer performance of CLIP-style models. Continual Post-Training of Language Models 131, TITLE: https://openreview.net/forum?id=m GDIItaI3o AUTHORS: Zixuan Ke, Haowei Lin, Yijia Shao, Tatsuya Konishi, Gyuhak Kim, Bing Liu HIGHLIGHT: Existing research has shown that post-training or adapting an LM using an unlabeled topical/domain corpus can improve the end-task performance in the domain. This paper proposes a novel method to continually post-train an LM with a sequence of unlabeled domain corpora to adapt the LMto these domains to improve their end-task performances. 132, TITLE: Learning to Generate Columns with Application to Vertex Coloring https://openreview.net/forum?id=JHW30A4DXtO AUTHORS: Yuan Sun, Andreas T Ernst, Xiaodong Li, Jake Weiner HIGHLIGHT: We present a new column generation approach based on Machine Learning (ML) for solving combinatorial optimization problems. 133, TITLE: Constraining Representations Yields Models That Know What They Don't Know https://openreview.net/forum?id=1w Amtk67X AUTHORS: Joao Monteiro, Pau Rodriguez, Pierre-Andre Noel, Issam H. Laradji, David Vazquez HIGHLIGHT: Such unsafe behaviour is particularly frequent when the use case slightly differs from the training context, and/or in the presence of an adversary. This work presents a novel direction to address these issues in a broad, general manner: imposing class-aware constraints on a model's internal activation patterns. 134, TITLE: Temporal Domain Generalization with Drift-Aware Dynamic Neural Networks https://openreview.net/forum?id=sWOsRj4nT1n AUTHORS: Guangji Bai, Chen Ling, Liang Zhao To address them, we propose a Temporal Domain Generalization with Drift-Aware Dynamic Neural Network HIGHLIGHT: (DRAIN) framework. 135, TITLE: Causal Estimation for Text Data with (Apparent) Overlap Violations https://openreview.net/forum?id=Ha2MnQM9Ph AUTHORS: Lin Gui, Victor Veitch

HIGHLIGHT: The purpose of this paper is to show how to handle causal identification and obtain robust causal estimation in the presence of apparent overlap violations.

136, TITLE: A Simple Approach for Visual Room Rearrangement: 3D Mapping and Semantic Search https://openreview.net/forum?id=1C6nCCaRe6p AUTHORS: Brandon Trabucco, Gunnar A Sigurdsson, Robinson Piramuthu, Gaurav S. Sukhatme, Ruslan Salakhutdinov HIGHLIGHT: Visual room rearrangement evaluates an agent's ability to rearrange objects in a room to a desired goal based solely on visual input. We propose a simple yet effective method for this problem: (1) search for and map which objects need to be rearranged, and (2) rearrange each object until the task is complete. 137. TITLE: Improved Training of Physics-Informed Neural Networks Using Energy-Based Priors: a Study on Electrical Impedance Tomography https://openreview.net/forum?id=zqkfJA6R1-r AUTHORS: Akarsh Pokkunuru, Pedram Rooshenas, Thilo Strauss, Anuj Abhishek, Taufiquar Khan HIGHLIGHT: In this work, we propose a Bayesian approach through a data-driven energy-based model (EBM) as a prior, to improve the overall accuracy and quality of tomographic reconstruction. 138. TITLE: ESD: Expected Squared Difference as a Tuning-Free Trainable Calibration Measure https://openreview.net/forum?id=bHW9njOSON AUTHORS: Hee Suk Yoon, Joshua Tian Jin Tee, Eunseop Yoon, Sunjae Yoon, Gwangsu Kim, Yingzhen Li, Chang D. Yoo HIGHLIGHT: We propose a tuning-free calibration obejctive loss Expected Squared Difference (ESD), where we view the calibration error from the perspective of the squared difference between two expectations. 139, TITLE: An Extensible Multi-modal Multi-task Object Dataset with Materials https://openreview.net/forum?id=n70oyIIS4g AUTHORS: Trevor Scott Standley, Ruohan Gao, Dawn Chen, Jiajun Wu, Silvio Savarese HIGHLIGHT: We present EMMa, an Extensible, Multimodal dataset of Amazon product listings that contains rich Material annotations. Does Zero-Shot Reinforcement Learning Exist? 140, TITLE: https://openreview.net/forum?id=MYEap OcQI AUTHORS: Ahmed Touati, Jérémy Rapin, Yann Ollivier HIGHLIGHT: Strategies for approximate zero-shot RL have been suggested using successor features (SFs) (Borsa et al., 2018) or forward-backward (FB) representations (Touati & Ollivier, 2021), but testing has been limited. After clarifying the relationships between these schemes, we introduce improved losses and new SF models, and test the viability of zero-shot RL schemes systematically on tasks from the Unsupervised RL benchmark (Laskin et al., 2021). 141. TITLE: Self-Stabilization: The Implicit Bias of Gradient Descent at the Edge of Stability https://openreview.net/forum?id=nhKHA59gXz AUTHORS: Alex Damian, Eshaan Nichani, Jason D. Lee We demonstrate that, far from being chaotic, the dynamics of gradient descent at the edge of stability can be HIGHLIGHT: captured by a cubic Taylor expansion: as the iterates diverge in direction of the top eigenvector of the Hessian due to instability, the cubic term in the local Taylor expansion of the loss function causes the curvature to decrease until stability is restored. 142. TITLE: Interactive Portrait Harmonization https://openreview.net/forum?id=AP0iZoaRaS AUTHORS: Jeya Maria Jose Valanarasu, HE Zhang, Jianming Zhang, Yilin Wang, Zhe Lin, Jose Echevarria, Yinglan Ma, Zijun Wei, Kalyan Sunkavalli, Vishal Patel HIGHLIGHT: To enable flexible interaction between user and harmonization, we introduce interactive harmonization, a new setting where the harmonization is performed with respect to a selected region in the reference image instead of the entire background.Furthermore, we also introduce a new dataset carefully curated for validating portrait harmonization. 143, TITLE: STREET: A MULTI-TASK STRUCTURED REASONING AND EXPLANATION BENCHMARK https://openreview.net/forum?id=1C kSW1-k0 AUTHORS: Danilo Neves Ribeiro, Shen Wang, Xiaofei Ma, Henghui Zhu, Rui Dong, Deguang Kong, Juliette Burger, Anjelica Ramos, zhiheng huang, William Yang Wang, George Karypis, Bing Xiang, Dan Roth HIGHLIGHT: We introduce STREET, a unified multi-task and multi-domain natural language reasoning and explanation benchmark. 144. TITLE: Hierarchical Sliced Wasserstein Distance https://openreview.net/forum?id=CUOaVn6mYEj AUTHORS: Khai Nguyen, Tongzheng Ren, Huy Nguyen, Litu Rout, Tan Minh Nguyen, Nhat Ho HIGHLIGHT: Therefore, for applications where the number of supports is relatively small compared with the dimension, e.g., several deep learning applications where the mini-batch approaches are utilized, the complexities from matrix multiplication of Radon Transform become the main computational bottleneck. To address this issue, we propose to derive projections by linearly and randomly combining a smaller number of projections which are named bottleneck projections. 145, TITLE: Restricted Strong Convexity of Deep Learning Models with Smooth Activations https://openreview.net/forum?id=PINRbk7h01 AUTHORS: Arindam Banerjee, Pedro Cisneros, Libin Zhu, Misha Belkin

- AUTHORS. Armdain Danerjee, Pedro Usneros, Libin Znu, Misna Beikin
- HIGHLIGHT: We consider the problem of optimization of deep learning models with smooth activation functions.
- 146, TITLE: Koopman Neural Operator Forecaster for Time-series with Temporal Distributional Shifts

https://openreview.net/forum?id=kUmdmHxK5N AUTHORS: Rui Wang, Yihe Dong, Sercan O Arik, Rose Yu HIGHLIGHT: In this paper, we propose a novel deep sequence model based on the Koopman theory for time series forecasting: Koopman Neural Forecaster (KNF) that leverages DNNs to learn the linear Koopman space and the coefficients of chosen measurement functions. Simplifying Model-based RL: Learning Representations, Latent-space Models, and Policies with One Objective 147, TITLE: https://openreview.net/forum?id=MQcmfgRxf7a AUTHORS: Raj Ghugare, Homanga Bharadhwaj, Benjamin Eysenbach, Sergey Levine, Russ Salakhutdinov HIGHLIGHT: In this work, we propose a single objective which jointly optimizes a latent-space model and policy to achieve high returns while remaining self-consistent. Minimum Description Length Control 148, TITLE: https://openreview.net/forum?id=oX3tGygjW1q AUTHORS: Ted Moskovitz, Ta-Chu Kao, Maneesh Sahani, Matthew Botvinick HIGHLIGHT: We propose a novel framework for multitask reinforcement learning based on the minimum description length (MDL) principle. 149, TITLE: Decoupled Training for Long-Tailed Classification With Stochastic Representations https://openreview.net/forum?id=bcYZwYo-0t AUTHORS: Giung Nam, Sunguk Jang, Juho Lee HIGHLIGHT: We propose a novel classifier re-training algorithm for long-tailed classification. 150. TITLE: Where to Begin? Exploring the Impact of Pre-Training and Initialization in Federated https://openreview.net/forum?id=Mpa3tRJFBb AUTHORS: John Nguyen, Jianyu Wang, Kshitiz Malik, Maziar Sanjabi, Michael Rabbat We empirically study the impact of starting from a pre-trained model in federated learning using four standard HIGHLIGHT: federated learning benchmark datasets. Unsurprisingly, starting from a pre-trained model reduces the training time required to reach a target error rate and enables the training of more accurate models (up to 40%) than is possible when starting from random initialization. Surprisingly, we also find that starting federated learning from a pre-trained initialization reduces the effect of both data and system heterogeneity. 151, TITLE: Martingale Posterior Neural Processes https://openreview.net/forum?id=-9PVqZ-IR AUTHORS: Hyungi Lee, Eunggu Yun, Giung Nam, Edwin Fong, Juho Lee HIGHLIGHT: In this work, we take a different approach based on the martingale posterior, a recently developed alternative to Bayesian inference. 152, TITLE: BigVGAN: A Universal Neural Vocoder with Large-Scale Training https://openreview.net/forum?id=iTtGCMDEzS AUTHORS: Sang-gil Lee, Wei Ping, Boris Ginsburg, Bryan Catanzaro, Sungroh Yoon HIGHLIGHT: In this work, we present BigVGAN, a universal vocoder that generalizes well for various out-of-distribution (OOD) scenarios without fine-tuning. 153. TITLE: Progressive Voronoi Diagram Subdivision Enables Accurate Data-free Class-Incremental Learning https://openreview.net/forum?id=zJXg Wmob03 AUTHORS: Chunwei Ma, Zhanghexuan Ji, Ziyun Huang, Yan Shen, Mingchen Gao, Jinhui Xu HIGHLIGHT: In this paper, we present \emph{iVoro}, a novel framework derived from computational geometry. 154, TITLE: MEDICAL IMAGE UNDERSTANDING WITH PRETRAINED VISION LANGUAGE MODELS: A COMPREHENSIVE STUDY https://openreview.net/forum?id=txlWziuCE5W AUTHORS: Ziyuan Qin, Hua Hui Yi, Qicheng Lao, Kang Li HIGHLIGHT: This paper discuss about how to leverage the trending vision language model to transfer to the medical domain, showing exciting performance on zero-shot and few-shot learning tasks. 155, TITLE: Approximate Bayesian Inference with Stein Functional Variational Gradient Descent https://openreview.net/forum?id=a2-aoqmeYM4 AUTHORS: Tobias Pielok, Bernd Bischl, David Rügamer HIGHLIGHT: We propose a general-purpose variational algorithm that forms a natural analogue of Stein variational gradient descent (SVGD) in function space. 156, TITLE: When and why Vision-Language Models behave like Bags-of-Words, and what to do about it? https://openreview.net/forum?id=KRLUvxh8uaX Mert Yuksekgonul, Federico Bianchi, Pratyusha Kalluri, Dan Jurafsky, James Zou AUTHORS: We present the settings where state-of-the-art VLMs behave like bags-of-words---i.e. when they have poor HIGHLIGHT: relational understanding, can blunder when linking objects to their attributes, and demonstrate a severe lack of order sensitivity. Here, we create the Attribution, Relation, and Order (ARO) benchmark to systematically evaluate the ability of VLMs to understand different types of relationships, attributes, and order information.

157, TITLE: Causal Imitation Learning via Inverse Reinforcement Learning https://openreview.net/forum?id=B-z41MBL tH AUTHORS: Kangrui Ruan, Junzhe Zhang, Xuan Di, Elias Bareinboim HIGHLIGHT: This paper studies imitation learning through causal lenses and extends the analysis and tools developed for behavior cloning (Zhang, Kumor, Bareinboim, 2020) to inverse reinforcement learning. The Surprising Computational Power of Nondeterministic Stack RNNs 158, TITLE: https://openreview.net/forum?id=o58JtGDs6v AUTHORS: Brian DuSell, David Chiang HIGHLIGHT: Nondeterminism is needed for recognizing all CFLs (not just deterministic CFLs), but in this paper, we show that nondeterminism and the neural controller interact to produce two more unexpected abilities. 159, TITLE: Ollivier-Ricci Curvature for Hypergraphs: A Unified Framework https://openreview.net/forum?id=sPCKNl5qDps AUTHORS: Corinna Coupette, Sebastian Dalleiger, Bastian Rieck HIGHLIGHT: We develop ORCHID, a flexible framework generalizing Ollivier-Ricci curvature to hypergraphs, and prove that the resulting curvatures have favorable theoretical properties. 160, TITLE: Hyperbolic Self-paced Learning for Self-supervised Skeleton-based Action Representations https://openreview.net/forum?id=3Bh6sRPKS3J AUTHORS: Luca Franco, Paolo Mandica, Bharti Munjal, Fabio Galasso HIGHLIGHT: We propose a novel HYperbolic Self-Paced model (HYSP) for learning skeleton-based action representations. 161. TITLE: Offline Congestion Games: How Feedback Type Affects Data Coverage Requirement https://openreview.net/forum?id=PXVGer7hmJ Haozhe Jiang, Qiwen Cui, Zhihan Xiong, Maryam Fazel, Simon Shaolei Du AUTHORS: Starting from the facility-level (a.k.a., semi-bandit) feedback, we propose a novel one-unit deviation coverage HIGHLIGHT: condition and show a pessimism-type algorithm that can recover an approximate NE. 162. TITLE: Auto-Encoding Goodness of Fit https://openreview.net/forum?id=JjCAdMUlu9v AUTHORS: Aaron Palmer, Zhiyi Chi, Derek Aguiar, Jinbo Bi HIGHLIGHT: In this work, we develop the Goodness of Fit Autoencoder (GoFAE), which incorporates hypothesis tests at two levels. 163, TITLE: Sparse tree-based Initialization for Neural Networks https://openreview.net/forum?id=78xgBm6ckZr AUTHORS: Patrick Lutz, Ludovic Arnould, Claire Boyer, Erwan Scornet In this work, we propose a new sparse initialization technique for (potentially deep) multilayer perceptrons HIGHLIGHT: (MLP): we first train a tree-based procedure to detect feature interactions and use the resulting information to initialize the network, which is subsequently trained via standard stochastic gradient strategies. Efficient Conditionally Invariant Representation Learning 164, TITLE: https://openreview.net/forum?id=dJruFeSRym1 AUTHORS: Roman Pogodin, Namrata Deka, Yazhe Li, Danica J. Sutherland, Victor Veitch, Arthur Gretton HIGHLIGHT: We introduce the Conditional Independence Regression CovariancE (CIRCE), a measure of conditional independence for multivariate continuous-valued variables. 165, TITLE: Can BERT Refrain from Forgetting on Sequential Tasks? A Probing Study https://openreview.net/forum?id=UazgYBMS9-W AUTHORS: Mingxu Tao, Yansong Feng, Dongyan Zhao HIGHLIGHT: We investigate the models on two typical kinds of NLP tasks, text classification and extractive question answering. 166, TITLE: Layer Grafted Pre-training: Bridging Contrastive Learning And Masked Image Modeling For Better Representations https://openreview.net/forum?id=jwdqNwyREyh Ziyu Jiang, Yinpeng Chen, Mengchen Liu, Dongdong Chen, Xiyang Dai, Lu Yuan, Zicheng Liu, Zhangyang AUTHORS: Wang HIGHLIGHT: In this paper, we start by making the empirical observation that a naive joint optimization of CL and MIM losses leads to conflicting gradient directions - more severe as the layers go deeper. 167, TITLE: DreamFusion: Text-to-3D using 2D Diffusion https://openreview.net/forum?id=FjNys5c7VyY Ben Poole, Ajay Jain, Jonathan T. Barron, Ben Mildenhall AUTHORS: Adapting this approach to 3D synthesis would require large-scale datasets of labeled 3D or multiview data and HIGHLIGHT: efficient architectures for denoising 3D data, neither of which currently exist. In this work, we circumvent these limitations by using a pretrained 2D text-to-image diffusion model to perform text-to-3D synthesis.

168, TITLE: Learning Input-agnostic Manipulation Directions in StyleGAN with Text Guidance

https://openreview.net/forum?id=47B ctC4pJ AUTHORS: Yoonjeon Kim, Hyunsu Kim, Junho Kim, Yunjey Choi, Eunho Yang HIGHLIGHT: Indeed, we show that it fails to discover a large portion of manipulation directions that can be found by existing methods, which manually manipulates latent space without texts. To alleviate this issue, we propose a novel method that learns a Dictionary, whose entry corresponds to the representation of a single channel, by taking into account the manipulation effect coming from the interaction with multiple other channels. 169. TITLE: Effective passive membership inference attacks in federated learning against overparameterized models https://openreview.net/forum?id=QsCSLPP55Ku AUTHORS: Jiacheng Li, Ninghui Li, Bruno Ribeiro This work considers the challenge of performing membership inference attacks in a federated learning setting --HIGHLIGHT: -for image classification--- where an adversary can only observe the communication between the central node and a single client (a passive white-box attack). 170. TITLE: Joint Edge-Model Sparse Learning is Provably Efficient for Graph Neural Networks https://openreview.net/forum?id=4UldFtZ CVF AUTHORS: Shuai Zhang, Meng Wang, Pin-Yu Chen, Sijia Liu, Songtao Lu, Miao Liu HIGHLIGHT: Encouraged by the empirical success of sparse learners in accelerating GNN training, this paper characterizes the impact of graph sampling and neuron pruning on the sample complexity and convergence rate for a desirable test accuracy quantitatively 171, TITLE: Tier Balancing: Towards Dynamic Fairness over Underlying Causal Factors https://openreview.net/forum?id=SZdfz5k7cd1 AUTHORS: Zevu Tang, Yatong Chen, Yang Liu, Kun Zhang In this paper, through causal modeling with a directed acyclic graph (DAG) on the decision-distribution HIGHLIGHT: interplay, we investigate the possibility of achieving long-term fairness from a dynamic perspective. CoRTX: Contrastive Framework for Real-time Explanation 172, TITLE: https://openreview.net/forum?id=L2MUOUp0beo AUTHORS: Yu-Neng Chuang, Guanchu Wang, Fan Yang, Quan Zhou, Pushkar Tripathi, Xuanting Cai, Xia Hu HIGHLIGHT: In this work, we propose a COntrastive Real-Time eXplanation (CoRTX) framework to learn the explanationoriented representation and relieve the intensive dependence of explainer training on explanation labels. Anamnesic Neural Differential Equations with Orthogonal Polynomial Projections 173, TITLE: https://openreview.net/forum?id=xYWqSjBcGM1 AUTHORS: Edward De Brouwer, Rahul G Krishnan We propose PolyODE, a Neural ODE that models the latent continuous-time process as a projection onto a HIGHLIGHT: basis of orthogonal polynomials. 174, TITLE: Large Language Models are Human-Level Prompt Engineers https://openreview.net/forum?id=92gvk82DE-AUTHORS: Yongchao Zhou, Andrei Ioan Muresanu, Ziwen Han, Keiran Paster, Silviu Pitis, Harris Chan, Jimmy Ba HIGHLIGHT: Inspired by classical program synthesis and the human approach to prompt engineering, we propose Automatic Prompt Engineer (APE) for automatic instruction generation and selection. 175, TITLE: AutoTransfer: AutoML with Knowledge Transfer - An Application to Graph Neural Networks https://openreview.net/forum?id=y81ppNf vg AUTHORS: Kaidi Cao, Jiaxuan You, Jiaju Liu, Jure Leskovec HIGHLIGHT: Here we propose AutoTransfer, an AutoML solution that improves search efficiency by transferring the prior architectural design knowledge to the novel task of interest. 176, TITLE: Explaining RL Decisions with Trajectories https://openreview.net/forum?id=5Egggz1q575 Shripad Vilasrao Deshmukh, Arpan Dasgupta, Chirag Agarwal, Nan Jiang, Balaji Krishnamurthy, Georgios AUTHORS: Theocharous, Jayakumar Subramanian HIGHLIGHT: In the literature, the explanation is often provided by saliency attribution to the features of the RL agent's state. In this work, we propose a complementary approach to these explanations, particularly for offline RL, where we attribute the policy decisions of a trained RL agent to the trajectories encountered by it during training. On Representing Linear Programs by Graph Neural Networks 177, TITLE: https://openreview.net/forum?id=cP2QVK-uvgd Ziang Chen, Jialin Liu, Xinshang Wang, Wotao Yin AUTHORS: HIGHLIGHT: While the literature has reported encouraging numerical results, this paper establishes the theoretical foundation of applying GNNs to solving LPs. On Representing Mixed-Integer Linear Programs by Graph Neural Networks 178. TITLE: https://openreview.net/forum?id=4gc3MGZra1d AUTHORS: Ziang Chen, Jialin Liu, Xinshang Wang, Wotao Yin HIGHLIGHT: This work discovers a fundamental limitation: there exist feasible and infeasible MILPs that all GNNs will, however, treat equally, indicating GNN's lacking power to express general MILPs. Then, we show that, by restricting the MILPs to

unfoldable ones or by adding random features, there exist GNNs that can reliably predict MILP feasibility, optimal objective values, and optimal solutions up to prescribed precision.

179, TITLE:	Efficient Discrete Multi Marginal Optimal Transport Regularization
https://openreview.net/	/forum?id=R98ZfMt-jE
AUTHORS:	Ronak Mehta, Jeffery Kline, Vishnu Suresh Lokhande, Glenn Fung, Vikas Singh
HIGHLIGHT:	In this paper, we leverage multi-marginal optimal transport (MMOT), where we take advantage of a procedure
that computes a genera	ilized earth mover's distance as a sub-routine.
180, TITLE: https://openreview.net/ AUTHORS: HIGHLIGHT: processing.	Graph Signal Sampling for Inductive One-Bit Matrix Completion: a Closed-form Solution /forum?id=G_HSyfLk0m Chao Chen, Haoyu Geng, Gang Zeng, Zhaobing Han, Hua Chai, Xiaokang Yang, Junchi Yan We propose a unified graph signal sampling framework which enjoys the benefits of graph signal analysis and
181, TITLE:	A New Hierarchy of Expressivity for Graph Neural Networks
https://openreview.net/	/forum?id=5cAI0qXxyv
AUTHORS:	Qing Wang, Dillon Ze Chen, Asiri Wijesinghe, Shouheng Li, Muhammad Farhan
HIGHLIGHT:	In this work we show that, contrary to the widely accepted view, the \$k\$-WL hierarchy is not well-suited for
measuring expressive (GNNs.
182, TITLE:	On Achieving Optimal Adversarial Test Error
https://openreview.net/	/forum?id=fVm3nZMZs9
AUTHORS:	Justin D. Li, Matus Telgarsky
HIGHLIGHT:	We first elucidate various fundamental properties of optimal adversarial predictors: the structure of optimal
adversarial convex pre	dictors in terms of optimal adversarial zero-one predictors, bounds relating the adversarial convex loss to the
adversarial zero-one lo	oss, and the fact that continuous predictors can get arbitrarily close to the optimal adversarial error for both
convex and zero-one lo	osses. Applying these results along with new Rademacher complexity bounds for adversarial training near
initialization, we prove	e that for general data distributions and perturbation sets, adversarial training on shallow networks with early
stopping and an idealiz	zed optimal adversary is able to achieve optimal adversarial test error.
183, TITLE:	Powderworld: A Platform for Understanding Generalization via Rich Task Distributions
https://openreview.net/	/forum?id=AWZgXGmsbA
AUTHORS:	Kevin Frans, Phillip Isola
HIGHLIGHT:	To take a step towards addressing this research bottleneck, this work presents Powderworld, a lightweight yet
expressive simulation of	environment running directly on the GPU.
184, TITLE: https://openreview.net/ AUTHORS: HIGHLIGHT: segmentation model w scale, fine-annotated 3 scenes.	3D Segmenter: 3D Transformer based Semantic Segmentation via 2D Panoramic Distillation (forum?id=4dZeBJ83oxk ZHENNAN WU, YANG LI, Yifei Huang, Lin Gu, Tatsuya Harada, Hiroyuki Sato Therefore, in this work, we propose the first 2D-to-3D knowledge distillation strategy to enhance 3D semantic ith knowledge embedded in the latent space of powerful 2D models.To facilitate our research, we create a large- D semantic segmentation benchmark, containing voxel-wise semantic labels and aligned panoramas of 5175
185, TITLE:	Fairness and Accuracy under Domain Generalization
https://openreview.net/	/forum?id=jBEXnEMdNOL
AUTHORS:	Thai-Hoang Pham, Xueru Zhang, Ping Zhang
HIGHLIGHT:	In this paper, we study the transfer of both fairness and accuracy under domain generalization where the data at
test time may be sampl	led from never-before-seen domains.
186, TITLE: https://openreview.net/ AUTHORS: HIGHLIGHT: deterministic oracles.	Text Summarization with Oracle Expectation /forum?id=HehQobsr0S Yumo Xu, Mirella Lapata In this work, we identify two flaws with the widely used greedy labeling approach: it delivers suboptimal and
187, TITLE:	Efficient Attention via Control Variates
https://openreview.net/	/forum?id=G-uNfHKrj46
AUTHORS:	Lin Zheng, Jianbo Yuan, Chong Wang, Lingpeng Kong
HIGHLIGHT:	We present a novel analysis of random feature attention based on control variates, which characterizes its gap to
full softmax attention a	and induces a novel efficient variant that significantly improves the approximation while remaining efficient.
188, TITLE:	Pitfalls of Gaussians as a noise distribution in NCE
https://openreview.net/	/forum?id=ovZE0KsbM3S
AUTHORS:	Holden Lee, Chirag Pabbaraju, Anish Prasad Sevekari, Andrej Risteski
HIGHLIGHT:	In practice, a common choice for \$q\$ is a Gaussian which matches the mean and covariance of the data. In this
paper, we show that su	ich a choice can result in an exponentially bad (in the ambient dimension) conditioning of the Hessian of the
loss - even for very sin	nple data distributions.

189, TITLE: HiCLIP: Contrastive Language-Image Pretraining with Hierarchy-aware Attention https://openreview.net/forum?id=0eTTKOOOQkV AUTHORS: Shijie Geng, Jianbo Yuan, Yu Tian, Yuxiao Chen, Yongfeng Zhang HIGHLIGHT: As a result, such hierarchical aggregation significantly improves the cross-modal alignment. To demonstrate the advantages of HiCLIP, we conduct qualitative analysis on its unsupervised hierarchy induction during inference, as well as extensive quantitative experiments on both visual recognition and vision-language downstream tasks. 190. TITLE: Sparse MoE with Random Routing as the New Dropout: Training Bigger and Self-Scalable Models https://openreview.net/forum?id=w1hwFUb 81 AUTHORS: Tianlong Chen, Zhenyu Zhang, AJAY KUMAR JAISWAL, Shiwei Liu, Zhangyang Wang HIGHLIGHT: To this end, we propose a new plug-and-play training framework, \$\textbf{SMoE-Dropout}\$ to enable scaling transformers to better accuracy in the full capacity setting without collapse. MIMT: Masked Image Modeling Transformer for Video Compression 191, TITLE: https://openreview.net/forum?id=j9m-mVnndbm AUTHORS: Jinxi Xiang, Kuan Tian, Jun Zhang HIGHLIGHT: We thus introduce an entropy model based on a masked image modeling transformer (MIMT) to learn the spatial-temporal dependencies. Sparsity May Cry: Let Us Fail (Current) Sparse Neural Networks Together! 192. TITLE: https://openreview.net/forum?id=J6F31Lg4Kdp AUTHORS: Shiwei Liu, Tianlong Chen, Zhenyu Zhang, Xuxi Chen, Tianjin Huang, AJAY KUMAR JAISWAL, Zhangyang Wang HIGHLIGHT: In pursuit of a more general evaluation and unveiling the true potential of sparse algorithms, we introduce "Sparsity May Cry" Benchmark (SMC-Bench), a collection of carefully curated 4 diverse tasks with 12 datasets, that accounts for capturing a wide-range of domain-specific knowledge. 193, TITLE: Interpretable Debiasing of Vectorized Language Representations with Iterative Orthogonalization https://openreview.net/forum?id=TkQ1sxd9P4 AUTHORS: Prince Osei Aboagye, Yan Zheng, Jack Shunn, Chin-Chia Michael Yeh, Junpeng Wang, Zhongfang Zhuang, Huiyuan Chen, Liang Wang, Wei Zhang, Jeff Phillips HIGHLIGHT: We propose a new mechanism to augment a word vector embedding representation that offers improved bias removal while retaining the key information-resulting in improved interpretability of the representation. 194, TITLE: The Surprising Effectiveness of Equivariant Models in Domains with Latent Symmetry https://openreview.net/forum?id=P4MUGRM4Acu AUTHORS: Dian Wang, Jung Yeon Park, Neel Sortur, Lawson L.S. Wong, Robin Walters, Robert Platt HIGHLIGHT: This paper discovers that equivariant models are surprisingly effective in domains with latent or partial symmetries. 195, TITLE: Unsupervised Model Selection for Time Series Anomaly Detection https://openreview.net/forum?id=gOZ pKANaPW AUTHORS: Mononito Goswami, Cristian Ignacio Challu, Laurent Callot, Lenon Minorics, Andrey Kan HIGHLIGHT: This paper answers the question -- Given an unlabeled dataset and a set of candidate time series anomaly detectors, how can we select the most accurate model? 196, TITLE: GoBigger: A Scalable Platform for Cooperative-Competitive Multi-Agent Interactive Simulation https://openreview.net/forum?id=NnOZT CR26Z AUTHORS: Ming Zhang, Shenghan Zhang, Zhenjie Yang, Lekai Chen, Jinliang Zheng, Chao Yang, Chuming Li, Hang Zhou, Yazhe Niu, Yu Liu HIGHLIGHT: Even though this has greatly promoted the development of multi-agent reinforcement learning (MARL), it is still not enough to support further exploration on the behavior of swarm intelligence between multiple teams, and cooperation between multiple agents due to their limited scalability. To alleviate this, we introduce GoBigger, a scalable platform for cooperativecompetition multi-agent interactive simulation. 197, TITLE: Representation Learning for Low-rank General-sum Markov Games https://openreview.net/forum?id=8FrovnZv4C Chengzhuo Ni, Yuda Song, Xuezhou Zhang, Zihan Ding, Chi Jin, Mengdi Wang AUTHORS: HIGHLIGHT: We leverage representation learning and present a model-based and a model-free approach to construct an effective representation from collected data. 198, TITLE: Exploring Low-Rank Property in Multiple Instance Learning for Whole Slide Image Classification https://openreview.net/forum?id=01KmhBsEPFO AUTHORS: Jinxi Xiang, Jun Zhang We leverage the properties of the apparent similarity in high-resolution WSIs, which essentially exhibit HIGHLIGHT: \textit{low-rank} structures in the data manifold, to develop a novel MIL with a boost in both feature embedding and feature aggregation.

199, TITLE: Win: Weight-Decay-Integrated Nesterov Acceleration for Adaptive Gradient Algorithms https://openreview.net/forum?id=CPdc77SQfQ5

AUTHORS: Pan Zhou, Xingyu Xie, Shuicheng YAN HIGHLIGHT: In this work, we explore the problem of "\textit {how to accelerate the convergence of adaptive gradient algorithms in a general manner}", and aim at providing practical insights to boost the training efficiency. 200. TITLE: Efficient Edge Inference by Selective Query https://openreview.net/forum?id=jpR98ZdIm2q AUTHORS: Anil Kag, Igor Fedorov, Aditya Gangrade, Paul Whatmough, Venkatesh Saligrama HIGHLIGHT: We propose a novel end-to-end hybrid learning framework that allows the edge to selectively query only those hard examples that the cloud can classify correctly. 201, TITLE: Learning topology-preserving data representations https://openreview.net/forum?id=llu-ixf-Tzf AUTHORS: Ilya Trofimov, Daniil Cherniavskii, Eduard Tulchinskii, Nikita Balabin, Serguei Barannikov, Evgeny Burnaev HIGHLIGHT: We propose a method for learning topology-preserving data representations (dimensionality reduction). 202. TITLE: Towards Understanding Why Mask Reconstruction Pretraining Helps in Downstream Tasks https://openreview.net/forum?id=PaEUQiY40Dk AUTHORS: Jiachun Pan, Pan Zhou, Shuicheng YAN HIGHLIGHT: To solve these problems, we first theoretically show that on an auto-encoder of a two/one-layered convolution encoder/decoder. MRP can capture all discriminative semantics of each potential semantic class in the pretraining dataset. Then considering the fact that the pretraining dataset is of huge size and high diversity and thus covers most semantics in downstream dataset, in fine-tuning phase, the pretrained encoder can capture as much semantics as it can in downstream datasets, and would not lost these semantics with theoretical guarantees. 203, TITLE: Leveraging Incompatibility to Defend Against Backdoor Poisoning https://openreview.net/forum?id=mkJm5Uy4HrQ Charles Jin, Melinda Sun, Martin Rinard AUTHORS: We identify an incompatibility property of the interaction of clean and poisoned data with the training HIGHLIGHT: algorithm, specifically that including poisoned data in the training dataset does not improve model accuracy on clean data and viceversa. Leveraging this property, we develop an algorithm that iteratively refines subsets of the poisoned dataset to obtain subsets that concentrate around either clean or poisoned data. 204, TITLE: Statistical Guarantees for Consensus Clustering https://openreview.net/forum?id=kQxry8Z6Fd9 AUTHORS: Zhixin Zhou, Gautam Dudeja, Arash A Amini We analyze the statistical performance of aggregation algorithms under a stochastic label perturbation model, HIGHLIGHT: and show that a \$K\$-means type algorithm followed by a local refinement step can achieve near optimal performance, with a rate that decays exponentially in \$N\$. 205, TITLE: More Centralized Training, Still Decentralized Execution: Multi-Agent Conditional Policy Factorization https://openreview.net/forum?id=znLlSgN-4S0 AUTHORS: Jiangxing Wang, Deheng Ye, Zongqing Lu HIGHLIGHT: To this end, we propose multi-agent conditional policy factorization (MACPF), which takes more centralized training but still enables decentralized execution. 206, TITLE: Calibrating Transformers via Sparse Gaussian Processes https://openreview.net/forum?id=jPVAFXHlbL AUTHORS: Wenlong Chen, Yingzhen Li HIGHLIGHT: Extending Transformer's success to safety-critical domains requires calibrated uncertainty estimation which remains under-explored. To address this, we propose Sparse Gaussian Process attention (SGPA), which performs Bayesian inference directly in the output space of multi-head attention blocks (MHAs) in transformer to calibrate its uncertainty. 207, TITLE: Red PANDA: Disambiguating Anomaly Detection by Removing Nuisance Factors https://openreview.net/forum?id=z37tDDHHgi AUTHORS: Niv Cohen, Jonathan Kahana, Yedid Hoshen HIGHLIGHT: Breaking from previous research, we present a new anomaly detection method that allows operators to exclude an attribute when detecting anomalies. STOCHASTIC NO-REGRET LEARNING FOR GENERAL GAMES WITH VARIANCE REDUCTION 208, TITLE: https://openreview.net/forum?id=oJZ8bPtCar AUTHORS: Yichi Zhou, Fang Kong, Shuai Li We show that a stochastic version of optimistic mirror descent (OMD), a variant of mirror descent with recency HIGHLIGHT: bias, converges fast in general games. The Dark Side of AutoML: Towards Architectural Backdoor Search 209, TITLE: https://openreview.net/forum?id=bsZUL1DGXe AUTHORS: Ren Pang, Changjiang Li, Zhaohan Xi, Shouling Ji, Ting Wang HIGHLIGHT: This paper asks the intriguing question: is it possible to exploit neural architecture search (NAS) as a new attack

HIGHLIGHT: It is paper asks the intriguing question: is it possible to exploit neural architecture search (NAS) as a new attack vector to launch previously improbable attacks? Specifically, we present EVAS, a new attack that leverages NAS to find neural architectures with inherent backdoors and exploits such vulnerability using input-aware triggers.

210, TITLE: Alternating Differentiation for Optimization Layers https://openreview.net/forum?id=KKBMz-EL4tD AUTHORS: Haixiang Sun, Ye Shi, Jingya Wang, Hoang Duong Tuan, H. Vincent Poor, Dacheng Tao HIGHLIGHT: In this paper, we developed a new framework, named Alternating Differentiation (Alt-Diff), that differentiates optimization problems (here, specifically in the form of convex optimization problems with polyhedral constraints) in a fast and recursive way. 211, TITLE: On The Relative Error of Random Fourier Features for Preserving Kernel Distance https://openreview.net/forum?id=qs2YCziX2o-AUTHORS: Kuan Cheng, Shaofeng H.-C. Jiang, Luojian Wei, Zhide Wei HIGHLIGHT: We show that for a significant range of kernels, including the well-known Laplacian kernels, RFF cannot approximate the kernel distance with small relative error using low dimensions. 212. TITLE: Decentralized Optimistic Hyperpolicy Mirror Descent: Provably No-Regret Learning in Markov Games https://openreview.net/forum?id=bn0GZZdDfI1 AUTHORS: Wenhao Zhan, Jason D. Lee, Zhuoran Yang HIGHLIGHT: In light of a recent hardness result (Liu et al., 2022), we focus on the setting where the opponent's previous policies are revealed to the agent for decision making. With such an information structure, we propose a new algorithm, Decentralized Optimistic hypeRpolicy mIrror deScent (DORIS), which achieves \$\sqrt{K}\$-regret in the context of general function approximation, where \$K\$ is the number of episodes. 213, TITLE: PAC Reinforcement Learning for Predictive State Representations https://openreview.net/forum?id=FVW7Mi2ph6C AUTHORS: Wenhao Zhan, Masatoshi Uehara, Wen Sun, Jason D. Lee HIGHLIGHT: In this paper we study online Reinforcement Learning (RL) in partially observable dynamical systems. 214, TITLE: Make-A-Video: Text-to-Video Generation without Text-Video Data https://openreview.net/forum?id=nJfylDvgzlq AUTHORS: Uriel Singer, Adam Polyak, Thomas Hayes, Xi Yin, Jie An, Songyang Zhang, Qiyuan Hu, Harry Yang, Oron Ashual, Oran Gafni, Devi Parikh, Sonal Gupta, Yaniv Taigman HIGHLIGHT: We propose Make-A-Video -- an approach for directly translating the tremendous recent progress in Text-to-Image (T2I) generation to Text-to-Video (T2V). 215. TITLE: Static Prediction of Runtime Errors by Learning to Execute Programs with External Resource Descriptions https://openreview.net/forum?id=lLp-C5nTdJG AUTHORS: David Bieber, Rishab Goel, Dan Zheng, Hugo Larochelle, Daniel Tarlow HIGHLIGHT: Here, we introduce a competitive programming dataset and task for predicting runtime errors, which we show is difficult for generic models like Transformers. We approach this task by developing an interpreter-inspired architecture with an inductive bias towards mimicking program executions, which models exception handling and "learns to execute" descriptions of external resources. 216, TITLE: MACTA: A Multi-agent Reinforcement Learning Approach for Cache Timing Attacks and Detection https://openreview.net/forum?id=CDIHZ78-Xzi AUTHORS: Jiaxun Cui, Xiaomeng Yang, Geunbae Lee, Mulong Luo, Peter Stone, Hsien-Hsin S. Lee, Benjamin Lee, G. Edward Suh, Wenjie Xiong, Yuandong Tian HIGHLIGHT: However, the current detection of cache timing attacks relies heavily on heuristics and expert knowledge, which can lead to brittleness and inability to adapt to new attacks. To mitigate these problems, we develop a two-player environment for cache-timing attacks and detection, and leverage the idea of population-based multi-agent reinforcement learning (MARL) to train both attackers and detectors.

217, TITLE: Quantized Compressed Sensing with Score-Based Generative Models

https://openreview.net/forum?id=OOWLRfAI V

AUTHORS: Xiangming Meng, Yoshiyuki Kabashima

HIGHLIGHT: Motivated by the power of score-based generative models (SGM, also known as diffusion models) in capturing the rich structure of natural signals beyond simple sparsity, we propose an unsupervised data-driven approach called quantized compressed sensing with SGM (QCS-SGM), where the prior distribution is modeled by a pre-trained SGM.

218, TITLE: SGDA with shuffling: faster convergence for nonconvex-P? minimax optimization

https://openreview.net/forum?id=6xXtM8bFFJ

AUTHORS: Hanseul Cho, Chulhee Yun

HIGHLIGHT: Most practical implementations of SGDA randomly reshuffle components and sequentially use them (i.e., without-replacement sampling); however, there are few theoretical results on this approach for minimax algorithms, especially outside the easier-to-analyze (strongly-)monotone setups. To narrow this gap, we study the convergence bounds of SGDA with random reshuffling (SGDA-RR) for smooth nonconvex-nonconcave objectives with Polyak-?ojasiewicz (P?) geometry.

219, TITLE: MOAT: Alternating Mobile Convolution and Attention Brings Strong Vision Models https://openreview.net/forum?id=H0HGljkxQFN

AUTHORS: Chenglin Yang, Siyuan Qiao, Qihang Yu, Xiaoding Yuan, Yukun Zhu, Alan Yuille, Hartwig Adam, Liang-Chieh Chen HIGHLIGHT: This paper presents MOAT, a family of neural networks that build on top of MObile convolution (i.e., inverted residual blocks) and ATtention.

220, TITLE:	View Synthesis with Sculpted Neural Points
https://openreview.ne	t/forum?id=0ypGZvm0er0
AUTHORS:	Yiming Zuo, Jia Deng
HIGHLIGHT:	In this work, we propose a new approach that performs view synthesis using point clouds.
221, TITLE: https://openreview.ne AUTHORS: HIGHLIGHT: where a large portion adjusted.	Extremely Simple Activation Shaping for Out-of-Distribution Detection t/forum?id=ndYXTEL6cZz Andrija Djurisic, Nebojsa Bozanic, Arjun Ashok, Rosanne Liu In contrast, in this work, we propose an extremely simple, post-hoc, on-the-fly activation shaping method, ASH, (e.g. 90%) of a sample's activation at a late layer is removed, and the rest (e.g. 10%) simplified or lightly
222, TITLE:	Temporal Disentanglement of Representations for Improved Generalisation in Reinforcement Learning
https://openreview.ne	t/forum?id=sPgP6alSLTD
AUTHORS:	Mhairi Dunion, Trevor McInroe, Kevin Sebastian Luck, Josiah P. Hanna, Stefano V Albrecht
HIGHLIGHT:	To learn more robust representations, we introduce TEmporal Disentanglement (TED), a self-supervised
auxiliary task that lea	ds to disentangled image representations exploiting the sequential nature of RL observations.
223, TITLE:	Dr.Spider: A Diagnostic Evaluation Benchmark towards Text-to-SQL Robustness
https://openreview.ne	t/forum?id=Wc5bmZZU9cy
AUTHORS:	Shuaichen Chang, Jun Wang, Mingwen Dong, Lin Pan, Henghui Zhu, Alexander Hanbo Li, Wuwei Lan, Sheng
Zhang, Jiarong Jiang,	Joseph Lilien, Steve Ash, William Yang Wang, Zhiguo Wang, Vittorio Castelli, Patrick Ng, Bing Xiang
HIGHLIGHT:	In this paper, we propose a comprehensive robustness benchmark based on Spider, a cross-domain text-to-SQL
benchmark, to diagno	se the model robustness.
224, TITLE:	HiT-MDP: Learning the SMDP option framework on MDPs with Hidden Temporal Variables
https://openreview.ne	t/forum?id=VuuDXDgujAc
AUTHORS:	Chang Li, Dongjin Song, Dacheng Tao
HIGHLIGHT:	To this end, we propose a novel Markov Decision Process (MDP), the Hidden Temporal MDP (HiT-MDP), and
prove that the option-	induced HiT-MDP is homomorphic equivalent to the option-induced SMDP.
225, TITLE:	Expressive Monotonic Neural Networks
https://openreview.ne	t/forum?id=w2P7fMy_RH
AUTHORS:	Niklas Nolte, Ouail Kitouni, Mike Williams
HIGHLIGHT:	In this work, we propose a weight-constrained architecture with a single residual connection to achieve exact
monotonic dependence	e in any subset of the inputs.
226, TITLE:	Information-Theoretic Analysis of Unsupervised Domain Adaptation
https://openreview.ne	t/forum?id=c5tbxWXU9-y
AUTHORS:	Ziqiao Wang, Yongyi Mao
HIGHLIGHT:	While our bounds for the first kind of error are in line with the traditional analysis and give similar insights, our
bounds on the second	kind of error are algorithm-dependent, which also provide insights into algorithm designs. Specifically, we
present two simple te	chniques for improving generalization in UDA and validate them experimentally.
227, TITLE:	Provably Efficient Lifelong Reinforcement Learning with Linear Representation
https://openreview.ne	t/forum?id=Qd0p0bl-A9t
AUTHORS:	Sanae Amani, Lin Yang, Ching-An Cheng
HIGHLIGHT:	We frame the problem as a linearly parameterized contextual Markov decision process (MDP), where each task
is specified by a conto	ext and the transition dynamics is context-independent, and we introduce a new completeness-style assumption
on the representation	which is sufficient to ensure the optimal multi-task policy is realizable under the linear representation. Under this
assumption, we propo	use an algorithm, called UCB Lifelong Value Distillation (UCBlvd), that provably achieves sublinear regret for
any sequence of tasks	while using only sublinear planning calls.
228, TITLE:	Valid P-Value for Deep Learning-driven Salient Region
https://openreview.ne	t/forum?id=qihMOPw4Sf_
AUTHORS:	Miwa Daiki, Vo Nguyen Le Duy, Ichiro Takeuchi
HIGHLIGHT:	In this study, we propose a method to quantify the reliability of a saliency region in the form of p-values.
229, TITLE: https://openreview.ne AUTHORS: HIGHLIGHT: first theoretical analys classification task.	A Theoretical Understanding of Vision Transformers: Learning, Generalization, and Sample Complexity t/forum?id=jClGv3Qjhb Hongkang Li, Meng Wang, Sijia Liu, Pin-Yu Chen Based on a data model characterizing both label-relevant and label-irrelevant tokens, this paper provides the sis of training a three-layer ViT, i.e., one self-attention layer followed by a two-layer perceptron, for a
220 TITLE.	Disentanglement of Correlated Factors via Hausdorff Factorized Support

AUTHORS: Karsten Roth, Mark Ibrahim, Zeynep Akata, Pascal Vincent, Diane Bouchacourt HIGHLIGHT: We develop a method that allows for disentangled representation learning not only under the assumption of independent factors of variation but instead fundamentally allows for much more realistic correlations during training. 231, TITLE: SCALE-UP: An Efficient Black-box Input-level Backdoor Detection via Analyzing Scaled Prediction Consistency https://openreview.net/forum?id=o0LFPcoFKnr AUTHORS: Junfeng Guo, Yiming Li, Xun Chen, Hanqing Guo, Lichao Sun, Cong Liu HIGHLIGHT: However, almost all of them cannot be adopted in MLaaS scenarios since they require getting access to or even modifying the suspicious models. In this paper, we propose a simple yet effective black-box input-level backdoor detection, called SCALE-UP, which requires only the predicted labels to alleviate this problem. Pink Noise Is All You Need: Colored Noise Exploration in Deep Reinforcement Learning 232. TITLE: https://openreview.net/forum?id=hQ9V5QN27eS AUTHORS: Onno Eberhard, Jakob Hollenstein, Cristina Pinneri, Georg Martius HIGHLIGHT: In this work, we perform a comprehensive experimental evaluation on MPO and SAC to explore the effectiveness of other colors of noise as action noise. 233, TITLE: Revisiting the Assumption of Latent Separability for Backdoor Defenses https://openreview.net/forum?id= wSHsgrVali AUTHORS: Xiangyu Qi, Tinghao Xie, Yiming Li, Saeed Mahloujifar, Prateek Mittal HIGHLIGHT: This question is central to understanding whether the assumption of latent separability provides a reliable foundation for defending against backdoor poisoning attacks. In this paper, we design adaptive backdoor poisoning attacks to present counter-examples against this assumption. 234, TITLE: Optimal Transport for Offline Imitation Learning https://openreview.net/forum?id=MhuFzFsrfvH AUTHORS: Yicheng Luo, zhengyao jiang, Samuel Cohen, Edward Grefenstette, Marc Peter Deisenroth In this paper, we introduce Optimal Transport Relabeling (OTR), an imitation learning algorithm that can HIGHLIGHT: automatically relabel offline data of mixed and unknown quality with rewards from a few good demonstrations. 235, TITLE: Mitigating Dataset Bias by Using Per-Sample Gradient https://openreview.net/forum?id=7mgUec-7GMv AUTHORS: Sumyeong Ahn, Seongyoon Kim, Se-Young Yun HIGHLIGHT: In this study, we propose a debiasing algorithm leveraging gradient called PGD (Per-sample Gradient-based Debiasing). Efficient Model Updates for Approximate Unlearning of Graph-Structured Data 236, TITLE: https://openreview.net/forum?id=fhcu4FBLciL AUTHORS: Eli Chien, Chao Pan, Olgica Milenkovic HIGHLIGHT: This work introduces the first known approach for \emph{approximate graph unlearning} with provable theoretical guarantees. MAST: Masked Augmentation Subspace Training for Generalizable Self-Supervised Priors 237, TITLE: https://openreview.net/forum?id=5KUPKjHYD-1 AUTHORS: Chen Huang, Hanlin Goh, Jiatao Gu, Joshua M. Susskind In this paper, we aim to learn self-supervised features that generalize well across a variety of downstream tasks HIGHLIGHT: (e.g., object classification, detection and instance segmentation) without knowing any task information beforehand. 238, TITLE: SmartFRZ: An Efficient Training Framework using Attention-Based Layer Freezing https://openreview.net/forum?id=i9UlAr1T xl AUTHORS: Sheng Li, Geng Yuan, Yue Dai, Youtao Zhang, Yanzhi Wang, Xulong Tang HIGHLIGHT: To this end, we propose a generic and efficient training framework (SmartFRZ). 239, TITLE: Sparse Random Networks for Communication-Efficient Federated Learning https://openreview.net/forum?id=k1FHgri5y3-AUTHORS: Berivan Isik, Francesco Pase, Deniz Gunduz, Tsachy Weissman, Zorzi Michele HIGHLIGHT: We propose an FL framework, where clients find a sparse random network using a stochastic strategy; and provide (1) lower communication cost, (2) higher accuracy, (3) faster convergence, and (4) at the end of the training, a compressed final model. PV3D: A 3D Generative Model for Portrait Video Generation 240, TITLE: https://openreview.net/forum?id=o3yygm3lnzS AUTHORS: Eric Zhongcong Xu, Jianfeng Zhang, Jun Hao Liew, Wenqing Zhang, Song Bai, Jiashi Feng, Mike Zheng Shou In this work, we propose PV3D, the first generative framework that can synthesize multi-view consistent HIGHLIGHT: portrait videos. 241, TITLE: S-NeRF: Neural Radiance Fields for Street Views https://openreview.net/forum?id=gx2yJS-ENqI AUTHORS: Ziyang Xie, Junge Zhang, Wenye Li, Feihu Zhang, Li Zhang

HIGHLIGHT: In this paper, we propose a new street-view NeRF (S-NeRF) that considers novel view synthesis of both the large-scale background scenes and the foreground moving vehicles jointly. 242. TITLE: The Influence of Learning Rule on Representation Dynamics in Wide Neural Networks https://openreview.net/forum?id=nZ2NtpolC5-AUTHORS: Blake Bordelon, Cengiz Pehlevan HIGHLIGHT: To gain insight into the relationship between learned features, function approximation, and the learning rule, we analyze infinite-width deep networks trained with gradient descent (GD) and biologically-plausible alternatives including feedback alignment (FA), direct feedback alignment (DFA), and error modulated Hebbian learning (Hebb), as well as gated linear networks (GLN). 243, TITLE: On the Data-Efficiency with Contrastive Image Transformation in Reinforcement Learning https://openreview.net/forum?id=-nm-rHXi5ga AUTHORS: Sicong Liu, Xi Sheryl Zhang, Yushuo Li, Yifan Zhang, Jian Cheng HIGHLIGHT: In the end, we propose a contrastive invariant transformation (CoIT), a simple yet promising learnable data augmentation combined with standard model-free algorithms to improve sample-efficiency. Dataless Knowledge Fusion by Merging Weights of Language Models 244. TITLE: https://openreview.net/forum?id=FCnohuR6AnM AUTHORS: Xisen Jin, Pengxiang Cheng, Daniel Preotiuc-Pietro, Xiang Ren HIGHLIGHT: In this paper, we study the problem of merging individual models built on different training data sets to obtain a single model that performs well both across all data set domains and can generalize on out-of-domain data. 245. TITLE: Long Range Language Modeling via Gated State Spaces https://openreview.net/forum?id=5MkYIYCbva AUTHORS: Harsh Mehta, Ankit Gupta, Ashok Cutkosky, Behnam Neyshabur HIGHLIGHT: In this work we focus on autoregressive sequence modeling over English books, Github source code and ArXiv mathematics articles. 246. TITLE: Generalization Bounds for Federated Learning: Fast Rates, Unparticipating Clients and Unbounded Losses https://openreview.net/forum?id=-EHqoysUYLx AUTHORS: Xiaolin Hu, Shaojie Li, Yong Liu HIGHLIGHT: This paper provides a theoretical analysis of generalization error of {federated learning}, which captures both heterogeneity and relatedness of the distributions. More ConvNets in the 2020s: Scaling up Kernels Beyond 51x51 using Sparsity 247, TITLE: https://openreview.net/forum?id=bXNl-myZkJl AUTHORS: Shiwei Liu, Tianlong Chen, Xiaohan Chen, Xuxi Chen, Qiao Xiao, Boqian Wu, Tommi Kärkkäinen, Mykola Pechenizkiy, Decebal Constantin Mocanu, Zhangyang Wang HIGHLIGHT: In this paper, we explore the possibility of training extreme convolutions larger than 31x31 and test whether the performance gap can be eliminated by strategically enlarging convolutions. Safe Exploration Incurs Nearly No Additional Sample Complexity for Reward-Free RL 248, TITLE: https://openreview.net/forum?id=wNUgn1n6esQ AUTHORS: Ruiquan Huang, Jing Yang, Yingbin Liang HIGHLIGHT: It remains unclear how such safe exploration requirement would affect the corresponding sample complexity in order to achieve the desired optimality of the obtained policy in planning. In this work, we make a first attempt to answer this question. 249, TITLE: CUTS: Neural Causal Discovery from Unstructured Time-Series Data https://openreview.net/forum?id=UG8bQcD3Emv AUTHORS: Cheng Yuxiao, Runzhao Yang, Tingxiong Xiao, Zongren Li, Jinli Suo, Kunlun He, Qionghai Dai HIGHLIGHT: However, most existing methods assume structured input data and degenerate greatly when encountering data with randomly missing entries or non-uniform sampling frequencies, which hampers their applications in real scenarios. To address this issue, here we present CUTS, a neural Granger causal discovery algorithm to jointly impute unobserved data points and build causal graphs, via plugging in two mutually boosting modules in an iterative framework: (i) Latent data prediction stage: designs a Delayed Supervision Graph Neural Network (DSGNN) to hallucinate and register unstructured data which might be of high dimension and with complex distribution; (ii) Causal graph fitting stage: builds a causal adjacency matrix with imputed data under sparse penalty. 250, TITLE: A Kernel Perspective of Skip Connections in Convolutional Networks https://openreview.net/forum?id=6H uOfcwiVh AUTHORS: Daniel Barzilai, Amnon Geifman, Meirav Galun, Ronen Basri HIGHLIGHT: Here we study their properties through their Gaussian Process and Neural Tangent kernels. 251, TITLE: RLx2: Training a Sparse Deep Reinforcement Learning Model from Scratch https://openreview.net/forum?id=DJEEqoAq7to AUTHORS: Yiqin Tan, Pihe Hu, Ling Pan, Jiatai Huang, Longbo Huang

HIGHLIGHT: In this work, we propose a novel sparse DRL training framework, "the Rigged Reinforcement Learning Lottery" (RLx2), which builds upon gradient-based topology evolution and is capable of training a sparse DRL model based entirely on a sparse network.

252, TITLE: NANSY++: Unified Voice Synthesis with Neural Analysis and Synthesis https://openreview.net/forum?id=elDEe8LYW7-AUTHORS: Hyeong-Seok Choi, Jinhyeok Yang, Juheon Lee, Hyeongju Kim HIGHLIGHT: To this end, we propose a unified framework of synthesizing and manipulating voice signals from analysis features, dubbed NANSY++. 253. TITLE: Equivariant Descriptor Fields: SE(3)-Equivariant Energy-Based Models for End-to-End Visual Robotic Manipulation Learning https://openreview.net/forum?id=dnjZSPGmY50 AUTHORS: Hyunwoo Ryu, Hong-in Lee, Jeong-Hoon Lee, Jongeun Choi HIGHLIGHT: In this paper, we present SE(3)-equivariant models for visual robotic manipulation from point clouds that can be trained fully end-to-end. 254, TITLE: Not All Tasks Are Born Equal: Understanding Zero-Shot Generalization https://openreview.net/forum?id=KGV-GBh8fb AUTHORS: Jing Zhou, Zongyu Lin, Yanan Zheng, Jian Li, Zhilin Yang HIGHLIGHT: Recent work has achieved remarkable zero-shot performance with multi-task prompted pretraining, but little has been understood. For the first time, we show that training on a small number of key tasks beats using all the training tasks, while removing these key tasks substantially hurts performance. Characterizing the spectrum of the NTK via a power series expansion 255. TITLE: https://openreview.net/forum?id=Tvms8xrZHyR AUTHORS: Michael Murray, Hui Jin, Benjamin Bowman, Guido Montufar HIGHLIGHT: Under mild conditions on the network initialization we derive a power series expansion for the Neural Tangent Kernel (NTK) of arbitrarily deep feedforward networks in the infinite width limit. 256, TITLE: Corrupted Image Modeling for Self-Supervised Visual Pre-Training https://openreview.net/forum?id=09hVcSDkea AUTHORS: Yuxin Fang, Li Dong, Hangbo Bao, Xinggang Wang, Furu Wei HIGHLIGHT: We introduce Corrupted Image Modeling (CIM) for self-supervised visual pre-training. 257, TITLE: Compositional Task Representations for Large Language Models https://openreview.net/forum?id=6axIMJA7ME3 AUTHORS: NAN SHAO, Zefan Cai, Hanwei xu, Chonghua Liao, Yanan Zheng, Zhilin Yang HIGHLIGHT: To this end, we propose a novel prompt-free approach, Compositional Task Representations (CTR), that employs multi-task training to learn a discrete, compositional codebook. Equivariant Hypergraph Diffusion Neural Operators 258, TITLE: https://openreview.net/forum?id=RiTjKoscnNd AUTHORS: Peihao Wang, Shenghao Yang, Yunyu Liu, Zhangyang Wang, Pan Li HIGHLIGHT: Inspired by hypergraph diffusion algorithms, this work proposes a new HNN architecture named ED-HNN, which provably approximates any continuous equivariant hypergraph diffusion operators that can model a wide range of higher-order relations. 259, TITLE: TextShield: Beyond Successfully Detecting Adversarial Sentences in NLP https://openreview.net/forum?id=xIWfWvKM7aQ AUTHORS: Lingfeng Shen, Ze Zhang, Haiyun Jiang, Ying Chen HIGHLIGHT: However, {the core limitation of previous detection methods is being incapable of giving correct predictions on adversarial sentences unlike defense methods from other paradigms.} To solve this issue, this paper proposes TextShield: (1) we discover a link between text attack and saliency information, and then we propose a saliency-based detector, which can effectively detect whether an input sentence is adversarial or not. (2) We design a saliency-based corrector, which converts the detected adversary sentences to benign ones. 260, TITLE: REVISITING PRUNING AT INITIALIZATION THROUGH THE LENS OF RAMANUJAN GRAPH https://openreview.net/forum?id=uVcDssQff AUTHORS: Duc N.M Hoang, Shiwei Liu, Radu Marculescu, Zhangyang Wang To better understand the underlying mechanism of PaI, we propose to interpret it through the lens of the HIGHLIGHT: Ramanujan Graph - a class of expander graphs that are sparse while being highly connected. 261, TITLE: Discovering Latent Knowledge in Language Models Without Supervision https://openreview.net/forum?id=ETKGuby0hcs AUTHORS: Collin Burns, Haotian Ye, Dan Klein, Jacob Steinhardt HIGHLIGHT: Specifically, we introduce a method for accurately answering yes-no questions given only unlabeled model activations. 262, TITLE: How to Exploit Hyperspherical Embeddings for Out-of-Distribution Detection? https://openreview.net/forum?id=aEFaE0W5pAd AUTHORS: Yifei Ming, Yiyou Sun, Ousmane Dia, Yixuan Li

HIGHLIGHT: In this work, we propose CIDER, a novel representation learning framework that exploits hyperspherical embeddings for OOD detection. 263, TITLE: Automatic Chain of Thought Prompting in Large Language Models https://openreview.net/forum?id=5NTt8GFjUHkr AUTHORS: Zhuosheng Zhang, Aston Zhang, Mu Li, Alex Smola HIGHLIGHT: To mitigate the effect of such mistakes, we investigate various principles for automatically constructing demonstrations and find that diversity matters. Inspired by these findings, we propose an automatic CoT prompting method called Auto-CoT. 264, TITLE: Parameter-Efficient Fine-Tuning Design Spaces https://openreview.net/forum?id=XSRSWxyJIC AUTHORS: Jiaao Chen, Aston Zhang, Xingjian Shi, Mu Li, Alex Smola, Diyi Yang HIGHLIGHT: Thus, we present a parameter-efficient fine-tuning design paradigm and discover design patterns that are applicable to different experimental settings. 265, TITLE: Learning Multimodal Data Augmentation in Feature Space https://openreview.net/forum?id=6SRDbbvU8s AUTHORS: Zichang Liu, Zhiqiang Tang, Xingjian Shi, Aston Zhang, Mu Li, Anshumali Shrivastava, Andrew Gordon Wilson HIGHLIGHT: In this paper, we introduce LeMDA, Learning Multimodal Data Augmentation, an easy-to-use method that automatically learns to jointly augment multimodal data in feature space, with no constraints on the identities of the modalities or the relationship between modalities. 266, TITLE: AIM: Adapting Image Models for Efficient Video Understanding https://openreview.net/forum?id=CIoSZ HKHS7 Taojiannan Yang, Yi Zhu, Yusheng Xie, Aston Zhang, Chen Chen, Mu Li AUTHORS: HIGHLIGHT: In this work, we propose a novel method to Adapt pre-trained Image Models (AIM) for efficient video understanding. 267, TITLE: Factorized Fourier Neural Operators https://openreview.net/forum?id=tmIiMPl4IPa AUTHORS: Alasdair Tran, Alexander Mathews, Lexing Xie, Cheng Soon Ong HIGHLIGHT: We propose the Factorized Fourier Neural Operator (F-FNO), a learning-based approach for simulating partial differential equations (PDEs). FaiREE: fair classification with finite-sample and distribution-free guarantee 268, TITLE: https://openreview.net/forum?id=shzu8d6 YAR AUTHORS: Puheng Li, James Zou, Linjun Zhang HIGHLIGHT: In this paper, we propose FaiREE, a fair classification algorithm which can satisfy group fairness constraints with finite-sample and distribution-free theoretical guarantees. 269, TITLE: Exponential Generalization Bounds with Near-Optimal Rates for \$L_q\$-Stable Algorithms https://openreview.net/forum?id=1 jtWjhSSkr AUTHORS: Xiaotong Yuan, Ping Li HIGHLIGHT: Therefore, a natural question we would like to address in this paper is whether it is possible to derive nearoptimal exponential generalization bounds for \$L q\$-stable learning algorithms. As the core contribution of the present work, we give an affirmative answer to this question by developing strict analogues of the near-optimal generalization and risk bounds of uniformly stable algorithms for \$L q\$-stable algorithms. Equal Improvability: A New Fairness Notion Considering the Long-term Impact 270, TITLE: https://openreview.net/forum?id=dhYUMMy0 Eg AUTHORS: Ozgur Guldogan, Yuchen Zeng, Jy-yong Sohn, Ramtin Pedarsani, Kangwook Lee HIGHLIGHT: To find a classifier that satisfies the EI requirement, we propose and study three different approaches that solve EI regularized optimization problems. 271. TITLE: Riemannian Metric Learning via Optimal Transport https://openreview.net/forum?id=v3y68gz-WEz AUTHORS: Christopher Scarvelis, Justin Solomon HIGHLIGHT: We introduce an optimal transport-based model for learning a metric tensor from cross-sectional samples of evolving probability measures on a common Riemannian manifold.

272, TITLE:MaskViT: Masked Visual Pre-Training for Video Predictionhttps://openreview.net/forum?id=QAV2CcLEDhAUTHORS:Agrim Gupta, Stephen Tian, Yunzhi Zhang, Jiajun Wu, Roberto Martín-Martín, Li Fei-FeiHIGHLIGHT:This work shows that we can create good video prediction models by pre-training transformers via maskedvisual modeling.Maski Martín Martín, Li Fei-Fei

273, TITLE: Prompting GPT-3 To Be Reliable https://openreview.net/forum?id=98p5x51L5af

AUTHORS: Chenglei Si, Zhe Gan, Zhengyuan Yang, Shuohang Wang, Jianfeng Wang, Jordan Lee Boyd-Graber, Lijuan Wang HIGHLIGHT: We establish simple and effective prompting methods to make GPT-3 reliable in terms of: robustness, fairness, calibration, factuality. 274, TITLE: Teacher Guided Training: An Efficient Framework for Knowledge Transfer https://openreview.net/forum?id=GVSf7Z7DbYL AUTHORS: Manzil Zaheer, Ankit Singh Rawat, Seungyeon Kim, Chong You, Himanshu Jain, Andreas Veit, Rob Fergus, Sanjiv Kumar HIGHLIGHT: In this paper, we propose the teacher-guided training (TGT) framework for training a high-quality compact model that leverages the knowledge acquired by pretrained generative models, while obviating the need to go through a large volume of data. 275, TITLE: Sparsity-Constrained Optimal Transport https://openreview.net/forum?id=yHY9NbQJ5BP AUTHORS: Tianlin Liu, Joan Puigcerver, Mathieu Blondel HIGHLIGHT: We propose in this paper a new approach for OT with explicit cardinality constraints on the transportation plan. 276, TITLE: Turning the Curse of Heterogeneity in Federated Learning into a Blessing for Out-of-Distribution Detection https://openreview.net/forum?id=mMNimwRb7Gr AUTHORS: Shuyang Yu, Junyuan Hong, Haotao Wang, Zhangyang Wang, Jiayu Zhou HIGHLIGHT On the other hand, a notorious challenge in FL is data heterogeneity where each client collects non-identically and independently distributed (non-iid) data. We propose to take advantage of such heterogeneity and turn the curse into a blessing that facilitates OoD detection in FL. 277, TITLE: Unbiased Stochastic Proximal Solver for Graph Neural Networks with Equilibrium States https://openreview.net/forum?id=j3cUWIMsFBN Mingjie Li, Yifei Wang, Yisen Wang, Zhouchen Lin AUTHORS: HIGHLIGHT: Such weakness limits the scalability of the implicit graph models. To tackle such limitations, we propose two unbiased stochastic proximal solvers inspired by the stochastic proximal gradient descent method and its variance reduction variant called USP and USP-VR solvers. 278, TITLE: Asynchronous Distributed Bilevel Optimization https://openreview.net/forum?id= i0-12XqVJZ AUTHORS: Yang Jiao, Kai Yang, Tiancheng Wu, Dongjin Song, Chengtao Jian HIGHLIGHT: As a remedy, we propose Asynchronous Distributed Bilevel Optimization (ADBO) algorithm. Relative Behavioral Attributes: Filling the Gap between Symbolic Goal Specification and Reward Learning 279, TITLE: from Human Preferences https://openreview.net/forum?id=lGz9u1ubUXE AUTHORS: Lin Guan, Karthik Valmeekam, Subbarao Kambhampati HIGHLIGHT: We propose two different parametric methods that can potentially encode any kind of behavioral attributes from ordered behavior clips. 280. TITLE: Neural Lagrangian Schr\" {0} dinger Bridge: Diffusion Modeling for Population Dynamics https://openreview.net/forum?id=d3QNWD_pcFv AUTHORS: Takeshi Koshizuka, Issei Sato HIGHLIGHT: To satisfy these requirements of the sample trajectories, we formulate the Lagrangian Schrödinger bridge (LSB) problem and propose to solve it approximately by modeling the advection-diffusion process with regularized neural SDE. One of the main difficulties in analyzing population dynamics is that we can only obtain observation data with coarse time intervals from fixedpoint observations due to experimental costs or measurement constraints. 281, TITLE: How to prepare your task head for finetuning https://openreview.net/forum?id=gVOXZproe-e AUTHORS: Yi Ren, Shangmin Guo, Wonho Bae, Danica J. Sutherland HIGHLIGHT: We analyze how the choice of task head controls feature adaptation and hence influences the downstream performance. 282, TITLE: Learning Fast and Slow for Time Series Forecasting https://openreview.net/forum?id=q-PbpHD3EOk AUTHORS: Quang Pham, Chenghao Liu, Doyen Sahoo, Steven Hoi In this work, inspired by the Complementary Learning Systems (CLS) theory, we propose Fast and Slow HIGHLIGHT: learning Network (FSNet) as a novel framework to address the challenges of online forecasting. 283, TITLE: Gradient-based optimization is not necessary for generalization in neural networks https://openreview.net/forum?id=QC10RmRbZy9 AUTHORS: Ping-yeh Chiang, Renkun Ni, David Yu Miller, Arpit Bansal, Jonas Geiping, Micah Goldblum, Tom Goldstein HIGHLIGHT: It is commonly believed that the implicit regularization of optimizers is needed for neural networks to

generalize in the overparameterized regime. In this paper, we observe experimentally that this implicit regularization behavior is {\em generic}, i.e. it does not depend strongly on the choice of optimizer.

 284, TITLE:
 Rhino: Deep Causal Temporal Relationship Learning with History-dependent Noise

 https://openreview.net/forum?id=i_1rbq8yFWC

 AUTHORS:
 Wenbo Gong, Joel Jennings, Cheng Zhang, Nick Pawlowski

 HIGHLIGHT:
 In this paper, we propose a structural equation model, called Rhino, which combines vector auto-regression,

 deep learning and variational inference to model non-linear relationships with instantaneous effects while allowing the noise

 distribution to be modulated by history observations.

285, TITLE: Mitigating Memorization of Noisy Labels via Regularization between Representations

https://openreview.net/forum?id=6qcYDVlVLnK

AUTHORS: Hao Cheng, Zhaowei Zhu, Xing Sun, Yang Liu

HIGHLIGHT: However, it is non-trivial to design a neural network with the best capacity given an arbitrary task. To circumvent this dilemma, instead of changing the model architecture, we decouple DNNs into an encoder followed by a linear classifier and propose to restrict the function space of a DNN by a representation regularizer.

286, TITLE: Backpropagation through Combinatorial Algorithms: Identity with Projection Works

https://openreview.net/forum?id=JZMR727O29

AUTHORS: Subham Sekhar Sahoo, Anselm Paulus, Marin Vlastelica, Vít Musil, Volodymyr Kuleshov, Georg Martius HIGHLIGHT: We propose a principled approach to exploit the geometry of the discrete solution space to treat the solver as a negative identity on the backward pass and further provide a theoretical justification.

287, TITLE: BSTT: A Bayesian Spatial-Temporal Transformer for Sleep Staging

https://openreview.net/forum?id=ZxdkjTgK_Dl

AUTHORS: Yuchen Liu, Ziyu Jia

HIGHLIGHT: In this paper, we propose a novel Bayesian spatial-temporal relation inference neural network, named Bayesian spatial-temporal transformer (BSTT), for sleep staging.

288, TITLE: Self-Guided Noise-Free Data Generation for Efficient Zero-Shot Learning

https://openreview.net/forum?id=h5OpjGd_lo6

AUTHORS: Jiahui Gao, Renjie Pi, LIN Yong, Hang Xu, Jiacheng Ye, Zhiyong Wu, WEIZHONG ZHANG, Xiaodan Liang, Zhenguo Li, Lingpeng Kong

HIGHLIGHT: In this paper, we propose a novel noise-robust re-weighting framework SunGen to automatically construct highquality data for zero-shot classification problems.

289, TITLE: H2RBox: Horizonal Box Annotation is All You Need for Oriented Object Detection https://openreview.net/forum?id=NPfDKT9OUJ3

AUTHORS: Xue Yang, Gefan Zhang, Wentong Li, Yue Zhou, Xuehui Wang, Junchi Yan

HIGHLIGHT: Oriented object detection emerges in many applications from aerial images to autonomous driving, while many existing detection benchmarks are annotated with horizontal bounding box only which is also less costive than fine-grained rotated box, leading to a gap between the readily available training corpus and the rising demand for oriented object detection. This paper proposes a simple yet effective oriented object detection approach called H2RBox merely using horizontal box annotation for weakly-supervised training, which closes the above gap and shows competitive performance even against those trained with rotated boxes.

290, TITLE: IDEAL: Query-Efficient Data-Free Learning from Black-Box Models

https://openreview.net/forum?id=ConT6H7MWL

AUTHORS: Jie Zhang, Chen Chen, Lingjuan Lyu

HIGHLIGHT: Nevertheless, these works require a large number of queries to the teacher model, which incurs significant monetary and computational costs. To address these problems, we propose a novel method called \emph{query-effIcient Data-free lEarning blAck-box modeLs} (IDEAL), which aims to query-efficiently learn from black-box model APIs to train a good student without any real data.

291, TITLE: Scaling Laws in Mean-Field Games

https://openreview.net/forum?id=fB4V-2QvCEm

AUTHORS: Pengdeng Li, Xinrun Wang, Shuxin Li, Hau Chan, Bo An

HIGHLIGHT: In this work, we attempt to bridge the two largely independently evolving fields of finite-agent and infiniteagent games, by studying the scaling laws in mean-field games.

292, TITLE: Towards Addressing Label Skews in One-Shot Federated Learning

https://openreview.net/forum?id=rzrqh85f4Sc

AUTHORS: Yiqun Diao, Qinbin Li, Bingsheng He

HIGHLIGHT: Due to the limited number of classes in each party, the local models misclassify the data from unseen classes into seen classes, which leads to very ineffective global models from voting. To address the label skew issue in one-shot FL, we propose a novel approach named FedOV which generates diverse outliers and introduces them as an additional unknown class in local training to improve the voting performance.

293, TITLE: Sequential Attention for Feature Selection

https://openreview.net/forum?id=TTLLGx3eet

AUTHORS:Taisuke Yasuda, Mohammadhossein Bateni, Lin Chen, Matthew Fahrbach, Gang Fu, Vahab MirrokniHIGHLIGHT:We propose a feature selection algorithm called Sequential Attention that achieves state-of-the-art empiricalresults for neural networks.

294. TITLE: Deep Transformers without Shortcuts: Modifying Self-attention for Faithful Signal Propagation https://openreview.net/forum?id=NPrsUQgMjKK AUTHORS: Bobby He, James Martens, Guodong Zhang, Aleksandar Botev, Andrew Brock, Samuel L Smith, Yee Whye Teh HIGHLIGHT: And so the question remains: \emph{is it possible to train deep vanilla transformers?} We answer this question in the affirmative by designing several approaches that use combinations of parameter initialisations, bias matrices and locationdependent rescaling to achieve faithful signal propagation in vanilla transformers. 295, TITLE: Approximation and non-parametric estimation of functions over high-dimensional spheres via deep ReLU networks https://openreview.net/forum?id=r90KYcuB7JS AUTHORS: Namjoon Suh, Tian-Yi Zhou, Xiaoming Huo HIGHLIGHT: We develop a new approximation and estimation analysis of deep feed-forward neural networks (FNNs) with the Rectified Linear Unit (ReLU) activation. 296, TITLE: Specformer: Spectral Graph Neural Networks Meet Transformers https://openreview.net/forum?id=0pdSt3oyJa1 AUTHORS: Deyu Bo, Chuan Shi, Lele Wang, Renjie Liao HIGHLIGHT: Furthermore, these filters are often constructed based on some fixed-order polynomials, which have limited expressiveness and flexibility. To tackle these issues, we introduce Specformer, which effectively encodes the set of all eigenvalues and performs self-attention in the spectral domain, leading to a learnable set-to-set spectral filter. 297. TITLE: MLPInit: Embarrassingly Simple GNN Training Acceleration with MLP Initialization https://openreview.net/forum?id=P8YIphWNEGO AUTHORS: Xiaotian Han, Tong Zhao, Yozen Liu, Xia Hu, Neil Shah HIGHLIGHT: To this end, we propose an embarrassingly simple, yet hugely effective initialization method for GNN training acceleration, called MLPInit. 298, TITLE: Empowering Graph Representation Learning with Test-Time Graph Transformation https://openreview.net/forum?id=Lnx15pr018 AUTHORS: Wei Jin, Tong Zhao, Jiayuan Ding, Yozen Liu, Jiliang Tang, Neil Shah HIGHLIGHT: Recent efforts have been made on tackling these issues from a modeling perspective which requires additional cost of changing model architectures or re-training model parameters. In this work, we provide a data-centric view to tackle these issues and propose a graph transformation framework named GTrans which adapts and refines graph data at test time to achieve better performance. Improving the Calibration of Fine-tuned Language Models via Denoising Variational Auto-Encoders 299, TITLE: https://openreview.net/forum?id=NI7StoWHJPT AUTHORS: Guande He, Jianfei Chen, Jun Zhu In this paper, we tackle the problem of calibrating fine-tuned language models. HIGHLIGHT: 300, TITLE: Softened Symbol Grounding for Neuro-symbolic Systems https://openreview.net/forum?id=HTJE5Krui0g AUTHORS: Zenan Li, Yuan Yao, Taolue Chen, Jingwei Xu, Chun Cao, Xiaoxing Ma, Jian L\" {u} HIGHLIGHT: This paper presents a novel, softened symbol grounding process, enabling the interactions of the two worlds in a mutually beneficial manner. 301, TITLE: Multi-task Self-supervised Graph Neural Networks Enable Stronger Task Generalization https://openreview.net/forum?id=1tHAZRqftM Mingxuan Ju, Tong Zhao, Qianlong Wen, Wenhao Yu, Neil Shah, Yanfang Ye, Chuxu Zhang AUTHORS: HIGHLIGHT: When applied to various downstream tasks, these frameworks rarely perform equally well for every task, because one philosophy may not span the extensive knowledge required for all tasks. In light of this, we introduce ParetoGNN, a multi-task SSL framework for node representation learning over graphs. 302, TITLE: Learning with Logical Constraints but without Shortcut Satisfaction https://openreview.net/forum?id=M2unceRvqhh Zenan Li, Zehua Liu, Yuan Yao, Jingwei Xu, Taolue Chen, Xiaoxing Ma, Jian L\"{u} AUTHORS: HIGHLIGHT: In this paper, we present a new framework for learning with logical constraints. 303, TITLE: Link Prediction with Non-Contrastive Learning https://openreview.net/forum?id=9Jaz4APHtWD AUTHORS: William Shiao, Zhichun Guo, Tong Zhao, Evangelos E. Papalexakis, Yozen Liu, Neil Shah HIGHLIGHT: In this work, we extensively evaluate the performance of existing non-contrastive methods for link prediction in both transductive and inductive settings. 304, TITLE: A Neural Mean Embedding Approach for Back-door and Front-door Adjustment https://openreview.net/forum?id=rLguqxYvYHB

AUTHORS: Liyuan Xu, Arthur Gretton

HIGHLIGHT: We consider the estimation of average and counterfactual treatment effects, under two settings: back-door adjustment and front-door adjustment.

305, TITLE: Can We Find Nash Equilibria at a Linear Rate in Markov Games? https://openreview.net/forum?id=eQzLwwGyQrb AUTHORS: Zhuoqing Song, Jason D. Lee, Zhuoran Yang HIGHLIGHT: We study decentralized learning in two-player zero-sum discounted Markov games where the goal is to design a policy optimization algorithm for either agent satisfying two properties. Weighted Ensemble Self-Supervised Learning 306, TITLE: https://openreview.net/forum?id=CL-sVR9pvF AUTHORS: Yangjun Ruan, Saurabh Singh, Warren Richard Morningstar, Alexander A Alemi, Sergey Ioffe, Ian Fischer, Joshua V. Dillon HIGHLIGHT: In this paper, we explore how ensemble methods can improve recent SSL techniques by developing a framework that permits data-dependent weighted cross-entropy losses. 307, TITLE: \$k\$NN Prompting: Learning Beyond the Context with Nearest Neighbor Inference https://openreview.net/forum?id=fe2S7736sNS Benfeng Xu, Quan Wang, Zhendong Mao, Yajuan Lyu, Qiaoqiao She, Yongdong Zhang AUTHORS: HIGHLIGHT: In-Context Learning, which formulates target tasks as prompt completion conditioned on in-context demonstrations, has become the prevailing and standard utilization of large language models. In this paper, we disclose an actual predicament for this typical usage that it can not scale up with training data due to context length restrictions. 308, TITLE: Logical Entity Representation in Knowledge-Graphs for Differentiable Rule Learning https://openreview.net/forum?id=JdgO-ht1uTN Chi Han, Qizheng He, Charles Yu, Xinya Du, Hanghang Tong, Heng Ji AUTHORS: This formulation overlooks additional contextual information from neighboring sub-graphs of entity variables x, HIGHLIGHT: y and z. Intuitively, there is a large gap here, as local sub-graphs have been found to provide important information for knowledge graph completion. Inspired by these observations, we propose Logical Entity RePresentation (LERP) to encode contextual information of entities in the knowledge graph. 309, TITLE: Divide to Adapt: Mitigating Confirmation Bias for Domain Adaptation of Black-Box Predictors https://openreview.net/forum?id=hVrXUps3LFA AUTHORS: Jianfei Yang, Xiangyu Peng, Kai Wang, Zheng Zhu, Jiashi Feng, Lihua Xie, Yang You HIGHLIGHT: Existing DABP approaches mostly rely on knowledge distillation (KD) from the black-box predictor, i.e., training the model with its noisy target-domain predictions, which however inevitably introduces the confirmation bias accumulated from the prediction noises and leads to degrading performance. To mitigate such bias, we propose a new strategy, \textit {divide-toadapt}, that purifies cross-domain knowledge distillation by proper domain division. 310, TITLE: A Simple Yet Powerful Deep Active Learning With Snapshots Ensembles https://openreview.net/forum?id=IVESH65r0Ar AUTHORS: Seohyeon Jung, Sanghyun Kim, Juho Lee HIGHLIGHT: In this paper, we highlight the effectiveness of snapshot ensembles for deep active learning. 311. TITLE: Domain-Indexing Variational Bayes for Domain Adaptation https://openreview.net/forum?id=pxStyaf2oJ5 AUTHORS: Zihao Xu, Hao He, Guang-Yuan Hao, Hao Wang HIGHLIGHT: However, such domain indices are not always available. To address this challenge, we first provide a formal definition of domain index from the probabilistic perspective, and then propose an adversarial variational Bayesian framework that infers domain indices from multi-domain data, thereby providing additional insight on domain relations and improving domain adaptation performance. 312, TITLE: Over-parameterized Model Optimization with Polyak-{\L}ojasiewicz Condition

https://openreview.net/forum?id=aBIpZvMdS56 AUTHORS: Yixuan Chen, Yubin Shi, Mingzhi Dong, Xiaochen Yang, Dongsheng Li, Yujiang Wang, Robert Dick, Qin Lv, Yingying Zhao, Fan Yang, Ning Gu, Li Shang HIGHLIGHT: This work proposes a new regularized risk minimization for over-parameterized models with a novel PL regularization and implements it via network pruning guided by PL-based condition number.

 313, TITLE:
 RGI: robust GAN-inversion for mask-free image inpainting and unsupervised pixel-wise anomaly detection

 https://openreview.net/forum?id=1UbNwQC89a

 AUTHORS:
 Shancong Mou, Xiaoyi Gu, Meng Cao, Haoping Bai, Ping Huang, Jiulong Shan, Jianjun Shi

 HIGHLIGHT:
 In this paper, we propose a Robust GAN-inversion (RGI) method with a provable robustness guarantee to achieve image restoration under unknown \textit {gross} corruptions, where a small fraction of pixels are completely corrupted.

 314, TITLE:
 Encoding Recurrence into Transformers

 https://openreview.net/forum?id=7YfHla7IxBJ
 AUTHORS:

 Feiqing Huang, Kexin Lu, Yuxi CAI, Zhen Qin, Yanwen Fang, Guangjian Tian, Guodong Li

 HIGHLIGHT:
 Encoding Recurrence into Transformers

315, TITLE: GAIN: On the Generalization of Instructional Action Understanding https://openreview.net/forum?id=R1PmWBiyp6w AUTHORS: Junlong Li, Guangyi Chen, Yansong Tang, Jinan Bao, Kun Zhang, Jie Zhou, Jiwen Lu HIGHLIGHT: In this paper, we introduce a benchmark, named GAIN, to analyze the GeneralizAbility of INstructional action understanding models Semantic Uncertainty: Linguistic Invariances for Uncertainty Estimation in Natural Language Generation 316, TITLE: https://openreview.net/forum?id=VD-AYtP0dve AUTHORS: Lorenz Kuhn, Yarin Gal, Sebastian Farquhar HIGHLIGHT: We introduce a method to measure uncertainty in large language models. 317, TITLE: Seeing Differently, Acting Similarly: Heterogeneously Observable Imitation Learning https://openreview.net/forum?id=3ULaIHxn9u7 AUTHORS: Xin-Qiang Cai, Yao-Xiang Ding, Zixuan Chen, Yuan Jiang, Masashi Sugiyama, Zhi-Hua Zhou HIGHLIGHT: However, in reality, the observation coexistence will be limited due to the high cost of acquiring expert observations. In this work, we study this challenging problem with limited observation coexistence under heterogeneous observations: Heterogeneously Observable Imitation Learning (HOIL). 318, TITLE: Pre-training via Denoising for Molecular Property Prediction https://openreview.net/forum?id=tYIMtogvee AUTHORS: Sheheryar Zaidi, Michael Schaarschmidt, James Martens, Hyunjik Kim, Yee Whye Teh, Alvaro Sanchez-Gonzalez, Peter Battaglia, Razvan Pascanu, Jonathan Godwin HIGHLIGHT: In this paper, we describe a pre-training technique based on denoising that achieves a new state-of-the-art in molecular property prediction by utilizing large datasets of 3D molecular structures at equilibrium to learn meaningful representations for downstream tasks. 319, TITLE: A new characterization of the edge of stability based on a sharpness measure aware of batch gradient distribution https://openreview.net/forum?id=bH-kCY6LdKg AUTHORS: Sungyoon Lee, Cheongjae Jang HIGHLIGHT: We propose a new sharpness measure (interaction-aware-sharpness) aware of the \emph{interaction} between the batch gradient distribution and the loss landscape geometry. Equivariant Energy-Guided SDE for Inverse Molecular Design 320, TITLE: https://openreview.net/forum?id=r0otLtOwYW AUTHORS: Fan Bao, Min Zhao, Zhongkai Hao, Peiyao Li, Chongxuan Li, Jun Zhu In this paper, we propose equivariant energy-guided stochastic differential equations (EEGSDE), a flexible HIGHLIGHT: framework for controllable 3D molecule generation under the guidance of an energy function in diffusion models. Mutual Partial Label Learning with Competitive Label Noise 321, TITLE: https://openreview.net/forum?id=EUrxG8IBCrC AUTHORS: Yan Yan, Yuhong Guo HIGHLIGHT: In this paper, we consider a more realistic PLL scenario with competitive noise labels that are more difficult to distinguish from the true label than the random noise labels. 322, TITLE: ImaginaryNet: Learning Object Detectors without Real Images and Annotations https://openreview.net/forum?id=9MbhFHqrti9 AUTHORS: Minheng Ni, Zitong Huang, Kailai Feng, Wangmeng Zuo HIGHLIGHT: In this paper, we define a novel paradigm as Imaginary-Supervised Object Detection (ISOD), where no real images and manual annotations are used for training object detectors. To resolve this challenge, we propose ImaginaryNet, a framework to learn object detectors by combining pretrained language model as well as text-to-image synthesis models. 323, TITLE: Delving into Semantic Scale Imbalance https://openreview.net/forum?id=07tc5kKRIo AUTHORS: Yanbiao Ma, Licheng Jiao, Fang Liu, Yuxin Li, Shuyuan Yang, Xu Liu HIGHLIGHT: In this work, we define and quantify the semantic scale of classes, which is equivalent to the feature diversity of classes. 324, TITLE: Agnostic Learning of General ReLU Activation Using Gradient Descent https://openreview.net/forum?id=EnrY5TOrbQ AUTHORS: Pranjal Awasthi, Alex Tang, Aravindan Vijayaraghavan HIGHLIGHT: We provide a convergence analysis of gradient descent for the problem of agnostically learning a single ReLU function under Gaussian distributions. Neural-based classification rule learning for sequential data 325, TITLE: https://openreview.net/forum?id=7tJyBmu9iCj AUTHORS: Marine Collery, Philippe Bonnard, François Fages, Remy Kusters HIGHLIGHT: In this paper, we propose a novel differentiable fully interpretable method to discover both local and global patterns (i.e. catching a relative or absolute temporal dependency) for rule-based binary classification.

326, TITLE: Max-Margin Works while Large Margin Fails: Generalization without Uniform Convergence https://openreview.net/forum?id=n-hKHMzBgy AUTHORS: Margalit Glasgow, Colin Wei, Mary Wootters, Tengyu Ma HIGHLIGHT: Nagarajan and Kolter (2019) show that in certain simple linear and neural-network settings, any uniform convergence bound will be vacuous, leaving open the question of how to prove generalization in settings where UC fails. Our main contribution is proving novel generalization bounds in two such settings, one linear, and one non-linear. Fast and Precise: Adjusting Planning Horizon with Adaptive Subgoal Search 327. TITLE: https://openreview.net/forum?id=7JsGYvjE88d AUTHORS: Micha? Zawalski, Micha? Tyrolski, Konrad Czechowski, Damian Stachura, Piotr Pi?kos, Tomasz Odrzygó?d?, Yuhuai Wu, ?ukasz Kuci?ski, Piotr Mi?o? HIGHLIGHT: Complex reasoning problems contain states that vary in the computational cost required to determine a good action plan. Taking advantage of this property, we propose Adaptive Subgoal Search (AdaSubS), a search method that adaptively adjusts the planning horizon. 328. TITLE: Exploring the Limits of Differentially Private Deep Learning with Group-wise Clipping https://openreview.net/forum?id=oze0clVGPeX AUTHORS: Jiyan He, Xuechen Li, Da Yu, Huishuai Zhang, Janardhan Kulkarni, Yin Tat Lee, Arturs Backurs, Nenghai Yu, Jiang Bian We explore whether further improvements along the two axes are possible and provide affirmative answers HIGHLIGHT: leveraging two instantiations of \emph{group-wise clipping}. 329, TITLE: Towards Minimax Optimal Reward-free Reinforcement Learning in Linear MDPs https://openreview.net/forum?id=U9HW6vvNClg AUTHORS: Pihe Hu, Yu Chen, Longbo Huang HIGHLIGHT: We propose a novel algorithm LSVI-RFE under the linear MDP setting, where the transition probability and reward functions are linear in a feature mapping. Localized Randomized Smoothing for Collective Robustness Certification 330, TITLE: https://openreview.net/forum?id=-k7Lvk0GpBl AUTHORS: Jan Schuchardt, Tom Wollschläger, Aleksandar Bojchevski, Stephan Günnemann HIGHLIGHT: We propose a more general collective robustness certificate for all types of models and further show that this approach is beneficial for the larger class of softly local models, where each output is dependent on the entire input but assigns different levels of importance to different input regions (e.g. based on their proximity in the image). 331, TITLE: Towards Open Temporal Graph Neural Networks https://openreview.net/forum?id=N9Pk5iSCzAn AUTHORS: Kaituo Feng, Changsheng Li, Xiaolu Zhang, JUN ZHOU In this paper, we propose a general and principled learning approach for open temporal graphs, called OTGNet, HIGHLIGHT: with the goal of addressing the above two challenges. Efficiently Controlling Multiple Risks with Pareto Testing 332. TITLE: https://openreview.net/forum?id=cyg2YXn_BqF AUTHORS: Bracha Laufer-Goldshtein, Adam Fisch, Regina Barzilay, Tommi S. Jaakkola HIGHLIGHT: Building on recent results in distribution-free, finite-sample risk control for general losses, we propose Pareto Testing: a two-stage process which combines multi-objective optimization with multiple hypothesis testing. 333, TITLE: Bridge the Inference Gaps of Neural Processes via Expectation Maximization https://openreview.net/forum?id=A7v2DqLjZdq AUTHORS: Qi Wang, Marco Federici, Herke van Hoof The topic of inference suboptimality and an analysis of the NP from the optimization objective perspective has HIGHLIGHT. hardly been studied in earlier work. To fix this issue, we propose a surrogate objective of the target log-likelihood of the meta dataset within the expectation maximization framework. 334, TITLE: Discovering Policies with DOMiNO https://openreview.net/forum?id=kjkdzBW3b8p AUTHORS: Tom Zahavy, Yannick Schroecker, Feryal Behbahani, Kate Baumli, Sebastian Flennerhag, Shaobo Hou, Satinder Singh HIGHLIGHT: In this work we propose a Reinforcement Learning (RL) agent that can discover complex behaviours in a rich environment with a simple reward function. 335, TITLE: Neural Architecture Design and Robustness: A Dataset https://openreview.net/forum?id=p8coElgiSDw AUTHORS: Steffen Jung, Jovita Lukasik, Margret Keuper HIGHLIGHT: Thus, the aim of this paper is to facilitate better streamlined research on architectural design choices with respect to their impact on robustness as well as, for example, the evaluation of surrogate measures for robustness. We evaluate all these networks on a range of common adversarial attacks and corruption types and introduce a database on neural architecture design and robustness evaluations.

336, TITLE: A Unified Framework of Soft Threshold Pruning

https://openreview.net/forum?id=cCFqcrq0d8 AUTHORS: Yanqi Chen, Zhaofei Yu, Wei Fang, Zhengyu Ma, Xiawu Zheng, Yonghong Tian HIGHLIGHT: In this work, we reformulate soft threshold pruning as an implicit optimization problem solved using the *Iterative Shrinkage-Thresholding* Algorithm (ISTA), a classic method from the fields of sparse recovery and compressed sensing. 337, TITLE: Improving Out-of-distribution Generalization with Indirection Representations https://openreview.net/forum?id=0f-0I6RFAch AUTHORS: Kha Pham, Hung Le, Man Ngo, Truyen Tran HIGHLIGHT: We propose a generic module named Indirection Layer (InLay), which leverages indirection and data internal relationships to effectively construct symbolic indirect representations to improve out-of-distribution generalization capabilities of various neural architectures. Accelerating Guided Diffusion Sampling with Splitting Numerical Methods 338. TITLE: https://openreview.net/forum?id=F0KTk2plQzO AUTHORS: Suttisak Wizadwongsa, Supasorn Suwajanakorn HIGHLIGHT: On the contrary, we discover that the same techniques do not work for guided sampling, and little has been explored about its acceleration. This paper explores the culprit of this problem and provides a solution based on operator splitting methods, motivated by our key finding that high-order numerical methods are unsuitable for the conditional function. Batch Multivalid Conformal Prediction 339. TITLE: https://openreview.net/forum?id=Dk7QQp8jHEo AUTHORS: Christopher Jung, Georgy Noarov, Ramya Ramalingam, Aaron Roth HIGHLIGHT: We develop fast distribution-free conformal prediction algorithms for obtaining multivalid coverage on exchangeable data in the batch setting. 340, TITLE: Long-Tailed Learning Requires Feature Learning https://openreview.net/forum?id=S-h1oFv-mq AUTHORS: Thomas Laurent, James von Brecht, Xavier Bresson HIGHLIGHT: We propose a simple data model inspired from natural data such as text or images, and use it to study the importance of learning features in order to achieve good generalization. 341, TITLE: The Onset of Variance-Limited Behavior for Networks in the Lazy and Rich Regimes https://openreview.net/forum?id=JLINxPOVTh7 AUTHORS: Alexander Atanasov, Blake Bordelon, Sabarish Sainathan, Cengiz Pehlevan However, at a critical sample size \$P^*\$, the finite-width network generalization begins to worsen compared to HIGHLIGHT: the infinite width performance. In this work, we empirically study the transition from the infinite width behavior to this variancelimited regime as a function of sample size \$P\$ and network width \$N\$. 342, TITLE: On Accelerated Perceptrons and Beyond https://openreview.net/forum?id=fYzLpCsGZVf AUTHORS: Guanghui Wang, Rafael Hanashiro, Etash Kumar Guha, Jacob Abernethy HIGHLIGHT: There have been several recent works that managed to improve this rate by a quadratic factor, to \$\Omega(\sqrt {\log n}/gamma)\$, with more sophisticated algorithms. In this paper, we unify these existing results under one framework by showing that they can all be described through the lens of solving min-max problems using modern acceleration techniques, mainly through \emph {optimistic} online learning. 343. TITLE: Sequential Latent Variable Models for Few-Shot High-Dimensional Time-Series Forecasting https://openreview.net/forum?id=7C9aRX2nBf2 AUTHORS: Xiajun Jiang, Ryan Missel, Zhiyuan Li, Linwei Wang HIGHLIGHT: We present the very first step toward few-shot high-dimensional sequence forecasting by a Bayesian metalearning model that learns the process of learning latent dynamics that changes with the small number of observations that are available. 344, TITLE: Continual Unsupervised Disentangling of Self-Organizing Representations https://openreview.net/forum?id=ih0uFRFhaZZ AUTHORS: Zhiyuan Li, Xiajun Jiang, Ryan Missel, Prashnna Kumar Gyawali, Nilesh Kumar, Linwei Wang HIGHLIGHT: We argue that this is because existing approaches treat continually-arrived data independently, without considering how they are related based on the underlying semantic factors. We address this by a new generative model describing a topologically-connected mixture of spike-and-slab distributions in the latent space, learned end-to-end in a continual fashion via principled variational inference. 345, TITLE: Learning to Decompose Visual Features with Latent Textual Prompts https://openreview.net/forum?id=wtcud6HroZr AUTHORS: Feng Wang, Manling Li, Xudong Lin, Hairong Lv, Alex Schwing, Heng Ji HIGHLIGHT: To address them, we propose Decomposed Feature Prompting (DeFo). 346, TITLE: Implicit Bias in Leaky ReLU Networks Trained on High-Dimensional Data https://openreview.net/forum?id=JpbLyEI5EwW AUTHORS: Spencer Frei, Gal Vardi, Peter Bartlett, Nathan Srebro, Wei Hu

HIGHLIGHT: In this work, we investigate the implicit bias of gradient flow and gradient descent in two-layer fully-connected neural networks with leaky ReLU activations when the training data are nearly-orthogonal, a common property of high-dimensional data.

347, TITLE: Is Attention All That NeRF Needs? https://openreview.net/forum?id=xE-LtsE-xx AUTHORS: Mukund Varma T, Peihao Wang, Xuxi Chen, Tianlong Chen, Subhashini Venugopalan, Zhangyang Wang HIGHLIGHT: We present Generalizable NeRF Transformer (GNT), a transformer-based architecture that reconstructs Neural Radiance Fields (NeRFs) and learns to renders novel views on the fly from source views. 348. TITLE: Squeeze Training for Adversarial Robustness https://openreview.net/forum?id=Z tmYu060Kr AUTHORS: Qizhang Li, Yiwen Guo, Wangmeng Zuo, Hao Chen In this paper, we highlight that some collaborative examples, nearly perceptually indistinguishable from both HIGHLIGHT: adversarial and benign examples yet show extremely lower prediction loss, can be utilized to enhance adversarial training. 349, TITLE: Domain Generalization via Heckman-type Selection Models https://openreview.net/forum?id=fk7RbGibe1 AUTHORS: Hyungu Kahng, Hyungrok Do, Judy Zhong HIGHLIGHT: In this paper, we formulate DG as a sample selection problem where each domain is sampled from a common underlying population through non-random sampling probabilities that correlate with both the features and the outcome. 350, TITLE: Context-enriched molecule representations improve few-shot drug discovery https://openreview.net/forum?id=XrMWUuEevr AUTHORS: Johannes Schimunek, Philipp Seidl, Lukas Friedrich, Daniel Kuhn, Friedrich Rippmann, Sepp Hochreiter, Günter Klambauer HIGHLIGHT: We introduce a new method for few-shot drug discovery. Do We Really Need Complicated Model Architectures For Temporal Networks? 351, TITLE: https://openreview.net/forum?id=avPPc0SvLv1 AUTHORS: Weilin Cong, Si Zhang, Jian Kang, Baichuan Yuan, Hao Wu, Xin Zhou, Hanghang Tong, Mehrdad Mahdavi HIGHLIGHT: In this paper, we propose \oure, a conceptually and technically simple architecture that consists of three components: \circled {1} a \emph {link-encoder} that is only based on multi-layer perceptrons (MLP) to summarize the information from temporal links, \circled {2} a \emph{node-encoder} that is only based on neighbor mean-pooling to summarize node information, and \circled {3} an MLP-based \emph{link classifier} that performs link prediction based on the outputs of the encoders. Accurate Neural Training with 4-bit Matrix Multiplications at Standard Formats 352, TITLE: https://openreview.net/forum?id=yTbNYYcopd AUTHORS: Brian Chmiel, Ron Banner, Elad Hoffer, Hilla Ben-Yaacov, Daniel Soudry HIGHLIGHT: Specifically, we examine the importance of having unbiased quantization in quantized neural network training, where to maintain it, and how to combine it with logarithmic quantization. 353, TITLE: Learning with Auxiliary Activation for Memory-Efficient Training https://openreview.net/forum?id=YgC62m4CY3r AUTHORS: Sunghyeon Woo, Dongsuk Jeon HIGHLIGHT: In this work, we propose a new learning rule which significantly reduces memory requirements while closely matching the performance of backpropagation. 354, TITLE: TRANSFORMER-PATCHER: ONE MISTAKE WORTH ONE NEURON https://openreview.net/forum?id=4oYUGeGBPm AUTHORS: Zeyu Huang, Yikang Shen, Xiaofeng Zhang, Jie Zhou, Wenge Rong, Zhang Xiong Thus a preferable solution is to rectify the mistakes as soon as they appear nonstop. Therefore, we extend the HIGHLIGHT: existing ME into the Sequential Model Editing (SME) to help develop more practical editing methods. 355, TITLE: An Additive Instance-Wise Approach to Multi-class Model Interpretation https://openreview.net/forum?id=5OygDd-4Eeh AUTHORS: Vy Vo, Van Nguyen, Trung Le, Quan Hung Tran, Reza Haf, Seyit Camtepe, Dinh Phung HIGHLIGHT: This work exploits the strengths of both methods and proposes a framework for learning local explanations simultaneously for multiple target classes. 356, TITLE: Guiding continuous operator learning through Physics-based boundary constraints https://openreview.net/forum?id=gfWNItGOES6 AUTHORS: Nadim Saad, Gaurav Gupta, Shima Alizadeh, Danielle C. Maddix HIGHLIGHT: In this work, we propose Boundary enforcing Operator Network (BOON) that enables the BC satisfaction of neural operators by making structural changes to the operator kernel. 357, TITLE: Making Substitute Models More Bayesian Can Enhance Transferability of Adversarial Examples https://openreview.net/forum?id=bjPPypbLre AUTHORS: Qizhang Li, Yiwen Guo, Wangmeng Zuo, Hao Chen

HIGHLIGHT: In this paper, by contrast, we opt for the diversity in substitute models and advocate to attack a Bayesian model for achieving desirable transferability. 358, TITLE: Sublinear Algorithms for Kernel Matrices via Kernel Density Estimation https://openreview.net/forum?id=74A-FDAyiL AUTHORS: Ainesh Bakshi, Piotr Indyk, Praneeth Kacham, Sandeep Silwal, Samson Zhou HIGHLIGHT: We give a framework for using recently developed tools for kernel density estimation to solve downstream kernel problems in sub-quadratic time. 359, TITLE: Choreographer: Learning and Adapting Skills in Imagination https://openreview.net/forum?id=PhkWyijGi5b AUTHORS: Pietro Mazzaglia, Tim Verbelen, Bart Dhoedt, Alexandre Lacoste, Sai Rajeswar HIGHLIGHT: Furthermore, it is unclear how to leverage the learned skill behaviors for adapting to downstream tasks in a data-efficient manner. We present Choreographer, a model-based agent that exploits its world model to learn and adapt skills in imagination. 360, TITLE: DropIT: Dropping Intermediate Tensors for Memory-Efficient DNN Training https://openreview.net/forum?id=Kn6i2BZW69w AUTHORS: Joya Chen, Kai Xu, Yuhui Wang, Yifei Cheng, Angela Yao HIGHLIGHT: The bulk of memory is occupied by caching intermediate tensors for gradient computation in the backward pass. We propose a novel method to reduce this footprint - Dropping Intermediate Tensors (DropIT). 361, TITLE: Exploring Temporally Dynamic Data Augmentation for Video Recognition https://openreview.net/forum?id=fxizKOdw9wb AUTHORS: Taeoh Kim, Jinhyung Kim, Minho Shim, Sangdoo Yun, Myunggu Kang, Dongyoon Wee, Sangyoun Lee HIGHLIGHT: These variations should be generated as diverse as possible using fewer additional hyper-parameters during training. Through this motivation, we propose a simple yet effective video data augmentation framework, DynaAugment. Computational Language Acquisition with Theory of Mind 362, TITLE: https://openreview.net/forum?id=C2ulri4duIs AUTHORS: Andy Liu, Hao Zhu, Emmy Liu, Yonatan Bisk, Graham Neubig Drawing inspiration from the modern operationalized versions of ToM implemented in Rabinowitz et al. (2018) HIGHLIGHT: and Zhu et al. (2021), we build language-learning agents equipped with ToM, and measure its effects on the learning process. 363, TITLE: Mind's Eye: Grounded Language Model Reasoning through Simulation https://openreview.net/forum?id=4rXMRuoJlai AUTHORS: Ruibo Liu, Jason Wei, Shixiang Shane Gu, Te-Yen Wu, Soroush Vosoughi, Claire Cui, Denny Zhou, Andrew M. Dai HIGHLIGHT: We present Mind's Eye, a paradigm to ground language model reasoning in the physical world. 364, TITLE: Language Models are Realistic Tabular Data Generators https://openreview.net/forum?id=cEygmQNOeI AUTHORS: Vadim Borisov, Kathrin Sessler, Tobias Leemann, Martin Pawelczyk, Gjergji Kasneci HIGHLIGHT: To this end, we propose GReaT (Generation of Realistic Tabular data), which exploits an auto-regressive generative LLM to sample synthetic and yet highly realistic tabular data. Aligning Model and Macaque Inferior Temporal Cortex Representations Improves Model-to-Human Behavioral 365, TITLE: Alignment and Adversarial Robustness https://openreview.net/forum?id=SMYdcXjJh1q AUTHORS: Joel Dapello, Kohitij Kar, Martin Schrimpf, Robert Baldwin Geary, Michael Ferguson, David Daniel Cox, James J. DiCarlo HIGHLIGHT: We conducted chronic, large-scale multi-electrode recordings across the IT cortex in six non-human primates (rhesus macaques). 366, TITLE: simpleKT: A Simple But Tough-to-Beat Baseline for Knowledge Tracing https://openreview.net/forum?id=9HiGqC9C-KA AUTHORS: Zitao Liu, Qiongqiong Liu, Jiahao Chen, Shuyan Huang, Weiqi Luo Furthermore, due to the lack of standardized evaluation protocol \citep {liu2022pykt}, there is no widely agreed HIGHLIGHT: KT baselines and published experimental comparisons become inconsistent and self-contradictory, i.e., the reported AUC scores of DKT on ASSISTments2009 range from 0.721 to 0.821 \citep {minn2018deep,yeung2018addressing}. Therefore, in this paper, we provide a strong but simple baseline method to deal with the KT task named \textsc{simpleKT}. 367, TITLE: Time Will Tell: New Outlooks and A Baseline for Temporal Multi-View 3D Object Detection https://openreview.net/forum?id=H3HcEJA2Um Jinhyung Park, Chenfeng Xu, Shijia Yang, Kurt Keutzer, Kris M. Kitani, Masayoshi Tomizuka, Wei Zhan AUTHORS: HIGHLIGHT: Building on our investigation, we propose to generate a cost volume from a long history of image observations, compensating for the coarse but efficient matching resolution with a more optimal multi-view matching setup.

368, TITLE: Massively Scaling Heteroscedastic Classifiers https://openreview.net/forum?id=sIoED-yPK9l AUTHORS: Mark Collier, Rodolphe Jenatton, Basil Mustafa, Neil Houlsby, Jesse Berent, Effrosyni Kokiopoulou HIGHLIGHT: We propose HET-XL, a heteroscedastic classifier whose parameter count when compared to a standard classifier scales independently of the number of classes. 369, TITLE: Interpretable Single/Multi-label Text Classification with Unsupervised Constituent-label alignments https://openreview.net/forum?id=MLJ5TF5FtXH AUTHORS: Xiang Hu, XinYu KONG, Kewei Tu HIGHLIGHT: Meanwhile, symbolic probabilistic models function with clear interpretability, but how to combine them with neural networks to enhance their performance remains to be explored. In this paper, we try to marry these two systems for text classification via structured language models. 370, TITLE: Transformer Meets Boundary Value Inverse Problems https://openreview.net/forum?id=HnlCZATopvr AUTHORS: Ruchi Guo, Shuhao Cao, Long Chen HIGHLIGHT: A Transformer-based deep direct sampling method is proposed for solving a class of boundary value inverse problem. DAG Learning via Sparse Relaxations 371. TITLE: https://openreview.net/forum?id=m9LCdYgN8-6 AUTHORS: Valentina Zantedeschi, Luca Franceschi, Jean Kaddour, Matt Kusner, Vlad Niculae HIGHLIGHT: We propose a continuous optimization framework for discovering a latent directed acyclic graph (DAG) from observational data. 372, TITLE: Soft Neighbors are Positive Supporters in Contrastive Visual Representation Learning https://openreview.net/forum?id=19vM PaUKz Chongjian GE, Jiangliu Wang, Zhan Tong, Shoufa Chen, Yibing Song, Ping Luo AUTHORS: In this paper, we rethink the instance discrimination framework and find the binary instance labeling HIGHLIGHT: insufficient to measure correlations between different samples. 373. TITLE: Finding the global semantic representation in GAN through Fréchet Mean https://openreview.net/forum?id=9ImtNIZ7bYx AUTHORS: Jaewoong Choi, Geonho Hwang, Hyunsoo Cho, Myungjoo Kang HIGHLIGHT: In other words, in this disentangled space, there exists the global semantic basis as a vector space where each basis component describes one attribute of generated images. In this paper, we propose an unsupervised method for finding this global semantic basis in the intermediate latent space in GANs. 374, TITLE: MARS: Meta-learning as Score Matching in the Function Space https://openreview.net/forum?id=WAgXmT8BeRj AUTHORS: Krunoslav Lehman Pavasovic, Jonas Rothfuss, Andreas Krause HIGHLIGHT: As a result, existing approaches resort to meta-learning restrictive diagonal Gaussian priors, severely limiting their expressiveness and performance. To circumvent these issues, we approach meta-learning through the lens of functional Bayesian neural network inference which views the prior as a stochastic process and performs inference in the function space. On the effectiveness of out-of-distribution data in self-supervised long-tail learning. 375, TITLE: https://openreview.net/forum?id=v8JIQdiN9Sh AUTHORS: Jianhong Bai, Zuozhu Liu, Hualiang Wang, Jin Hao, YANG FENG, Huanpeng Chu, Haoji Hu HIGHLIGHT: In this paper, we propose an alternative but easy-to-use and effective solution, \textbf{C} ontrastive with \textbf{O}ut-of-distribution (OOD) data for \textbf{L}ong-\textbf{T}ail learning (COLT), which can effectively exploit OOD data to dynamically re-balance the feature space. Faster Gradient-Free Methods for Escaping Saddle Points 376, TITLE: https://openreview.net/forum?id=KDhFkA6MQsW AUTHORS: Hualin Zhang, Bin Gu HIGHLIGHT: In this paper, we study the case when calculations of explicit gradients are expensive or even infeasible, and only function values are accessible. 377. TITLE: A View From Somewhere: Human-Centric Face Representations https://openreview.net/forum?id=mMaInr0r0c AUTHORS: Jerone Theodore Alexander Andrews, Przemyslaw Joniak, Alice Xiang HIGHLIGHT: We propose to implicitly learn a set of continuous face-varying dimensions, without ever asking an annotator to explicitly categorize a person. 378, TITLE: Dynamical systems embedding with a physics-informed convolutional network https://openreview.net/forum?id=z9C5dGip90 AUTHORS: Matt Ricci, Noa Moriel, Zoe Piran, Mor Nitzan HIGHLIGHT: Here, we propose, \texttt{phase2vec}, an embedding method that learns high-quality, physically-meaningful representations of dynamical systems without supervision.

379, TITLE: Mind the Pool: Convolutional Neural Networks Can Overfit Input Size https://openreview.net/forum?id=cWmtUcsYC3V
AUTHORS: Bilal Alsallakh, David Yan, Narine Kokhlikyan, Vivek Miglani, Orion Reblitz-Richardson, Pamela Bhattacharva HIGHLIGHT: This issue is inherent to pooling arithmetic, with standard downsampling layers playing a major role in favoring certain input sizes and skewing the weights accordingly. We present a solution to this problem by depriving these layers from the arithmetic cues they use to overfit the input size. Training-Free Structured Diffusion Guidance for Compositional Text-to-Image Synthesis 380, TITLE: https://openreview.net/forum?id=PUIqjT4rzq7 AUTHORS: Weixi Feng, Xuehai He, Tsu-Jui Fu, Varun Jampani, Arjun Reddy Akula, Pradyumna Narayana, Sugato Basu, Xin Eric Wang, William Yang Wang HIGHLIGHT: In this work, we strive to improve the compositional skills of existing large-scale T2I models, specifically more accurate attribute binding and better image compositions. A Theory of Dynamic Benchmarks 381, TITLE: https://openreview.net/forum?id=i8L9qoeZOS AUTHORS: Ali Shirali, Rediet Abebe, Moritz Hardt HIGHLIGHT: In contrast to an extensive theoretical and empirical study of the static setting, the dynamic setting lags behind due to limited empirical studies and no apparent theoretical foundation to date. Responding to this deficit, we initiate a theoretical study of dynamic benchmarking. 382, TITLE: LAVA: Data Valuation without Pre-Specified Learning Algorithms https://openreview.net/forum?id=JJuP86nBl4g AUTHORS: Hoang Anh Just, Feiyang Kang, Tianhao Wang, Yi Zeng, Myeongseob Ko, Ming Jin, Ruoxi Jia HIGHLIGHT: This work leapfrogs over the current limits of data valuation methods by introducing a new framework that can value training data in a way that is oblivious to the downstream learning algorithm. Multi-level Protein Structure Pre-training via Prompt Learning 383, TITLE: https://openreview.net/forum?id=XGagtiJ8XC Zeyuan Wang, Qiang Zhang, Shuang-Wei HU, Haoran Yu, Xurui Jin, Zhichen Gong, Huajun Chen AUTHORS: HIGHLIGHT: Considering protein sequences can determine multi-level structures, in this paper, we aim to realize the comprehensive potential of protein sequences for function prediction. 384, TITLE: Towards Robustness Certification Against Universal Perturbations https://openreview.net/forum?id=7GEvPKxjtt AUTHORS: Yi Zeng, Zhouxing Shi, Ming Jin, Feiyang Kang, Lingjuan Lyu, Cho-Jui Hsieh, Ruoxi Jia In this paper, we investigate the problem of certifying neural network robustness against universal perturbations HIGHLIGHT: (UPs), which have been widely used in universal adversarial attacks and backdoor attacks. 385, TITLE: AutoGT: Automated Graph Transformer Architecture Search https://openreview.net/forum?id=GcM7qfl5zY AUTHORS: Zizhao Zhang, Xin Wang, Chaoyu Guan, Ziwei Zhang, Haoyang Li, Wenwu Zhu HIGHLIGHT: In this paper, we study the problem of automated graph Transformer, for the first time. Blurring Diffusion Models 386, TITLE: https://openreview.net/forum?id=OjDkC57x5sz AUTHORS: Emiel Hoogeboom, Tim Salimans HIGHLIGHT: Here, we show that blurring can equivalently be defined through a Gaussian diffusion process with nonisotropic noise. 387, TITLE: Adversarial Imitation Learning with Preferences https://openreview.net/forum?id=bhfp5GlDtGe AUTHORS: Aleksandar Taranovic, Andras Gabor Kupcsik, Niklas Freymuth, Gerhard Neumann HIGHLIGHT: In this paper we propose a novel method for policy learning that incorporates two different feedback types, namely \emph{demonstrations} and \emph{preferences}. 388, TITLE: Almost Linear Constant-Factor Sketching for \$\ell 1\$ and Logistic Regression https://openreview.net/forum?id=gu-SC0dpkvw Alexander Munteanu, Simon Omlor, David Woodruff AUTHORS: HIGHLIGHT: We give the first constant factor approximate sketches for 11 and logistic regression in a turnstile stream with almost linear sketching dimension that result in an efficient optimization problem in the sketch space. 389, TITLE: On the Soft-Subnetwork for Few-Shot Class Incremental Learning https://openreview.net/forum?id=z57WK51GeHd AUTHORS: Haeyong Kang, Jaehong Yoon, Sultan Rizky Hikmawan Madjid, Sung Ju Hwang, Chang D. Yoo Inspired by Regularized Lottery Ticket Hypothesis, which states that competitive smooth (non-binary) HIGHLIGHT: subnetworks exist within a dense network, we propose a few-shot class-incremental learning method referred to as Soft-SubNetworks (SoftNet).

390, TITLE: Efficient Offline Policy Optimization with a Learned Model https://openreview.net/forum?id=Yt-yM-JbYFO

AUTHORS: HIGHLIGHT: benchmark.	Zichen Liu, Siyi Li, Wee Sun Lee, Shuicheng YAN, Zhongwen Xu We propose a regularized one-step model-based method that outperforms MuZero Unplugged on Atari
391, TITLE:	Bispectral Neural Networks
https://openreview.net	et/forum?id=xnsg4pfKb7
AUTHORS:	Sophia Sanborn, Christian A Shewmake, Bruno Olshausen, Christopher J. Hillar
HIGHLIGHT:	We present a neural network architecture, Bispectral Neural Networks (BNNs) for learning representations that
are invariant to the act	etions of compact commutative groups on the space over which a signal is defined.
392, TITLE:	Learning Group Importance using the Differentiable Hypergeometric Distribution
https://openreview.ne	t/forum?id=75O7S_L4oY
AUTHORS:	Thomas M. Sutter, Laura Manduchi, Alain Ryser, Julia E Vogt
HIGHLIGHT:	In this work, we propose the differentiable hypergeometric distribution.
393, TITLE:	TTN: A Domain-Shift Aware Batch Normalization in Test-Time Adaptation
https://openreview.net	et/forum?id=EQfeudmWLQ
AUTHORS:	Hyesu Lim, Byeonggeun Kim, Jaegul Choo, Sungha Choi
HIGHLIGHT:	In this paper, we identify that CBN and TBN are in a trade-off relationship and present a new test-time
normalization (TTN)	method that interpolates the statistics by adjusting the importance between CBN and TBN according to the
domain-shift sensitiv	ity of each BN layer.
394, TITLE:	ManiSkill2: A Unified Benchmark for Generalizable Manipulation Skills
https://openreview.net	t/forum?id=b_CQDy9vrD1
AUTHORS:	Jiayuan Gu, Fanbo Xiang, Zhan Ling, Xinyue Wei, Xiqiang Liu, Xuanlin Li, Rui Chen, Stone Tao, Tongzhou
Mu, Pengwei Xie, Yu	unchao Yao, Yihe Tang, Xiaodi Yuan, Zhiao Huang, Hao Su
HIGHLIGHT:	To this end, we present ManiSkill2, the next generation of the SAPIEN ManiSkill benchmark, to address
critical pain points of	ten encountered by researchers when using benchmarks for generalizable manipulation skills.
395, TITLE:	MixPro: Data Augmentation with MaskMix and Progressive Attention Labeling for Vision Transformer
https://openreview.net	et/forum?id=dRjWsd3gwsm
AUTHORS:	Qihao Zhao, Yangyu Huang, Wei Hu, Fan Zhang, Jun Liu
HIGHLIGHT:	To address the aforementioned issues, we propose MaskMix and Progressive Attention Labeling (PAL) in
image and label space	e, respectively.
396, TITLE:	Flow Matching for Generative Modeling
https://openreview.net	t/forum?id=PqvMRDCJT9t
AUTHORS:	Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, Matthew Le
HIGHLIGHT:	We introduce a new paradigm for generative modeling built on Continuous Normalizing Flows (CNFs),
allowing us to train C	cNFs at unprecedented scale.
397, TITLE:	Compositional Prompt Tuning with Motion Cues for Open-vocabulary Video Relation Detection
https://openreview.net	et/forum?id=mE91GkXYipg
AUTHORS:	Kaifeng Gao, Long Chen, Hanwang Zhang, Jun Xiao, Qianru Sun
HIGHLIGHT:	In this paper, we propose compositional prompt tuning with motion cues: an extended prompt tuning paradigm
for compositional pre	edictions of video data.
398, TITLE:	Out-of-Distribution Detection and Selective Generation for Conditional Language Models
https://openreview.net	et/forum?id=kJUS5nD0vPB
AUTHORS:	Jie Ren, Jiaming Luo, Yao Zhao, Kundan Krishna, Mohammad Saleh, Balaji Lakshminarayanan, Peter J Liu
HIGHLIGHT:	Furthermore, the space of potential low-quality outputs is larger as arbitrary text can be generated and it is
important to know w	hen to trust the generated output. We present a highly accurate and lightweight OOD detection method for CLMs,
and demonstrate its e	ffectiveness on abstractive summarization and translation.
399, TITLE:	Budgeted Training for Vision Transformer
https://openreview.ne	et/forum?id=sVzBN-DIJRi
AUTHORS:	zhuofan xia, Xuran Pan, Xuan Jin, Yuan He, Hui Xue', Shiji Song, Gao Huang
HIGHLIGHT:	In this paper, we address the problem by proposing a framework that enables the training process under any
training budget, while	e achieving competitive model performances.
400, TITLE:	ODAM: Gradient-based Instance-Specific Visual Explanations for Object Detection
https://openreview.net	et/forum?id=kJWcI39kXY
AUTHORS:	Chenyang ZHAO, Antoni B. Chan
HIGHLIGHT:	We propose the Gradient-weighted Object Detector Activation Mapping (Grad-ODAM), a visualized
explanation techniqu	e for interpreting the predictions of object detectors.
401, TITLE: https://openreview.net AUTHORS: Zhang	Regression with Label Differential Privacy t/forum?id=h9O0wsmL-cT Badih Ghazi, Pritish Kamath, Ravi Kumar, Ethan Leeman, Pasin Manurangsi, Avinash Varadarajan, Chiyuan

HIGHLIGHT: We prove that the optimal mechanism takes the form of a "randomized response on bins", and propose an efficient algorithm for finding the optimal bin values. 402, TITLE: Boosting Adversarial Transferability using Dynamic Cues https://openreview.net/forum?id=SZynfVLGd5 AUTHORS: Muzammal Naseer, Ahmad Mahmood, Salman Khan, Fahad Khan HIGHLIGHT: In this work, we induce dynamic cues within the image models without sacrificing their original performance on images. 403, TITLE: Towards Inferential Reproducibility of Machine Learning Research https://openreview.net/forum?id=li4GQCQWkv AUTHORS: Michael Hagmann, Philipp Meier, Stefan Riezler HIGHLIGHT: In this paper, we propose to shift from the goal of duplicating a SOTA training result without any changes to a new type of reproducibility called inferential reproducibility that treats performance variation depending on data characteristics, metaparameter settings, and their interactions as an inherent and interesting feature of non-deterministic deep learning, not as a bug that needs to be resolved. 404, TITLE: Mole-BERT: Rethinking Pre-training Graph Neural Networks for Molecules https://openreview.net/forum?id=jevY-DtiZTR Jun Xia, Chengshuai Zhao, Bozhen Hu, Zhangyang Gao, Cheng Tan, Yue Liu, Siyuan Li, Stan Z. Li AUTHORS: HIGHLIGHT: We explain the negative transfer in molecular graph pre-training and develop two novel pre-training strategies to alleviate this issue. 405, TITLE: The KFIoU Loss for Rotated Object Detection https://openreview.net/forum?id=qUKsCztWlKq AUTHORS: Xue Yang, Yue Zhou, Gefan Zhang, Jirui Yang, Wentao Wang, Junchi Yan, XIAOPENG ZHANG, Qi Tian HIGHLIGHT: In this paper, we propose an effective approximate SkewIoU loss based on Gaussian modeing and Kalman filter, which mainly consists of two items. 406. TITLE: PowerOuant: Automorphism Search for Non-Uniform Ouantization https://openreview.net/forum?id=s1KljJpAukm AUTHORS: Edouard YVINEC, Arnaud Dapogny, Matthieu Cord, Kevin Bailly HIGHLIGHT: In this paper, we identity the uniformity of the quantization operator as a limitation of existing approaches, and propose a data-free non-uniform method. 407, TITLE: Rethinking skip connection model as a learnable Markov chain https://openreview.net/forum?id=yQdBtFfleh6 AUTHORS: Chen Dengsheng, Jie Hu, Wenwen Qiang, Xiaoming Wei, Enhua Wu In this work, we deep dive into the model's behaviors with skip connections which can be formulated as a HIGHLIGHT: learnable Markov chain. 408. TITLE: Effects of Graph Convolutions in Multi-layer Networks https://openreview.net/forum?id=P-73JPgRs0R AUTHORS: Aseem Baranwal, Kimon Fountoulakis, Aukosh Jagannath HIGHLIGHT: We present a rigorous theoretical understanding of the effects of graph convolutions in multi-layer networks. 409, TITLE: Addressing Parameter Choice Issues in Unsupervised Domain Adaptation by Aggregation https://openreview.net/forum?id=M95oDwJXayG AUTHORS: Marius-Constantin Dinu, Markus Holzleitner, Maximilian Beck, Hoan Duc Nguyen, Andrea Huber, Hamid Eghbal-zadeh, Bernhard A. Moser, Sergei Pereverzyev, Sepp Hochreiter, Werner Zellinger HIGHLIGHT: While several heuristics exist that follow this strategy, methods are still missing that rely on thorough theories for bounding the target error. In this turn, we propose a method that extends weighted least squares to vector-valued functions, e.g., deep neural networks. 410, TITLE: Bayesian semi-supervised learning with a principled likelihood from a generative model of data curation https://openreview.net/forum?id=zOHQGKO3WGY AUTHORS: Stoil Krasimirov Ganev, Laurence Aitchison We are thus able to introduce Bayesian SSL, which gives considerable improvements over standard SSL in the HIGHLIGHT: setting of 40 labelled points on CIFAR-10, with performance of \$92.2\pm 0.3\%\$ vs \$88.6\%\$ in the original FixMatch paper. 411, TITLE: Spherical Sliced-Wasserstein https://openreview.net/forum?id=jXQ0ipgMdU AUTHORS: Clément Bonet, Paul Berg, Nicolas Courty, François Septier, Lucas Drumetz, Minh Tan Pham HIGHLIGHT: We propose a SW discrepancy on the sphere using only tools intrinsic to the manifold. Scenario-based Question Answering with Interacting Contextual Properties 412, TITLE: https://openreview.net/forum?id=tPrRs6YB2P AUTHORS: Haitian Sun, William W. Cohen, Ruslan Salakhutdinov HIGHLIGHT: Although understanding the relationship between conditions is crucial for solving this challenging QA task, limited work has been done so far in modeling this. In this paper, we propose the T-Reasoner model, which solves this problem with

three jointly learned modules: an entailment module which checks whether a condition has been satisfied by the scenario, a decoding module which locates eligible answers from documents, and a reasoning module which infers the relationship between conditions and performs a reasoning step to determine the logically consistent answers and identify missing conditions.

413. TITLE: CktGNN: Circuit Graph Neural Network for Electronic Design Automation https://openreview.net/forum?id=NE2911Kq1sp AUTHORS: Zehao Dong, Weidong Cao, Muhan Zhang, Dacheng Tao, Yixin Chen, Xuan Zhang HIGHLIGHT: By recognizing the graph nature of circuits, this paper presents a Circuit Graph Neural Network (CktGNN) that simultaneously automates the circuit topology generation and device sizing based on the encoder-dependent optimization subroutines. 414, TITLE: Prompt-to-Prompt Image Editing with Cross-Attention Control https://openreview.net/forum?id=_CDixzkzeyb AUTHORS: Amir Hertz, Ron Mokady, Jay Tenenbaum, Kfir Aberman, Yael Pritch, Daniel Cohen-or HIGHLIGHT: In this paper, we pursue an intuitive prompt-to-prompt editing framework, where the edits are controlled by text only. 415, TITLE: Efficient Out-of-Distribution Detection based on In-Distribution Data Patterns Memorization with Modern Hopfield Energy https://openreview.net/forum?id=KkazG4lgKL AUTHORS: Jinsong Zhang, Qiang Fu, Xu Chen, Lun Du, Zelin Li, Gang Wang, xiaoguang Liu, Shi Han, Dongmei Zhang HIGHLIGHT: Unlike existing OOD methods refining the confidence estimation procedure from output logits with handpicked hyperparameters, we propose a new store-then-compare paradigm. 416, TITLE: CodeT: Code Generation with Generated Tests https://openreview.net/forum?id=ktrw68Cmu9c AUTHORS: Bei Chen, Fengji Zhang, Anh Nguyen, Daoguang Zan, Zeqi Lin, Jian-Guang Lou, Weizhu Chen HIGHLIGHT: In this paper, we propose a novel method, CodeT, that leverages the same pre-trained language models to automatically generate test cases for the code samples, thus reducing the human effort and increasing the coverage of the test scenarios. 417, TITLE: Does Deep Learning Learn to Abstract? A Systematic Probing Framework https://openreview.net/forum?id=QB1dMPEXau5 AUTHORS: Shengnan An, Zeqi Lin, Bei Chen, Qiang Fu, Nanning Zheng, Jian-Guang Lou HIGHLIGHT: In this paper, we introduce a systematic probing framework to explore the abstraction capability of deep learning models from a transferability perspective. 418, TITLE: Exact Group Fairness Regularization via Classwise Robust Optimization https://openreview.net/forum?id=Q-WfHzmiG9m AUTHORS: Sangwon Jung, Taeeon Park, Sanghyuk Chun, Taesup Moon HIGHLIGHT: To that end, we propose a principled method that indeed can incorporate an \$\textit {exact}\$ form of a welljustified group fairness metric, Difference of Conditional Accuracy (DCA), as a regularizer using a \$\textit{classwise}\$ distributionally robust optimization (DRO) framework. 419, TITLE: Weighted Clock Logic Point Process https://openreview.net/forum?id=YfUICnZMwk7 AUTHORS: Ruixuan Yan, Yunshi Wen, Debarun Bhattacharjya, Ronny Luss, Tengfei Ma, Achille Fokoue, Anak Agung Julius HIGHLIGHT: We present a novel framework for modeling temporal point processes called clock logic neural networks (CLNN) which learn weighted clock logic (wCL) formulas as interpretable temporal rules by which some events promote or inhibit other events. 420, TITLE: DiffEdit: Diffusion-based semantic image editing with mask guidance https://openreview.net/forum?id=31ge0p5o-M-AUTHORS: Guillaume Couairon, Jakob Verbeek, Holger Schwenk, Matthieu Cord HIGHLIGHT: In this article, we propose DiffEdit, a method to take advantage of text-conditioned diffusion models for the task of semantic image editing, where the goal is to edit an image based on a text query. 421, TITLE: Human alignment of neural network representations https://openreview.net/forum?id=ReDQ1OUQR0X AUTHORS: Lukas Muttenthaler, Jonas Dippel, Lorenz Linhardt, Robert A. Vandermeulen, Simon Kornblith HIGHLIGHT: In this paper, we investigate the factors that affect alignment between the representations learned by neural networks and human concept representations. 422, TITLE: SMART: Self-supervised Multi-task pretrAining with contRol Transformers https://openreview.net/forum?id=9piH3Hg8QEf AUTHORS: Yanchao Sun, Shuang Ma, Ratnesh Madaan, Rogerio Bonatti, Furong Huang, Ashish Kapoor HIGHLIGHT: The challenge becomes combinatorially more complex if we want to pretrain representations amenable to a large variety of tasks. To tackle this problem, in this work, we formulate a general pretraining-finetuning pipeline for sequential

decision making, under which we propose a generic pretraining framework \textit {Self-supervised Multi-task pretrAining with contRol Transformer (SMART)}.

423, TITLE: Are More Layers Beneficial to Graph Transformers? https://openreview.net/forum?id=uagC-X9XMi8 AUTHORS: Haiteng Zhao, Shuming Ma, Dongdong Zhang, Zhi-Hong Deng, Furu Wei HIGHLIGHT: To this end, we propose a novel graph transformer model named DeepGraph that explicitly employs substructure tokens in the encoded representation, and applies local attention on related nodes to obtain substructure based attention encoding. 424, TITLE: A Universal 3D Molecular Representation Learning Framework https://openreview.net/forum?id=6K2RM6wVqKu AUTHORS: Gengmo Zhou, Zhifeng Gao, Qiankun Ding, Hang Zheng, Hongteng Xu, Zhewei Wei, Linfeng Zhang, Guolin Ke HIGHLIGHT: In this paper, we propose a universal 3D MRL framework that significantly enlarges the representation ability and application scope of MRL schemes. 425. TITLE: Accurate Bayesian Meta-Learning by Accurate Task Posterior Inference https://openreview.net/forum?id=sb-IkS8DQw2 AUTHORS: Michael Volpp, Philipp Dahlinger, Philipp Becker, Christian Daniel, Gerhard Neumann HIGHLIGHT: Prior work studies a range of architectural modifications to boost performance, such as attentive computation paths or improved context aggregation schemes, while the influence of the VI scheme remains under-explored. We aim to bridge this gap by introducing GMM-NP, a novel BML model, which builds on recent work that enables highly accurate, full-covariance Gaussian mixture (GMM) TP approximations by combining VI with natural gradients and trust regions. 426. TITLE: Q-Pensieve: Boosting Sample Efficiency of Multi-Objective RL Through Memory Sharing of Q-Snapshots https://openreview.net/forum?id=AwWaBXLIJE AUTHORS: Wei Hung, Bo Kai Huang, Ping-Chun Hsieh, Xi Liu We show that \$Q\$-Pensieve can be naturally integrated with soft policy iteration with convergence guarantee. HIGHLIGHT: To substantiate this concept, we propose the technique of \$Q\$ replay buffer, which stores the learned \$Q\$-networks from the past iterations, and arrive at a practical actor-critic implementation. 427, TITLE: Single-shot General Hyper-parameter Optimization for Federated Learning https://openreview.net/forum?id=3RhuF8foyPW AUTHORS: Yi Zhou, Parikshit Ram, Theodoros Salonidis, Nathalie Baracaldo, Horst Samulowitz, Heiko Ludwig HIGHLIGHT: We introduce Federated Loss SuRface Aggregation (FLoRA), a general FL-HPO solution framework that can address use cases of tabular data and any Machine Learning (ML) model including gradient boosting training algorithms, SVMs, neural networks, among others and thereby further expands the scope of FL-HPO. AE-FLOW: Autoencoders with Normalizing Flows for Medical Images Anomaly Detection 428, TITLE: https://openreview.net/forum?id=90mCr1q54Z AUTHORS: Yuzhong Zhao, Qiaoqiao Ding, Xiaoqun Zhang More specifically, we propose a normalizing flow based autoencoder for an efficient and tractable HIGHLIGHT: representation of normal medical images. 429, TITLE: What Is Missing in IRM Training and Evaluation? Challenges and Solutions https://openreview.net/forum?id=MjsDeTcDEy AUTHORS: Yihua Zhang, Pranay Sharma, Parikshit Ram, Mingyi Hong, Kush R. Varshney, Sijia Liu HIGHLIGHT: Therefore, a series of advanced IRM algorithms have been developed that show practical improvement over IRMV1. In this work, we revisit these recent IRM advancements and identify and resolve three practical limitations in IRM training and evaluation. Distributional Meta-Gradient Reinforcement Learning 430, TITLE: https://openreview.net/forum?id=LGkmUauBUL AUTHORS: Haiyan Yin, Shuicheng YAN, Zhongwen Xu HIGHLIGHT: All the existing algorithms adhere to the same reward learning regime, where the adaptive return is simply formulated in the form of expected cumulative rewards, upon which the policy and critic update rules are specified under well adopted distance metrics. In this paper, we present a novel algorithm which builds on the success of meta-gradient RL algorithms and effectively improves such algorithms by following a simple recipe, i.e., going beyond the expected return to formulate and learn the return in a more expressive form, value distributions. 431, TITLE: Min-Max Multi-objective Bilevel Optimization with Applications in Robust Machine Learning https://openreview.net/forum?id=PvDY71zKsvP AUTHORS: Alex Gu, Songtao Lu, Parikshit Ram, Tsui-Wei Weng We consider a generic min-max multi-objective bilevel optimization problem with applications in robust HIGHLIGHT: machine learning such as representation learning and hyperparameter optimization. Linearly Mapping from Image to Text Space 432, TITLE: https://openreview.net/forum?id=8tYRqb05pVn AUTHORS: Jack Merullo, Louis Castricato, Carsten Eickhoff, Ellie Pavlick

HIGHLIGHT: Specifically, we show that the image representations from vision models can be transferred as continuous prompts to frozen LMs by training only a single linear projection.

433, TITLE: Evidential Uncertainty and Diversity Guided Active Learning for Scene Graph Generation https://openreview.net/forum?id=xI1ZTtVOtlz AUTHORS: Shuzhou Sun, Shuaifeng Zhi, Janne Heikkilä, Li Liu HIGHLIGHT: However, directly porting current AL methods to the SGG task poses the following challenges: 1) unreliable uncertainty estimates, and 2) data bias problems. To deal with these challenges, we propose EDAL (textbf{E}vidential Uncertainty and \textbf{D}iversity Guided Deep \textbf{A}ctive \textbf{L}earning), a novel AL framework tailored for the SGG task. 434, TITLE: StableDR: Stabilized Doubly Robust Learning for Recommendation on Data Missing Not at Random https://openreview.net/forum?id=3VO1y5N7K1H AUTHORS: Haoxuan Li, Chunyuan Zheng, Peng Wu HIGHLIGHT: Moreover, the fact that DR relies more on extrapolation will lead to suboptimal performance. To address the above limitations while retaining double robustness, we propose a stabilized doubly robust (StableDR) learning approach with a weaker reliance on extrapolation. 435, TITLE: Variational Latent Branching Model for Off-Policy Evaluation https://openreview.net/forum?id=3VFQfAG3vwi AUTHORS: Qitong Gao, Ge Gao, Min Chi, Miroslav Pajic HIGHLIGHT: In this work, we propose the variational latent branching model (VLBM) to learn the transition function of MDPs by formulating the environmental dynamics as a compact latent space, from which the next states and rewards are then sampled. TDR-CL: Targeted Doubly Robust Collaborative Learning for Debiased Recommendations 436, TITLE: https://openreview.net/forum?id=EIgLnNx 1C Haoxuan Li, Yan Lyu, Chunyuan Zheng, Peng Wu AUTHORS: HIGHLIGHT: In this paper, we propose a principled approach that can effectively reduce the bias and variance simultaneously for existing DR estimators when the error-imputation model is misspecified. Improving Deep Policy Gradients with Value Function Search 437, TITLE: https://openreview.net/forum?id=6qZC7pfenQm AUTHORS: Enrico Marchesini, Christopher Amato In this paper, we focus on improving value approximation and analyzing the effects on Deep Policy Gradient HIGHLIGHT: primitives such as value prediction, variance reduction, and correlation of gradient estimates with the true gradient. 438, TITLE: LMSeg: Language-guided Multi-dataset Segmentation https://openreview.net/forum?id=P44WPn1 aJV AUTHORS: Qiang Zhou, Yuang Liu, Chaohui Yu, Jingliang Li, Zhibin Wang, Fan Wang HIGHLIGHT: In this paper, we investigate the multi-dataset segmentation and propose a scalable Language-guided Multidataset Segmentation framework, dubbed LMSeg, which supports both semantic and panoptic segmentation. 439, TITLE: Graph Neural Networks for Link Prediction with Subgraph Sketching https://openreview.net/forum?id=m1oqEOAozQU Benjamin Paul Chamberlain, Sergey Shirobokov, Emanuele Rossi, Fabrizio Frasca, Thomas Markovich, Nils AUTHORS: Yannick Hammerla, Michael M. Bronstein, Max Hansmire HIGHLIGHT: Based on our analysis, we propose a novel full-graph GNN called ELPH (Efficient Link Prediction with Hashing) that passes subgraph sketches as messages to approximate the key components of SGNNs without explicit subgraph construction. 440, TITLE: Extracting Robust Models with Uncertain Examples https://openreview.net/forum?id=cMAjKYftNwx Guanlin Li, Guowen Xu, Shangwei Guo, Han Qiu, Jiwei Li, Tianwei Zhang AUTHORS: HIGHLIGHT: However, how to extract a robust model with similar resilience against adversarial attacks is never investigated. This paper presents the first study toward this goal. 441, TITLE: Combinatorial Pure Exploration of Causal Bandits https://openreview.net/forum?id=pBBsrPzq7aF AUTHORS: Nuoya Xiong, Wei Chen We provide the first gap-dependent and fully adaptive pure exploration algorithms on two types of causal HIGHLIGHT: models --- the binary generalized linear model (BGLM) and general graphs. 442, TITLE: Continuous-time identification of dynamic state-space models by deep subspace encoding https://openreview.net/forum?id= 4n3k3d1ob AUTHORS: Gerben I. Beintema, Maarten Schoukens, Roland Tóth However, even with numerous recent developments, the CT nonlinear state-space (NL-SS) model identification HIGHLIGHT: problem remains to be solved in full, considering common experimental aspects such as the presence of external inputs, measurement noise, latent states, and general robustness. This paper presents a novel estimation method that addresses all these aspects and that can obtain state-of-the-art results on multiple benchmarks with compact fully connected neural networks capturing the CT dynamics. 443, TITLE: Better Generative Replay for Continual Federated Learning

443, 111LE: Better Generative Replay for Continual Federated Learning https://openreview.net/forum?id=cRxYWKiTan

AUTHORS: Daiqing Qi, Handong Zhao, Sheng Li HIGHLIGHT: By analyzing the behaviors of clients during training, we find the unstable training process caused by distributed training on non-IID data leads to a notable performance degradation. To address this problem, we propose our FedCIL model with two simple but effective solutions: 1. 444, TITLE: Sub-Task Decomposition Enables Learning in Sequence to Sequence Tasks https://openreview.net/forum?id=BrJATVZDWEH AUTHORS: Noam Wies, Yoav Levine, Amnon Shashua HIGHLIGHT: Recently, several works have demonstrated high gains by taking a straightforward approach for incorporating intermediate supervision in compounded natural language problems: the sequence-to-sequence LM is fed with an augmented input, in which the decomposed tasks' labels are simply concatenated to the original input. In this paper, we prove a positive learning result that motivates these recent efforts. On the Word Boundaries of Emergent Languages Based on Harris's Articulation Scheme 445, TITLE: https://openreview.net/forum?id=b4t9 XASt6G AUTHORS: Ryo Ueda, Taiga Ishii, Yusuke Miyao HIGHLIGHT: However, it is not obvious whether such a simulated language would have the same properties as natural language. In this paper, we test if they satisfy HAS. 446. TITLE: Generative Modelling with Inverse Heat Dissipation https://openreview.net/forum?id=4PJUBT9f2Ol AUTHORS: Severi Rissanen, Markus Heinonen, Arno Solin HIGHLIGHT: Inspired by diffusion models and the empirical success of coarse-to-fine modelling, we propose a new model that generates images through iteratively inverting the heat equation, a PDE that locally erases fine-scale information when run over the 2D plane of the image. 447. TITLE: Self-supervision through Random Segments with Autoregressive Coding (RandSAC) https://openreview.net/forum?id=Ubc74gTVo3 AUTHORS: Tianyu Hua, Yonglong Tian, Sucheng Ren, Michalis Raptis, Hang Zhao, Leonid Sigal HIGHLIGHT: Inspired by the success of self-supervised autoregressive representation learning in natural language (GPT and its variants), and advances in recent visual architecture design with Vision Transformers (ViTs), in this paper, we explore the effects various design choices have on the success of applying such training strategies for visual feature learning. 448. TITLE: Ask Me Anything: A simple strategy for prompting language models https://openreview.net/forum?id=bhUPJnS2g0X AUTHORS: Simran Arora, Avanika Narayan, Mayee F Chen, Laurel Orr, Neel Guha, Kush Bhatia, Ines Chami, Christopher Re HIGHLIGHT: We propose a prompting strategy based on aggregating the predictions of multiple prompts, which enables a 6B parameter model to exceed the few-shot performance of GPT3-175B on 15/20 popular benchmarks. 449, TITLE: DAVA: Disentangling Adversarial Variational Autoencoder https://openreview.net/forum?id=CW6KmU5wPh Benjamin Estermann, Roger Wattenhofer AUTHORS: HIGHLIGHT: To address the issue, we introduce DAVA, a novel training procedure for variational auto-encoders that alleviates the issue of hyperparameter selection at the cost of a comparatively small overhead. 450, TITLE: Temperature Schedules for self-supervised contrastive methods on long-tail data https://openreview.net/forum?id=ejHUr4nfHhD AUTHORS: Anna Kukleva, Moritz Böhle, Bernt Schiele, Hilde Kuehne, Christian Rupprecht HIGHLIGHT: In this paper, we analyse the behaviour of one of the most popular variants of SSL, i.e. contrastive methods, on imbalanced data. 451, TITLE: From \$t\$-SNE to UMAP with contrastive learning https://openreview.net/forum?id=B8a1FcY0vi AUTHORS: Sebastian Damrich, Niklas Böhm, Fred A Hamprecht, Dmitry Kobak HIGHLIGHT: In this work, we uncover their conceptual connection via a new insight into contrastive learning methods. 452, TITLE: Sharper Bounds for Uniformly Stable Algorithms with Stationary \$\varphi\$-mixing Process https://openreview.net/forum?id=8E5Yazboyh AUTHORS: Shi Fu, Yunwen Lei, Qiong Cao, Xinmei Tian, Dacheng Tao HIGHLIGHT: In this paper, we use algorithmic stability to study the generalization performance of learning algorithms with \$\varphi\$-mixing data, where the dependency between observations weakens over time. 453, TITLE: Pareto-Efficient Decision Agents for Offline Multi-Objective Reinforcement Learning https://openreview.net/forum?id=Ki4ocDm364 AUTHORS: Baiting Zhu, Meihua Dang, Aditya Grover In this work, we propose a new data-driven setup for offline MORL, where we wish to learn a preference-HIGHLIGHT: agnostic policy agent using only a finite dataset of offline demonstrations of other agents and their preferences.

454, TITLE: Neural Collapse Inspired Feature-Classifier Alignment for Few-Shot Class-Incremental Learning

https://openreview.net/forum?id=y5W8tpojhtJ AUTHORS: Yibo Yang, Haobo Yuan, Xiangtai Li, Zhouchen Lin, Philip Torr, Dacheng Tao HIGHLIGHT: In this paper, we deal with this misalignment dilemma in FSCIL inspired by the recently discovered phenomenon named neural collapse, which reveals that the last-layer features of the same class will collapse into a vertex, and the vertices of all classes are aligned with the classifier prototypes, which are formed as a simplex equiangular tight frame (ETF). Efficient recurrent architectures through activity sparsity and sparse back-propagation through time 455, TITLE: https://openreview.net/forum?id=lJdOlWg8td AUTHORS: Anand Subramoney, Khaleelulla Khan Nazeer, Mark Schöne, Christian Mayr, David Kappel HIGHLIGHT: We propose a solution inspired by biological neuron dynamics that makes the communication between RNN units sparse and discrete. Neural ePDOs: Spatially Adaptive Equivariant Partial Differential Operator Based Networks 456, TITLE: https://openreview.net/forum?id=D1Iqfm7WTkk AUTHORS: Lingshen He, Yuxuan Chen, Zhengyang Shen, Yibo Yang, Zhouchen Lin HIGHLIGHT: In this work, we propose a novel nonlinear PDOs scheme that is both spatially adaptive and translation equivariant. 457, TITLE: Learning to Segment from Noisy Annotations: A Spatial Correction Approach https://openreview.net/forum?id=Qc OopMEBnC AUTHORS: Jiachen Yao, Yikai Zhang, Songzhu Zheng, Mayank Goswami, Prateek Prasanna, Chao Chen HIGHLIGHT: In this paper, we propose a novel noise model for segmentation problems that encodes spatial correlation and bias, which are prominent in segmentation annotations. 458, TITLE: Edge Guided GANs with Contrastive Learning for Semantic Image Synthesis https://openreview.net/forum?id=qcJmsP3oE9 Hao Tang, XIAOJUAN QI, Guolei Sun, Dan Xu, Nicu Sebe, Radu Timofte, Luc Van Gool AUTHORS: HIGHLIGHT: We propose a novel \underline {e}dge guided \underline {g} enerative \underline {a} dversarial \underline {n}etwork with \underline {c} ontrastive learning (ECGAN) for the challenging semantic image synthesis task. 459, TITLE: Wasserstein Auto-encoded MDPs: Formal Verification of Efficiently Distilled RL Policies with Many-sided Guarantees https://openreview.net/forum?id=JLLTtEdh1ZY AUTHORS: Florent Delgrange, Ann Nowe, Guillermo Perez We introduce the Wasserstein auto-encoded MDP (WAE-MDP), a latent space model that fixes those issues by HIGHLIGHT: minimizing a penalized form of the optimal transport between the behaviors of the agent executing the original policy and the distilled policy, for which the formal guarantees apply. 460, TITLE: STaSy: Score-based Tabular data Synthesis https://openreview.net/forum?id=1mNssCWt v AUTHORS: Jayoung Kim, Chaejeong Lee, Noseong Park HIGHLIGHT: In this paper, we present a new model named \$\textbf{S}\$core-based \$\textbf{Ta}\$bular data \$\textbf{Sy}\$nthesis (\$\texttt{STaSy}\$) and its training strategy based on the paradigm of score-based generative modeling. 461. TITLE: De Novo Molecular Generation via Connection-aware Motif Mining https://openreview.net/forum?id=Q Jex18-qDi AUTHORS: Zijie Geng, Shufang Xie, Yingce Xia, Lijun Wu, Tao Qin, Jie Wang, Yongdong Zhang, Feng Wu, Tie-Yan Liu HIGHLIGHT: In this work, we propose MiCaM to generate molecules based on mined connection-aware motifs. 462, TITLE: When Source-Free Domain Adaptation Meets Learning with Noisy Labels https://openreview.net/forum?id=u2Pd6x794I AUTHORS: Li Yi, Gezheng Xu, Pengcheng Xu, Jiaqi Li, Ruizhi Pu, Charles Ling, Ian McLeod, Boyu Wang HIGHLIGHT: In this paper, we study SFDA from the perspective of learning with label noise (LLN). 463, TITLE: Test-Time Adaptation via Self-Training with Nearest Neighbor Information https://openreview.net/forum?id=EzLtB4M1SbM AUTHORS: Minguk Jang, Sae-Young Chung, Hye Won Chung However, under test-time domain shift, accuracy of the pseudo labels cannot be guaranteed, and thus the TTA HIGHLIGHT: methods often encounter performance degradation at the adapted classifier. To overcome this limitation, we propose a novel test-time adaptation method, called Test-time Adaptation via Self-Training with nearest neighbor information (TAST), which is composed of the following procedures: (1) adds trainable adaptation modules on top of the trained feature extractor; (2) newly defines a pseudolabel distribution for the test data by using the nearest neighbor information; (3) trains these modules only a few times during test time to match the nearest neighbor-based pseudo label distribution and a prototype-based class distribution for the test data; and (4) predicts the label of test data using the average predicted class distribution from these modules. Federated Neural Bandits 464, TITLE: https://openreview.net/forum?id=38m4h8HcNRL

 AUTHORS:
 Zhongxiang Dai, Yao Shu, Arun Verma, Flint Xiaofeng Fan, Bryan Kian Hsiang Low, Patrick Jaillet

 HIGHLIGHT:
 So, this paper introduces the federated neural-upper confidence bound (FN-UCB) algorithm.

465, TITLE:	Measuring axiomatic identifiability of counterfactual image models
https://openreview.ne	et/forum?id=lZOUQQvwI3q
AUTHORS:	Miguel Monteiro, Fabio De Sousa Ribeiro, Nick Pawlowski, Daniel C. Castro, Ben Glocker
HIGHLIGHT:	We present a general framework for evaluating image counterfactuals.
466, TITLE: https://openreview.ne AUTHORS: HIGHLIGHT: policy evaluation.	In-sample Actor Critic for Offline Reinforcement Learning t/forum?id=dfDv0WU853R Hongchang Zhang, Yixiu Mao, Boyuan Wang, Shuncheng He, Yi Xu, Xiangyang Ji We propose In-sample Actor Critic (IAC) which utilizes sampling-importance resampling to execute in-sample
467, TITLE:	Planning Goals for Exploration
https://openreview.ne	tt/forum?id=6qeBuZSo7Pr
AUTHORS:	Edward S. Hu, Richard Chang, Oleh Rybkin, Dinesh Jayaraman
HIGHLIGHT:	We propose "planning exploratory goals" (PEG), a method that sets goals for each training episode to directly
optimize an intrinsic	exploration reward.
468, TITLE:	Kernel Neural Optimal Transport
https://openreview.ne	tt/forum?id=Zuc_MHtUma4
AUTHORS:	Alexander Korotin, Daniil Selikhanovych, Evgeny Burnaev
HIGHLIGHT:	We study the Neural Optimal Transport (NOT) algorithm which uses the general optimal transport formulation
and learns stochastic	transport plans.
469, TITLE:	Targeted Hyperparameter Optimization with Lexicographic Preferences Over Multiple Objectives
https://openreview.ne	tt/forum?id=0Ij9_q567Ma
AUTHORS:	Shaokun Zhang, Feiran Jia, Chi Wang, Qingyun Wu
HIGHLIGHT:	We propose to do targeted hyperparameter optimization with lexicographic preference over multiple objectives,
motivated by various	practical applications.
470, TITLE: https://openreview.ne AUTHORS: HIGHLIGHT: algorithm.	A Non-monotonic Self-terminating Language Model tt/forum?id=vw-5EgYbJZr Cheolhyoung Lee, Eugene Choi, Kyunghyun Cho In this paper, we focus on the problem of non-terminating sequences resulting from an incomplete decoding
471, TITLE:	Noise-Robust De-Duplication at Scale
https://openreview.ne	tt/forum?id=bAz2DBS35i
AUTHORS:	Emily Silcock, Luca D'Amico-Wong, Jinglin Yang, Melissa Dell
HIGHLIGHT:	This study uses the unique timeliness of historical news wires to create a 27,210 document dataset, with
122,876 positive dup	licate pairs, for studying noise-robust de-duplication.
472, TITLE: https://openreview.ne AUTHORS: Aaron Courville HIGHLIGHT: generalization proper	Simplicial Embeddings in Self-Supervised Learning and Downstream Classification tt/forum?id=RWtGreRpovS Samuel Lavoie, Christos Tsirigotis, Max Schwarzer, Ankit Vani, Michael Noukhovitch, Kenji Kawaguchi, We use softmax to embed representations in a collection of simplices in SSL models, which offers improved ties for downstream classification.
473, TITLE:	Policy Pre-training for Autonomous Driving via Self-supervised Geometric Modeling
https://openreview.ne	tt/forum?id=X5SUR7g2vVw
AUTHORS:	Penghao Wu, Li Chen, Hongyang Li, Xiaosong Jia, Junchi Yan, Yu Qiao
HIGHLIGHT:	To this end, we propose PPGeo (Policy Pre-training via Geometric modeling), an intuitive and straightforward
fully self-supervised	framework curated for the policy pre-training in visuomotor driving.
474, TITLE:	Adaptive Budget Allocation for Parameter-Efficient Fine-Tuning
https://openreview.ne	tt/forum?id=lq62uWRJjiY
AUTHORS:	Qingru Zhang, Minshuo Chen, Alexander Bukharin, Pengcheng He, Yu Cheng, Weizhu Chen, Tuo Zhao
HIGHLIGHT:	As a consequence, the fine-tuning performance is suboptimal. To bridge this gap, we propose MARVEL, which
adaptively allocates t	he parameter budget among weight matrices according to their importance score.
475, TITLE:	Treeformer: Dense Gradient Trees for Efficient Attention Computation
https://openreview.ne	tt/forum?id=DWn1TEb2fK
AUTHORS:	Lovish Madaan, Srinadh Bhojanapalli, Himanshu Jain, Prateek Jain
HIGHLIGHT:	In this work, we view attention computation as that of nearest neighbor retrieval, and use decision tree based
hierarchical navigatio	on to reduce the retrieval cost per query token from linear in sequence length to nearly logarithmic.
476, TITLE:	GPTQ: Accurate Quantization for Generative Pre-trained Transformers
https://openreview.ne	tt/forum?id=tcbBPnfwxS
AUTHORS:	Elias Frantar, Saleh Ashkboos, Torsten Hoefler, Dan Alistarh

HIGHLIGHT: While there is emerging work on relieving this pressure via model compression, the applicability and performance of existing compression techniques is limited by the scale and complexity of GPT models. In this paper, we address this challenge, and propose GPTQ, a new one-shot weight quantization method based on approximate second-order information, that is both highly-accurate and highly-efficient.

477, TITLE: Neural Optimal Transport

https://openreview.net/forum?id=d8CBRIWNkqH

AUTHORS: Alexander Korotin, Daniil Selikhanovych, Evgeny Burnaev

HIGHLIGHT: We present a novel neural-networks-based algorithm to compute optimal transport maps and plans for strong and weak transport costs.

478, TITLE: DEP-RL: Embodied Exploration for Reinforcement Learning in Overactuated and Musculoskeletal Systems https://openreview.net/forum?id=C-xa_D3oTj6

AUTHORS: Pierre Schumacher, Daniel Haeufle, Dieter Büchler, Syn Schmitt, Georg Martius

HIGHLIGHT: We identify differential extrinsic plasticity (DEP), a method from the domain of self-organization, as being able to induce state-space covering exploration within seconds of interaction.

479, TITLE: Optimal Activation Functions for the Random Features Regression Model

https://openreview.net/forum?id=ltWade-cpK

AUTHORS: Jianxin Wang, José Bento

HIGHLIGHT: We find scenarios under which the optimal AFs are linear, saturated linear functions, or expressible in terms of Hermite polynomials.

480, TITLE: Unsupervised Object-Centric Learning with Bi-level Optimized Query Slot Attention https://openreview.net/forum?id=_-FN9mJsgg

AUTHORS: Baoxiong Jia, Yu Liu, Siyuan Huang

HIGHLIGHT: These methods, however, have been exceedingly difficult to train without supervision and are ambiguous in the notion of object, especially for complex natural scenes. In this paper, we propose to address these issues by (1) initializing Slot-Attention modules with learnable queries and (2) optimizing the model with bi-level optimization.

481, TITLE: Augmentation with Projection: Towards an Effective and Efficient Data Augmentation Paradigm for Distillation

https://openreview.net/forum?id=kPPVmUF6bM

AUTHORS:Ziqi Wang, Yuexin Wu, Frederick Liu, Daogao Liu, Le Hou, Hongkun Yu, Jing Li, Heng JiHIGHLIGHT:To this end, we propose AugPro (Augmentation with Projection), an effective and efficient data augmentationmethod for distillation.

482, TITLE: Learning an Invertible Output Mapping Can Mitigate Simplicity Bias in Neural Networks https://openreview.net/forum?id=zH9GcZ3ZGXu

 AUTHORS:
 Sravanti Addepalli, Anshul Nasery, Venkatesh Babu Radhakrishnan, Praneeth Netrapalli, Prateek Jain

 HIGHLIGHT:
 This phenomenon, we term \emph{Feature Replication Hypothesis}, coupled with the \emph{Implicit Bias} of

 SGD to converge to maximum margin solutions in the feature space, leads the models to rely mostly on the simple features for

 classification. To mitigate this bias, we propose \emph{Feature Reconstruction Regularizer (FRR)} to ensure that the learned features

 can be reconstructed back from the logits.

483, TITLE: Imbalanced Semi-supervised Learning with Bias Adaptive Classifier https://openreview.net/forum?id=rVM8wD2G7Dy

AUTHORS: Renzhen Wang, Xixi Jia, Quanziang Wang, Yichen Wu, Deyu Meng

HIGHLIGHT: However, such an assumption is far from realistic scenarios and thus severely limits the performance of current pseudo-labeling methods under the context of class-imbalance. To alleviate this problem, we design a bias adaptive classifier that targets the imbalanced SSL setups.

 484, TITLE:
 On Compositional Uncertainty Quantification for Seq2seq Graph Parsing

 https://openreview.net/forum?id=rJcLocAJpA6

 AUTHORS:
 Zi Lin, Jeremiah Zhe Liu, Du Phan, Panupong Pasupat, Jingbo Shang

 UCIVIEW:
 Di Lin, Jeremiah Zhe Liu, Du Phan, Panupong Pasupat, Jingbo Shang

HIGHLIGHT: In this paper, we aim to quantify and evaluate compositional uncertainty for seq2seq graph parsing by proposing a simple probabilistic framework and rigorous evaluation metrics.

485, TITLE: EUCLID: Towards Efficient Unsupervised Reinforcement Learning with Multi-choice Dynamics Model https://openreview.net/forum?id=xQAjSr64PTc

AUTHORS: Yifu Yuan, Jianye HAO, Fei Ni, Yao Mu, YAN ZHENG, Yujing Hu, Jinyi Liu, Yingfeng Chen, Changjie Fan HIGHLIGHT: To this end, we propose an Efficient Unsupervised Reinforcement Learning Framework with Multi-choice Dynamics model (EUCLID), which introduces a novel model-fused paradigm to jointly pre-train the dynamics model and unsupervised exploration policy in the pre-training phase, thus better leveraging the environmental samples and improving the downstream task sampling efficiency.

486, TITLE: A General Framework for Sample-Efficient Function Approximation in Reinforcement Learning https://openreview.net/forum?id=dqITIpZ5Z4b

AUTHORS: Zixiang Chen, Chris Junchi Li, Huizhuo Yuan, Quanquan Gu, Michael Jordan

HIGHLIGHT: In this paper, we propose a general framework that unifies model-based and model-free RL, and an Admissible Bellman Characterization (ABC) class that subsumes nearly all Markov decision process (MDP) models in the literature for tractable RL.

487, TITLE: Sequence to sequence text generation with diffusion models https://openreview.net/forum?id=jQj- rLVXsj AUTHORS: Shansan Gong, Mukai Li, Jiangtao Feng, Zhiyong Wu, Lingpeng Kong HIGHLIGHT: Despite the success in domains using continuous signals such as vision and audio, adapting diffusion models to natural language is difficult due to the discrete nature of text. We tackle this challenge by proposing DiffuSeq: a diffusion model designed for sequence-to-sequence (Seq2Seq) text generation tasks. Measure the Predictive Heterogeneity 488. TITLE: https://openreview.net/forum?id=g2oB k-18b AUTHORS: Jiashuo Liu, Jiayun Wu, Renjie Pi, Renzhe Xu, Xingxuan Zhang, Bo Li, Peng Cui HIGHLIGHT: We prove that it can be reliably estimated from finite data with PAC bounds even in high dimensions. 489, TITLE: InPL: Pseudo-labeling the Inliers First for Imbalanced Semi-supervised Learning https://openreview.net/forum?id=m6ahb1mpwwX AUTHORS: Zhuoran Yu, Yin Li, Yong Jae Lee In this work, we present a new perspective of pseudo-labeling for imbalanced SSL. HIGHLIGHT: 490, TITLE: PandA: Unsupervised Learning of Parts and Appearances in the Feature Maps of GANs https://openreview.net/forum?id=iUdSB2kK9GY AUTHORS: James Oldfield, Christos Tzelepis, Yannis Panagakis, Mihalis Nicolaou, Ioannis Patras However, existing methods are often tailored to specific GAN architectures and are limited to either discovering HIGHLIGHT: global semantic directions that do not facilitate localized control, or require some form of supervision through manually provided regions or segmentation masks. In this light, we present an architecture-agnostic approach that jointly discovers factors representing spatial parts and their appearances in an entirely unsupervised fashion. 491. TITLE: Unsupervised visualization of image datasets using contrastive learning https://openreview.net/forum?id=nI2HmVA0hvt AUTHORS: Niklas Böhm, Philipp Berens, Dmitry Kobak HIGHLIGHT: Here, we present a new method, called t-SimCNE, for unsupervised visualization of image data. 492, TITLE: Hebbian and Gradient-based Plasticity Enables Robust Memory and Rapid Learning in RNNs https://openreview.net/forum?id=2WklawyeI08 Yu Duan, Zhongfan Jia, Qian Li, Yi Zhong, Kaisheng Ma AUTHORS: While the underlying neural mechanisms are not fully understood, various evidence supports that synaptic HIGHLIGHT: plasticity plays a critical role in memory formation and fast learning. Inspired by these results, we equip Recurrent Neural Networks (RNNs) with plasticity rules to enable them to adapt their parameters according to ongoing experiences. In addition to the traditional local Hebbian plasticity, we propose a global, gradient-based plasticity rule, which allows the model to evolve towards its selfdetermined target. Learned Index with Dynamic \$\epsilon\$ 493, TITLE: https://openreview.net/forum?id=UiaUEICawgw AUTHORS: Daoyuan Chen, Wuchao Li, Yaliang Li, Bolin Ding, Kai Zeng, Defu Lian, Jingren Zhou HIGHLIGHT: In this paper, we propose a mathematically-grounded learned index framework with dynamic \$\epsilon\$, which is efficient and pluggable to existing learned index methods. 494, TITLE: ViT-Adapter: Exploring Plain Vision Transformer for Accurate Dense Predictions https://openreview.net/forum?id=plKu2GByCNW AUTHORS: Zhe Chen, Yuchen Duan, Wenhai Wang, Junjun He, Tong Lu, Jifeng Dai, Yu Qiao HIGHLIGHT: Unlike recently advanced variants that incorporate vision-specific inductive biases into their architectures, the plain ViT suffers inferior performance on dense predictions due to weak prior assumptions. To address this issue, we propose the ViT-Adapter, which allows plain ViT to achieve comparable performance to vision-specific transformers. 495. TITLE: Pareto Invariant Risk Minimization https://openreview.net/forum?id=esFxSb_0pSL Yongqiang Chen, Kaiwen Zhou, Yatao Bian, Binghui Xie, Bingzhe Wu, Yonggang Zhang, MA KAILI, Han AUTHORS: Yang, Peilin Zhao, Bo Han, James Cheng HIGHLIGHT: Consequently, these compromises could easily lead to suboptimal performance of either the ERM or OOD objective. To address these issues, we introduce a multi-objective optimization (MOO) perspective to understand the OOD optimization process, and propose a new optimization scheme called PAreto Invariant Risk Minimization (PAIR). ROSCOE: A Suite of Metrics for Scoring Step-by-Step Reasoning

496, TITLE:

https://openreview.net/forum?id=xYlJRpzZtsY

AUTHORS: Olga Golovneva, Moya Peng Chen, Spencer Poff, Martin Corredor, Luke Zettlemover, Maryam Fazel-Zarandi, Asli Celikyilmaz

HIGHLIGHT: In this work, we present ROSCOE, a suite of interpretable, unsupervised automatic scores that improve and extend previous text generation evaluation metrics.

497, TITLE:	Variational Information Pursuit for Interpretable Predictions
https://openreview.net	/forum?id=77ISWa-Tm3Z
AUTHORS:	Aditya Chattopadhyay, Kwan Ho Ryan Chan, Benjamin David Haeffele, Donald Geman, Rene Vidal
HIGHLIGHT:	In this work, we propose Variational Information Pursuit (V-IP), a variational characterization of IP which
bypasses the need for	learning generative models.
498, TITLE:	Deep Learning on Implicit Neural Representations of Shapes
https://openreview.net	/forum?id=OoOIW-3uadi
AUTHORS:	Luca De Luigi, Adriano Cardace, Riccardo Spezialetti, Pierluigi Zama Ramirez, Samuele Salti, Luigi di Stefano
HIGHLIGHT:	In this paper, we put forward this research problem and propose inr2vec, a framework that can compute a
compact latent represe	entation for an input INR in a single inference pass.
499, TITLE:	Image to Sphere: Learning Equivariant Features for Efficient Pose Prediction
https://openreview.net	/forum?id=_2bDpAtr7PI
AUTHORS:	David Klee, Ondrej Biza, Robert Platt, Robin Walters
HIGHLIGHT:	Instead, we propose a novel mapping of features from the image domain to the 3D rotation manifold.
500, TITLE: https://openreview.net AUTHORS: HIGHLIGHT: regular ReLU network generalization.	Generalization and Estimation Error Bounds for Model-based Neural Networks /forum?id=9F_xlC7sk9 Avner Shultzman, Eyar Azar, Miguel R. D. Rodrigues, Yonina C. Eldar We show that the generalization abilities of model-based networks for sparse recovery outperform those of cs, and derive practical design rules that allow to construct model-based networks with guaranteed high
501, TITLE:	Consolidator: Mergable Adapter with Group Connections for Vision Transformer
https://openreview.net	/forum?id=J_Cja7cpgW
AUTHORS:	Tianxiang Hao, Hui Chen, Yuchen Guo, Guiguang Ding
HIGHLIGHT:	Moreover, some of them bring heavy inference cost though benefiting storage. To tackle these problems, we
propose consolidator t	to achieve efficient transfer learning for vision transformers.
502, TITLE: https://openreview.net AUTHORS: HIGHLIGHT: disentangled temporal dynamics.	Multivariate Time-series Imputation with Disentangled Temporal Representations //forum?id=rdjeCNUS6TG SHUAI LIU, Xiucheng Li, Gao Cong, Yile Chen, YUE JIANG Different from existing approaches, we propose TIDER, a novel matrix factorization-based method with representations that account for multiple factors, namely trend, seasonality, and local bias, to model complex
503, TITLE: https://openreview.net AUTHORS: HIGHLIGHT: GCNs. To fill this gap formulated an influence attributed graph.	Characterizing the Influence of Graph Elements /forum?id=51GXyzOKOp Zizhang Chen, Peizhao Li, Hongfu Liu, Pengyu Hong Since the nodes/edges in a graph are interdependent in GCNs, it is challenging to derive influence functions for we started with the simple graph convolution (SGC) model that operates on an attributed graph, and ce function to approximate the changes of model parameters when a node or an edge is removed from an
504, TITLE:	LipsFormer: Introducing Lipschitz Continuity to Vision Transformers
https://openreview.net	/forum?id=cHf1DcCwcH3
AUTHORS:	Xianbiao Qi, Jianan Wang, Yihao Chen, Yukai Shi, Lei Zhang
HIGHLIGHT:	We present a Lipschitz continuous Transformer, called LipsFormer, to pursue training stability both
theoretically and empiri	irically for Transformer-based models.
505, TITLE:	Neuro-Symbolic Procedural Planning with Commonsense Prompting
https://openreview.net	/forum?id=iOc57X9KM54
AUTHORS:	Yujie Lu, Weixi Feng, Wanrong Zhu, Wenda Xu, Xin Eric Wang, Miguel Eckstein, William Yang Wang
HIGHLIGHT:	In contrast, this paper proposes a neuro-symbolic procedural PLANner (PLAN) that elicits procedural planning
knowledge from the L	LMs with commonsense-infused prompting.
506, TITLE:	Robust Scheduling with GFlowNets
https://openreview.net	/forum?id=ZBUthI6wK9h
AUTHORS:	David W Zhang, Corrado Rainone, Markus Peschl, Roberto Bondesan
HIGHLIGHT:	In this work, we propose a new approach to scheduling by sampling proportionally to the proxy metric using a
novel GFlowNet meth	lod.
507, TITLE:	On the Performance of Temporal Difference Learning With Neural Networks
https://openreview.net	/forum?id=6JMXLWX68Kj
AUTHORS:	HAOXING TIAN, Ioannis Paschalidis, Alex Olshevsky

HIGHLIGHT: In this paper we provide a convergence analysis of Neural TD Learning with a projection onto \$B(\theta_0, \omega)\$, a ball of fixed radius \$\omega\$ around the initial point \$\theta_0\$.

508, TITLE: WikiWhy: Answering and Explaining Cause-and-Effect Questions https://openreview.net/forum?id=vaxnu-Utr41 AUTHORS: Matthew Ho, Aditya Sharma, Justin Chang, Michael Saxon, Sharon Levy, Yujie Lu, William Yang Wang HIGHLIGHT: We introduce WikiWhy, a QA dataset built around a novel auxiliary task: explaining why an answer is true in natural language. Data augmentation alone can improve adversarial training 509, TITLE: https://openreview.net/forum?id=y4uc4NtTWaq AUTHORS: Lin Li, Michael W. Spratling HIGHLIGHT: This work proves that, contrary to previous findings, data augmentation alone can significantly boost accuracy and robustness in adversarial training. Spikformer: When Spiking Neural Network Meets Transformer 510, TITLE: https://openreview.net/forum?id=frE4fUwz h AUTHORS: Zhaokun Zhou, Yuesheng Zhu, Chao He, Yaowei Wang, Shuicheng YAN, Yonghong Tian, Li Yuan HIGHLIGHT: In this paper, we consider leveraging both self-attention capability and biological properties of SNNs, and propose a novel Spiking Self Attention (SSA) as well as a powerful framework, named Spiking Transformer (Spikformer). 511, TITLE: NERDS: A General Framework to Train Camera Denoisers from Single Noisy Images https://openreview.net/forum?id=NO0ThzteQdI AUTHORS: Heewon Kim, Kyoung Mu Lee HIGHLIGHT: To utilize this property, we can adopt noisy/clean image synthesis at low-resolution to train camera denoisers. On this basis, we propose a new solution pipeline -- NERDS that estimates camera noises and synthesizes noisy-clean image pairs from only noisy images. 512, TITLE: Modeling the Data-Generating Process is Necessary for Out-of-Distribution Generalization https://openreview.net/forum?id=uyqks-LILZX AUTHORS: Jivat Neet Kaur, Emre Kiciman, Amit Sharma Based on the relationship between spurious attributes and the classification label, we obtain realizations of the HIGHLIGHT: canonical causal graph that characterize common distribution shifts and show that each shift entails different independence constraints over observed variables. 513, TITLE: Strong inductive biases provably prevent harmless interpolation https://openreview.net/forum?id=7i6OZa7oij AUTHORS: Michael Aerni, Marco Milanta, Konstantin Donhauser, Fanny Yang This paper argues that the degree to which interpolation is harmless hinges upon the strength of an estimator's HIGHLIGHT: inductive bias, i.e., how heavily the estimator favors solutions with a certain structure: while strong inductive biases prevent harmless interpolation, weak inductive biases can even require fitting noise to generalize well. 514, TITLE: Certified Training: Small Boxes are All You Need https://openreview.net/forum?id=7oFuxtJtUMH AUTHORS: Mark Niklas Mueller, Franziska Eckert, Marc Fischer, Martin Vechev We propose the novel certified training method, SABR, which outperforms existing methods across HIGHLIGHT: perturbation magnitudes on MNIST, CIFAR-10, and TinyImageNet, in terms of both standard and certifiable accuracies. 515, TITLE: Efficient Certified Training and Robustness Verification of Neural ODEs https://openreview.net/forum?id=KyoVpYvWWnK AUTHORS: Mustafa Zeqiri, Mark Niklas Mueller, Marc Fischer, Martin Vechev HIGHLIGHT: However, despite significant progress in robustness verification for standard feed-forward architectures, the verification of high dimensional NODEs remains an open problem. In this work we address this challenge and propose GAINS, an analysis framework for NODEs combining three key ideas: (i) a novel class of ODE solvers, based on variable but discrete time steps, (ii) an efficient graph representation of solver trajectories, and (iii) a novel abstraction algorithm operating on this graph representation. 516, TITLE: Confidence Estimation Using Unlabeled Data https://openreview.net/forum?id=sOXU-PEJSgQ AUTHORS: Chen Li, Xiaoling Hu, Chao Chen HIGHLIGHT: In this paper, we propose the first confidence estimation method for a semi-supervised setting, when most training labels are unavailable. 517, TITLE: Neural Episodic Control with State Abstraction https://openreview.net/forum?id=C2fsSj3ZGiU AUTHORS: Zhuo Li, Derui Zhu, Yujing Hu, Xiaofei Xie, Lei Ma, YAN ZHENG, Yan Song, Yingfeng Chen, Jianjun Zhao HIGHLIGHT: This work introduces Neural Episodic Control with State Abstraction (NECSA), a simple but effective state abstraction-based episodic control containing a more comprehensive episodic memory, a novel state evaluation, and a multi-step state analysis. 518, TITLE: Leveraging Large Language Models for Multiple Choice Question Answering https://openreview.net/forum?id=yKbprarjc5B

AUTHORS: Joshua Robinson, David Wingate

HIGHLIGHT: We show that a model with high MCSB ability performs much better with the natural approach than with the traditional approach across 20 diverse tasks and largely closes the gap with the SOTA, suggesting that the MCQA ability of LLMs has been previously underestimated.

519, TITLE:	Relative representations enable zero-shot latent space communication
https://openreview.net	//forum?id=SrC-nwieGJ
AUTHORS:	Luca Moschella, Valentino Maiorca, Marco Fumero, Antonio Norelli, Francesco Locatello, Emanuele Rodola In this work, we prepess to adopt pairwise similarities as an alternative data correspondention, that can be used to
enforce the desired inv	variance without any additional training.
520, TITLE:	ILA-DA: Improving Transferability of Intermediate Level Attack with Data Augmentation
https://openreview.net	/forum?id=OM7doLjQbOQ
AUTHORS:	Chiu Wai Yan, Tsz-Him Cheung, Dit-Yan Yeung
on these two observation	ions, we propose ILA-DA, which employs three novel augmentation techniques to enhance ILA.
521, TITLE:	Contrastive Alignment of Vision to Language Through Parameter-Efficient Transfer Learning
https://openreview.net	/forum?id=x0BPR9iXc1
AUTHORS:	Zaid Khan, Yun Fu Wa danai ka ang ina famaning tang atau tang tang tang tang tang tang tang tang
efficient training and t	that parameter-efficient scaling scales with model and dataset size.
522, TITLE:	Real-time variational method for learning neural trajectory and its dynamics
https://openreview.net	/forum?id=M_MvkWgQSt
HIGHLIGHT	Maunew Dowling, Yuan Zhao, Il Memming Park However, despite the potential of real-time alternatives to give immediate feedback to experimentalists, and
enhance experimental	design, they have received markedly less attention. In this work, we introduce the exponential family variational
Kalman filter (eVKF)	, an online recursive Bayesian method aimed at inferring latent trajectories while simultaneously learning the
dynamical system gen	erating them.
523, TITLE:	Minimalistic Unsupervised Learning with the Sparse Manifold Transform
https://openreview.net	//torum?id=nN_nBVKAhhD Vubai Chan Zavu Vun Vi Ma Bruno Olshausan Vann LeCun
HIGHLIGHT:	We describe a minimalistic and interpretable method for unsupervised learning, without resorting to data
augmentation, hyperpa	arameter tuning, or other engineering designs, that achieves performance close to the SOTA SSL methods.
524, TITLE:	Self-Ensemble Protection: Training Checkpoints Are Good Data Protectors
https://openreview.net	/forum?id=9MO7bjoAfIA Siste Chen, Cana Yuan, Xiawan Chana, Vitan Cana, Minakai Qin, Vanshi Wana, Xiaalin Uyana
HIGHLIGHT	In this paper, we propose a self-ensemble protection (SEP) method to take advantage of intermediate
checkpoints in a single	e training process for data protection.
525, TITLE:	On the duality between contrastive and non-contrastive self-supervised learning
https://openreview.net	//torum?id=kDEL91Dutpa Quantin Garrida, Vuhai Chan, Adrian Pardas, Laurant Naiman, Vann LaCun
HIGHLIGHT:	We show that contrastive and non-contrastive self-supervised methods can be shown to be closely related, and
then study how implei	mentation details impact performance. We validate empirically our findings and significantly improve known
behaviours.	
526, TITLE:	AGRO: Adversarial discovery of error-prone Groups for Robust Optimization
AUTHORS:	Bhargavi Paranjane. Pradeen Dasigi. Vivek Srikumar, Luke Zettlemover, Hannaneh Hajishirzi
HIGHLIGHT:	We propose AGROAdversarial Group discovery for Distributionally Robust Optimizationan end-to-end
approach that jointly i	dentifies error-prone groups and improves accuracy on them.
527, TITLE:	Harnessing Mixed Offline Reinforcement Learning Datasets via Trajectory Weighting
https://openreview.net	/torum?id=OhUAblg2/z Zhang Wai Hong Dami Tachat das Combas Dulkit Agrawal Domain Larocha
HIGHLIGHT:	We show that in mixed datasets consisting of mostly low-return trajectories and minor high-return trajectories.
state-of-the-art offline	RL algorithms are overly restrained by low-return trajectories and fail to exploit high-performing trajectories to
the fullest. To overcor weighted to induce an	ne this issue, we show that, in deterministic MDPs with stochastic initial states, the dataset sampling can be re- artificial dataset whose behavior policy has a higher return.
528, TITLE:	Proto-Value Networks: Scaling Representation Learning with Auxiliary Tasks
https://openreview.net	t/forum?id=oGDKSt9JrZi
AUTHORS:	Jesse Farebrother, Joshua Greaves, Rishabh Agarwal, Charline Le Lan, Ross Goroshin, Pablo Samuel Castro,
Marc G Bellemare	This is nothing symptotic sitten that many synthic wy tasks are defined anoso downly. and hence say he to start does

HIGHLIGHT: This is perhaps surprising given that many auxiliary tasks are defined procedurally, and hence can be treated as an essentially infinite source of information about the environment. Based on this observation, we study the effectiveness of auxiliary

tasks for learning rich representations, focusing on the setting where the number of tasks and the size of the agent's network are simultaneously increased.

529. TITLE: Self-Consistency Improves Chain of Thought Reasoning in Language Models https://openreview.net/forum?id=1PL1NIMMrw AUTHORS: Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc V Le, Ed H. Chi, Sharan Narang, Aakanksha Chowdhery, Denny Zhou HIGHLIGHT: In this paper, we propose a new decoding strategy, self-consistency, to replace the naive greedy decoding used in chain-of-thought prompting. 530, TITLE: Investigating Multi-task Pretraining and Generalization in Reinforcement Learning https://openreview.net/forum?id=sSt9fROSZRO AUTHORS: Adrien Ali Taiga, Rishabh Agarwal, Jesse Farebrother, Aaron Courville, Marc G Bellemare HIGHLIGHT: In this work we propose to investigate the generalization capabilities of a popular actor-critic method, IMPALA. 531, TITLE: ChordMixer: A Scalable Neural Attention Model for Sequences with Different Length https://openreview.net/forum?id=E8mzu3JbdR AUTHORS: Ruslan Khalitov, Tong Yu, Lei Cheng, Zhirong Yang HIGHLIGHT: Here we propose a simple neural network building block called ChordMixer which can model the attention for long sequences with variable lengths. 532, TITLE: Personalized Federated Learning with Feature Alignment and Classifier Collaboration https://openreview.net/forum?id=SXZr8aDKia AUTHORS: Jian Xu, Xinyi Tong, Shao-Lun Huang HIGHLIGHT: In this work, we conduct explicit local-global feature alignment by leveraging global semantic knowledge for learning a better representation. EA-HAS-Bench: Energy-aware Hyperparameter and Architecture Search Benchmark 533, TITLE: https://openreview.net/forum?id=n-bvaLSCC78 AUTHORS: Shuguang Dou, XINYANG JIANG, Cai Rong Zhao, Dongsheng Li Specifically, we present the first large-scale energy-aware benchmark that allows studying AutoML methods to HIGHLIGHT: achieve better trade-offs between performance and search energy consumption, named EA-HAS-Bench. 534, TITLE: Distributionally Robust Recourse Action https://openreview.net/forum?id=E3ip6qBLF7 AUTHORS: Duy Nguyen, Ngoc Bui, Viet Anh Nguyen However, this assumption does not always hold in practice because of data distribution shifts, and in this case, HIGHLIGHT: the recourse action may become invalid. To redress this shortcoming, we propose the Distributionally Robust Recourse Action (DiRRAc) framework, which generates a recourse action that has high probability of being valid under a mixture of model shifts. A probabilistic framework for task-aligned intra- and inter-area neural manifold estimation 535. TITLE: https://openreview.net/forum?id=kt-dcBQcSA AUTHORS: Edoardo Balzani, Jean-Paul G Noel, Pedro Herrero-Vidal, Dora E Angelaki, Cristina Savin HIGHLIGHT: Here we propose a novel probabilistic framework that allows for interpretable partitioning of population variability within and across areas in the context of naturalistic behavior. 536, TITLE: Block and Subword-Scaling Floating-Point (BSFP) : An Efficient Non-Uniform Quantization For Low Precision Inference https://openreview.net/forum?id=VWm4o413V9e AUTHORS: Yun-Chen Lo, Tse-Kuang Lee, Ren-Shuo Liu HIGHLIGHT: In this paper, we propose Block and Subword-Scaling Floating-Point (BSFP), a non-uniform quantization scheme for the skewed and non-uniform distribution of weight vectors in neural networks. 537, TITLE: Simple Yet Effective Graph Contrastive Learning for Recommendation https://openreview.net/forum?id=FKXVK9dyMM AUTHORS: Xuheng Cai, Chao Huang, Lianghao Xia, Xubin Ren In this paper, we propose a simple yet effective graph contrastive learning paradigm LightGCL that mitigates HIGHLIGHT: these issues that negatively impact the generality and robustness of CL-based recommenders. 538, TITLE: Unbiased Supervised Contrastive Learning https://openreview.net/forum?id=Ph5cJSfD2XN AUTHORS: Carlo Alberto Barbano, Benoit Dufumier, Enzo Tartaglione, Marco Grangetto, Pietro Gori HIGHLIGHT: In this work, we tackle the problem of learning representations that are robust to biases. 539, TITLE: SQA3D: Situated Question Answering in 3D Scenes https://openreview.net/forum?id=IDJx97BC38 AUTHORS: Xiaojian Ma, Silong Yong, Zilong Zheng, Qing Li, Yitao Liang, Song-Chun Zhu, Siyuan Huang HIGHLIGHT: We propose a new task to benchmark scene understanding of embodied agents: Situated Question Answering in 3D Scenes (SQA3D).

540, TITLE: Data Valuation Without Training of a Model https://openreview.net/forum?id=XIzO8zr-WbM AUTHORS: Ki Nohyun, Hoyong Choi, Hye Won Chung HIGHLIGHT: In this paper, we provide a training-free data valuation score, called complexity-gap score, which is a datacentric score to quantify the influence of individual instances in generalization of two-layer overparameterized neural networks. 541, TITLE: HotProtein: A Novel Framework for Protein Thermostability Prediction and Editing https://openreview.net/forum?id=YDJRFWBMNby AUTHORS: Tianlong Chen, Chengyue Gong, Daniel Jesus Diaz, Xuxi Chen, Jordan Tyler Wells, qiang liu, Zhangyang Wang, Andrew Ellington, Alex Dimakis, Adam Klivans HIGHLIGHT: We present \$\texttf {HotProtein}\$, a large-scale protein dataset with \textif {growth temperature} annotations of thermostability, containing \$182\$K amino acid sequences and \$3\$K folded structures from \$230\$ different species with a wide temperature range $-20^{\text{circ}} \times C^{\text{circ}} \times C^{\text{circ}} \times C^{\text{circ}}$ 542, TITLE: Switch-NeRF: Learning Scene Decomposition with Mixture of Experts for Large-scale Neural Radiance Fields https://openreview.net/forum?id=PQ2zoIZqvm AUTHORS: Zhenxing MI, Dan Xu HIGHLIGHT: To tackle these issues, in this paper, we propose Switch-NeRF, a novel end-to-end large-scale NeRF with learning-based scene decomposition. Measuring Forgetting of Memorized Training Examples 543. TITLE: https://openreview.net/forum?id=7bJizxLKrR AUTHORS: Matthew Jagielski, Om Thakkar, Florian Tramer, Daphne Ippolito, Katherine Lee, Nicholas Carlini, Eric Wallace, Shuang Song, Abhradeep Guha Thakurta, Nicolas Papernot, Chiyuan Zhang HIGHLIGHT: When models are trained on large datasets, we show that privacy attacks become less effective on examples seen early in training, and investigate why. 544, TITLE: Fundamental Limits in Formal Verification of Message-Passing Neural Networks https://openreview.net/forum?id=WlbG820mRH-AUTHORS: Marco Sälzer, Martin Lange We show that in the context of Message Passing Neural Networks (MPNN), a common Graph Neural Network HIGHLIGHT: (GNN) model, formal verification is impossible. 545, TITLE: Part-Based Models Improve Adversarial Robustness https://openreview.net/forum?id=bAMTaeqluh4 AUTHORS: Chawin Sitawarin, Kornrapat Pongmala, Yizheng Chen, Nicholas Carlini, David Wagner We show that combining human prior knowledge with end-to-end learning can improve the robustness of deep HIGHLIGHT: neural networks by introducing a part-based model for object classification. 546, TITLE: GLM-130B: An Open Bilingual Pre-trained Model https://openreview.net/forum?id=-Aw0rrrPUF AUTHORS: Aohan Zeng, Xiao Liu, Zhengxiao Du, Zihan Wang, Hanyu Lai, Ming Ding, Zhuoyi Yang, Yifan Xu, Wendi Zheng, Xiao Xia, Weng Lam Tam, Zixuan Ma, Yufei Xue, Jidong Zhai, Wenguang Chen, Zhiyuan Liu, Peng Zhang, Yuxiao Dong, Jie Tang HIGHLIGHT: We introduce GLM-130B, a bilingual (English and Chinese) pre-trained language model with 130 billion parameters. 547, TITLE: How robust is unsupervised representation learning to distribution shift? https://openreview.net/forum?id=LiXDW7CF94J Yuge Shi, Imant Daunhawer, Julia E Vogt, Philip Torr, Amartya Sanyal AUTHORS: HIGHLIGHT: We posit that the input-driven objectives of unsupervised algorithms lead to representations that are more robust to distribution shift than the target-driven objective of SL. We verify this by extensively evaluating the performance of SSL and AE on both synthetic and realistic distribution shift datasets. 548, TITLE: 3D Equivariant Diffusion for Target-Aware Molecule Generation and Affinity Prediction https://openreview.net/forum?id=kJqXEPXMsE0 Jiaqi Guan, Wesley Wei Qian, Xingang Peng, Yufeng Su, Jian Peng, Jianzhu Ma AUTHORS: HIGHLIGHT: However, current 3D target-aware models either rely on the voxelized atom densities or the autoregressive sampling process, which are not equivariant to rotation or easily violate geometric constraints resulting in unrealistic structures. In this work, we develop a 3D equivariant diffusion model to solve the above challenges. 549, TITLE: Differentially Private Adaptive Optimization with Delayed Preconditioners https://openreview.net/forum?id=j1zQGmQQOX1 Tian Li, Manzil Zaheer, Ken Liu, Sashank J. Reddi, Hugh Brendan McMahan, Virginia Smith AUTHORS: HIGHLIGHT: In this work, we explore techniques to estimate and efficiently adapt to gradient geometry in private adaptive

550, TITLE: On the Perils of Cascading Robust Classifiers https://openreview.net/forum?id=tQG-o3SeipT

optimization without auxiliary data.

AUTHORS: Ravi Mangal, Zifan Wang, Chi Zhang, Klas Leino, Corina Pasareanu, Matt Fredrikson HIGHLIGHT: We present a new attack against cascading ensembles and show that: (1) there exists an adversarial input for up to 88\% of the samples where the ensemble claims to be certifiably robust and accurate; and (2) the accuracy of a cascading ensemble under our attack is as low as 11\% when it claims to be certifiably robust and accurate on 97\% of the test set. Graph Contrastive Learning for Skeleton-based Action Recognition 551, TITLE: https://openreview.net/forum?id=PLUXnnxUdr4 AUTHORS: Xiaohu Huang, Hao Zhou, Bin Feng, Xinggang Wang, Wenyu Liu, Jian Wang, Haocheng Feng, Junyu Han, Errui Ding, Jingdong Wang HIGHLIGHT: In this paper, we propose a graph contrastive learning framework for skeleton-based action recognition (\$\textit{SkeletonGCL}\$) to explore the \$\textit{global}\$ context across all sequences. An Image is Worth One Word: Personalizing Text-to-Image Generation using Textual Inversion 552, TITLE: https://openreview.net/forum?id=NAQvF08TcvG AUTHORS: Rinon Gal, Yuval Alaluf, Yuval Atzmon, Or Patashnik, Amit Haim Bermano, Gal Chechik, Daniel Cohen-or HIGHLIGHT: In other words, we ask: how can we use language-guided models to turn *our* cat into a painting, or imagine a new product based on *our* favorite toy? Here we present a simple approach that allows such creative freedom. 553, TITLE: Learning Cut Selection for Mixed-Integer Linear Programming via Hierarchical Sequence Model https://openreview.net/forum?id=Zob4P9bRNcK AUTHORS: Zhihai Wang, Xijun Li, Jie Wang, Yufei Kuang, Mingxuan Yuan, Jia Zeng, Yongdong Zhang, Feng Wu HIGHLIGHT: Moreover, we observe from extensive empirical results that (P3) what order of selected cuts should be preferred has a significant impact on the efficiency of solving MILPs as well. To address this challenge, we propose a novel hierarchical sequence model (HEM) to learn cut selection policies via reinforcement learning. 554, TITLE: Visual Classification via Description from Large Language Models https://openreview.net/forum?id=jlAjNL8z5cs AUTHORS: Sachit Menon, Carl Vondrick HIGHLIGHT: We present an alternative framework for classification with VLMs, which we call classification by description. 555, TITLE: CLIPSep: Learning Text-queried Sound Separation with Noisy Unlabeled Videos https://openreview.net/forum?id=H-T3F0dMbyj AUTHORS: Hao-Wen Dong, Naoya Takahashi, Yuki Mitsufuji, Julian McAuley, Taylor Berg-Kirkpatrick HIGHLIGHT: In this work, we aim to approach text-queried universal sound separation by using only unlabeled data. 556, TITLE: Lossless Adaptation of Pretrained Vision Models For Robotic Manipulation https://openreview.net/forum?id=5IND3TXJRb-AUTHORS: Mohit Sharma, Claudio Fantacci, Yuxiang Zhou, Skanda Koppula, Nicolas Heess, Jon Scholz, Yusuf Aytar HIGHLIGHT: Unfortunately, fine-tuning disrupts the pretrained visual representation, and causes representational drift towards the fine-tuned task thus leading to a loss of the versatility of the original model. We introduce a method for lossless adaptation to address this shortcoming of classical fine-tuning. 557, TITLE: The Modality Focusing Hypothesis: Towards Understanding Crossmodal Knowledge Distillation https://openreview.net/forum?id=w0QXrZ3N-s AUTHORS: Zihui Xue, Zhengqi Gao, Sucheng Ren, Hang Zhao HIGHLIGHT: In this paper, we present a thorough understanding of crossmodal KD. 558, TITLE: FedFA: Federated Feature Augmentation https://openreview.net/forum?id=U9vFP90jU0 AUTHORS: Tianfei Zhou, Ender Konukoglu HIGHLIGHT: The primary goal of this paper is to develop a robust federated learning algorithm to address feature shift in clients' samples, which can be caused by various factors, e.g., acquisition differences in medical imaging. 559, TITLE: Adversarial Training descends without descent: Finding actual descent directions based on Danskin's theorem https://openreview.net/forum?id=I3HCE7Ro78H AUTHORS: Fabian Latorre, Igor Krawczuk, Leello Tadesse Dadi, Thomas Pethick, Volkan Cevher HIGHLIGHT: More precisely, we provide a counterexample to a corollary of Danskin's Theorem presented in the seminal paper of Madry et al. (2018) which states that a solution of the inner maximization problem can yield a descent direction for the adversarially robust loss. 560, TITLE: Visual Recognition with Deep Nearest Centroids https://openreview.net/forum?id=CsKwavjr7A AUTHORS: Wenguan Wang, Cheng Han, Tianfei Zhou, Dongfang Liu HIGHLIGHT: We devise deep nearest centroids (DNC), a conceptually elegant yet surprisingly effective network for largescale visual recognition, by revisiting Nearest Centroids, one of the most classic and simple classifiers. 561, TITLE: Dual Diffusion Implicit Bridges for Image-to-Image Translation https://openreview.net/forum?id=5HLoTvVGDe AUTHORS: Xuan Su, Jiaming Song, Chenlin Meng, Stefano Ermon

HIGHLIGHT: We present Dual Diffusion Implicit Bridges (DDIBs), an image translation method based on diffusion models, that circumvents training on domain pairs. 562, TITLE: CROM: Continuous Reduced-Order Modeling of PDEs Using Implicit Neural Representations https://openreview.net/forum?id=FUORz1tG8Og AUTHORS: Peter Yichen Chen, Jinxu Xiang, Dong Heon Cho, Yue Chang, G A Pershing, Henrique Teles Maia, Maurizio M Chiaramonte, Kevin Thomas Carlberg, Eitan Grinspun HIGHLIGHT: We propose to accelerate PDE solvers using reduced-order modeling (ROM). 563, TITLE: Universal Vision-Language Dense Retrieval: Learning A Unified Representation Space for Multi-Modal Retrieval https://openreview.net/forum?id=PQOlkgsBsik AUTHORS: Zhenghao Liu, Chenyan Xiong, Yuanhuiyi Lv, Zhiyuan Liu, Ge Yu This paper presents Universal Vision-Language Dense Retrieval (UniVL-DR), which builds a unified model for HIGHLIGHT: multi-modal retrieval. DINO as a von Mises-Fisher mixture model 564, TITLE: https://openreview.net/forum?id=cMJo1FTwBTQ AUTHORS: Hariprasath Govindarajan, Per Sidén, Jacob Roll, Fredrik Lindsten HIGHLIGHT: Given the fact that the learned representations are -normalized, we show that DINO can be interpreted as a mixture model of von Mises-Fisher components. 565, TITLE: Average Sensitivity of Decision Tree Learning https://openreview.net/forum?id=boik01vhssB AUTHORS: Satoshi Hara, Yuichi Yoshida HIGHLIGHT: Specifically, we adopt the notion of average sensitivity as a stability measure, and design an algorithm with low average sensitivity that outputs a decision tree whose accuracy is nearly equal to the optimal decision tree. 566, TITLE: Relational Attention: Generalizing Transformers for Graph-Structured Tasks https://openreview.net/forum?id=cFuMmbWiN6 AUTHORS: Cameron Diao, Ricky Loynd HIGHLIGHT: As set processors, transformers are at a disadvantage in reasoning over more general graph-structured data where nodes represent entities and edges represent relations between entities. To address this shortcoming, we generalize transformer attention to consider and update edge vectors in each transformer layer. 567, TITLE: Holistic Adversarially Robust Pruning https://openreview.net/forum?id=sAJDi9lD06L AUTHORS: Qi Zhao, Christian Wressnegger We propose a novel method, HARP, that copes with aggressive pruning significantly better than prior work. HIGHLIGHT: 568, TITLE: Classically Approximating Variational Quantum Machine Learning with Random Fourier Features https://openreview.net/forum?id=ymFhZxw70uz AUTHORS: Jonas Landman, Slimane Thabet, Constantin Dalyac, Hela Mhiri, Elham Kashefi HIGHLIGHT: In our work, we propose a classical sampling method that can closely approximate most VQCs with Hamiltonian encoding, given only the description of their architecture. 569, TITLE: Distilling Model Failures as Directions in Latent Space https://openreview.net/forum?id=99RpBVpLiX AUTHORS: Saachi Jain, Hannah Lawrence, Ankur Moitra, Aleksander Madry HIGHLIGHT: This can make these methods labor-intensive and dataset-specific. To address these shortcomings, we present a scalable method for automatically distilling a model's failure modes. Theoretical Characterization of the Generalization Performance of Overfitted Meta-Learning 570, TITLE: https://openreview.net/forum?id=Jifob4dSh99 AUTHORS: Peizhong Ju, Yingbin Liang, Ness Shroff However, the theoretical understanding of when and why overparameterized models such as DNNs can HIGHLIGHT: generalize well in meta-learning is still limited. As an initial step towards addressing this challenge, this paper studies the generalization performance of overfitted meta-learning under a linear regression model with Gaussian features. 571, TITLE: Stable Target Field for Reduced Variance Score Estimation https://openreview.net/forum?id=WmIwYTd0YTF AUTHORS: Yilun Xu, Shangyuan Tong, Tommi S. Jaakkola We propose to remedy the problem by incorporating a reference batch for minibatch updates where the HIGHLIGHT: reference batch is used to calculate weighted conditional scores as the more stable training targets. Unveiling the sampling density in non-uniform geometric graphs 572, TITLE: https://openreview.net/forum?id=mnVf1W6ipGm AUTHORS: Raffaele Paolino, Aleksandar Bojchevski, Stephan Günnemann, Gitta Kutyniok, Ron Levie

HIGHLIGHT: We introduce geometric graphs with hubs, an effective model for real-world graphs, and retrieve the sampling density by which those graphs are sampled from continuous latent spaces, to achieve various tasks.

573. TITLE: Interneurons accelerate learning dynamics in recurrent neural networks for statistical adaptation https://openreview.net/forum?id=3mlITJRYYbs AUTHORS: David Lipshutz, Cengiz Pehlevan, Dmitri Chklovskii HIGHLIGHT: In this work, we explore the computational benefits of mediating recurrent communication via interneurons compared with direct recurrent connections. 574, TITLE: Scaling Laws for a Multi-Agent Reinforcement Learning Model https://openreview.net/forum?id=ZrEbzL9eQ3W AUTHORS: Oren Neumann, Claudius Gros HIGHLIGHT: In this paper we present an extensive study of performance scaling for a cornerstone reinforcement learning algorithm, AlphaZero. Combinatorial-Probabilistic Trade-Off: P-Values of Community Properties Test in the Stochastic Block Models 575, TITLE: https://openreview.net/forum?id=8qjSA5QACb40 AUTHORS: Shuting Shen, Junwei Lu HIGHLIGHT: We propose an inferential framework testing the general community combinatorial properties of the stochastic block model. 576. TITLE: Learning the Positions in CountSketch https://openreview.net/forum?id=iV9Cs8s8keU AUTHORS: Yi Li, Honghao Lin, Simin Liu, Ali Vakilian, David Woodruff HIGHLIGHT: In this work, we propose the first learning-based algorithms that also optimize the locations of the non-zero entries. 577, TITLE: Language Modelling with Pixels https://openreview.net/forum?id=FkSp8VW8RjH AUTHORS: Phillip Rust, Jonas F. Lotz, Emanuele Bugliarello, Elizabeth Salesky, Miryam de Lhoneux, Desmond Elliott HIGHLIGHT: Tackling this bottleneck results in a trade-off between what can be represented in the embedding matrix and computational issues in the output layer. This paper introduces PIXEL, the Pixel-based Encoder of Language, which suffers from neither of these issues. 578, TITLE: Active Learning for Object Detection with Evidential Deep Learning and Hierarchical Uncertainty Aggregation https://openreview.net/forum?id=MnEjsw-vj-X AUTHORS: Younghyun Park, Soyeong Kim, Wonjeong Choi, Dong-Jun Han, Jaekyun Moon HIGHLIGHT: In this paper, we propose a new active learning strategy for object detection that overcomes the shortcomings of prior works. 579, TITLE: Your Contrastive Learning Is Secretly Doing Stochastic Neighbor Embedding https://openreview.net/forum?id=XFSCKELP3bp AUTHORS: Tianyang Hu, Zhili LIU, Fengwei Zhou, Wenjia Wang, Weiran Huang HIGHLIGHT: Contrastive learning, especially self-supervised contrastive learning (SSCL), has achieved great success in extracting powerful features from unlabeled data. In this work, we contribute to the theoretical understanding of SSCL and uncover its connection to the classic data visualization method, stochastic neighbor embedding (SNE), whose goal is preserving pairwise distances. 580, TITLE: DexDeform: Dexterous Deformable Object Manipulation with Human Demonstrations and Differentiable Physics https://openreview.net/forum?id=LIV7- 7pYPl AUTHORS: Sizhe Li, Zhiao Huang, Tao Chen, Tao Du, Hao Su, Joshua B. Tenenbaum, Chuang Gan HIGHLIGHT. In this work, we aim to learn dexterous manipulation of deformable objects using multi-fingered hands.Concretely, we first collect a small set of human demonstrations using teleoperation. 581, TITLE: Brain-like representational straightening of natural movies in robust feedforward neural networks https://openreview.net/forum?id=mCmerkTCG2S AUTHORS: Tahereh Toosi, Elias Issa HIGHLIGHT: Here, we show robustness to noise can produce representational straightening in feedforward neural networks. 582, TITLE: Warping the Space: Weight Space Rotation for Class-Incremental Few-Shot Learning https://openreview.net/forum?id=kPLzOfPfA21 AUTHORS: Do-Yeon Kim, Dong-Jun Han, Jun Seo, Jaekyun Moon In this paper, we introduce WaRP, the \textit {weight space rotation process}, which transforms the original HIGHLIGHT: parameter space into a new space so that we can push most of the previous knowledge compactly into only a few important parameters. Amortised Invariance Learning for Contrastive Self-Supervision 583, TITLE: https://openreview.net/forum?id=nXOhmfFu5n AUTHORS: Ruchika Chavhan, Jan Stuehmer, Calum Heggan, Mehrdad Yaghoobi, Timothy Hospedales HIGHLIGHT: We introduce the notion of amortized invariance learning for contrastive self supervision.

584, TITLE: Autoregressive Conditional Neural Processes https://openreview.net/forum?id=OAsXFPBfTBh AUTHORS: Wessel Bruinsma, Stratis Markou, James Requeima, Andrew Y. K. Foong, Anna Vaughan, Tom Andersson, Anthony Buonomo, Scott Hosking, Richard E Turner HIGHLIGHT: In this work, we instead propose to change how CNPs are deployed at test time, without any modifications to the model or training procedure. Control Graph as Unified IO for Morphology-Task Generalization 585, TITLE: https://openreview.net/forum?id=HcUf-QwZeFh AUTHORS: Hiroki Furuta, Yusuke Iwasawa, Yutaka Matsuo, Shixiang Shane Gu HIGHLIGHT: In this work, we explore a method for learning a single policy that manipulates various forms of agents to solve various tasks by distilling a large amount of proficient behavioral data. Contrastive Corpus Attribution for Explaining Representations 586, TITLE: https://openreview.net/forum?id=eWKfMBL5to AUTHORS: Chris Lin, Hugh Chen, Chanwoo Kim, Su-In Lee Although this enabled explanations of unsupervised models, the interpretation of this approach can still be HIGHLIGHT: opaque because similarity to the explicand's representation may not be meaningful to humans. To address this, we propose contrastive corpus similarity, a novel and semantically meaningful scalar explanation output based on a reference corpus and a contrasting foil set of samples. 587, TITLE: ACMP: Allen-Cahn Message Passing with Attractive and Repulsive Forces for Graph Neural Networks https://openreview.net/forum?id=4fZc 79Lrqs Yuelin Wang, Kai Yi, Xinliang Liu, Yu Guang Wang, Shi Jin AUTHORS: HIGHLIGHT: Neural message passing is a basic feature extraction unit for graph-structured data considering neighboring node features in network propagation from one layer to the next. We model such process by an interacting particle system with attractive and repulsive forces and the Allen-Cahn force arising in the modeling of phase transition. Discovering Generalizable Multi-agent Coordination Skills from Multi-task Offline Data 588. TITLE: https://openreview.net/forum?id=53FvUAdP7d AUTHORS: Fuxiang Zhang, Chengxing Jia, Yi-Chen Li, Lei Yuan, Yang Yu, Zongzhang Zhang HIGHLIGHT: In this paper, we propose a novel Offline MARL algorithm to Discover coordination Skills (ODIS) from multitask data. 589, TITLE: Unified Discrete Diffusion for Simultaneous Vision-Language Generation https://openreview.net/forum?id=8JqINxA-2a AUTHORS: Minghui Hu, Chuanxia Zheng, Heliang Zheng, Tat-Jen Cham, Chaoyue Wang, Zuopeng Yang, Dacheng Tao, Ponnuthurai N. Suganthan HIGHLIGHT: The recently developed discrete diffusion model performs extraordinarily well in generation tasks, especially in the text-to-image task, showing great potential for modeling multimodal signals. In this paper, we leverage these properties and present a unified multimodal generation model, which can perform text-based, image-based, and even vision-language simultaneous generation using a single model. CrAM: A Compression-Aware Minimizer 590, TITLE: https://openreview.net/forum?id= eTZBs-yedr AUTHORS: Alexandra Peste, Adrian Vladu, Eldar Kurtic, Christoph H Lampert, Dan Alistarh HIGHLIGHT: In this work we propose a new compression-aware minimizer dubbed CrAM that modifies the optimization step in a principled way, in order to produce models whose local loss behavior is stable under compression operations such as pruning. 591, TITLE: Winning Both the Accuracy of Floating Point Activation and the Simplicity of Integer Arithmetic https://openreview.net/forum?id=z921By1ehjI AUTHORS: Yulhwa Kim, Jaeyong Jang, Jehun Lee, Jihoon Park, Jeonghoon Kim, Byeongwook Kim, Baeseong park, Se Jung Kwon, Dongsoo Lee, jae-joon kim HIGHLIGHT: In this paper, to simultaneously achieve the accuracy of FP activation and the simplicity of integer arithmetic, we present a method for replacing FP arithmetic with integer one without changing FP activations in the storage format while weights are quantized. Variance-Aware Sparse Linear Bandits 592, TITLE: https://openreview.net/forum?id=tkwP32nsEq AUTHORS: Yan Dai, Ruosong Wang, Simon Shaolei Du HIGHLIGHT: In this paper, we present the first variance-aware regret guarantee for sparse linear bandits: \$\widetilde{\mathcal O}\left(\sqrt{d\sum_{t=1}^T \sigma_t^2} + 1\right)\$, where $\frac{t^2}{s}$ is the variance of the noise at the \$t\$-th round. 593, TITLE: DDM\$^2\$: Self-Supervised Diffusion MRI Denoising with Generative Diffusion Models https://openreview.net/forum?id=0vgjc50HfcC Tiange Xiang, Mahmut Yurt, Ali B Syed, Kawin Setsompop, Akshay Chaudhari AUTHORS: Here, we propose Denoising Diffusion Models for Denoising Diffusion MRI (DDM^2), a self-supervised HIGHLIGHT: denoising method for MRI denoising using diffusion denoising generative models. 594, TITLE: Voxurf: Voxel-based Efficient and Accurate Neural Surface Reconstruction

https://openreview.net/forum?id=DSy8tP4WctmZ AUTHORS: Tong Wu, Jiaqi Wang, Xingang Pan, Xudong XU, Christian Theobalt, Ziwei Liu, Dahua Lin HIGHLIGHT: In this work, we present Voxurf, a voxel-based surface reconstruction approach that is both efficient and accurate. 595, TITLE: Guiding Energy-based Models via Contrastive Latent Variables https://openreview.net/forum?id=CZmHHj9MgkP AUTHORS: Hankook Lee, Jongheon Jeong, Sejun Park, Jinwoo Shin HIGHLIGHT: In this paper, we propose a novel and effective framework for improving EBMs via contrastive representation learning (CRL). 596. TITLE: The Curious Case of Benign Memorization https://openreview.net/forum?id=4C8ChYvMYBn AUTHORS: Sotiris Anagnostidis, Gregor Bachmann, Lorenzo Noci, Thomas Hofmann HIGHLIGHT: We demonstrate that deep models have the surprising ability to separate noise from signal by distributing the task of memorization and feature learning to different layers. 597. TITLE: Non-parametric Outlier Synthesis https://openreview.net/forum?id=JHklpEZqduQ AUTHORS: Leitian Tao, Xuefeng Du, Jerry Zhu, Yixuan Li HIGHLIGHT: In this paper, we propose a novel framework, non-parametric outlier synthesis (NPOS), which generates artificial OOD training data and facilitates learning a reliable decision boundary between ID and OOD data. 598. TITLE: Deep Ranking Ensembles for Hyperparameter Optimization https://openreview.net/forum?id= ruvo2KCL2x AUTHORS: Abdus Salam Khazi, Sebastian Pineda Arango, Josif Grabocka HIGHLIGHT: As a result, we present a novel method that meta-learns neural network surrogates optimized for ranking the configurations' performances while modeling their uncertainty via ensembling. 599. TITLE: Unmasking the Lottery Ticket Hypothesis: What's Encoded in a Winning Ticket's Mask? https://openreview.net/forum?id=xSsW2Am-ukZ AUTHORS: Mansheej Paul, Feng Chen, Brett W. Larsen, Jonathan Frankle, Surya Ganguli, Gintare Karolina Dziugaite HIGHLIGHT: And why is iterative pruning needed, i.e. why can't we prune to very high sparsities in one shot? We develop answers to these questions in terms of the geometry of the error landscape. 600, TITLE: Building a Subspace of Policies for Scalable Continual Learning https://openreview.net/forum?id=UKr0MwZM6fL AUTHORS: Jean-Baptiste Gaya, Thang Doan, Lucas Caccia, Laure Soulier, Ludovic Denoyer, Roberta Raileanu HIGHLIGHT: In this work, we aim to strike a better balance between scalability and performance by designing a method whose size grows adaptively depending on the task sequence. Voint Cloud: Multi-View Point Cloud Representation for 3D Understanding 601. TITLE: https://openreview.net/forum?id=IpGgfpMucHj AUTHORS: Abdullah Hamdi, Silvio Giancola, Bernard Ghanem To this end, we introduce the concept of the multi-view point cloud (Voint cloud), representing each 3D point HIGHLIGHT: as a set of features extracted from several view-points. 602, TITLE: Constructive TT-representation of the tensors given as index interaction functions with applications https://openreview.net/forum?id=yLzLfM-Esnu AUTHORS: Gleb Ryzhakov, Ivan Oseledets HIGHLIGHT: This paper presents a method to build explicit tensor-train (TT) representations. 603, TITLE: Real-Time Image Demoir\$\acute {e}\$ing on Mobile Devices https://openreview.net/forum?id=PmP sf3JkrH AUTHORS: Yuxin Zhang, Mingbao Lin, Xunchao Li, Han Liu, Guozhi Wang, Fei Chao, Ren Shuai, Yafei Wen, Xiaoxin Chen, Rongrong Ji HIGHLIGHT: In this paper, we launch the first study on accelerating demoir\$\acute{e}\$ing networks and propose a dynamic demoir\$\acute{e}\$ing acceleration method (DDA) towards a real-time deployment on mobile devices. 604, TITLE: Simple and Scalable Nearest Neighbor Machine Translation https://openreview.net/forum?id=uu1GBD9SILe AUTHORS: Yuhan Dai, Zhirui Zhang, Qiuzhi Liu, Qu Cui, Weihua Li, Yichao Du, Tong Xu HIGHLIGHT: In this paper, we propose a simple and scalable nearest neighbor machine translation framework to drastically promote the decoding and storage efficiency of \$k\$NN-based models while maintaining the translation performance. Near-optimal Coresets for Robust Clustering 605, TITLE: https://openreview.net/forum?id=Nc1ZkRW8Vde AUTHORS: Lingxiao Huang, Shaofeng H.-C. Jiang, Jianing Lou, Xuan Wu

 $\label{eq:highlight} HIGHLIGHT: We consider robust clustering problems in \mathbb{R}^d\$, specifically \$k\$-clustering problems (e.g., \$k\$-Median and \$k\$-Means) with \$m\$ \emph{outliers}, where the cost for a given center set \$C \subset \mathbb{R}^d\ aggregates the distances from \$C\$ to all but the furthest \$m\$ data points, instead of all points as in classical clustering. \\

606. TITLE: Temporal Dependencies in Feature Importance for Time Series Prediction https://openreview.net/forum?id=C0q9oBc3n4 AUTHORS: Kin Kwan Leung, Clayton Rooke, Jonathan Smith, Saba Zuberi, Maksims Volkovs HIGHLIGHT: Time series data introduces two key challenges for explainability methods: firstly, observations of the same feature over subsequent time steps are not independent, and secondly, the same feature can have varying importance to model predictions over time. In this paper, we propose Windowed Feature Importance in Time (WinIT), a feature removal based explainability approach to address these issues. Reliability of CKA as a Similarity Measure in Deep Learning 607. TITLE: https://openreview.net/forum?id=8HRvyxc606 AUTHORS: MohammadReza Davari, Stefan Horoi, Amine Natik, Guillaume Lajoie, Guy Wolf, Eugene Belilovsky HIGHLIGHT: In this work we present analysis that formally characterizes CKA sensitivity to a large class of simple transformations, which can naturally occur in the context of modern machine learning. 608, TITLE: Identifiability Results for Multimodal Contrastive Learning https://openreview.net/forum?id=U 2kugoTcB AUTHORS: Imant Daunhawer, Alice Bizeul, Emanuele Palumbo, Alexander Marx, Julia E Vogt HIGHLIGHT: In this work, we present new identifiability results for multimodal contrastive learning, showing that it is possible to recover shared factors in a more general setup than the multi-view setting studied previously. 609, TITLE: Prototypical Calibration for Few-shot Learning of Language Models https://openreview.net/forum?id=nUsP91FADUF Zhixiong Han, Yaru Hao, Li Dong, Yutao Sun, Furu Wei AUTHORS: HIGHLIGHT: In this work, we propose prototypical calibration to adaptively learn a more robust decision boundary for zeroand few-shot classification, instead of greedy decoding. 610, TITLE: Rotamer Density Estimators are Unsupervised Learners of the Effect of Mutations on Protein-Protein Interaction https://openreview.net/forum?id= X9Y11K2mD AUTHORS: Shitong Luo, Yufeng Su, Zuofan Wu, Chenpeng Su, Jian Peng, Jianzhu Ma HIGHLIGHT: In this work, we demonstrate that mutational effects on binding can be predicted by the change in conformational flexibility of the protein-protein interface.

611, TITLE: Exploring Active 3D Object Detection from a Generalization Perspective

https://openreview.net/forum?id=2RwXVje1rAh

AUTHORS: Yadan Luo, Zhuoxiao Chen, Zijian Wang, Xin Yu, Zi Huang, Mahsa Baktashmotlagh HIGHLIGHT: Our empirical study, however, suggests that mainstream uncertainty-based and diversity-based active learning policies are not effective when applied in the 3D detection task, as they fail to balance the trade-off between point cloud informativeness and box-level annotation costs. To overcome this limitation, we jointly investigate three novel criteria in our framework CRB for point cloud acquisition - label conciseness, feature representativeness and geometric balance, which hierarchically filters out the point clouds of redundant 3D bounding box labels, latent features and geometric characteristics (e.g., point cloud density) from the unlabeled sample pool and greedily selects informative ones with fewer objects to annotate.

612, TITLE: Copy is All You Need

https://openreview.net/forum?id=CROIOA9Nd8C

AUTHORS: Tian Lan, Deng Cai, Yan Wang, Heyan Huang, Xian-Ling Mao

HIGHLIGHT: In this paper, we formulate text generation as progressively copying text segments (e.g., words or phrases) from an existing text collection.

613, TITLE: Latent State Marginalization as a Low-cost Approach to Improving Exploration https://openreview.net/forum?id=b0UksKFcTOL

AUTHORS:Dinghuai Zhang, Aaron Courville, Yoshua Bengio, Qinqing Zheng, Amy Zhang, Ricky T. Q. ChenHIGHLIGHT:In this work, we propose the adoption of latent variable policies within the MaxEnt framework, which we canprovably approximate any policy distribution, and additionally, naturally emerges under the use of world models with a latent beliefstate.

614, TITLE: Generative Augmented Flow Networks

https://openreview.net/forum?id=urF_CBK5XC0

 AUTHORS:
 Ling Pan, Dinghuai Zhang, Aaron Courville, Longbo Huang, Yoshua Bengio

 HIGHLIGHT:
 Indeed, intermediate rewards play a critical role in learning, for example from intrinsic motivation to provide

 intermediate feedback even in particularly challenging sparse reward tasks. Inspired by this, we propose Generative Augmented Flow

 Networks (GAFlowNets), a novel learning framework to incorporate intermediate rewards into GFlowNets.

615, TITLE: What Makes Convolutional Models Great on Long Sequence Modeling? https://openreview.net/forum?id=TGJSPbRpJX-AUTHORS: Yuhong Li, Tianle Cai, Yi Zhang, Deming Chen, Debadeepta Dey

HIGHLIGHT: Based on the two principles, we propose a simple yet effective convolutional model called Structured Global Convolution (SGConv). 616, TITLE: Learning Math Reasoning from Self-Sampled Correct and Partially-Correct Solutions https://openreview.net/forum?id=4D4TSJE6-K AUTHORS: Ansong Ni, Jeevana Priya Inala, Chenglong Wang, Alex Polozov, Christopher Meek, Dragomir Radev, Jianfeng Gao HIGHLIGHT: This way, the finetuned models are biased towards the limited reference solutions, which limits their generalization to unseen examples. To mitigate this issue, we propose to let the model perform sampling during training and learn from both self-sampled fully-correct solutions, which yield the correct answer upon execution, and partially-correct solutions, whose intermediate state matches an intermediate state of a known correct solution. GFlowNets and variational inference 617. TITLE: https://openreview.net/forum?id=uKiE0VIluA-AUTHORS: Nikolay Malkin, Salem Lahlou, Tristan Deleu, Xu Ji, Edward J Hu, Katie E Everett, Dinghuai Zhang, Yoshua Bengio HIGHLIGHT: This paper builds bridges between two families of probabilistic algorithms: (hierarchical) variational inference (VI), which is typically used to model distributions over continuous spaces, and generative flow networks (GFlowNets), which have been used for distributions over discrete structures such as graphs. 618, TITLE: Unsupervised 3d object learning through neuron activity aware plasticity https://openreview.net/forum?id=mXPoBtnpMnuy AUTHORS: Beomseok Kang, Biswadeep Chakraborty, Saibal Mukhopadhyay HIGHLIGHT: We present an unsupervised deep learning model for 3D object classification. 619, TITLE: Heterogeneous Neuronal and Synaptic Dynamics for Spike-Efficient Unsupervised Learning: Theory and **Design Principles** https://openreview.net/forum?id=QIRtAqoXwj AUTHORS: Biswadeep Chakraborty, Saibal Mukhopadhyay HIGHLIGHT: This paper shows that the heterogeneity in neuronal and synaptic dynamics reduces the spiking activity of a Recurrent Spiking Neural Network (RSNN) while improving prediction performance, enabling spike-efficient (unsupervised) learning. Associative Memory Augmented Asynchronous Spatiotemporal Representation Learning for Event-based 620, TITLE: Perception https://openreview.net/forum?id=ZCStthyW-TD AUTHORS: Uday Kamal, Saurabh Dash, Saibal Mukhopadhyay We propose \$\textit {EventFormer}\$, a computationally efficient event-based representation learning framework HIGHLIGHT: for asynchronously processing event camera data. 621, TITLE: Backstepping Temporal Difference Learning https://openreview.net/forum?id=YPChvOgRXRA AUTHORS: Han-Dong Lim, Donghwan Lee HIGHLIGHT: To overcome the divergent behavior, several off-policy TD learning algorithms have been developed until now. In this work, we provide a unified view of such algorithms from a purely control-theoretic perspective. 622, TITLE: TrojText: Test-time Invisible Textual Trojan Insertion https://openreview.net/forum?id=ja4Lpp5mqc2 AUTHORS: Yepeng Liu, Bo Feng, Qian Lou HIGHLIGHT: In this paper, we propose TrojText to study whether the invisible textual Trojan attack can be efficiently performed without the presence of training data in a more realistic and cost-efficient manner. Cross-Layer Retrospective Retrieving via Layer Attention 623, TITLE: https://openreview.net/forum?id=pvgEL1yS3Ql Yanwen Fang, Yuxi CAI, Jintai Chen, Jingyu Zhao, Guangjian Tian, Guodong Li AUTHORS: HIGHLIGHT: More and more evidence has shown that strengthening layer interactions can enhance the representation power of a deep neural network, while self-attention excels at learning interdependencies by retrieving query-activated information. Motivated by this, we devise a cross-layer attention mechanism, called multi-head recurrent layer attention (MRLA), that sends a query representation of the current layer to all previous layers to retrieve query-related information from different levels of receptive fields. 624, TITLE: Provable Defense Against Geometric Transformations https://openreview.net/forum?id=ThXqBsRI-cY AUTHORS: Rem Yang, Jacob Laurel, Sasa Misailovic, Gagandeep Singh HIGHLIGHT: However, no prior work has been able to incorporate the objective of deterministic certified robustness against geometric transformations into the training procedure, as existing verifiers are exceedingly slow. To address these challenges, we propose the first provable defense for deterministic certified geometric robustness.

625, TITLE: Improving Deep Regression with Ordinal Entropy https://openreview.net/forum?id=raU07GpP0P

AUTHORS:	Shihao Zhang, Linlin Yang, Michael Bi Mi, Xiaoxu Zheng, Angela Yao
HIGHLIGHT:	Based on the analysis, we propose an ordinal entropy loss to encourage higher-entropy feature spaces while
maintaining ordinal r	elationships to improve the performance of regression tasks.
626, TITLE: https://openreview.nd AUTHORS: HIGHLIGHT: text pair data.	Learning Object-Language Alignments for Open-Vocabulary Object Detection et/forum?id=mjHlitXvReu Chuang Lin, Peize Sun, Yi Jiang, Ping Luo, Lizhen Qu, Gholamreza Haffari, Zehuan Yuan, Jianfei Cai In this paper, we propose a novel open-vocabulary object detection framework directly learning from image-
627, TITLE:	SP2 : A Second Order Stochastic Polyak Method
https://openreview.ne	et/forum?id=5mqFra2ZSuf
AUTHORS:	Shuang Li, William Joseph Swartworth, Martin Taká?, Deanna Needell, Robert M. Gower
HIGHLIGHT:	We show SP2 is very competitive on matrix completion, non-convex test problems and logistic regression.
628, TITLE:	ExpressivE: A Spatio-Functional Embedding For Knowledge Graph Completion
https://openreview.nd	et/forum?id=xkev3_np08z
AUTHORS:	Aleksandar Pavlovic, Emanuel Sallinger
HIGHLIGHT:	Embedding models have yielded promising results for KGC, yet any current KGC embedding model is
incapable of: (1) fully	y capturing vital inference patterns (e.g., composition), (2) capturing prominent logical rules jointly (e.g.,
hierarchy and compo	sition), and (3) providing an intuitive interpretation of captured patterns. In this work, we propose ExpressivE, a
fully expressive spati	io-functional embedding model that solves all these challenges simultaneously.
629, TITLE:	Pareto-Optimal Diagnostic Policy Learning in Clinical Applications via Semi-Model-Based Deep
Reinforcement Learn	hing
https://openreview.ne	et/forum?id=0WVNuEnqVu
AUTHORS:	Zheng Yu, Yikuan Li, Joseph Chahn Kim, Kaixuan Huang, Yuan Luo, Mengdi Wang
HIGHLIGHT:	In this work, we use reinforcement learning (RL) to find a dynamic policy that selects lab test panels
sequentially based or	a previous observations, ensuring accurate testing at a low cost.
630, TITLE:	Memory Gym: Partially Observable Challenges to Memory-Based Agents
https://openreview.net	et/forum?id=jHc8dCx6DDr
AUTHORS:	Marco Pleines, Matthias Pallasch, Frank Zimmer, Mike Preuss
HIGHLIGHT:	Memory Gym is a novel benchmark for challenging Deep Reinforcement Learning agents to memorize events
across long sequence	es, be robust to noise, and generalize.
631, TITLE:	Multi-lingual Evaluation of Code Generation Models
https://openreview.nd	et/forum?id=Bo7eeXm6An8
AUTHORS:	Ben Athiwaratkun, Sanjay Krishna Gouda, Zijian Wang, Xiaopeng Li, Yuchen Tian, Ming Tan, Wasi Uddin
Ahmad, Shiqi Wang,	Qing Sun, Mingyue Shang, Sujan Kumar Gonugondla, Hantian Ding, Varun Kumar, Nathan Fulton, Arash
Farahani, Siddhartha	Jain, Robert Giaquinto, Haifeng Qian, Murali Krishna Ramanathan, Ramesh Nallapati, Baishakhi Ray,
Parminder Bhatia, Su	idipta Sengupta, Dan Roth, Bing Xiang
HIGHLIGHT:	We present MBXP, an execution-based code completion benchmark in 10+ programming languages.
632, TITLE:	Neural Design for Genetic Perturbation Experiments
https://openreview.nd	et/forum?id=TUBpc5rqGA
AUTHORS:	Aldo Pacchiano, Drausin Wulsin, Robert A Barton, Luis Voloch
HIGHLIGHT:	This work provides a theoretically sound framework for iteratively exploring the space of perturbations in
pooled batches in ord	ler to maximize a target phenotype under an experimental budget. Inspired by this application domain, we study
the problem of batch	query bandit optimization and introduce the Optimistic Arm Elimination (\$\mathrm{OAE}\$) principle designed
to find an almost opti	imal arm under different functional relationships between the queries (arms) and the outputs (rewards).
633, TITLE:	MPCFORMER: FAST, PERFORMANT AND PRIVATE TRANSFORMER INFERENCE WITH MPC
https://openreview.nd	et/forum?id=CWmvjOEhgH-
AUTHORS:	Dacheng Li, Hongyi Wang, Rulin Shao, Han Guo, Eric Xing, Hao Zhang
HIGHLIGHT:	In this paper, we design the framework MPCFORMER using secure multi-party computation (MPC) and
Knowledge Distillati	on (KD).
634, TITLE:	Searching Lottery Tickets in Graph Neural Networks: A Dual Perspective
https://openreview.nd	et/forum?id=Dvs-a3aymPe
AUTHORS:	Kun Wang, Yuxuan Liang, Pengkun Wang, Xu Wang, Pengfei Gu, Junfeng Fang, Yang Wang
HIGHLIGHT:	In this paper, we explore the searching of graph lottery tickets from a complementary perspective
transforming a rando	m ticket into a graph lottery ticket, which allows us to more comprehensively explore the relationships between
the original network/	graph and their sparse counterpart.
635, TITLE: Transition	Learning Adversarial Linear Mixture Markov Decision Processes with Bandit Feedback and Unknown
https://openreview.ne	et/forum?id=sVU54nyaA9K
AUTHORS:	Canzhe Zhao, Ruofeng Yang, Baoxiang Wang, Shuai Li

We propose an efficient algorithm LSUOB-REPS which achieves HIGHLIGHT:

 $\operatorname{S}(dS^2) = \operatorname{S}(dS^2) + \operatorname{S$ feature mapping, \$S\$ is the size of the state space, \$A\$ is the size of the action space, \$H\$ is the episode length and \$K\$ is the number of episodes.

636, TITLE: Partial Label Unsupervised Domain Adaptation with Class-Prototype Alignment https://openreview.net/forum?id=jpq0qHggw3t

AUTHORS: Yan Yan, Yuhong Guo

HIGHLIGHT: However, this assumption does not hold in many real-world scenarios where the training and test data come from different distributions. In this paper, we formalize this learning scenario as a new problem called partial label unsupervised domain adaptation (PLUDA).

ResAct: Reinforcing Long-term Engagement in Sequential Recommendation with Residual Actor 637. TITLE: https://openreview.net/forum?id=HmPOzJQhbwg

AUTHORS: Wanqi Xue, Qingpeng Cai, Ruohan Zhan, Dong Zheng, Peng Jiang, Kun Gai, Bo An

HIGHLIGHT: In this paper, we propose ResAct which seeks a policy that is close to, but better than, the online-serving policy. In this way, we can collect sufficient data near the learned policy so that state-action values can be properly estimated, and there is no need to perform online exploration.

Tailoring Language Generation Models under Total Variation Distance 638. TITLE:

https://openreview.net/forum?id=VELL0PlWfc

Haozhe Ji, Pei Ke, Zhipeng Hu, Rongsheng Zhang, Minlie Huang AUTHORS:

HIGHLIGHT: Moreover, in the attempt to cover the low-probability regions in the data distribution, the model systematically overestimates the probability of corrupted text sequences, which we conjecture is one of the main reasons for text degeneration during autoregressive decoding. To remedy this problem, we leverage the total variation distance (TVD) with its robustness to outliers, and develop practical bounds to apply it to language generation.

639, TITLE: Toward Adversarial Training on Contextualized Language Representation

https://openreview.net/forum?id=xZD10GhCvM

AUTHORS: Hongqiu Wu, Yongxiang Liu, Hanwen Shi, hai zhao, Min Zhang

Based on the observations, we propose simple yet effective \textit {Contextualized representation-Adversarial HIGHLIGHT: Training} (CreAT), in which the attack is explicitly optimized to deviate the contextualized representation and obtains the global worst-case adversarial examples.

640, TITLE: Composite Slice Transformer: An Efficient Transformer with Composition of Multi-Scale Multi-Range Attentions

https://openreview.net/forum?id=nWTzIsgrYNN

AUTHORS: Mingu Lee, Saurabh Pitre, Tianyu Jiang, Pierre-David Letourneau, Matthew J Morse, Kanghwan Jang, Joseph Soriaga, Parham Noorzad, Hsin-Pai Cheng, Christopher Lott

HIGHLIGHT: In this paper, we propose Composite Slice Transformer (CST), a Transformer-based network equipped with a composition of multi-scale multi-range attentions, boosting both efficiency and modeling capability.

User-Interactive Offline Reinforcement Learning 641, TITLE:

https://openreview.net/forum?id=a4COps0uokg

AUTHORS: Phillip Swazinna, Steffen Udluft, Thomas Runkler HIGHLIGHT: At the same time, offline RL algorithms are not able to tune their most important hyperparameter - the proximity of the learned policy to the original policy. We propose an algorithm that allows the user to tune this hyperparameter at runtime, thereby addressing both of the above mentioned issues simultaneously.

642, TITLE: Monocular Scene Reconstruction with 3D SDF Transformers https://openreview.net/forum?id=-iADdfa4GKH AUTHORS: Weihao Yuan, Xiaodong Gu, Heng Li, Zilong Dong, Siyu Zhu HIGHLIGHT: In this work, we propose an SDF transformer network, which replaces the role of 3D CNN for better 3D feature aggregation. DENSE RGB SLAM WITH NEURAL IMPLICIT MAPS 643, TITLE: https://openreview.net/forum?id=OUK1ExlbbA Heng Li, Xiaodong Gu, Weihao Yuan, luwei yang, Zilong Dong, Ping Tan AUTHORS: HIGHLIGHT: In this paper, we present a dense RGB SLAM method with neural implicit map representation. 644, TITLE: Federated Learning as Variational Inference: A Scalable Expectation Propagation Approach https://openreview.net/forum?id=dZrQR7OR11 AUTHORS: Han Guo, Philip Greengard, Hongyi Wang, Andrew Gelman, Eric Xing, Yoon Kim HIGHLIGHT: This paper extends the inference view and describes a variational inference formulation of federated learning where the goal is to find a global variational posterior that well-approximates the true posterior. 645, TITLE: Near-optimal Policy Identification in Active Reinforcement Learning https://openreview.net/forum?id=3OR2tbtnYC-

AUTHORS: Xiang Li, Viraj Mehta, Johannes Kirschner, Ian Char, Willie Neiswanger, Jeff Schneider, Andreas Krause, Ilija Bogunovic

HIGHLIGHT: We propose the AE-LSVI algorithm for best policy identification, a novel variant of the kernelized least-squares value iteration (LSVI) algorithm that combines optimism with pessimism for active exploration (AE).

646, TITLE: Semi-supervised Community Detection via Structural Similarity Metrics

https://openreview.net/forum?id=cxvEGLCHpgl AUTHORS:

Yicong Jiang, Tracy Ke

HIGHLIGHT: We propose a fast algorithm that computes a 'structural similarity metric' between the new node and each of the \$K\$ communities, aggregating information in labeled and unlabeled data.

Diffusion Models for Causal Discovery via Topological Ordering 647. TITLE:

https://openreview.net/forum?id=Idusfje4-Wq

AUTHORS: Pedro Sanchez, Xiao Liu, Alison Q O'Neil, Sotirios A. Tsaftaris

HIGHLIGHT: However, existing computational methods for obtaining the Hessian still do not scale as the number of variables and the number of samples are increased. Therefore, inspired by recent innovations in diffusion probabilistic models (DPMs), we propose \emph{DiffAN}, a topological ordering algorithm that leverages DPMs.

648, TITLE: FoSR: First-order spectral rewiring for addressing oversquashing in GNNs https://openreview.net/forum?id=3YjQfCLdrzz AUTHORS: Kedar Karhadkar, Pradeep Kr. Banerjee, Guido Montufar HIGHLIGHT: We propose a computationally efficient algorithm that prevents oversquashing by systematically adding edges

to the graph based on spectral expansion.

649, TITLE: One-Pixel Shortcut: On the Learning Preference of Deep Neural Networks

https://openreview.net/forum?id=p7G8t5FVn2h

AUTHORS: Shutong Wu, Sizhe Chen, Cihang Xie, Xiaolin Huang

In this paper, we resolve this problem from a novel perspective by perturbing only one pixel in each image.By HIGHLIGHT: OPS, we introduce an unlearnable dataset called CIFAR-10-S, which is indistinguishable from CIFAR-10 by humans but induces the trained model to extremely low accuracy.

650, TITLE: Defending against Adversarial Audio via Diffusion Model https://openreview.net/forum?id=5-Df3tljit7

AUTHORS: Shutong Wu, Jiongxiao Wang, Wei Ping, Weili Nie, Chaowei Xiao

HIGHLIGHT: In this paper, we propose an adversarial purification-based defense pipeline, AudioPure, for acoustic systems via off-the-shelf diffusion models.

651, TITLE: Calibration Matters: Tackling Maximization Bias in Large-scale Advertising Recommendation Systems https://openreview.net/forum?id=wzlWiO_WY4

AUTHORS: Yewen Fan, Nian Si, Kun Zhang

HIGHLIGHT: It persists even if unbiased predictions are achieved on every datapoint and worsens when covariate shifts exist between the training and test sets. To mitigate this problem, we quantify maximization bias and propose a variance-adjusting debiasing (VAD) meta-algorithm in this paper.

652, TITLE: Scalable Estimation of Nonparametric Markov Networks with Mixed-Type Data

https://openreview.net/forum?id=qBvBycTqVJ

AUTHORS: Yujia Zheng, Ignavier Ng, Yewen Fan, Kun Zhang

HIGHLIGHT: In this work, we generalize the characterization of the conditional independence structure to handle general distributions for all data types (i.e., continuous, discrete, and mixed-type) with general functional relations among variables, thus giving rise to a Markov network structure learning algorithm in one of the most general settings.

653, TITLE: FIT: A Metric for Model Sensitivity https://openreview.net/forum?id=PDG4-Y3aboN AUTHORS: Ben Zandonati, Adrian Alan Pol, Maurizio Pierini, Olya Sirkin, Tal Kopetz HIGHLIGHT: We find that FIT can estimate the final performance of a network without retraining. 654, TITLE: Avoiding spurious correlations via logit correction https://openreview.net/forum?id=5BaqCFVh5qL AUTHORS: Sheng Liu, Xu Zhang, Nitesh Sekhar, Yue Wu, Prateek Singhal, Carlos Fernandez-Granda HIGHLIGHT: In this work, we explicitly consider the presence of the potential spurious correlations exist in the majority of training data. 655, TITLE: Outcome-directed Reinforcement Learning by Uncertainty \& Temporal Distance-Aware Curriculum Goal Generation https://openreview.net/forum?id=v69itrHLEu AUTHORS: Daesol Cho, Seungjae Lee, H. Jin Kim HIGHLIGHT: To alleviate it, we propose an uncertainty \& temporal distance-aware curriculum goal generation method for

the outcome-directed RL via solving a bipartite matching problem.

656, TITLE: Packed Ensembles for efficient uncertainty estimation https://openreview.net/forum?id=XXTyv1zD9zD

AUTHORS: Olivier Laurent, Adrien Lafage, Enzo Tartaglione, Geoffrey Daniel, Jean-marc Martinez, Andrei Bursuc, Gianni Franchi HIGHLIGHT: We introduce Packed-Ensembles (PE), a strategy to design and train lightweight structured ensembles by carefully modulating the dimension of their encoding space. 657, TITLE: Calibrating the Rigged Lottery: Making All Tickets Reliable https://openreview.net/forum?id=KdwnGErdT6 AUTHORS: Bowen Lei, Ruqi Zhang, Dongkuan Xu, Bani Mallick HIGHLIGHT: In this paper, we propose a new sparse training method to produce sparse models with improved confidence calibration. 658, TITLE: Dynamic Update-to-Data Ratio: Minimizing World Model Overfitting https://openreview.net/forum?id=ZIkHSXzd9O7 AUTHORS: Nicolai Dorka, Tim Welschehold, Wolfram Burgard HIGHLIGHT: As a solution, we propose a new general method that dynamically adjusts the update to data (UTD) ratio during training based on under- and overfitting detection on a small subset of the continuously collected experience not used for training. 659, TITLE: Selection-Inference: Exploiting Large Language Models for Interpretable Logical Reasoning https://openreview.net/forum?id=3Pf3Wg6o-A4 AUTHORS: Antonia Creswell, Murray Shanahan, Irina Higgins HIGHLIGHT: We show that language models tend to perform fairly well at single step inference or entailment tasks, but struggle to chain together multiple reasoning steps to solve more complex problems. In light of this, we propose a Selection-Inference (SI) framework that exploits pre-trained LLMs as general processing modules, and alternates between selection and inference to generate a series of interpretable, casual reasoning steps leading to the final answer. 660, TITLE: Iterative Patch Selection for High-Resolution Image Recognition https://openreview.net/forum?id=QCrw0u9LQ7 AUTHORS: Benjamin Bergner, Christoph Lippert, Aravindh Mahendran HIGHLIGHT We propose a simple method, Iterative Patch Selection (IPS), which decouples the memory usage from the input size and thus enables the processing of arbitrarily large images under tight hardware constraints. 661, TITLE: Neural Networks Efficiently Learn Low-Dimensional Representations with SGD https://openreview.net/forum?id=6taykzqcPD AUTHORS: Alireza Mousavi-Hosseini, Sejun Park, Manuela Girotti, Ioannis Mitliagkas, Murat A Erdogdu HIGHLIGHT: We study the problem of training a two-layer neural network (NN) of arbitrary width using stochastic gradient descent (SGD) where the input $\boldsymbol{x} = \frac{x^2}{x^2} + \frac{x^2}{x^2}$ $index model, i.e., \$y=g(\langle\boldsymbol \{u_1\}\boldsymbol \{x\}\rangle,...,\langle\boldsymbol \{u_k\}\boldsymbol \{x\}\rangle)\$ with a noisy link function \$g\$. 662, TITLE: Contextual bandits with concave rewards, and an application to fair ranking https://openreview.net/forum?id=UT- SVOyD1H AUTHORS: Virginie Do, Elvis Dohmatob, Matteo Pirotta, Alessandro Lazaric, Nicolas Usunier HIGHLIGHT: We present the first algorithm with provably vanishing regret for CBCR without restrictions on the policy space, whereas prior works were restricted to finite policy spaces or tabular representations. 663, TITLE: MA-BERT: Towards Matrix Arithmetic-only BERT Inference by Eliminating Complex Non-linear Functions https://openreview.net/forum?id=HtAfbHa7LAL AUTHORS: Neo Wei Ming, Zhehui Wang, Cheng Liu, Rick Siow Mong Goh, Tao Luo HIGHLIGHT: Specifically, we propose four correlated techniques that include approximating softmax with a two-layer neural network, replacing GELU with ReLU, fusing normalization layers with adjacent linear layers, and leveraging knowledge transfer from baseline models. 664, TITLE: Stochastic Multi-Person 3D Motion Forecasting https://openreview.net/forum?id= s1N-DnxdyT AUTHORS: Sirui Xu, Yu-Xiong Wang, Liangyan Gui We propose a dual-level generative modeling framework that separately models independent individual HIGHLIGHT: movements at the local level and social interactions at the global level. Maximizing Spatio-Temporal Entropy of Deep 3D CNNs for Efficient Video Recognition 665, TITLE: https://openreview.net/forum?id=lj1Eb1OPeNw AUTHORS: Junyan Wang, Zhenhong Sun, Yichen Qian, Dong Gong, Xiuyu Sun, Ming Lin, Maurice Pagnucco, Yang Song HIGHLIGHT: In this work, we propose to automatically design efficient 3D CNN architectures via a novel training-free neural architecture search approach tailored for 3D CNNs considering the model complexity. 666, TITLE: Training language models for deeper understanding improves brain alignment https://openreview.net/forum?id=KzkLAE49H9b AUTHORS: Khai Loong Aw, Mariya Toneva HIGHLIGHT:

HIGHLIGHT: However, it is still an open question whether these models are learning a deeper understanding of the text, or if the models are simply learning a heuristic to complete the task. This work investigates this further by turning to the one language processing system that truly understands complex language: the human brain.

667. TITLE: Draft, Sketch, and Prove: Guiding Formal Theorem Provers with Informal Proofs https://openreview.net/forum?id=SMa9EAovKMC AUTHORS: Albert Qiaochu Jiang, Sean Welleck, Jin Peng Zhou, Timothee Lacroix, Jiacheng Liu, Wenda Li, Mateja Jamnik, Guillaume Lample, Yuhuai Wu HIGHLIGHT: While previous studies to automate formalization focused on powerful search algorithms, no attempts were made to take advantage of available informal proofs. In this work, we introduce Draft, Sketch, and Prove (DSP), a method that maps informal proofs to formal proof sketches, and uses the sketches to guide an automated prover by directing its search to easier subproblems. 668, TITLE: Mini-batch \$k\$-means terminates within \$O(d/epsilon)\$ iterations https://openreview.net/forum?id=jREF4bkfi S AUTHORS: Gregory Schwartzman We answer the question: "Does \emph{local} progress (on batches) imply \emph{global} progress (on the HIGHLIGHT: entire dataset) for mini-batch \$k\$-means?" 669, TITLE: Active Learning in Bayesian Neural Networks with Balanced Entropy Learning Principle https://openreview.net/forum?id=ZTMuZ68B1g AUTHORS: Jae Oh Woo In this paper, we design and propose a new uncertainty measure, Balanced Entropy Acquisition (BalEntAcq), HIGHLIGHT: which captures the information balance between the uncertainty of underlying softmax probability and the label variable. 670, TITLE: Explaining Temporal Graph Models through an Explorer-Navigator Framework https://openreview.net/forum?id=BR ZhvcYbGJ AUTHORS: Wenwen Xia, Mincai Lai, Caihua Shan, Yao Zhang, Xinnan Dai, Xiang Li, Dongsheng Li HIGHLIGHT: To bridge the gap, in this paper, we propose T-GNNExplainer for temporal graph model explanation. 671, TITLE: FedExP: Speeding up Federated Averaging via Extrapolation https://openreview.net/forum?id=IPrzNbddXV AUTHORS: Divyansh Jhunjhunwala, Shiqiang Wang, Gauri Joshi HIGHLIGHT: In this work, we present FedExP, a method to adaptively determine the server step size in FL based on dynamically varying pseudo-gradients throughout the FL process. Planning with Large Language Models for Code Generation 672, TITLE: https://openreview.net/forum?id=Lr8cOOtYbfL AUTHORS: Shun Zhang, Zhenfang Chen, Yikang Shen, Mingyu Ding, Joshua B. Tenenbaum, Chuang Gan HIGHLIGHT: In this work, we propose a novel Transformer decoding algorithm, Planning-Guided Transformer Decoding (PG-TD), that uses a planning algorithm to do lookahead search and guide the Transformer to generate better programs. Hyper-Decision Transformer for Efficient Online Policy Adaptation 673, TITLE: https://openreview.net/forum?id=AatUEvC-Wjv AUTHORS: Mengdi Xu, Yuchen Lu, Yikang Shen, Shun Zhang, Ding Zhao, Chuang Gan Decision Transformers (DT) have demonstrated strong performances in offline reinforcement learning settings, HIGHLIGHT: but quickly adapting to unseen novel tasks remains challenging. To address this challenge, we propose a new framework, called Hyper-Decision Transformer (HDT), that can generalize to novel tasks from a handful of demonstrations in a data- and parameterefficient manner. 674, TITLE: Multimodal Analogical Reasoning over Knowledge Graphs https://openreview.net/forum?id=NRHajbzg8y0P AUTHORS: Ningyu Zhang, Lei Li, Xiang Chen, Xiaozhuan Liang, Shumin Deng, Huajun Chen HIGHLIGHT: To this end, we introduce the new task of multimodal analogical reasoning over a knowledge graph, which requires multimodal reasoning ability with the help of background knowledge. Specifically, we construct a Multimodal Analogical Reasoning dataSet (MARS) and a multimodal knowledge graph MarKG. 675, TITLE: Provable Memorization Capacity of Transformers https://openreview.net/forum?id=8JCg5xJCTPR AUTHORS: Junghwan Kim, Michelle Kim, Barzan Mozafari In this work, we present the first study of the memorization capacity of the Transformer architecture. HIGHLIGHT: 676, TITLE: Countinuous pseudo-labeling from the start https://openreview.net/forum?id=m3twGT2bAug AUTHORS: Dan Berrebbi, Ronan Collobert, Samy Bengio, Navdeep Jaitly, Tatiana Likhomanenko We believe this has the potential for over-fitting to the labeled dataset in low resource settings and that ST from HIGHLIGHT: the start of training should reduce over-fitting. In this paper we show how we can do this by dynamically controlling the evolution of PLs during the training process in ASR. 677, TITLE: Recursive Time Series Data Augmentation https://openreview.net/forum?id=51gD4vU-124s AUTHORS: Amine Mohamed Aboussalah, Minjae Kwon, Raj G Patel, Cheng Chi, Chi-Guhn Lee

HIGHLIGHT: Training on available realizations, where data is limited, often induces severe over-fitting thereby preventing generalization. To address this issue, we introduce a general recursive framework for time series augmentation, which we call the Recursive Interpolation Method (RIM).

678, TITLE: InCoder: A Generative Model for Code Infilling and Synthesis https://openreview.net/forum?id=hQwb-lbM6EL AUTHORS: Daniel Fried, Armen Aghajanyan, Jessy Lin, Sida Wang, Eric Wallace, Freda Shi, Ruiqi Zhong, Scott Yih, Luke Zettlemoyer, Mike Lewis HIGHLIGHT: We introduce InCoder, a unified generative model that can perform program synthesis (via left-to-right generation) as well as editing (via masking and infilling). Learning where and when to reason in neuro-symbolic inference 679. TITLE: https://openreview.net/forum?id=en9V5F8PR-AUTHORS: Cristina Cornelio, Jan Stuehmer, Shell Xu Hu, Timothy Hospedales HIGHLIGHT: However, all the existing methods usually either impose the constraints in a "weak" form at training time, with no guarantees at inference, or fail to provide a general framework that supports different tasks and constraint types. We tackle this open problem from a neuro-symbolic perspective. 680, TITLE: A Time Series is Worth 64 Words: Long-term Forecasting with Transformers https://openreview.net/forum?id=Jbdc0vTOcol AUTHORS: Yuqi Nie, Nam H Nguyen, Phanwadee Sinthong, Jayant Kalagnanam HIGHLIGHT: We propose an efficient design of Transformer-based models for multivariate time series forecasting and selfsupervised representation learning. 681, TITLE: Diversify and Disambiguate: Out-of-Distribution Robustness via Disagreement https://openreview.net/forum?id=RVTOp3MwT3n AUTHORS: Yoonho Lee, Huaxiu Yao, Chelsea Finn HIGHLIGHT: We propose DivDis, a simple two-stage framework for identifying and resolving ambiguity in data. 682, TITLE: Surgical Fine-Tuning Improves Adaptation to Distribution Shifts https://openreview.net/forum?id=APuPRxjHvZ AUTHORS: Yoonho Lee, Annie S Chen, Fahim Tajwar, Ananya Kumar, Huaxiu Yao, Percy Liang, Chelsea Finn HIGHLIGHT: A common approach to transfer learning under distribution shift is to fine-tune the last few layers of a pretrained model, preserving learned features while also adapting to the new task. This paper shows that in such settings, selectively finetuning a subset of layers (which we term surgical fine-tuning) matches or outperforms commonly used fine-tuning approaches. 683, TITLE: Predictive Inference with Feature Conformal Prediction https://openreview.net/forum?id=0uRm1YmFTu AUTHORS: Jiaye Teng, Chuan Wen, Dinghuai Zhang, Yoshua Bengio, Yang Gao, Yang Yuan HIGHLIGHT: In this paper, we propose feature conformal prediction, which extends the scope of conformal prediction to semantic feature spaces by leveraging the inductive bias of deep representation learning. 684, TITLE: On the Edge of Benign Overfitting: Label Noise and Overparameterization Level https://openreview.net/forum?id=UrEwJebCxk AUTHORS: Kaiyue Wen, Jiaye Teng, Jingzhao Zhang HIGHLIGHT: In this work, we examine whether overfitting is truly benign in real-world classification tasks. 685, TITLE: Sampling with Mollified Interaction Energy Descent https://openreview.net/forum?id=zWy7dqOcel AUTHORS: Lingxiao Li, qiang liu, Anna Korba, Mikhail Yurochkin, Justin Solomon HIGHLIGHT: In this paper, we present a new optimization-based method for sampling called mollified interaction energy descent (MIED). 686, TITLE: A non-asymptotic analysis of oversmoothing in Graph Neural Networks https://openreview.net/forum?id=CJd-BtnwtXq Xinyi Wu, Zhengdao Chen, William Wei Wang, Ali Jadbabaie AUTHORS: HIGHLIGHT: Specifically, we distinguish between two different effects when applying graph convolutions—an undesirable mixing effect that homogenizes node representations in different classes, and a desirable denoising effect that homogenizes node representations in the same class. 687, TITLE: Systematic Rectification of Language Models via Dead-end Analysis https://openreview.net/forum?id=k8_yVW3Wqln AUTHORS: Meng Cao, Mehdi Fatemi, Jackie CK Cheung, Samira Shabanian HIGHLIGHT: To this end, we formally extend the dead-end theory from the recent reinforcement learning (RL) literature to also cover uncertain outcomes. 688, TITLE: CLARE: Conservative Model-Based Reward Learning for Offline Inverse Reinforcement Learning https://openreview.net/forum?id=5aT4ganOd98 AUTHORS: Sheng Yue, Guanbo Wang, Wei Shao, Zhaofeng Zhang, Sen Lin, Ju Ren, Junshan Zhang

HIGHLIGHT: This work aims to tackle a major challenge in offline Inverse Reinforcement Learning (IRL), namely the reward extrapolation error, where the learned reward function may fail to explain the task correctly and misguide the agent in unseen environments due to the intrinsic covariate shift.

689, TITLE: https://openreview.net	Agent-based Graph Neural Networks //forum?id=8WTAh0tj2jC
AUTHORS:	Karolis Martinkus, Pál András Papp, Benedikt Schesch, Roger Wattenhofer
HIGHLIGHT:	We present a novel graph neural network we call AgentNet, which is designed specifically for graph-level tasks.
690, TITLE: https://openreview.net	Compositional Task Generalization with Discovered Successor Feature Modules //forum?id=DrtSx1z40Ib
AUTHORS:	Wilka Torrico Carvalho, Angelos Filos, Richard Lewis, Honglak Lee, Satinder Singh
HIGHLIGHT	In this work, we present a novel neural network architecture, "Modular Successor Feature Approximators"
(MSFA), where modu	les both discover what is useful to predict, and learn their own predictive representations.
691, TITLE: https://openreview.net	Spike Calibration: Bridging the Gap between ANNs and SNNs in ANN-SNN Conversion /forum?id=PFbzoWZyZRX
AUTHORS:	Zecheng Hao, Jianhao Ding, Tong Bu, Tiejun Huang, Zhaofei Yu
HIGHLIGHT:	However, the performance degrades severely under low time-steps, which hampers the practical applications of
SNNs on neuromorph define offset spike to a	ic chips. In this paper, instead of evaluating different conversion errors and then eliminating these errors, we measure the deviation degree of actual and desired firing rates of SNNs.
692, TITLE: https://openreview.net	An Equal-Size Hard EM Algorithm for Diverse Dialogue Generation t/forum?id=k5PEHHY4spM
AUTHORS:	Yuqiao Wen, Yongchang Hao, Yanshuai Cao, Lili Mou
HIGHLIGHT:	In this work, we propose an Equal-size Hard ExpectationMaximization (EqHard-EM) algorithm to train a
multi-decoder model f	for diverse dialogue generation.
693, TITLE:	DepthFL : Depthwise Federated Learning for Heterogeneous Clients
https://openreview.net	t/forum?id=pf8RIZTMU58
AUTHORS:	Miniae Kim, Sangyoon Yu, Suhvun Kim, Soo-Mook Moon
HIGHLIGHT:	This paper proposes a new approach based on depth scaling called DepthFL.
694, TITLE:	Neural Agents Struggle to Take Turns in Bidirectional Emergent Communication
https://openreview.net	
AUTHORS:	Valentin Taillandier, Dieuwke Hupkes, Benoit Sagot, Emmanuel Dupoux, Paul Michel
HIGHLIGHT	In this paper, we investigate the conditions under which artificial agents may naturally develop turn-taking
conventions in a simp	le language game.
695, TITLE: https://openreview.net	MaskFusion: Feature Augmentation for Click-Through Rate Prediction via Input-adaptive Mask Fusion
	Chao Lino Lino Linuan Linuan Lin Vi Cuo Chonger Song
HIGHI IGHT.	Chao Liao, Jianchao Tan, Jiyuan Jia, 11 Guo, Chenglu Song
halanaa hatuyaan aana	However, these works entire some non-more many other near the metaction modeling using Division ignore the
faither faither from the	ranzation and memorization during the recommendation. To murgate these problems, we propose an adaptive
existing deep part stru	cture of deep CTR models dynamically, besides the common feature interactions between the input feature and the
696, TITLE: Separation	Correlative Information Maximization Based Biologically Plausible Neural Networks for Correlated Source
https://openreview.net	t/forum?id=8JsaP7j1cL0
AUTHORS:	Bariscan Bozkurt, Ate? ?sfendiyaro?lu, Cengiz Pehleyan, Alper Tunga Erdogan
HIGHLIGHT:	Most prior attempts at this problem proposed neural networks that implement independent component analysis.
which works under the	e limitation that latent elements are mutually independent. Here, we relax this limitation and propose a
biologically plausible	neural network that extracts correlated latent sources by exploiting information about their domains.
697, TITLE:	Multifactor Sequential Disentanglement via Structured Koopman Autoencoders
https://openreview.net	Viorum/id=6ruPie9rbnC
AUTHORS:	Nimrod Berman, Ilan Naiman, Omri Azencol
multifactor disentangl	overall, we propose a simple and easy to code new deep model that is fully unsupervised and it supports ement.
698, TITLE:	Learnable Behavior Control: Breaking Atari Human World Records via Sample-Efficient Behavior Selection
https://openreview.net	t/forum?id=FeWvD0L_a4
AUTHORS: Xia	Jiajun Fan, Yuzheng Zhuang, Yuecheng Liu, Jianye HAO, Bin Wang, Jiangcheng Zhu, Hao Wang, Shu-Tao

HIGHLIGHT: In this paper, we propose a general framework called Learnable Behavioral Control (LBC) to address the limitation, which a) enables a significantly enlarged behavior selection space via formulating a hybrid behavior mapping from all policies; b) constructs a unified goal-directed learnable process for behavior selection.

699. TITLE: Variance Reduction is an Antidote to Byzantines: Better Rates, Weaker Assumptions and Communication Compression as a Cherry on the Top https://openreview.net/forum?id=pfuqQQCB34 AUTHORS: Eduard Gorbunov, Samuel Horváth, Peter Richtárik, Gauthier Gidel HIGHLIGHT: We derive theoretical convergence guarantees for Byz-VR-MARINA outperforming previous state-of-the-art for general non-convex and Polyak-Lojasiewicz loss functions. 700. TITLE: Weakly Supervised Explainable Phrasal Reasoning with Neural Fuzzy Logic https://openreview.net/forum?id=Hu4r-dedqR0 AUTHORS: Zijun Wu, Zi Xuan Zhang, Atharva Naik, Zhijian Mei, Mauajama Firdaus, Lili Mou HIGHLIGHT: In this work, we address the explainability for NLI by weakly supervised logical reasoning, and propose an Explainable Phrasal Reasoning (EPR) approach. 701, TITLE: Certified Defences Against Adversarial Patch Attacks on Semantic Segmentation https://openreview.net/forum?id=b0JxQC7JLWh AUTHORS: Maksym Yatsura, Kaspar Sakmann, N. Grace Hua, Matthias Hein, Jan Hendrik Metzen HIGHLIGHT: Adversarial patch attacks are an emerging security threat for real world deep learning applications. We present Demasked Smoothing, the first approach (up to our knowledge) to certify the robustness of semantic segmentation models against this threat model. 702, TITLE: Function-space regularized Rényi divergences https://openreview.net/forum?id=89GT-S49mGd AUTHORS: Jeremiah Birrell, Yannis Pantazis, Paul Dupuis, Luc Rey-Bellet, Markos Katsoulakis HIGHLIGHT: We propose a new family of regularized Renyi divergences parametrized not only by the order \$\alpha\$ but also by a variational function space. Novel View Synthesis with Diffusion Models 703. TITLE: https://openreview.net/forum?id=HtoA0oT30jC AUTHORS: Daniel Watson, William Chan, Ricardo Martin Brualla, Jonathan Ho, Andrea Tagliasacchi, Mohammad Norouzi HIGHLIGHT: We present 3DiM (pronounced "three-dim"), a diffusion model for 3D novel view synthesis from as few as a single image. Guarded Policy Optimization with Imperfect Online Demonstrations 704, TITLE: https://openreview.net/forum?id=O5rKg7IRQIO AUTHORS: Zhenghai Xue, Zhenghao Peng, Quanyi Li, Zhihan Liu, Bolei Zhou HIGHLIGHT: In this work we relax the assumption of a well-performing teacher and develop a new method that can incorporate arbitrary teacher policies with modest or inferior performance. LilNetX: Lightweight Networks with EXtreme Model Compression and Structured Sparsification 705, TITLE: https://openreview.net/forum?id=NVZvalzCLg AUTHORS: Sharath Girish, Kamal Gupta, Saurabh Singh, Abhinav Shrivastava We introduce LilNetX, an end-to-end trainable technique for neural networks that enables learning models with HIGHLIGHT: specified accuracy-rate-computation trade-off. 706, TITLE: Clean-image Backdoor: Attacking Multi-label Models with Poisoned Labels Only https://openreview.net/forum?id=rFQfjDC9Mt AUTHORS: Kangjie Chen, Xiaoxuan Lou, Guowen Xu, Jiwei Li, Tianwei Zhang HIGHLIGHT: However, all existing backdoor attacks exclusively require to modify training inputs (e.g., images), which may be impractical in real-world applications. In this paper, we aim to break this wall and propose the first clean-image backdoor attack, which only poisons the training labels without touching the training samples. 707, TITLE: Markup-to-Image Diffusion Models with Scheduled Sampling https://openreview.net/forum?id=81VJDmOE2o1 AUTHORS: Yuntian Deng, Noriyuki Kojima, Alexander M Rush HIGHLIGHT: Building on recent advances in image generation, we present a fully data-driven approach to rendering markup into images. 708, TITLE: How Much Space Has Been Explored? Measuring the Chemical Space Covered by Databases and Machine-Generated Molecules https://openreview.net/forum?id=Yo06F8kfMa1 Yutong Xie, Ziqiao Xu, Jiaqi Ma, Qiaozhu Mei AUTHORS: HIGHLIGHT: In this paper, we propose a novel evaluation framework for measures of the chemical space based on two analyses: an axiomatic analysis with two intuitive axioms that a good measure should obey, and an empirical analysis on the correlation between a measure and a proxy gold standard. Understanding DDPM Latent Codes Through Optimal Transport 709, TITLE: https://openreview.net/forum?id=6PIrhAx1j4i

AUTHORS: Valentin Khrulkov, Gleb Ryzhakov, Andrei Chertkov, Ivan Oseledets

HIGHLIGHT: While having important practical applications, such as the estimation of the likelihood, the theoretical properties of this map are not yet fully understood. In the present work, we partially address this question for the popular case of the VP-SDE (DDPM) approach.

710, TITLE:	Achieve the Minimum Width of Neural Networks for Universal Approximation	
https://openreview.ne	t/forum?id=hfUJ4ShyDEU	
AUTHORS:	Yongqiang Cai	
HIGHLIGHT:	In this paper, we consider neural networks with an arbitrary set of activation functions.	
711, TITLE: https://openreview.ne	UNIFIED-IO: A Unified Model for Vision, Language, and Multi-modal Tasks	
AUTHORS:Jiasen Lu, Christopher Clark, Rowan Zellers, Roozbeh Mottaghi, Aniruddha KembhaviHIGHLIGHT:We propose Unified-IO, a model that performs a large variety of AI tasks spanning classical computer visiontasks, including pose estimation, object detection, depth estimation and image generation, vision-and-language tasks such as regioncaptioning and referring expression, to natural language processing tasks such as question answering and paraphrasing.		
712, TITLE:	Distributed Differential Privacy in Multi-Armed Bandits	
https://openreview.ne	t/forum?id=cw8FeirkIfU	
AUTHORS:	Sayak Ray Chowdhury, Xingyu Zhou	
HIGHLIGHT:	In this work, we aim to obtain a pure-DP guarantee under distributed trust model while sacrificing no more	
regret than that under	central trust model.	
713, TITLE:	Learning Low Dimensional State Spaces with Overparameterized Recurrent Neural Networks	
https://openreview.ne	t/forum?id=k9CF4h3muD	
AUTHORS:	Edo Cohen-Karlik, Itamar Menuhin-Gruman, Nadav Cohen, Raja Giryes, Amir Globerson	
HIGHLIGHT:	In this paper, we analyze the extrapolation properties of GD when applied to overparameterized linear RNNs.	
714, TITLE: https://openreview.ne	Betty: An Automatic Differentiation Library for Multilevel Optimization	
AUTHORS:	Sang Keun Choe, Willie Neiswanger, Pengtao Xie, Eric Xing	
HIGHLIGHT:	However, gradients in MLO, which are obtained by composing best-response Jacobians via the chain rule, are	
notoriously difficult t	o implement and memory/compute intensive. We take an initial step towards closing this gap by introducing	
Betty, a software libra	ary for large-scale MLO.	
715, TITLE:	Near-Optimal Adversarial Reinforcement Learning with Switching Costs	
https://openreview.ne	tt/forum?id=i9ogGQHYbkY Ming Shi, Yingbin Liang, Ness Shroff	
HIGHLIGHT: the transition function function is unknown.	Moreover, we propose two novel switching-reduced algorithms with regrets that match our lower bound when n is known, and match our lower bound within a small factor of $\tilde{O} (H^{1/3})$ when the transition Our regret analysis demonstrates the near-optimal performance of them.	
716, TITLE:	NAGphormer: A Tokenized Graph Transformer for Node Classification in Large Graphs	
https://openreview.ne	t/forum?id=8KYeilT3Ow	
AUTHORS:	Jinsong Chen, Kaiyuan Gao, Gaichao Li, Kun He In this work, we observe that existing graph Transformers treat nodes as independent tokens and construct a	
single long sequence	composed of all node tokens so as to train the Transformer model, causing it hard to scale to large graphs due to	
the quadratical compl	lexity on the number of nodes for the self-attention computation.	
717, TITLE:	TiAda: A Time-scale Adaptive Algorithm For Nonconvex Minimax Optimization	
https://openreview.ne	tr/forum?id=zClyiZ5V6sL	
AUTHORS:	Xiang Li, Junchi YANG, Niao He	
HIGHLIGHT:	In this work, we propose a single-loop adaptive GDA algorithm called TiAda for nonconvex minimax	
optimization that auto	omatically adapts to the time-scale separation.	
718, TITLE:	Collaborative Pure Exploration in Kernel Bandit	
https://openreview.ne	t/forum?id=hLbeJ6jObDD	
AUTHORS:	Yihan Du, Wei Chen, Yuko Kuroki, Longbo Huang	
HIGHLIGHT:	In this paper, we propose a novel Collaborative Pure Exploration in Kernel Bandit model (CoPE-KB), where	
multiple agents collab	porate to complete different but related tasks with limited communication.	
719, TITLE:	Clifford Neural Layers for PDE Modeling	
https://openreview.ne	t/forum?id=okwxL c4x84	
AUTHORS:	Johannes Brandstetter, Rianne van den Berg, Max Welling, Jayesh K Gupta	
HIGHLIGHT:	Their algebraic properties, such as multiplication, addition and other arithmetic operations can be described by	
Clifford algebras. To	our knowledge, this paper presents the first usage of such multivector representations together with Clifford	
convolutions and Clif	fford Fourier transforms in the context of deep learning.	
720, TITLE:	How Informative is the Approximation Error from Tensor Decomposition for Neural Network Compression?	
https://openreview.ne	t/forum?id=sKHqgFOaFXI	
AUTHORS:	Jetze Schuurmans, kim batselier, Julian Kooij	

HIGHLIGHT: We show empirically an approximation error resulting from compressing a network layer with tensor decomposition is correlated with the classification error, enabling the choice of layer, decomposition and rank to be based on the approximation error. 721, TITLE: CUDA: Curriculum of Data Augmentation for Long-tailed Recognition https://openreview.net/forum?id=RgUPdudkWlN AUTHORS: Sumyeong Ahn, Jongwoo Ko, Se-Young Yun HIGHLIGHT: In this study, we first investigate the correlation between the degree of augmentation and class-wise performance, and find that the proper degree of augmentation must be allocated for each class to mitigate class imbalance problems. Motivated by this finding, we propose a simple and efficient novel curriculum, which is designed to find the appropriate per-class strength of data augmentation, called CUDA: CUrriculum of Data Augmentation for long-tailed recognition. Revisiting Graph Adversarial Attack and Defense From a Data Distribution Perspective 722. TITLE: https://openreview.net/forum?id=dSYoPjM5J W AUTHORS: Kuan Li, Yang Liu, Xiang Ao, Qing He HIGHLIGHT: In this work, we discover an interesting phenomenon: the adversarial edges are not uniformly distributed on the graph. 723, TITLE: Understanding new tasks through the lens of training data via exponential tilting https://openreview.net/forum?id=DBMttEEoLbw AUTHORS: Subha Maity, Mikhail Yurochkin, Moulinath Banerjee, Yuekai Sun HIGHLIGHT: We consider the problem of reweighing the training samples to gain insights into the distribution of the target task. 724, TITLE: A VAE for Transformers with Nonparametric Variational Information Bottleneck https://openreview.net/forum?id=60kjC cs03X AUTHORS: James Henderson, Fabio James Fehr HIGHLIGHT: We propose a Variational AutoEncoder (VAE) for Transformers by developing a Variational Information Bottleneck (VIB) regulariser for Transformer embeddings. 725, TITLE: Behavior Prior Representation learning for Offline Reinforcement Learning https://openreview.net/forum?id=hQ4K9Bf4G2B AUTHORS: Hongyu Zang, Xin Li, Jie Yu, Chen Liu, Riashat Islam, Remi Tachet des Combes, Romain Laroche HIGHLIGHT: In this work, we introduce a simple, yet effective approach for learning state representations. 726, TITLE: LPT: Long-tailed Prompt Tuning for Image Classification https://openreview.net/forum?id=8pOVAeo8ie AUTHORS: Bowen Dong, Pan Zhou, Shuicheng Yan, Wangmeng Zuo Though promising, fine-tuning the whole pretrained model tends to suffer from high cost in computation and HIGHLIGHT: deployment of different models for different tasks, as well as weakened generalization capability for overfitting to certain features of long-tailed data. To alleviate these issues, we propose an effective Long-tailed Prompt Tuning (LPT) method for long-tailed classification tasks. Towards Real-Time Neural Image Compression With Mask Decay 727, TITLE: https://openreview.net/forum?id=XUxad2Gj40n AUTHORS: Wang Guo-Hua, Jiahao Li, Bin Li, Yan Lu HIGHLIGHT: In this paper, we propose an efficient single-model variable-bit-rate network, which is able to run at 30 FPS with 768x512 input images and still outperforms VVC for the RD performance. 728, TITLE: Policy Expansion for Bridging Offline-to-Online Reinforcement Learning https://openreview.net/forum?id=-Y34L45JR6z AUTHORS: Haichao Zhang, Wei Xu, Haonan Yu HIGHLIGHT: One natural approach is to initialize the policy for online learning with the one trained offline. In this work, we introduce a policy expansion scheme for this task. 729, TITLE: Learning differentiable solvers for systems with hard constraints https://openreview.net/forum?id=vdv6CmGksr0 AUTHORS: Geoffrey Négiar, Michael W. Mahoney, Aditi Krishnapriyan HIGHLIGHT: We introduce a practical method to enforce linear partial differential equation (PDE) constraints for functions defined by neural networks (NNs), up to a desired tolerance. 730, TITLE: Confidence-Based Feature Imputation for Graphs with Partially Known Features https://openreview.net/forum?id=YPKBIILy-Kt AUTHORS: Daeho Um, Jiwoong Park, Seulki Park, Jin young Choi HIGHLIGHT: However, in cases of high rates of missing features, they were unable to avoid significant performance degradation. To overcome this limitation, we introduce a novel concept of channel-wise confidence in a node feature, which is assigned to each imputed channel feature of a node for reflecting the certainty of the imputation. 731, TITLE: LiftedCL: Lifting Contrastive Learning for Human-Centric Perception https://openreview.net/forum?id=WHlt5tLz12T

AUTHORS:	Ziwei Chen, Qiang Li, Xiaofeng Wang, Wankou Yang
HIGHLIGHT:	In this paper, we propose the Lifting Contrastive Learning (LiftedCL) to obtain 3D-aware human-centric
representations which	a absorb 3D human structure information.
732, TITLE:	Learning Controllable Adaptive Simulation for Multi-scale Physics
https://openreview.ne	et/forum?id=PbfgkZ2HdbE
AUTHORS:	Tailin Wu, Takashi Maruyama, Qingqing Zhao, Gordon Wetzstein, Jure Leskovec
HIGHLIGHT:	In this work, we introduce Learning controllable Adaptive simulation for Multi-scale Physics (LAMP) as the
first full deep learning	g-based surrogate model that jointly learns the evolution model and optimizes appropriate spatial resolutions that
devote more compute	e to the highly dynamic regions.
733, TITLE:	FedDAR: Federated Domain-Aware Representation Learning
https://openreview.ne	et/forum?id=6P9Y25Pljl6
AUTHORS:	Aoxiao Zhong, Hao He, Zhaolin Ren, Na Li, Quanzheng Li
HIGHLIGHT:	In this work, we focus on a special non-iid FL problem, called Domain-mixed FL, where each client's data
distribution is assume	ed to be a mixture of several predefined domains.
734, TITLE:	Learning to Estimate Shapley Values with Vision Transformers
https://openreview.ne	t/forum?id=5ktFNz_pJLK
AUTHORS:	Ian Connick Covert, Chanwoo Kim, Su-In Lee
HIGHLIGHT:	In this work, we aim to make Shapley values practical for vision transformers (ViTs).
735, TITLE:	Fuzzy Alignments in Directed Acyclic Graph for Non-Autoregressive Machine Translation
https://openreview.ne	tt/forum?id=LSz-gQyd0zE
AUTHORS:	Zhengrui Ma, Chenze Shao, Shangtong Gui, Min Zhang, Yang Feng
HIGHLIGHT:	In this paper, we hold the view that all paths in the graph are fuzzily aligned with the reference sentence.
736, TITLE:	Prompt Learning with Optimal Transport for Vision-Language Models
https://openreview.ne	et/forum?id=zqwryBoXYnh
AUTHORS:	Guangyi Chen, Weiran Yao, Xiangchen Song, Xinyue Li, Yongming Rao, Kun Zhang
HIGHLIGHT:	Unlike conventional methods of only learning one single prompt, we propose to learn multiple comprehensive
prompts to describe d	liverse characteristics of categories such as intrinsic attributes or extrinsic contexts.
737, TITLE:	Masked Frequency Modeling for Self-Supervised Visual Pre-Training
https://openreview.ne	t/forum?id=9-umxtNPx5E
AUTHORS:	Jiahao Xie, Wei Li, Xiaohang Zhan, Ziwei Liu, Yew-Soon Ong, Chen Change Loy
HIGHLIGHT:	We present Masked Frequency Modeling (MFM), a unified frequency-domain-based approach for self-
supervised pre-trainir	ng of visual models.
738, TITLE:	A framework for benchmarking Class-out-of-distribution detection and its application to ImageNet
https://openreview.ne	et/forum?id=luubb9W6Jtk
AUTHORS:	Ido Galil, Mohammed Dabbah, Ran El-Yaniv
HIGHLIGHT:	In this paper we present a novel technique to benchmark image classifiers' ability to detect class-out-of-
distribution instances	i (i.e., instances whose true labels the model does not recognize) at various levels of detection difficulty.
739, TITLE: https://openreview.net AUTHORS: HIGHLIGHT: forgetting of learned a framework that sim alleviation (FA).	Twofer: Tackling Continual Domain Shift with Simultaneous Domain Generalization and Adaptation et/forum?id=L8iZdgeKmI6 Chenxi Liu, Lixu Wang, Lingjuan Lyu, Chen Sun, Xiao Wang, Qi Zhu However, existing DG works are ineffective for continually changing domains due to severe catastrophic knowledge. To overcome these limitations of DA or DG in tackling continual domain shifts, we propose Twofer, ultaneously achieves target domain generalization (TDG), target domain adaptation (TDA), and forgetting
740, TITLE:	VA-DepthNet: A Variational Approach to Single Image Depth Prediction
https://openreview.ne	et/forum?id=xjxUjHa_Wpa
AUTHORS:	Ce Liu, Suryansh Kumar, Shuhang Gu, Radu Timofte, Luc Van Gool
HIGHLIGHT:	The paper's main contribution is to reveal the benefit of classical and well-founded variational constraints in the
neural network design	n for the SIDP task.
741, TITLE:	Human Motion Diffusion Model
https://openreview.ne	t/forum?id=SJ1kSyO2jwu
AUTHORS:	Guy Tevet, Sigal Raab, Brian Gordon, Yoni Shafir, Amit Haim Bermano, Daniel Cohen-or
HIGHLIGHT:	In this paper, we introduce Motion Diffusion Model (MDM), a carefully adapted classifier-free diffusion-based
generative model for	human motion data.
742, TITLE:	Human MotionFormer: Transferring Human Motions with Vision Transformers
https://openreview.ne	et/forum?id=lQVpasnQS62

AUTHORS: Hongyu Liu, Xintong Han, Chenbin Jin, Lihui Qian, Huawei Wei, Zhe Lin, Faqiang Wang, Haoye Dong, Yibing Song, Jia Xu, Qifeng Chen

HIGHLIGHT: In this paper, we propose Human MotionFormer, a hierarchical ViT framework for motion transfer between two human subjects.

743, TITLE: On the Convergence of AdaGrad on \$\mathbd{R}^d\$: Beyond Convexity, Non-Asymptotic Rate and Acceleration

https://openreview.net/forum?id=ULnHxczCBaE

AUTHORS: Zijian Liu, Ta Duy Nguyen, Alina Ene, Huy Nguyen

HIGHLIGHT: Furthermore, in the stochastic setting, only a modified version of AdaGrad, different from the one commonly used in practice, in which the latest gradient is not used to update the stepsize, has been analyzed. Our paper aims at bridging these gaps and developing a deeper understanding of AdaGrad and its variants in the standard setting of smooth convex functions as well as the more general setting of quasar convex functions.

744, TITLE: Semi Parametric Inducing Point Networks

https://openreview.net/forum?id=FE99-fDrWd5

AUTHORS: Richa Rastogi, Yair Schiff, Alon Hacohen, Zhaozhi Li, Ian Lee, Yuntian Deng, Mert R. Sabuncu, Volodymyr Kuleshov

HIGHLIGHT: We introduce semi-parametric inducing point networks (SPIN), a general-purpose architecture that can query the training set at inference time in a compute-efficient manner.

745, TITLE: Breaking the Curse of Dimensionality in Multiagent State Space: A Unified Agent Permutation Framework https://openreview.net/forum?id=OxNQXyZK-K8

AUTHORS: Jianye HAO, Xiaotian Hao, Hangyu Mao, Weixun Wang, Yaodong Yang, Dong Li, YAN ZHENG, Zhen Wang HIGHLIGHT: Such a curse of dimensionality results in poor scalability and low sample efficiency, inhibiting MARL for decades. To break this curse, we propose a unified agent permutation framework that exploits the permutation invariance (PI) and permutation equivariance (PE) inductive biases to reduce the multiagent state space.

746, TITLE: Using Language to Extend to Unseen Domains

https://openreview.net/forum?id=eR2dG8yjnQ

AUTHORS: Lisa Dunlap, Clara Mohri, Devin Guillory, Han Zhang, Trevor Darrell, Joseph E. Gonzalez, Aditi Raghunathan, Anna Rohrbach

HIGHLIGHT: We instead consider how simply verbalizing the training domain (e.g. ``photos of birds") as well as domains we want to extend to but do not have data for (e.g. ``paintings of birds") can improve robustness.

 747, TITLE:
 SIMPLE: Specialized Model-Sample Matching for Domain Generalization

 https://openreview.net/forum?id=BqrPeZ_e5P

 AUTHORS:
 Ziyue Li, Kan Ren, XINYANG JIANG, Yifei Shen, Haipeng Zhang, Dongsheng Li

 HIGHLIGHT:
 In this paper, we propose an alternative direction, i.e., to efficiently leverage a pool of pretrained models

 without fine-tuning.
 Interval

748, TITLE: New Insights for the Stability-Plasticity Dilemma in Online Continual Learning https://openreview.net/forum?id=fxC7kJYwA_a

AUTHORS: Dahuin Jung, Dongjin Lee, Sunwon Hong, Hyemi Jang, Ho Bae, Sungroh Yoon

HIGHLIGHT: To overcome the stability-plasticity dilemma in online continual learning, we propose an online continual learning framework named multi-scale feature adaptation network (MuFAN) that utilizes a richer context encoding extracted from different levels of a pre-trained network.

749, TITLE: DFlow: Learning to Synthesize Better Optical Flow Datasets via a Differentiable Pipeline https://openreview.net/forum?id=5O2uzDusEN5

AUTHORS: Kwon Byung-Ki, Nam Hyeon-Woo, Ji-Yun Kim, Tae-Hyun Oh

HIGHLIGHT: However, manually identifying and verifying all such necessary properties are intractable mainly due to the requirement of large-scale trial-and-error experiments with iteratively generating whole synthetic datasets. To tackle this challenge, we propose a differentiable optical flow data generation pipeline and a loss function to drive the pipeline, called DFlow.

 750, TITLE:
 Ensuring DNN Solution Feasibility for Optimization Problems with Linear Constraints

 https://openreview.net/forum?id=QVcDQJdFTG

 AUTHORS:
 Tianyu Zhao, Xiang Pan, Minghua Chen, Steven Low

 HIGHLIGHT:
 We propose preventive learning as the first framework to guarantee Deep Neural Network (DNN) solution

feasibility for optimization problems with linear constraints without post-processing.

751, TITLE: Toeplitz Neural Network for Sequence Modeling

https://openreview.net/forum?id=IxmWsm4xrua

AUTHORS: Zhen Qin, Xiaodong Han, Weixuan Sun, Bowen He, Dong Li, Dongxu Li, Yuchao Dai, Lingpeng Kong, Yiran Zhong

HIGHLIGHT: While showing good performance, the transformer models are inefficient to scale to long input sequences, mainly due to the quadratic space-time complexity of attention. To overcome this inefficiency, we propose to model sequences with a relative position encoded Toeplitz matrix and use a Toeplitz matrix-vector production trick to reduce the space-time complexity of the sequence modeling to log linear.

752, TITLE: Model ensemble instead of prompt fusion: a sample-specific knowledge transfer method for few-shot prompt tuning

https://openreview.net/forum?id=p0yrSRbN5Bu AUTHORS: XIANGYU PENG, Chen Xing, Prafulla Kumar Choubey, Chien-Sheng Wu, Caiming Xiong HIGHLIGHT: In this work, we focus on improving the few-shot performance of prompt tuning by transferring knowledge from soft prompts of source tasks with abundant training samples. Neural Image-based Avatars: Generalizable Radiance Fields for Human Avatar Modeling 753, TITLE: https://openreview.net/forum?id=-ng-FXFlzgK AUTHORS: YoungJoong Kwon, Dahun Kim, Duygu Ceylan, Henry Fuchs HIGHLIGHT: We present a method that enables synthesizing novel views and novel poses of arbitrary human performers from sparse multi-view images. Retrieval-based Controllable Molecule Generation 754. TITLE: https://openreview.net/forum?id=vDFA1tpuLvk AUTHORS: Zichao Wang, Weili Nie, Zhuoran Qiao, Chaowei Xiao, Richard Baraniuk, Anima Anandkumar HIGHLIGHT: In this work, we propose a new retrieval-based framework for controllable molecule generation. 755, TITLE: Deep Generative Symbolic Regression https://openreview.net/forum?id=o7koEEMA1bR AUTHORS: Samuel Holt, Zhaozhi Qian, Mihaela van der Schaar HIGHLIGHT: We make the observation that closed-form equations often have structural characteristics and invariances (e.g. the commutative law) that could be further exploited to build more effective symbolic regression solutions. Motivated by this observation, our key contribution is to leverage pre-trained deep generative models to capture the intrinsic regularities of equations, thereby providing a solid foundation for subsequent optimization steps. 756, TITLE: Order Matters: Agent-by-agent Policy Optimization https://openreview.net/forum?id=Q-neeWNVv1 Xihuai Wang, Zheng Tian, Ziyu Wan, Ying Wen, Jun Wang, Weinan Zhang AUTHORS: In this paper, we propose the \textbf{A}gent-by-\textbf{a}gent \textbf{P}olicy \textbf{O}ptimization (A2PO) HIGHLIGHT: algorithm to improve the sample efficiency and retain the guarantees of monotonic improvement for each agent during training. 757, TITLE: A GNN-Guided Predict-and-Search Framework for Mixed-Integer Linear Programming https://openreview.net/forum?id=pHMpgT5xWaE AUTHORS: Qingyu Han, Linxin Yang, Qian Chen, Xiang Zhou, Dong Zhang, Akang Wang, Ruoyu Sun, Xiaodong Luo HIGHLIGHT: In this work, we combine ML with optimization and propose a novel predict-and-search framework for efficiently identifying high-quality feasible solutions. Improved Convergence of Differential Private SGD with Gradient Clipping 758, TITLE: https://openreview.net/forum?id=FRLswckPXQ5 Huang Fang, Xiaoyun Li, Chenglin Fan, Ping Li AUTHORS: HIGHLIGHT: As a by-product, we propose a new clipping technique, called value clipping, to mitigate the computational overhead caused by the classic gradient clipping. Solving Constrained Variational Inequalities via a First-order Interior Point-based Method 759, TITLE: https://openreview.net/forum?id=RQY2AXFMRiu AUTHORS: Tong Yang, Michael Jordan, Tatjana Chavdarova HIGHLIGHT: We develop an interior-point approach to solve constrained variational inequality (cVI) problems. 760, TITLE: Generating Diverse Cooperative Agents by Learning Incompatible Policies https://openreview.net/forum?id=UkU05GOH7 6 AUTHORS: Rujikorn Charakorn, Poramate Manoonpong, Nat Dilokthanakul HIGHLIGHT: In this work, we propose to learn diverse behaviors via policy compatibility. Logical Message Passing Networks with One-hop Inference on Atomic Formulas 761, TITLE: https://openreview.net/forum?id=SoyOsp7i 1 AUTHORS: Zihao Wang, Yangqiu Song, Ginny Wong, Simon See HIGHLIGHT: In this paper, we propose a simple framework for complex query answering that decomposes the KG embeddings from neural set operators. 762, TITLE: Rethinking the Effect of Data Augmentation in Adversarial Contrastive Learning https://openreview.net/forum?id=0qmwFNJyxCL Rundong Luo, Yifei Wang, Yisen Wang AUTHORS: HIGHLIGHT: Motivated by this observation, we revisit existing self-AT and discover an inherent dilemma that affects self-AT robustness: either strong or weak data augmentations are harmful to self-AT, and a medium strength is insufficient to bridge the gap. To resolve this dilemma, we propose a simple remedy named DynACL (Dynamic Adversarial Contrastive Learning). Data-Free One-Shot Federated Learning Under Very High Statistical Heterogeneity 763. TITLE: https://openreview.net/forum?id= hb4vM3jspB

AUTHORS: Emilio Luz-Ricca, Clare Elizabeth Heinbaugh, Huajie Shao
HIGHLIGHT: However, one-shot FL methods often degrade under high statistical heterogeneity, fail to promote pipeline security, or require an auxiliary public dataset. To address these limitations, we propose two novel data-free one-shot FL methods: FedCVAE-Ens and its extension FedCVAE-KD.

764, TITLE:	Code Translation with Compiler Representations
AUTHORS:	Marc Szafraniec, Baptiste Roziere, Hugh James Leather, Patrick Labatut, Francois Charton, Gabriel Synnaeve
HIGHLIGHT:	In this paper, we leverage low-level compiler intermediate representations (IR) code translation.
765, TITLE: Classifiers?	What Can we Learn From The Selective Prediction And Uncertainty Estimation Performance Of 523 Imagenet
https://openreview.ne	t/forum?id=p66AzKi6Xim
AUTHORS:	Ido Galil, Mohammed Dabbah, Ran El-Yaniv
of 523 existing pretra	ined deep ImageNet classifiers that are available in popular repositories.
766, TITLE: https://openreview.ne	Predictor-corrector algorithms for stochastic optimization under gradual distribution shift t/forum?id=2SV2dlfBuE3
AUTHORS:	Subha Maity, Debarghya Mukherjee, Moulinath Banerjee, Yuekai Sun
HIGHLIGHI:	Often, the underlying process that drives the distribution shift is continuous in nature. We exploit this
in the underlying data	generating process through a predictor-corrector term in the update rule.
767, TITLE:	The In-Sample Softmax for Offline Reinforcement Learning
https://openreview.ne	t/forum?id=u-RuvyDYqCM Cheniun Xiao, Han Wang, Vangchen Pan, Adam White, Martha White
HIGHLIGHT:	We highlight a simple fact: it is more straightforward to approximate an in-sample softmax using only actions
in the dataset.	
768, TITLE: https://openreview.ne	Replay Memory as An Empirical MDP: Combining Conservative Estimation with Experience Replay
AUTHORS:	Hongming Zhang, Chenjun Xiao, Han Wang, Jun Jin, bo xu, Martin Müller
HIGHLIGHT: Replay Memo	In this work, we further exploit the information in the replay memory by treating it as an empirical ry MDP (RM-MDP)}.
760 TITLE.	A Unified Aleshnois Danamentiya an Lingshitz Neural Networks
https://openreview.ne	t/forum?id=k71IGLC8cfc
AUTHORS:	Alexandre Araujo, Aaron J Havens, Blaise Delattre, Alexandre Allauzen, Bin Hu
AOL and CPL can be	We present a novel algebraic perspective unifying various types of 1-Lipschitz neural networks, and show that re-derived and generalized using exactly the same semidefinite programming (SDP) condition.
770, TITLE:	Curriculum-based Co-design of Morphology and Control of Voxel-based Soft Robots
https://openreview.ne	t/forum?id=r9fX833CsuN Vuving Wang, Shuang Wu, Haobo Eu, OIANG EU, Tiantian Zhang, Vongzhe Chang, Yuegian Wang,
HIGHLIGHT:	In this paper, we present a Curriculum-based Co-design (CuCo) method for learning to design and control
VSRs through an easy	y-to-difficult process.
771, TITLE:	When Data Geometry Meets Deep Function: Generalizing Offline Reinforcement Learning
https://openreview.ne	t/forum?id=IMO7TC7cuuh Jianxiong Li, Xianyuan Zhan, Haoran Xu, Xianggu Zhu, Jingjing Liu, Va Oin Zhang
HIGHLIGHT:	In our study, one interesting observation is that deep \textit {Q} functions approximate well inside the convex
hull of training data. I	Inspired by this, we propose a new method, \textit {DOGE (Distance-sensitive Offline RL with better
Generalization)}.	
772, TITLE: https://openreview.ne	Impossibly Good Experts and How to Follow Them
AUTHORS:	Aaron Walsman, Muru Zhang, Sanjiban Choudhury, Dieter Fox, Ali Farhadi
HIGHLIGHT:	We consider the sequential decision making problem of learning from an expert that has access to more
in the reduced inform	earner. We provide a set of necessary criteria on the expert that will allow a learner to recover the optimal policy ation space from the expert's advice alone.
773, TITLE:	Memorization-Dilation: Modeling Neural Collapse Under Noise
https://openreview.ne	t/forum?id=cJWxqmmDL2b
AUTHORS: HIGHLIGHT	Duc Anh Nguyen, Ron Levie, Julian Lienen, Eyke Hüllermeier, Gitta Kutyniok In this work we study a more realistic variant of the layer peaked model, which takes the positivity of the
features into account.	In this work we study a more realistic variant of the layer-pected model, which takes the positivity of the
774, TITLE:	SLTUNET: A Simple Unified Model for Sign Language Translation
https://openreview.ne	t/forum?id=EBS4C77p_5S
AUTHORS:	Biao Zhang, Mathias Müller, Rico Sennrich

We propose SLTUNET, a simple unified neural model designed to support multiple SLTrelated tasks jointly, HIGHLIGHT: such as sign-to-gloss, gloss-to-text and sign-to-text translation. 775, TITLE: Mind the Gap: Offline Policy Optimizaiton for Imperfect Rewards https://openreview.net/forum?id=WumysvcMvV6 AUTHORS: Jianxiong Li, Xiao Hu, Haoran Xu, Jingjing Liu, Xianyuan Zhan, Qing-Shan Jia, Ya-Qin Zhang HIGHLIGHT: In this study, we propose a unified offline policy optimization approach, \textit {RGM} (Reward Gap Minimization), which can smartly handle diverse types of imperfect rewards. Meta Learning to Bridge Vision and Language Models for Multimodal Few-Shot Learning 776, TITLE: https://openreview.net/forum?id=3oWo92cQyxL AUTHORS: Ivona Najdenkoska, Xiantong Zhen, Marcel Worring HIGHLIGHT: To make the whole process learnable, we introduce a multimodal meta-learning approach. 777, TITLE: Masked Image Modeling with Denoising Contrast https://openreview.net/forum?id=1fZd4owfJP6 AUTHORS: Kun Yi, Yixiao Ge, Xiaotong Li, Shusheng Yang, Dian Li, Jianping Wu, Ying Shan, Xiaohu Qie HIGHLIGHT: We first treat masked patch prediction as denoising contrastive learning in self-supervised image pre-training, achieving state-of-the-art results. 778, TITLE: Learning on Large-scale Text-attributed Graphs via Variational Inference https://openreview.net/forum?id=q0nmYciuuZN AUTHORS: Jianan Zhao, Meng Qu, Chaozhuo Li, Hao Yan, Qian Liu, Rui Li, Xing Xie, Jian Tang HIGHLIGHT: In this paper, we propose an efficient and effective solution to learning on large text-attributed graphs by fusing graph structure and language learning with a variational Expectation-Maximization (EM) framework, called GLEM. Distributionally Robust Post-hoc Classifiers under Prior Shifts 779, TITLE: https://openreview.net/forum?id=3KUfbI9 DQE AUTHORS: Jiaheng Wei, Harikrishna Narasimhan, Ehsan Amid, Wen-Sheng Chu, Yang Liu, Abhishek Kumar HIGHLIGHT: We present an extremely lightweight post-hoc approach that performs scaling adjustments to predictions from a pre-trained model, with the goal of minimizing a distributionally robust loss around a chosen target distribution. 780, TITLE: Learning Sparse and Low-Rank Priors for Image Recovery via Iterative Reweighted Least Squares Minimization https://openreview.net/forum?id=TXPN6MtdSE4 AUTHORS: Stamatios Lefkimmiatis, Iaroslav Sergeevich Koshelev HIGHLIGHT: In this work we introduce a novel optimization algorithm for image recovery under learned sparse and low-rank constraints, which are parameterized with weighted extensions of the \$\ell p^p\$-vector and \$\mathcal{S} p^p\$ Schatten-matrix quasi-norms for \$0\! Adversarial Training of Self-supervised Monocular Depth Estimation against Physical-World Attacks 781, TITLE: https://openreview.net/forum?id=LfdEuhjR5GV AUTHORS: Zhiyuan Cheng, James Chenhao Liang, Guanhong Tao, Dongfang Liu, Xiangyu Zhang HIGHLIGHT: In this work, we propose a novel adversarial training method for self-supervised MDE models based on view synthesis without using the depth ground truth. 782. TITLE: Sparse Upcycling: Training Mixture-of-Experts from Dense Checkpoints https://openreview.net/forum?id=T5nUQDrM4u AUTHORS: Aran Komatsuzaki, Joan Puigcerver, James Lee-Thorp, Carlos Riquelme Ruiz, Basil Mustafa, Joshua Ainslie, Yi Tay, Mostafa Dehghani, Neil Houlsby HIGHLIGHT: In this work, we propose sparse upcycling -- a simple way to reuse sunk training costs by initializing a sparsely activated Mixture-of-Experts model from a dense checkpoint. **Diffusion Probabilistic Fields** 783, TITLE: https://openreview.net/forum?id=ik91mY-2GN AUTHORS: Peiye Zhuang, Samira Abnar, Jiatao Gu, Alex Schwing, Joshua M. Susskind, Miguel Ángel Bautista HIGHLIGHT: In this paper we introduce Diffusion Probabilistic Fields (DPF), a diffusion model that can learn distributions over continuous functions defined over metric spaces, commonly known as fields. 784, TITLE: SpeedyZero: Mastering Atari with Limited Data and Time https://openreview.net/forum?id=Mg5CLXZgvLJ Yixuan Mei, Jiaxuan Gao, Weirui Ye, Shaohuai Liu, Yang Gao, Yi Wu AUTHORS: We develop SpeedyZero, a distributed RL system built upon a state-of-the-art model-based RL method, HIGHLIGHT: EfficientZero, with a dedicated system design for fast distributed computation. Learning Zero-Shot Cooperation with Humans, Assuming Humans Are Biased 785, TITLE: https://openreview.net/forum?id=TrwE819aJzs AUTHORS: Chao Yu, Jiaxuan Gao, Weilin Liu, Botian Xu, Hao Tang, Jiaqi Yang, Yu Wang, Yi Wu HIGHLIGHT: We propose a more general framework, Hidden-Utility Self-Play (HSP), which explicitly models human biases

as hidden reward functions in the self-play objective.

786. TITLE: What Do Self-Supervised Vision Transformers Learn? https://openreview.net/forum?id=azCKuYyS74 AUTHORS: Namuk Park, Wonjae Kim, Byeongho Heo, Taekyung Kim, Sangdoo Yun HIGHLIGHT: We present comparative studies on how and why contrastive learning (CL) and masked image modeling (MIM) differ in their representations and performance on downstream tasks. 787, TITLE: Symmetric Pruning in Quantum Neural Networks https://openreview.net/forum?id=K96AogLDT2K AUTHORS: Xinbiao Wang, Junyu Liu, Tongliang Liu, Yong Luo, Yuxuan Du, Dacheng Tao HIGHLIGHT: Empirical evidence showed that QNNs with handcraft symmetric ans\"atze generally experience better trainability than those with asymmetric ans\"atze, while theoretical explanations remain vague. To fill this knowledge gap, here we propose the effective quantum neural tangent kernel (EQNTK) and connect this concept with over-parameterization theory to quantify the convergence of QNNs towards the global optima. 788, TITLE: Learning Continuous Normalizing Flows For Faster Convergence To Target Distribution via Ascent Regularizations https://openreview.net/forum?id=6iEoTr-jeB7 AUTHORS: Shuangshuang Chen, Sihao Ding, Yiannis Karayiannidis, Mårten Björkman HIGHLIGHT: In this work, we propose a new class of continuous NFs, ascent continuous normalizing flows (ACNFs), that makes a base distribution converge faster to a target distribution. 789, TITLE: Beyond calibration: estimating the grouping loss of modern neural networks https://openreview.net/forum?id=6w1k-IixnL8 AUTHORS: Alexandre Perez-Lebel, Marine Le Morvan, Gael Varoquaux HIGHLIGHT: Here, we propose an estimator to approximate the grouping loss. 790, TITLE: Tuning Frequency Bias in Neural Network Training with Nonuniform Data https://openreview.net/forum?id=oLIZ2jGTiv AUTHORS: Annan Yu, Yunan Yang, Alex Townsend HIGHLIGHT: Using the Neural Tangent Kernel (NTK), one can provide a theoretically rigorous analysis for training where data are drawn from constant or piecewise-constant probability densities. Since most training data sets are not drawn from such distributions, we use the NTK model and a data-dependent quadrature rule to theoretically quantify the frequency biasing of NN training given fully nonuniform data. 791, TITLE: Quantifying Memorization Across Neural Language Models https://openreview.net/forum?id=TatRHT_1cK AUTHORS: Nicholas Carlini, Daphne Ippolito, Matthew Jagielski, Katherine Lee, Florian Tramer, Chiyuan Zhang We describe three log-linear relationships that quantify the degree to which LMs emit memorized training data. HIGHLIGHT: 792, TITLE: Temporal Coherent Test Time Optimization for Robust Video Classification https://openreview.net/forum?id=-t4D61w4zvQ Chenyu Yi, SIYUAN YANG, Yufei Wang, Haoliang Li, Yap-peng Tan, Alex Kot AUTHORS: HIGHLIGHT: In this work, we propose a Temporal Coherent Test-time Optimization framework (TeCo) to utilize spatiotemporal information in test-time optimization for robust video classification. 793, TITLE: MultiViz: Towards Visualizing and Understanding Multimodal Models https://openreview.net/forum?id=i2 TvOFmEml AUTHORS: Paul Pu Liang, Yiwei Lyu, Gunjan Chhablani, Nihal Jain, Zihao Deng, Xingbo Wang, Louis-Philippe Morency, Ruslan Salakhutdinov How can we visualize the internal modeling of multimodal interactions in these models? Our paper aims to fill HIGHLIGHT this gap by proposing MultiViz, a method for analyzing the behavior of multimodal models by scaffolding the problem of interpretability into 4 stages: (1) unimodal importance: how each modality contributes towards downstream modeling and prediction, (2) cross-modal interactions: how different modalities relate with each other, (3) multimodal representations: how unimodal and crossmodal interactions are represented in decision-level features, and (4) multimodal prediction: how decision-level features are composed to make a prediction. 794, TITLE: Learning Locality and Isotropy in Dialogue Modeling https://openreview.net/forum?id=dPs6BGO2QT0 AUTHORS: Han Wu, Haochen Tan, Mingjie Zhan, Gangming Zhao, Shaoqing Lu, Ding Liang, Linqi Song HIGHLIGHT: In this paper, we find that the generated representations are also not conversational, losing the conversation structure information during the context modeling stage. 795, TITLE: SimPer: Simple Self-Supervised Learning of Periodic Targets

https://openreview.net/forum?id=EKpMeEV0hOo

AUTHORS:Yuzhe Yang, Xin Liu, Jiang Wu, Silviu Borac, Dina Katabi, Ming-Zher Poh, Daniel McDuffHIGHLIGHT:In this paper, we present SimPer, a simple contrastive SSL regime for learning periodic information in data.

796, TITLE: Simplicial Hopfield networks

https://openreview.net/forum?id=_QLsH8gatwx

AUTHORS: Thomas F Burns, Tomoki Fukai HIGHLIGHT: Inspired by setwise connectivity in biology, we extend Hopfield networks by adding setwise connections and embedding these connections in a simplicial complex. 797. TITLE: A critical look at evaluation of GNNs under heterophily: Are we really making progress? https://openreview.net/forum?id=tJbbQfw-5wv AUTHORS: Oleg Platonov, Denis Kuznedelev, Michael Diskin, Artem Babenko, Liudmila Prokhorenkova In this work, we challenge this evaluation setting. Then, we propose a set of heterophilous graphs of varying HIGHLIGHT: properties that we believe can serve as a better benchmark for testing the performance of GNNs under heterophily. 798. TITLE: Recon: Reducing Conflicting Gradients From the Root For Multi-Task Learning https://openreview.net/forum?id=ivwZO-HnzG AUTHORS: Guangyuan SHI, Qimai Li, Wenlong Zhang, Jiaxin Chen, Xiao-Ming Wu HIGHLIGHT: In this paper, we take a different approach to reduce conflicting gradients from the root. 799, TITLE: A Model or 603 Exemplars: Towards Memory-Efficient Class-Incremental Learning https://openreview.net/forum?id=S07feAlQHgM AUTHORS: Da-Wei Zhou, Qi-Wei Wang, Han-Jia Ye, De-Chuan Zhan By analyzing the effect of different layers in the network, we find that shallow and deep layers have different HIGHLIGHT: characteristics in CIL. Motivated by this, we propose a simple yet effective baseline, denoted as MEMO for Memory-efficient Expandable MOdel. Reversible Column Networks 800, TITLE: https://openreview.net/forum?id=Oc2vIWU0iFY AUTHORS: Yuxuan Cai, Yizhuang Zhou, Qi Han, Jianjian Sun, Xiangwen Kong, Jun Li, Xiangyu Zhang HIGHLIGHT: We propose a new neural network design paradigm Reversible Column Networks (RevCols). \$\rm A^2Q\$: Aggregation-Aware Quantization for Graph Neural Networks 801, TITLE: https://openreview.net/forum?id=7L2mgi0TNEP AUTHORS: Zeyu Zhu, Fanrong Li, Zitao Mo, Qinghao Hu, Gang Li, Zejian Liu, Xiaoyao Liang, Jian Cheng HIGHLIGHT: Through an in-depth analysis of the topology of GNNs, we observe that the topology of the graph leads to significant differences between nodes, and most of the nodes in a graph appear to have a small aggregation value. Motivated by this, in this paper, we propose the Aggregation-Aware mixed-precision Quantization (\$\rm A^2Q\$) for GNNs, where an appropriate bitwidth is automatically learned and assigned to each node in the graph. 802, TITLE: Wav2Tok: Deep Sequence Tokenizer for Audio Retrieval https://openreview.net/forum?id=v8Mi8KU6056 AUTHORS: Adhiraj Banerjee, Vipul Arora HIGHLIGHT: This paper proposes Wav2Tok, a deep sequence tokenizer for audio that converts continuous-valued audio sequences to sequences of discrete tokens that are easier to retrieve via sequence queries. In-Situ Text-Only Adaptation of Speech Models with Low-Overhead Speech Imputations 803. TITLE: https://openreview.net/forum?id=T2Ncx PN2K AUTHORS: Ashish Mittal, Sunita Sarawagi, Preethi Jyothi We propose a new approach (TOLSTOI) that imputes speech representations internal to a baseline RNN-T, HIGHLIGHT: starting from text-only inputs, and performs in-situ adaptation that results in higher adaptation accuracy without any runtime overheads during decoding. 804, TITLE: The Tilted Variational Autoencoder: Improving Out-of-Distribution Detection https://openreview.net/forum?id=YlGsTZODyjz AUTHORS: Griffin Floto, Stefan Kremer, Mihai Nica HIGHLIGHT: Therefore, a small volume in the high-density region of the prior is problematic because it restricts the separation of latent points. To ameliorate this, we propose a simple generalization of the Gaussian distribution, called the tilted Gaussian, which has a maximum probability density occurring on a sphere instead of a single point. 805, TITLE: E3Bind: An End-to-End Equivariant Network for Protein-Ligand Docking https://openreview.net/forum?id=sO1OiAftOFv Yangtian Zhang, Huiyu Cai, Chence Shi, Jian Tang AUTHORS: HIGHLIGHT: This work focuses on blind flexible self-docking, where we aim to predict the positions, orientations and conformations of docked molecules. 806, TITLE: Sampling-free Inference for Ab-Initio Potential Energy Surface Networks https://openreview.net/forum?id=Tuk3Pgaizx AUTHORS: Nicholas Gao, Stephan Günnemann HIGHLIGHT: In this work, we address the inference shortcomings by proposing the Potential learning from ab-initio Networks (PlaNet) framework, in which we simultaneously train a surrogate model in addition to the neural wave function. 807, TITLE: Cycle-consistent Masked AutoEncoder for Unsupervised Domain Generalization

https://openreview.net/forum?id=wC98X1qpDBA

AUTHORS: Haiyang Yang, SHIXIANG TANG, Xiaotong Li, Feng Zhu, Yizhou Wang, Meilin Chen, LEI BAI, Rui Zhao, Wanli Ouvang

HIGHLIGHT: In this paper, we propose a cycle cross-domain reconstruction task for unsupervised domain generalization in the absence of paired images.

 808, TITLE:
 Towards Interpretable Deep Reinforcement Learning with Human-Friendly Prototypes

 https://openreview.net/forum?id=hWwY_Jq0xsN

 AUTHORS:
 Eoin M. Kenny, Mycal Tucker, Julie Shah

HIGHLIGHT: An "interpretable-by-design" deep reinforcement learning agent is proposed which uses prototypes for decision making.

809, TITLE: Self-Distillation for Further Pre-training of Transformers

https://openreview.net/forum?id=kj6oK_Hj40

AUTHORS: Seanie Lee, Minki Kang, Juho Lee, Sung Ju Hwang, Kenji Kawaguchi

HIGHLIGHT: However, all of them solely focus on language models and we empirically find that a Vision Transformer is vulnerable to overfitting as we continue to pretrain the model on target unlabeled data. In order to tackle this limitation, we propose self-distillation as a regularization for a further pre-training stage.

810, TITLE: KnowDA: All-in-One Knowledge Mixture Model for Data Augmentation in Low-Resource NLP https://openreview.net/forum?id=2nocgE1m0A

AUTHORS: Yufei Wang, Jiayi Zheng, Can Xu, Xiubo Geng, Tao Shen, Chongyang Tao, Daxin Jiang

HIGHLIGHT: Consequently, they have trivial task-specific knowledge and are limited to yielding low-quality synthetic data. To combat this issue, we propose Knowledge Mixture Data Augmentation Model (KnowDA), a Seq2Seq language model pretrained on a mixture of diverse NLP tasks under a novel framework of Knowledge Mixture Training (KoMT).

811, TITLE: DELTA: DEBIASED FULLY TEST-TIME ADAPTATION

https://openreview.net/forum?id=eGm22rqG93

AUTHORS: Bowen Zhao, Chen Chen, Shu-Tao Xia

HIGHLIGHT: In this paper, we provide a plug-in solution called DELTA for debiased fully test-time adaptation, which consists of two components: (i) Test-time batch renormalization (TBR), introduced to alleviate the bias in normalization statistics.

812, TITLE: Self-Supervised Set Representation Learning for Unsupervised Meta-Learning https://openreview.net/forum?id=kIAx30hYi p

AUTHORS: Dong Bok Lee, Seanie Lee, Kenji Kawaguchi, Yunji Kim, Jihwan Bang, Jung-Woo Ha, Sung Ju Hwang HIGHLIGHT: One notable aspect of metric-based meta-learning, however, is that it is widely interpreted as a set-level problem since the inference of discriminative class prototypes (or set representations) from few examples is crucial for the performance of downstream tasks. Motivated by this, we propose Set-SimCLR, a novel self-supervised set representation learning framework for targeting UML problem.

813, TITLE: Learning Human-Compatible Representations for Case-Based Decision Support https://openreview.net/forum?id=r0xte-t401

AUTHORS: Han Liu, Yizhou Tian, Chacha Chen, Shi Feng, Yuxin Chen, Chenhao Tan

HIGHLIGHT: In this work, we incorporate ideas from metric learning with supervised learning to examine the importance of alignment for effective decision support.

814, TITLE: LexIMAE: Lexicon-Bottlenecked Pretraining for Large-Scale Retrieval https://openreview.net/forum?id=PfpEtB3-csK

AUTHORS: Tao Shen, Xiubo Geng, Chongyang Tao, Can Xu, Xiaolong Huang, Binxing Jiao, Linjun Yang, Daxin Jiang HIGHLIGHT: Despite it deeply exploiting the lexicon-representing capability of pre-trained language models, a crucial gap remains between language modeling and lexicon-weighting retrieval -- the former preferring certain or low-entropy words whereas the latter favoring pivot or high-entropy words -- becoming the main barrier to lexicon-weighting performance for large-scale retrieval. To bridge this gap, we propose a brand-new pre-training framework, lexicon-bottlenecked masked autoencoder (LexMAE), to learn importance-aware lexicon representations.

815, TITLE: HypeR: Multitask Hyper-Prompted Training Enables Large-Scale Retrieval Generalization https://openreview.net/forum?id=kUf4BcWXGJr

AUTHORS:ZeFeng Cai, Chongyang Tao, Tao Shen, Can Xu, Xiubo Geng, Xin Alex Lin, Liang He, Daxin JiangHIGHLIGHT:To this end, we propose HypeR, a hyper-prompted training mechanism to enable uniform retrieval across tasksof different domains.

 816, TITLE:
 Spiking Convolutional Neural Networks for Text Classification

 https://openreview.net/forum?id=pgU3k7QXu20

 AUTHORS:
 Changze Lv, Jianhan Xu, Xiaoqing Zheng

 HIGHLIGHT:
 This work presents a "conversion + fine-tuning" two-step method for t

HIGHLIGHT: This work presents a "conversion + fine-tuning" two-step method for training SNN for text classification and proposes a simple but effective way to encode pre-trained word embeddings as spike trains.

817, TITLE: Federated Learning from Small Datasets

https://openreview.net/forum?id=hDDV11sRV8

AUTHORS: Michael Kamp, Jonas Fischer, Jilles Vreeken

HIGHLIGHT: Bad local models can arbitrarily deteriorate the aggregate model quality, causing federating learning to fail in these settings. We propose a novel approach that avoids this problem by interleaving model aggregation and permutation steps. 818, TITLE: A Study of Causal Confusion in Preference-Based Reward Learning https://openreview.net/forum?id=R0Xxvr X3ZA AUTHORS: Jeremy Tien, Jerry Zhi-Yang He, Zackory Erickson, Anca Dragan, Daniel S. Brown HIGHLIGHT: We find that the presence of non-causal distractor features, noise in the stated preferences, partial state observability, and larger model capacity can all exacerbate causal confusion. Robust Graph Dictionary Learning 819. TITLE: https://openreview.net/forum?id=qxRscesArBZ AUTHORS: Weijie Liu, Jiahao Xie, Chao Zhang, Makoto Yamada, Nenggan Zheng, Hui Qian HIGHLIGHT: This paper proposes a robust graph dictionary learning method based on a novel robust variant of GWD. 820, TITLE: Ordered GNN: Ordering Message Passing to Deal with Heterophily and Over-smoothing https://openreview.net/forum?id=wKPmPBHSnT6 AUTHORS: Yunchong Song, Chenghu Zhou, Xinbing Wang, Zhouhan Lin HIGHLIGHT: In this work, we propose to order the messages passing into the node representation, with specific blocks of neurons targeted for message passing within specific hops. 821, TITLE: Omnigrok: Grokking Beyond Algorithmic Data https://openreview.net/forum?id=zDiHoIWa0q1 AUTHORS: Ziming Liu, Eric J Michaud, Max Tegmark HIGHLIGHT: We aim to understand grokking by analyzing the loss landscapes of neural networks, identifying the mismatch between training and test losses as the cause for grokking. ManyDG: Many-domain Generalization for Healthcare Applications 822, TITLE: https://openreview.net/forum?id=lcSfirnflpW AUTHORS: Chaoqi Yang, M Brandon Westover, Jimeng Sun HIGHLIGHT: In this paper, considering the diversity of patient covariates, we propose a new setting by treating each patient as a separate domain (leading to many domains). 823, TITLE: D4AM: A General Denoising Framework for Downstream Acoustic Models https://openreview.net/forum?id=5fvXH49wk2 AUTHORS: Chi-Chang Lee, Yu Tsao, Hsin-Min Wang, Chu-Song Chen HIGHLIGHT: In this study, we propose a general denoising framework for various downstream acoustic models, called D4AM. 824, TITLE: General Neural Gauge Fields https://openreview.net/forum?id=XWkWK2UagFR AUTHORS: Fangneng Zhan, Lingjie Liu, Adam Kortylewski, Christian Theobalt HIGHLIGHT: This begs a question: can we learn the gauge transformation along with the neural scene representations in an end-to-end manner? In this work, we extend this problem to a general paradigm with a taxonomy of discrete and continuous cases, and develop an end-to-end training framework to jointly optimize the gauge transformation and radiance fields. 825, TITLE: Fooling SHAP with Stealthily Biased Sampling https://openreview.net/forum?id=J4mJjotSauh AUTHORS: gabriel laberge, Ulrich Aïvodji, Satoshi Hara, Mario Marchand, Foutse Khomh HIGHLIGHT: In this paper, we propose a complementary family of attacks that leave the model intact and manipulate SHAP explanations using stealthily biased sampling of the data points used to approximate expectations w.r.t the background distribution. 826, TITLE: Inequality phenomenon in \$1 {\infty}\$-adversarial training, and its unrealized threats https://openreview.net/forum?id=4t9q35BxGr Ranjie Duan, YueFeng Chen, Yao Zhu, Xiaojun Jia, Rong Zhang, Hui Xue' AUTHORS: HIGHLIGHT: To validate our hypothesis, we proposed two simple attacks that either perturb or replace important features with noise or occlusion. Continual Transformers: Redundancy-Free Attention for Online Inference 827. TITLE: https://openreview.net/forum?id=PolHquob8M7 AUTHORS: Lukas Hedegaard, Arian Bakhtiarnia, Alexandros Iosifidis In this work, we propose novel formulations of the Scaled Dot-Product Attention, which enable Transformers to HIGHLIGHT: perform efficient online token-by-token inference on a continual input stream. 828, TITLE: Efficient Federated Domain Translation https://openreview.net/forum?id=uhLAcrAZ9cJ AUTHORS: Zeyu Zhou, Sheikh Shams Azam, Christopher Brinton, David I. Inouye While most existing FL algorithms focus on the conventional non-IID setting of class imbalance or missing HIGHLIGHT: classes across clients, in practice, the distribution differences could be more complex, e.g., changes in class conditional (domain) distributions. In this paper, we consider this complex case in FL wherein each client has access to only one domain distribution.

829, TITLE: Meta Knowledge Condensation for Federated Learning https://openreview.net/forum?id=TDf-XFAwc79 AUTHORS: Ping Liu, Xin Yu, Joey Tianyi Zhou HIGHLIGHT: Unlike existing paradigms, we introduce an alternative perspective to significantly decrease the federate learning communication cost without leaking original data.

Unicom: Universal and Compact Representation Learning for Image Retrieval 830, TITLE:

https://openreview.net/forum?id=3YFDsSRSxB-

AUTHORS: Xiang An, Jiankang Deng, Kaicheng Yang, Jaiwei Li, Ziyong Feng, Jia Guo, Jing Yang, Tongliang Liu HIGHLIGHT: In this paper, we first cluster the large-scale LAION dataset into one million pseudo classes based on the joint textual and visual features extracted by the CLIP model.

Self-supervised Geometric Correspondence for Category-level 6D Object Pose Estimation in the Wild 831, TITLE: https://openreview.net/forum?id=ZKDUlVMqG O

AUTHORS: Kaifeng Zhang, Yang Fu, Shubhankar Borse, Hong Cai, Fatih Porikli, Xiaolong Wang

Current approaches are restricted by leveraging annotations from simulation or collected from humans. In this HIGHLIGHT: paper, we overcome this barrier by introducing a self-supervised learning approach trained directly on large-scale real-world object videos for category-level 6D pose estimation in the wild.

Task-Aware Information Routing from Common Representation Space in Lifelong Learning 832. TITLE: https://openreview.net/forum?id=-M0TNnyWFT5 AUTHORS:

Prashant Shivaram Bhat, Bahram Zonooz, Elahe Arani

HIGHLIGHT: Thus, inspired by Global Workspace Theory of conscious information access in the brain, we propose TAMiL, a continual learning method that entails task-attention modules to capture task-specific information from the common representation space.

833, TITLE: Error Sensitivity Modulation based Experience Replay: Mitigating Abrupt Representation Drift in Continual Learning

https://openreview.net/forum?id=zlbci7019Z3

AUTHORS: Fahad Sarfraz, Elahe Arani, Bahram Zonooz

HIGHLIGHT: To this end, we propose ESMER which employs a principled mechanism for modulating the error sensitivity in a dual-memory rehearsal-based system.

Towards Robust Object Detection Invariant to Real-World Domain Shifts 834, TITLE:

https://openreview.net/forum?id=vqSyt8D3ny

AUTHORS: Oi Fan, Mattia Segu, Yu-Wing Tai, Fisher Yu, Chi-Keung Tang, Bernt Schiele, Dengxin Dai HIGHLIGHT: Existing classification domain generalization (DG) methods cannot effectively solve the robust object detection problem, because they either rely on multiple source domains with large style variance or destroy the content structures of the original images. In this paper, we analyze and investigate effective solutions to overcome domain style overfitting for robust object detection without the above shortcomings.

835. TITLE: PaLI: A Jointly-Scaled Multilingual Language-Image Model

https://openreview.net/forum?id=mWVoBz4W0u

AUTHORS: Xi Chen, Xiao Wang, Soravit Changpinyo, AJ Piergiovanni, Piotr Padlewski, Daniel Salz, Sebastian Goodman, Adam Grycner, Basil Mustafa, Lucas Beyer, Alexander Kolesnikov, Joan Puigcerver, Nan Ding, Keran Rong, Hassan Akbari, Gaurav Mishra, Linting Xue, Ashish V Thapliyal, James Bradbury, Weicheng Kuo, Mojtaba Seyedhosseini, Chao Jia, Burcu Karagol Ayan, Carlos Riquelme Ruiz, Andreas Peter Steiner, Anelia Angelova, Xiaohua Zhai, Neil Houlsby, Radu Soricut

HIGHLIGHT: We present PaLI, a model that extends this approach to the joint modeling of language and vision.

836, TITLE: Learnable Topological Features For Phylogenetic Inference via Graph Neural Networks https://openreview.net/forum?id=hVVUY7p64WL AUTHORS: Cheng Zhang HIGHLIGHT: In this paper, we propose a novel structural representation method for phylogenetic inference based on learnable topological features.

837, TITLE: Flow Annealed Importance Sampling Bootstrap

https://openreview.net/forum?id=XCTVFJwS9LJ

Laurence Illing Midgley, Vincent Stimper, Gregor N. C. Simm, Bernhard Schölkopf, José Miguel Hernández-AUTHORS: Lobato

HIGHLIGHT: We train normalizing flows to fit multi-modal target distributions by generating samples where the flow is a poor approximation of the target using an annealed importance sampling bootstrap procedure.

838, TITLE: NeRN: Learning Neural Representations for Neural Networks https://openreview.net/forum?id=9gfir3fSy3J Maor Ashkenazi, Zohar Rimon, Ron Vainshtein, Shir Levi, Elad Richardson, Pinchas Mintz, Eran Treister AUTHORS:

HIGHLIGHT: Similarly to the spatial smoothness of visual scenes, we show that incorporating a smoothness constraint over the original network's weights aids NeRN towards a better reconstruction.

839. TITLE: LS-IQ: Implicit Reward Regularization for Inverse Reinforcement Learning https://openreview.net/forum?id=o3Q4m8jg4BR

AUTHORS: Firas Al-Hafez, Davide Tateo, Oleg Arenz, Guoping Zhao, Jan Peters HIGHLIGHT: Previous works show that a squared norm regularization on the implicit reward function is effective, but do not provide a theoretical analysis of the resulting properties of the algorithms. In this work, we show that using this regularizer under a mixture distribution of the policy and the expert provides a particularly illuminating perspective: the original objective can be understood as squared Bellman error minimization, and the corresponding optimization problem minimizes the \$\chi^2\$-Divergence between the expert and the mixture distribution. 840, TITLE: Fairness-aware Contrastive Learning with Partially Annotated Sensitive Attributes https://openreview.net/forum?id=woa783QMul AUTHORS: Fengda Zhang, Kun Kuang, Long Chen, Yuxuan Liu, Chao Wu, Jun Xiao HIGHLIGHT: In this paper, we investigate a novel and practical problem of Fair Unsupervised Representation Learning with Partially annotated Sensitive labels (FURL-PS). In this way, we construct a balanced and unbiased dataset. A Laplace-inspired Distribution on SO(3) for Probabilistic Rotation Estimation 841. TITLE: https://openreview.net/forum?id=Mvetq8DO05O AUTHORS: Yingda Yin, Yang Wang, He Wang, Baoquan Chen HIGHLIGHT: In this paper, we draw inspiration from multivariate Laplace distribution and propose a novel Rotation Laplace distribution on SO(3). In-context Reinforcement Learning with Algorithm Distillation 842. TITLE: https://openreview.net/forum?id=hy0a5MMPUv AUTHORS: Michael Laskin, Luyu Wang, Junhyuk Oh, Emilio Parisotto, Stephen Spencer, Richie Steigerwald, DJ Strouse, Steven Stenberg Hansen, Angelos Filos, Ethan Brooks, maxime gazeau, Himanshu Sahni, Satinder Singh, Volodymyr Mnih HIGHLIGHT: We propose Algorithm Distillation (AD), a method for distilling reinforcement learning (RL) algorithms into neural networks by modeling their training histories with a causal sequence model. Robust Attention for Contextual Biased Visual Recognition 843, TITLE: https://openreview.net/forum?id=8XqDnrmZQNF AUTHORS: Ruyang Liu, Jingjia Huang, Ge Li, Thomas H. Li HIGHLIGHT: In this paper, to learn causal object features robust for contextual bias, we propose a novel attention module named Interventional Dual Attention (IDA) for visual recognition. 844, TITLE: Minimax Optimal Kernel Operator Learning via Multilevel Training https://openreview.net/forum?id=zEn1BhaNYsC AUTHORS: Jikai Jin, Yiping Lu, Jose Blanchet, Lexing Ying HIGHLIGHT: In this paper, we study the statistical limit of learning a Hilbert-Schmidt operator between two infinitedimensional Sobolev reproducing kernel Hilbert spaces. 845, TITLE: EAGLE: Large-scale Learning of Turbulent Fluid Dynamics with Mesh Transformers https://openreview.net/forum?id=mfIX4QpsARJ AUTHORS: Steeven JANNY, Aurélien Bénéteau, Madiha Nadri, Julie Digne, Nicolas THOME, Christian Wolf HIGHLIGHT: We introduce a new large-scale dataset for learning non-steady fluid mechanics and a method based on selfattention on graphs Matching receptor to odorant with protein language and graph neural networks 846, TITLE: https://openreview.net/forum?id=q9VherQJd8 AUTHORS: Matej Hladiš, Maxence Lalis, Sebastien Fiorucci, Jérémie Topin HIGHLIGHT: In this work, we combine [CLS] token from protBERT with a molecular graph and propose a tailored GNN architecture incorporating inductive biases from the protein-molecule binding. Asynchronous Gradient Play in Zero-Sum Multi-agent Games 847. TITLE: https://openreview.net/forum?id=vPXp7K Yhre AUTHORS: Ruicheng Ao, Shicong Cen, Yuejie Chi In this paper, we make progress by studying asynchronous gradient plays in zero-sum polymatrix games under HIGHLIGHT: delayed feedbacks. 848. TITLE: Concept-level Debugging of Part-Prototype Networks https://openreview.net/forum?id=oiwXWPDTyNk AUTHORS: Andrea Bontempelli, Stefano Teso, Katya Tentori, Fausto Giunchiglia, Andrea Passerini HIGHLIGHT: However, like other models, they are prone to picking up confounders and shortcuts from the data, thus suffering from compromised prediction accuracy and limited generalization. We propose ProtoPDebug, an effective concept-level debugger for ProtoPNets in which a human supervisor, guided by the model's explanations, supplies feedback in the form of what part-prototypes must be forgotten or kept, and the model is fine-tuned to align with this supervision. 849, TITLE: Dual Algorithmic Reasoning https://openreview.net/forum?id=hhvkdRdWt1F Danilo Numeroso, Davide Bacciu, Petar Veli?kovi? AUTHORS: HIGHLIGHT: Here, we propose to learn algorithms by exploiting duality of the underlying algorithmic problem.

Sampling-based inference for large linear models, with application to linearised Laplace

850, TITLE:

https://openreview.net/forum?id=aoDyX6vSqsd AUTHORS: Javier Antoran, Shreyas Padhy, Riccardo Barbano, Eric Nalisnick, David Janz, José Miguel Hernández-Lobato HIGHLIGHT: Alas, the computational cost associated with Bayesian linear models constrains this method's application to small networks, small output spaces and small datasets. We address this limitation by introducing a scalable sample-based Bayesian inference method for conjugate Gaussian multi-output linear models, together with a matching method for hyperparameter (regularisation) selection. 851. TITLE: Video Scene Graph Generation from Single-Frame Weak Supervision https://openreview.net/forum?id=KLrGINoxzb4 AUTHORS: Siqi Chen, Long Chen, Jun Xiao HIGHLIGHT: In this paper, we propose the first weakly-supervised VidSGG task with only single-frame weak supervision: SF-VidSGG. Personalized Reward Learning with Interaction-Grounded Learning (IGL) 852, TITLE: https://openreview.net/forum?id=wGvzQWFyUB Jessica Maghakian, Paul Mineiro, Kishan Panaganti, Mark Rucker, Akanksha Saran, Cheng Tan AUTHORS: HIGHLIGHT: We propose applying the recent Interaction Grounded Learning (IGL) paradigm to address the challenge of learning representations of diverse user communication modalities. 853. TITLE: Optimizing Bi-Encoder for Named Entity Recognition via Contrastive Learning https://openreview.net/forum?id=9EAQVEINuum AUTHORS: Sheng Zhang, Hao Cheng, Jianfeng Gao, Hoifung Poon HIGHLIGHT: We present a bi-encoder framework for named entity recognition (NER), which applies contrastive learning to map candidate text spans and entity types into the same vector representation space. 854, TITLE: On the Saturation Effect of Kernel Ridge Regression https://openreview.net/forum?id=tFvr-kYWs Y AUTHORS: Yicheng Li, Haobo Zhang, Qian Lin The saturation effect has been widely observed in practices and a saturation lower bound of KRR has been HIGHLIGHT: conjectured for decades. In this paper, we provide a proof of this long-standing conjecture. 855, TITLE: Visually-Augmented Language Modeling https://openreview.net/forum?id=8IN-qLkl215 AUTHORS: Weizhi Wang, Li Dong, Hao Cheng, Haoyu Song, Xiaodong Liu, Xifeng Yan, Jianfeng Gao, Furu Wei HIGHLIGHT: However, current large-scale pre-trained language models rely on the text-only self-supervised training with massive text data, which precludes them from utilizing relevant visual information when necessary. To address this, we propose a novel pre-training framework, named VaLM, to Visually-augment text tokens with retrieved relevant images for Language Modeling. 856, TITLE: DiGress: Discrete Denoising diffusion for graph generation https://openreview.net/forum?id=UaAD-Nu86WX AUTHORS: Clement Vignac, Igor Krawczuk, Antoine Siraudin, Bohan Wang, Volkan Cevher, Pascal Frossard HIGHLIGHT: This work introduces DiGress, a discrete denoising diffusion model for generating graphs with categorical node and edge attributes. Diffusion Probabilistic Modeling of Protein Backbones in 3D for the motif-scaffolding problem 857. TITLE: https://openreview.net/forum?id=6TxBxqNME1Y AUTHORS: Brian L. Trippe, Jason Yim, Doug Tischer, David Baker, Tamara Broderick, Regina Barzilay, Tommi S. Jaakkola HIGHLIGHT: We propose to learn a distribution over diverse and longer protein backbone structures via an E(3)-equivariant graph neural network. 858, TITLE: Advancing Radiograph Representation Learning with Masked Record Modeling https://openreview.net/forum?id=w-x7U26GM7j AUTHORS: Hong-Yu Zhou, Chenyu Lian, Liansheng Wang, Yizhou Yu HIGHLIGHT: Modern studies in radiograph representation learning (R\$^2\$L) rely on either self-supervision to encode invariant semantics or associated radiology reports to incorporate medical expertise, while the complementarity between them is barely noticed. To explore this, we formulate the self- and report-completion as two complementary objectives and present a unified framework based on masked record modeling (MRM). 859, TITLE: Time to augment visual self-supervised learning https://openreview.net/forum?id=08xdgmwCP81 AUTHORS: Arthur Aubret, Markus R. Ernst, Céline Teulière, Jochen Triesch HIGHLIGHT: This gives access to "augmentations" not commonly used in SSL, like watching the same object from multiple viewpoints or against different backgrounds. Here, we systematically investigate and compare the potential benefits of such timebased augmentations for learning object categories. 860, TITLE: Parametrizing Product Shape Manifolds by Composite Networks https://openreview.net/forum?id=F_EhNDSamN

AUTHORS: Josua Sassen, Klaus Hildebrandt, Benedikt Wirth, Martin Rumpf

HIGHLIGHT: This, however, often comes with high computational costs, which raises the question if one can learn an efficient neural network approximation. We show that this is indeed possible for shape spaces with a special product structure, namely those smoothly approximable by a direct sum of low-dimensional manifolds.

861, TITLE:	Moderate Coreset: A Universal Method of Data Selection for Real-world Data-efficient Deep Learning
https://openreview.net	t/forum?id=7D5EECbOaf9
AUTHORS:	Xiaobo Xia, Jiale Liu, Jun Yu, Xu Shen, Bo Han, Tongliang Liu
HIGHLIGHT:	In this paper, to address the issue, a concept of the moderate coreset is discussed.
862, TITLE:	Generative Modeling Helps Weak Supervision (and Vice Versa)
https://openreview.net	t/forum?id=3OaBBATwsvP
AUTHORS:	Benedikt Boecking, Nicholas Roberts, Willie Neiswanger, Stefano Ermon, Frederic Sala, Artur Dubrawski
HIGHLIGHT:	While these techniques would seem to be usable in concert, improving one another, how to build an interface
between them is not w	vell-understood. In this work, we propose a model fusing programmatic weak supervision and generative
adversarial networks a	and provide theoretical justification motivating this fusion.
863, 111LE:	Modeling Multimodal Aleatoric Uncertainty in Segmentation with Mixture of Stochastic Experts
https://openreview.net	t/torum?id=KE_wJD2RK4
AUTHORS:	Zhitong Gao, Yucong Chen, Chuyu Zhang, Xuming He
HIGHLIGHT:	In this work, we focus on capturing the data-inherent uncertainty (aka aleatoric uncertainty) in segmentation,
typically when ambig	uities exist in input images.
864, TTTLE:	Weakly-supervised HOI Detection via Prior-guided Bi-level Representation Learning
nttps://openreview.net	Viorum/id=resApvicqSB
AUTHORS:	Bo Wan, Yongtei Liu, Desen Zhou, Tinne Tuytelaars, Xuming He
HIGHLIGHI:	In contrast, we develop a CLIP-guided HOI representation capable of incorporating the prior knowledge at both
image level and HOI i	instance level, and adopt a self-taught mechanism to prune incorrect human-object associations.
865, IIILE:	Boosting Causal Discovery via Adaptive Sample Reweighting
https://openreview.net	//orum?id=LNpMtk15AS4
AUTHORS:	An Zhang, Fangtu Liu, Wenchang Ma, Zhibo Cai, Xiang Wang, Tat-Seng Chua
HIGHLIGHT	We propose a simple yet effective model-agnostic framework to boost causal discovery performance by
dynamically learning	the adaptive weights for the Reweighted Score function, Rescore for short, where the learned weights tailors
quantitatively to the in	nportant degree of each samples.
866, 111LE:	Near-Optimal Deployment Efficiency in Reward-Free Reinforcement Learning with Linear Function
Approximation	
https://openreview.net	Viorum?id=SNwH0dDGI/
AUTHORS:	Dan Qiao, Yu-Xiang Wang
HIGHLIGHT:	Under the linear MDP setting with feature dimension \$45 and planning norizon \$15, we propose a new
algorithm that collects	s at most $\frac{1}{1}$ s $\frac{1}{1}$ (c) $\frac{1}{1$
\$\epsilon\$-optimal po	licy for any (possibly data-dependent) choice of reward functions.
967 TITI E.	An Adaptive Policy to Employ Sharpage Aware Minimization
607, IIILE.	An Adaptive Folicy to Employ Sharpless-Aware Minimization
AUTUODS.	Waion Jing Harsi Yang Yu Zhang James Kwak
UICUI ICUT.	weisen Janig, Hansi Tang, Tu Zhang, Janes Kwok
IIIOIILIOIII.	in this paper, we design an adaptive poncy to employ SAM based on the loss failuscape geometry.
868 TITI E.	Conturing the Motion of Every Joint, 3D Human Pose and Shane Estimation with Independent Tokens
https://openreview.net	Capturing in envolution of Every sonit. 5D fruman fose and shape Estimation with independent forens
AUTUODS.	Son Varg War Hang Gong Lin, CHOZHONG LHO, Wankay Yang, Gong VII
HIGHLIGHT	Set Fails, well fields, Gailg Liu, GOOZHONG LOO, walked Fails, Gailg 10
IIIOIILIOIII.	in this paper we present a novel method to estimate 5D numan pose and shape norm monocular videos.
860 TITI E.	PatchDCT: Patch Refinement for High Quality Instance Segmentation
https://openreview.net	Factorize 1.7 and Action for Figh Quarty instance Segmentation
AUTHORS:	Onrou Wen Jimi Yang Yue Yang Kewei Liang
HIGHLIGHT.	Thus we propose a simple and power method named PatchDCT which separates the mask decoded from a DCT
vector into several nat	Thus, we propose a simple and novel method handed r action r, when separates the mask decoded from a Definition of the second refines each patch by the designed classifier and represent
vector into several par	tenes and remes caen paten by the designed classifier and regressor.
870 TITI F.	Jointly Learning Visual and Auditory Speech Representations from Raw Data
https://openreview.net	/forum?id=BPwlovf5iO
AUTHORS	Alexandros Haliassos, Pingchuan Ma, Rodrigo Mira, Stavros Petridis, Maia Pantic
HIGHLIGHT	We present RAVEn a self-supervised multi-modal approach to jointly learn visual and auditory speech
representations.	
1	
871, TITLE:	Reparameterization through Spatial Gradient Scaling
https://openreview.net	t/forum?id=Kpdewuy7RU6
AUTHORS:	Alexander Detkov, Mohammad Salameh, Muhammad Fetrat, Jialin Zhang, Robin Luwei, SHANGLING JUI.
D: Nin	

HIGHLIGHT: In this paper, we present a novel spatial gradient scaling method to redistribute learning focus among weights in convolutional neural networks. 872, TITLE: LMC: Fast Training of GNNs via Subgraph Sampling with Provable Convergence https://openreview.net/forum?id=5VBBA91N6n AUTHORS: Zhihao Shi, Xize Liang, Jie Wang HIGHLIGHT: This poses significant challenges to their convergence analysis and convergence speeds, which seriously limits their reliable real-world applications. To address this challenge, we propose a novel subgraph-wise sampling method with a convergence guarantee, namely Local Message Compensation (LMC). 873, TITLE: Gray-Box Gaussian Processes for Automated Reinforcement Learning https://openreview.net/forum?id=rmoMvptXK7M AUTHORS: Gresa Shala, André Biedenkapp, Frank Hutter, Josif Grabocka HIGHLIGHT: In this paper, we propose a novel gray-box Bayesian Optimization technique for HPO in RL, that enriches Gaussian Processes with reward curve estimations based on generalized logistic functions. 874, TITLE: Continuized Acceleration for Quasar Convex Functions in Non-Convex Optimization https://openreview.net/forum?id=yYbhKqdi7Hz AUTHORS: Jun-Kun Wang, Andre Wibisono HIGHLIGHT: In this work, we show that a recently proposed continuized Nesterov acceleration can be applied to minimizing quasar convex functions and achieves the optimal bound with a high probability. 875, TITLE: Towards Understanding GD with Hard and Conjugate Pseudo-labels for Test-Time Adaptation https://openreview.net/forum?id=FJXf1FXN8C AUTHORS: Jun-Kun Wang, Andre Wibisono HIGHLIGHT: In this work, we aim at theoretically understanding GD with hard and conjugate labels for a binary classification problem. Accelerating Hamiltonian Monte Carlo via Chebyshev Integration Time 876, TITLE: https://openreview.net/forum?id=FbRY1XVfwK AUTHORS: Jun-Kun Wang, Andre Wibisono In this work, we consider accelerating the process of sampling from a distribution $\phi(x) = \phi(x)$ HIGHLIGHT: f(x))\$ via HMC via time-varying integration time. 877. TITLE: MEDFAIR: BENCHMARKING FAIRNESS FOR MEDICAL IMAGEING https://openreview.net/forum?id=6ve2CkeQe5S AUTHORS: Yongshuo Zong, Yongxin Yang, Timothy Hospedales HIGHLIGHT In this work, we introduce MEDFAIR, a framework to benchmark the fairness of machine learning models for medical imaging. 878, TITLE: Progress measures for grokking via mechanistic interpretability https://openreview.net/forum?id=9XFSbDPmdW AUTHORS: Neel Nanda, Lawrence Chan, Tom Lieberum, Jess Smith, Jacob Steinhardt HIGHLIGHT: In this work, we argue that progress measures can be found via mechanistic interpretability---that is, by reverse engineering learned models into components and measuring the progress of each component over the course of training. 879, TITLE: Decision Transformer under Random Frame Dropping https://openreview.net/forum?id=NmZXv4467ai AUTHORS: Kaizhe Hu, Ray Chen Zheng, Yang Gao, Huazhe Xu HIGHLIGHT: To devise a robust and deployable algorithm, we propose Decision Transformer under Random Frame Dropping(DeFog), an offline RL algorithm that enables agents to act robustly in frame dropping scenarios without online interaction. 880, TITLE: Modeling content creator incentives on algorithm-curated platforms https://openreview.net/forum?id=l6CpxixmUg AUTHORS: Jiri Hron, Karl Krauth, Michael Jordan, Niki Kilbertus, Sarah Dean HIGHLIGHT: To this end, we propose tools for numerically finding equilibria in exposure games, and illustrate results of an audit on the MovieLens and LastFM datasets. 881, TITLE: DM-NeRF: 3D Scene Geometry Decomposition and Manipulation from 2D Images https://openreview.net/forum?id=C PRLz8bEJx AUTHORS: Bing WANG, Lu Chen, Bo Yang HIGHLIGHT: In this paper, we study the problem of 3D scene geometry decomposition and manipulation from 2D views. 882, TITLE: An efficient encoder-decoder architecture with top-down attention for speech separation https://openreview.net/forum?id=fzberKYWKsI Kai Li, Runxuan Yang, Xiaolin Hu AUTHORS: In this paper, we provide a bio-inspired efficient encoder-decoder architecture by mimicking the brain's top-HIGHLIGHT: down attention, called TDANet, with decreased model complexity without sacrificing performance.

883, TITLE: Bayesian Oracle for bounding information gain in neural encoding models

https://openreview.net/forum?id=iYC5hOMqUg

 AUTHORS:
 Konstantin-Klemens Lurz, Mohammad Bashiri, Edgar Y. Walker, Fabian H. Sinz

 HIGHLIGHT:
 In this work, we generalize the jack-knife oracle estimator for the mean---commonly used for correlation

 metrics---to a flexible Bayesian oracle estimator for IG based on posterior predictive distributions.

884, TITLE: Improving the imputation of missing data with Markov Blanket discovery

https://openreview.net/forum?id=GrpU6dxFmMN

AUTHORS: Yang Liu, Anthony Constantinou

HIGHLIGHT: In this paper, we propose a novel Markov Blanket discovery approach to determine the optimal feature set for a given variable by considering both observed variables and missingness of partially observed variables to account for systematic missingness.

885, TITLE: Simple Emergent Action Representations from Multi-Task Policy Training

https://openreview.net/forum?id=NUl0ylt7SM

AUTHORS: Pu Hua, Yubei Chen, Huazhe Xu

HIGHLIGHT: While sensory representations have been widely studied, the representations of actions that form motor skills are yet under exploration. In this work, we find that when a multi-task policy network takes as input states and task embeddings, a space based on the task embeddings emerges to contain meaningful action representations with moderate constraints. Within this space, interpolated or composed embeddings can serve as a high-level interface to instruct the agent to perform meaningful action sequences.

886, TITLE: Malign Overfitting: Interpolation and Invariance are Fundamentally at Odds

https://openreview.net/forum?id=dQNL7Zsta3

AUTHORS: Yoav Wald, Gal Yona, Uri Shalit, Yair Carmon

HIGHLIGHT: This suggests that the phenomenon of ``benign overfitting", in which models generalize well despite interpolating, might not favorably extend to settings in which robustness or fairness are desirable. In this work we provide a theoretical justification for these observations.

887, TITLE: Empirical Study of Pre-training a Backbone for 3D Human Pose and Shape Estimation

https://openreview.net/forum?id=8U4joMeLRF

AUTHORS: Hongsuk Choi, Hyeongjin Nam, Taeryung Lee, Gyeongsik Moon, Kyoung Mu Lee

HIGHLIGHT: However, its effects on 3DHPSE are open to question, whose target is fixed to a single class, the human. In this regard, we inspect the effectiveness of SSL on 3DHPSE and investigate two other pre-training approaches that have received relatively less attention.

888, TITLE: Selective Frequency Network for Image Restoration

https://openreview.net/forum?id=tyZ1ChGZIKO

AUTHORS: Yuning Cui, Yi Tao, Zhenshan Bing, Wenqi Ren, Xinwei Gao, Xiaochun Cao, Kai Huang, Alois Knoll HIGHLIGHT: In this paper, we exploit a multi-branch and content-aware module to decompose the feature into separate frequency subbands dynamically and locally, and then accentuate the useful ones via the channel-wise attention mechanism.

889, TITLE: Proposal-Contrastive Pretraining for Object Detection from Fewer Data

https://openreview.net/forum?id=gm0VZ-h-hPy

AUTHORS: Quentin Bouniot, Romaric Audigier, Angelique Loesch, Amaury Habrard

HIGHLIGHT: To address this problem, we are interested in transformer-based object detectors that have recently gained traction in the community with good performance and with the particularity of generating many diverse object proposals. In this work, we present Proposal Selection Contrast (ProSeCo), a novel unsupervised overall pretraining approach that leverages this property.

890, TITLE: Mass-Editing Memory in a Transformer

https://openreview.net/forum?id=MkbcAHIYgyS

AUTHORS: Kevin Meng, Arnab Sen Sharma, Alex J Andonian, Yonatan Belinkov, David Bau

HIGHLIGHT: We develop MEMIT, a method for directly updating a language model with many memories, demonstrating experimentally that it can scale up to thousands of associations for GPT-J (6B) and GPT-NeoX (20B), exceeding prior work by an order of magnitude.

891, TITLE: Unsupervised Semantic Segmentation with Self-supervised Object-centric Representations https://openreview.net/forum?id=1_jFneF07YC

 AUTHORS:
 Andrii Zadaianchuk, Matthaeus Kleindessner, Yi Zhu, Francesco Locatello, Thomas Brox

 HIGHLIGHT:
 In this paper, we show that recent advances in self-supervised representation learning enable unsupervised

 object discovery and semantic segmentation with a performance that matches the state of the field on supervised semantic segmentation 10 years ago.

 892, TITLE:
 CANIFE: Crafting Canaries for Empirical Privacy Measurement in Federated Learning

 https://openreview.net/forum?id=Kf7Yyf4O0u

 AUTHORS:
 Samuel Maddock, Alexandre Sablayrolles, Pierre Stock

 HIGHLIGHT:
 In this paper, we aim to achieve a tighter measurement of the model exposure by considering a realistic threat model.

893, TITLE: Effective Self-supervised Pre-training on Low-compute networks without Distillation https://openreview.net/forum?id=cbpRzMy-UZH

AUTHORS: Fuwen Tan, Fatemeh Sadat Saleh, Brais Martinez Most prior works attribute this poor performance to the capacity bottleneck of the low-compute networks and HIGHLIGHT: opt to bypass the problem through the use of knowledge distillation (KD). In this work, we revisit SSL for efficient neural networks, taking a closer at what are the detrimental factors causing the practical limitations, and whether they are intrinsic to the self-supervised low-compute setting. 894, TITLE: DualAfford: Learning Collaborative Visual Affordance for Dual-gripper Manipulation https://openreview.net/forum?id=I YZANaz5X AUTHORS: Yan Zhao, Ruihai Wu, Zhehuan Chen, Yourong Zhang, Qingnan Fan, Kaichun Mo, Hao Dong HIGHLIGHT: In this work, we propose a novel learning framework, DualAfford, to learn collaborative affordance for dualgripper manipulation tasks. We will release code and data upon acceptance. Instance-wise Batch Label Restoration via Gradients in Federated Learning 895. TITLE: https://openreview.net/forum?id=FIrQfNSOoTr AUTHORS: Kailang Ma, Yu Sun, Jian Cui, Dawei Li, Zhenyu Guan, Jianwei Liu HIGHLIGHT: An analytic method is proposed to perform instance-wise batch label restoration from only the gradient of the final layer. 896, TITLE: Perfectly Secure Steganography Using Minimum Entropy Coupling https://openreview.net/forum?id=HQ67mj5rJdR AUTHORS: Christian Schroeder de Witt, Samuel Sokota, J Zico Kolter, Jakob Nicolaus Foerster, Martin Strohmeier HIGHLIGHT In this work, we show that a steganography procedure is perfectly secure under Cachin (1998)'s information theoretic-model of steganography if and only if it is induced by a coupling. 897, TITLE: Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model https://openreview.net/forum?id=mRieQgMtNTQ Yinhuai Wang, Jiwen Yu, Jian Zhang AUTHORS: In this work, we propose the Denoising Diffusion Null-Space Model (DDNM), a novel zero-shot framework for HIGHLIGHT: arbitrary linear IR problems, including but not limited to image super-resolution, colorization, inpainting, compressed sensing, and deblurring. 898, TITLE: UniKGQA: Unified Retrieval and Reasoning for Solving Multi-hop Question Answering Over Knowledge Graph https://openreview.net/forum?id=Z63RvyAZ2Vh AUTHORS: Jinhao Jiang, Kun Zhou, Xin Zhao, Ji-Rong Wen HIGHLIGHT: In this paper, we propose UniKGOA, a novel approach for multi-hop KGOA task, by unifying retrieval and reasoning in both model architecture and parameter learning. 899, TITLE: ImageNet-X: Understanding Model Mistakes with Factor of Variation Annotations https://openreview.net/forum?id=HXz7Vcm3VgM AUTHORS: Badr Youbi Idrissi, Diane Bouchacourt, Randall Balestriero, Ivan Evtimov, Caner Hazirbas, Nicolas Ballas, Pascal Vincent, Michal Drozdzal, David Lopez-Paz, Mark Ibrahim While existing benchmarks surface examples challenging for models, they do not explain why such mistakes HIGHLIGHT: arise. To address this need, we introduce ImageNet-X-a set of sixteen human annotations of factors such as pose, background, or lighting the entire ImageNet-1k validation set as well as a random subset of 12k training images. 900. TITLE: AnyDA: Anytime Domain Adaptation https://openreview.net/forum?id=yyLvxYBJV1B AUTHORS: Omprakash Chakraborty, Aadarsh Sahoo, Rameswar Panda, Abir Das HIGHLIGHT: In this paper, we introduce a simple yet effective framework for anytime domain adaptation that is executable with dynamic resource constraints to achieve accuracy-efficiency trade-offs under domain-shifts. 901, TITLE: Towards Effective and Interpretable Human-Agent Collaboration in MOBA Games: A Communication Perspective https://openreview.net/forum?id=q3F0UBAruO AUTHORS: Yiming Gao, Feiyu Liu, Liang Wang, Zhenjie Lian, Weixuan Wang, Siqin Li, Xianliang Wang, Xianhan Zeng, Rundong Wang, jiawei wang, QIANG FU, Yang Wei, Lanxiao Huang, Wei Liu In this paper, we propose to enable humans and agents to collaborate through explicit communication by HIGHLIGHT: designing an efficient and interpretable Meta-Command Communication-based framework, dubbed MCC, for accomplishing effective human-agent collaboration in MOBA games. 902, TITLE: A Higher Precision Algorithm for Computing the \$1\$-Wasserstein Distance https://openreview.net/forum?id=aMXD8gqsIiC AUTHORS: Pankaj K Agarwal, Sharath Raghvendra, Pouyan Shirzadian, Rachita Sowle We propose an algorithm that runs in \$O(T(n,\varepsilon/d) \log n)\$ time and boosts the accuracy of estimating HIGHLIGHT: $\operatorname{W}(\operatorname{un},\operatorname{u}) \to \operatorname{d}(\operatorname{un},\operatorname{u}) \to \operatorname{d}(\operatorname{un},\operatorname{un}) \to \operatorname{d}(\operatorname{un},\operatorname{un}) \to \operatorname{un}, \operatorname{un}) \to \operatorname{un}, \operatorname{un} \to \operatorname{un} \to \operatorname{un}, \operatorname{un} \to \operatorname{un} \to$ 903, TITLE: Interpretations of Domain Adaptations via Layer Variational Analysis https://openreview.net/forum?id=YtntjusJV6

AUTHORS: Huan-Hsin Tseng, Hsin-Yi Lin, Kuo-Hsuan Hung, Yu Tsao

HIGHLIGHT: This study establishes both formal derivations and heuristic analysis to formulate the theory of transfer learning in deep learning. 904, TITLE: A Self-Attention Ansatz for Ab-initio Quantum Chemistry https://openreview.net/forum?id=xveTeHVIF7j AUTHORS: Ingrid von Glehn, James S Spencer, David Pfau HIGHLIGHT: We present a novel neural network architecture using self-attention, the Wavefunction Transformer (PsiFormer), which can be used as an approximation (or "Ansatz") for solving the many-electron Schrödinger equation, the fundamental equation for quantum chemistry and material science. 905, TITLE: Learning What and Where - Unsupervised Disentangling Location and Identity Tracking https://openreview.net/forum?id=NeDc-Ak-H AUTHORS: Manuel Traub, Sebastian Otte, Tobias Menge, Matthias Karlbauer, Jannik Thuemmel, Martin V. Butz HIGHLIGHT: Here we introduce a self-supervised LOCation and Identity tracking system (Loci), which excels on the CATER tracking challenge. 906, TITLE: Socratic Models: Composing Zero-Shot Multimodal Reasoning with Language https://openreview.net/forum?id=G2Q2Mh3avow AUTHORS: Andy Zeng, Maria Attarian, brian ichter, Krzysztof Marcin Choromanski, Adrian Wong, Stefan Welker, Federico Tombari, Aveek Purohit, Michael S Ryoo, Vikas Sindhwani, Johnny Lee, Vincent Vanhoucke, Pete Florence HIGHLIGHT: We investigate how multimodal prompt engineering can use language as the intermediate representation to combine complementary knowledge from different pretrained (potentially multimodal) language models for a variety of tasks. 907. TITLE: Multi-skill Mobile Manipulation for Object Rearrangement https://openreview.net/forum?id=Z3IClM bzvP AUTHORS: Jiayuan Gu, Devendra Singh Chaplot, Hao Su, Jitendra Malik To this end, we propose that the manipulation skills should include mobility to have flexibility in interacting HIGHLIGHT: with the target object from multiple locations and at the same time the navigation skill could have multiple end points which lead to successful manipulation. 908, TITLE: On The Inadequacy of Optimizing Alignment and Uniformity in Contrastive Learning of Sentence Representations https://openreview.net/forum?id=MxvHVNukama AUTHORS: Zhijie Nie, Richong Zhang, Yongyi Mao HIGHLIGHT: In this work, we consider the suitability of the decoupled form of contrastive loss, i.e., alignment and uniformity, in SRL. 909, TITLE: BALTO: efficient tensor program optimization with diversity-based active learning https://openreview.net/forum?id=CN223OXgyb5 AUTHORS: Jun Bi, Xiaqing Li, Qi Guo, Rui Zhang, Yuanbo Wen, Xing Hu, Zidong Du, Xinkai Song, Yifan Hao, Yunji Chen HIGHLIGHT: In this paper, we propose BALTO, a fast TPO approach with biased-diversity-based active learning, aiming at reducing much lower training costs under similar optimization accuracy. The key insight is that random sampling of existing approaches suffers from a heavy redundancy of low-performance programs, which incurs tremendous duplicated time-consuming measurements. 910. TITLE: Decision S4: Efficient Sequence-Based RL via State Spaces Layers https://openreview.net/forum?id=kqHkCVS7wbj AUTHORS: Shmuel Bar David, Itamar Zimerman, Eliya Nachmani, Lior Wolf HIGHLIGHT: In this work, we present two main algorithms: (i) an off-policy training procedure that works with trajectories, while still maintaining the training efficiency of the S4 model. 911, TITLE: DecAF: Joint Decoding of Answers and Logical Forms for Question Answering over Knowledge Bases https://openreview.net/forum?id=XHc5zRPxqV9 AUTHORS: Donghan Yu, Sheng Zhang, Patrick Ng, Henghui Zhu, Alexander Hanbo Li, Jun Wang, Yiqun Hu, William Yang Wang, Zhiguo Wang, Bing Xiang HIGHLIGHT: We propose a novel KBQA framework that jointly generates both direct answers and logical forms, and then combines them to obtain the final answers. 912, TITLE: ZiCo: Zero-shot NAS via inverse Coefficient of Variation on Gradients https://openreview.net/forum?id=rwo-ls5GqGn AUTHORS: Guihong Li, Yuedong Yang, Kartikeya Bhardwaj, Radu Marculescu HIGHLIGHT: To improve this state of affairs, as the main theoretical contribution, we first reveal how some specific gradient properties across different samples impact the convergence rate of neural networks. Based on this theoretical analysis, we propose a new zero-shot proxy, ZiCo, the first proxy that works consistently better than #Params. 913, TITLE: Hierarchical Relational Learning for Few-Shot Knowledge Graph Completion https://openreview.net/forum?id=zlwBI2gQL3K

 AUTHORS:
 Han Wu, Jie Yin, Bala Rajaratnam, Jianyuan Guo

 HIGHLIGHT:
 In this paper, we propose a hierarchical relational learning method (HiRe) for few-shot KG completion.

914. TITLE: Integrating Symmetry into Differentiable Planning with Steerable Convolutions https://openreview.net/forum?id=n7CPzMPKQ1 AUTHORS: Linfeng Zhao, Xupeng Zhu, Lingzhi Kong, Robin Walters, Lawson L.S. Wong HIGHLIGHT: Motivated by equivariant convolution networks, we treat the path planning problem as \textit {signals} over grids. 915, TITLE: Spectral Augmentation for Self-Supervised Learning on Graphs https://openreview.net/forum?id=DjzBCrMBJ p AUTHORS: Lu Lin, Jinghui Chen, Hongning Wang HIGHLIGHT: In this work, we aim to find a principled way for topology augmentations by exploring the invariance of graphs from the spectral perspective. Can We Faithfully Represent Absence States to Compute Shapley Values on a DNN? 916, TITLE: https://openreview.net/forum?id=YV8tP7bW6Kt AUTHORS: Jie Ren, Zhanpeng Zhou, Qirui Chen, Quanshi Zhang HIGHLIGHT: However, there are no studies investigating how to represent the absence of input variables and verify the faithfulness of baseline values. Therefore, we revisit the feature representation of a deep model in terms of causality, and propose to use causal patterns to examine whether the masking method faithfully removes information encoded in the input variable. 917, TITLE: OTOv2: Automatic, Generic, User-Friendly https://openreview.net/forum?id=7ynoX1ojPMt Tianyi Chen, Luming Liang, Tianyu DING, Zhihui Zhu, Ilya Zharkov AUTHORS: HIGHLIGHT: We propose the second generation of Only-Train-Once (OTOv2), which trains and compresses an arbitrary DNN only once from scratch to produce a more compact model with competitive performance without fine-tuning. 918, TITLE: RPM: Generalizable Behaviors for Multi-Agent Reinforcement Learning https://openreview.net/forum?id=HnSceSzlfrY AUTHORS: Wei Qiu, Xiao Ma, Bo An, Svetlana Obraztsova, Shuicheng YAN, Zhongwen Xu HIGHLIGHT: In this work, we model the problem with Markov Games and propose a simple yet effective method, ranked policy memory (RPM), to collect diverse multi-agent trajectories for training MARL policies with good generalizability. 919. TITLE: Is the Performance of My Deep Network Too Good to Be True? A Direct Approach to Estimating the Bayes Error in Binary Classification https://openreview.net/forum?id=FZdJQgy05rz AUTHORS: Takashi Ishida, Ikko Yamane, Nontawat Charoenphakdee, Gang Niu, Masashi Sugiyama HIGHLIGHT: We propose a simple and direct Bayes error estimator, where we just take the mean of the labels that show \emph{uncertainty} of the classes. 920, TITLE: Compositional Law Parsing with Latent Random Functions https://openreview.net/forum?id=PEuxUXIMLIA AUTHORS: Fan Shi, Bin Li, Xiangyang Xue HIGHLIGHT: In this paper, we propose a deep latent variable model for Compositional LAw Parsing (CLAP). Weakly Supervised Knowledge Transfer with Probabilistic Logical Reasoning for Object Detection 921. TITLE: https://openreview.net/forum?id=4yqxDCbzS98 AUTHORS: Martijn Oldenhof, Adam Arany, Yves Moreau, Edward De Brouwer HIGHLIGHT: In this work, we propose ProbKT, a framework based on probabilistic logical reasoning to train object detection models with arbitrary types of weak supervision. 922, TITLE: 3EF: Class-Incremental Learning via Efficient Energy-Based Expansion and Fusion https://openreview.net/forum?id=iP77 axu0h3 AUTHORS: Fu-Yun Wang, Da-Wei Zhou, Liu Liu, Yatao Bian, Han-Jia Ye, De-Chuan Zhan, Peilin Zhao HIGHLIGHT: In this work, we propose a unifying energy-based theory and framework called Efficient Energy-Based Expansion and Fusion (3EF) to analyze and achieve the goal of CIL. 923. TITLE: Neural Bregman Divergences for Distance Learning https://openreview.net/forum?id=nJ3Vx78Nf7p AUTHORS: Fred Lu, Edward Raff, Francis Ferraro HIGHLIGHT: We propose a new approach to learning arbitrary Bergman divergences in a differentiable manner via input convex neural networks and show that it overcomes significant limitations of previous works. 924, TITLE: ChiroDiff: Modelling chirographic data with Diffusion Models https://openreview.net/forum?id=1ROAstc9jv AUTHORS: Ayan Das, Yongxin Yang, Timothy Hospedales, Tao Xiang, Yi-Zhe Song Consequently, temporal data has been modelled as discrete token sequences of fixed sampling rate instead of HIGHLIGHT: capturing the true underlying concept. In this paper, we introduce a powerful model-class namely "Denoising Diffusion Probabilistic Models" or DDPMs for chirographic data that specifically addresses these flaws.

925, TITLE: Learning Harmonic Molecular Representations on Riemannian Manifold

https://openreview.net/forum?id=ySCL-NG I3 AUTHORS: Yiqun Wang, Yuning Shen, Shi Chen, Lihao Wang, Fei YE, Hao Zhou HIGHLIGHT: In this work, we propose a Harmonic Molecular Representation learning (HMR) framework, which represents a molecule using the Laplace-Beltrami eigenfunctions of the molecular surface. 926, TITLE: Scaleformer: Iterative Multi-scale Refining Transformers for Time Series Forecasting https://openreview.net/forum?id=sCrnllCtjoE AUTHORS: Mohammad Amin Shabani, Amir H. Abdi, Lili Meng, Tristan Sylvain HIGHLIGHT: In this paper, we propose a general multi-scale framework that can be applied to state-of-the-art transformerbased time series forecasting models (FEDformer, Autoformer, etc.). 927, TITLE: Mitigating Gradient Bias in Multi-objective Learning: A Provably Convergent Approach https://openreview.net/forum?id=dLAYGdKTi2 AUTHORS: Heshan Devaka Fernando, Han Shen, Miao Liu, Subhajit Chaudhury, Keerthiram Murugesan, Tianyi Chen HIGHLIGHT: To this end, we develop a stochastic multi-objective gradient correction (MoCo) method for multi-objective optimization. 928, TITLE: Understanding Neural Coding on Latent Manifolds by Sharing Features and Dividing Ensembles https://openreview.net/forum?id=1UCaQYUdE_o AUTHORS: Martin Bierke, Lukas Schott, Kristopher T Jensen, Claudia Battistin, David A, Klindt, Benjamin Adric Dunn HIGHLIGHT: We propose \$\textit {feature sharing}\$ across neural tuning curves, which significantly improves performance and leads to better-behaved optimization.

929, TITLE: Meta-Learning in Games

https://openreview.net/forum?id=uHaWaNhCvZD

 AUTHORS:
 Keegan Harris, Ioannis Anagnostides, Gabriele Farina, Mikhail Khodak, Steven Wu, Tuomas Sandholm

 HIGHLIGHT:
 In practice, however, strategic interactions—ranging from routing problems to online advertising auctions—

 evolve dynamically, thereby leading to many similar games to be solved. To address this gap, we introduce meta-learning for

 equilibrium finding and learning to play games.

930, TITLE: On Emergence of Activation Sparsity in Trained Transformers

https://openreview.net/forum?id=TJ2nxciYCk-

AUTHORS: Zonglin Li, Chong You, Srinadh Bhojanapalli, Daliang Li, Ankit Singh Rawat, Sashank J. Reddi, Ke Ye, Felix Chern, Felix Yu, Ruiqi Guo, Sanjiv Kumar

HIGHLIGHT: This paper reveals a curious observation that modern large-scale machine learning models with Transformer architectures have sparse activation maps.

931, TITLE: Adaptive Robust Evidential Optimization For Open Set Detection from Imbalanced Data

https://openreview.net/forum?id=3yJ-hcJBqe AUTHORS: Hitesh Sapkota, Qi Yu

HIGHLIGHT: In this paper, we propose Adaptive Robust Evidential Optimization (AREO) that offers a principled way to quantify sample uncertainty through evidential learning while optimally balancing the model training over all classes in the close set through adaptive distributively robust optimization (DRO).

932, TITLE: A view of mini-batch SGD via generating functions: conditions of convergence, phase transitions, benefit from negative momenta.

https://openreview.net/forum?id=bzaPGEllsjE

AUTHORS: Maksim Velikanov, Denis Kuznedelev, Dmitry Yarotsky

HIGHLIGHT: In this paper we develop a new analytic framework to analyze noise-averaged properties of mini-batch SGD for linear models at constant learning rates, momenta and sizes of batches.

933, TITLE: On the Usefulness of Embeddings, Clusters and Strings for Text Generation Evaluation https://openreview.net/forum?id=bvpkw7UIRdU

AUTHORS: Tiago Pimentel, Clara Isabel Meister, Ryan Cotterell

HIGHLIGHT: We provide a theoretical and empirical analysis of why a recently-proposed automatic evaluation metric for language generators correlates well with human judgments. We identify its use of embeddings from pretrained language models as the main reason.

 934, TITLE:
 Is Adversarial Training Really a Silver Bullet for Mitigating Data Poisoning?

 https://openreview.net/forum?id=zKvm1ETDOq

 AUTHORS:
 Rui Wen, Zhengyu Zhao, Zhuoran Liu, Michael Backes, Tianhao Wang, Yang Zhang

 HIGHLIGHT:
 We propose an indiscriminative feature-based poisoning approach to substantially degrade adversarial training,

 which was previously considered to be impossible.
 Backes, Tianhao Wang, Yang Zhang

935, TITLE: The Devil is in the Wrongly-classified Samples: Towards Unified Open-set Recognition https://openreview.net/forum?id=xLr0I xYGAs

AUTHORS:Jun CEN, Di Luan, Shiwei Zhang, Yixuan Pei, Yingya Zhang, Deli Zhao, Shaojie Shen, Qifeng ChenHIGHLIGHT:Recently, Unified Open-set Recognition (UOSR) has been proposed to reject not only unknown samples butalso known but wrongly classified samples, which tends to be more practical in real-world applications. In this paper, we deeplyanalyze the UOSR task under different training and evaluation settings to shed light on this promising research direction.

936, TITLE:	DFPC: Data flow driven pruning of coupled channels without data.
https://openreview.net	et/forum?id=mhnHqRqcjYU
AUTHORS:	Tanay Narshana, Chaitanya Murti, Chiranjib Bhattacharyya
HIGHLIGHT:	We propose a novel data-free algorithm to accelerate neural networks via pruning coupled channels.
937, TITLE: without fine tuning https://openreview.nd AUTHORS: HIGHLIGHT: original distribution a	TVSPrune - Pruning Non-discriminative filters via Total Variation separability of intermediate representations et/forum?id=sZI1Oj9KBKy Chaitanya Murti, Tanay Narshana, Chiranjib Bhattacharyya In this work, we address the challenge of pruning filters with only access to random samples drawn from the and without access to the original training set or loss function.
938, TITLE:	Deep Variational Implicit Processes
https://openreview.nd	et/forum?id=8aeSJNbmbQq
AUTHORS:	Luis A. Ortega, Simon Rodriguez Santana, Daniel Hernández-Lobato
HIGHLIGHT:	We describe a scalable variational inference algorithm for training DVIP and show that it outperforms previous
IP-based methods and	d also deep GPs.
939, TITLE:	Contrastive Learning for Unsupervised Domain Adaptation of Time Series
https://openreview.ne	et/forum?id=xPkJYRsQGM
AUTHORS:	Yilmazcan Ozyurt, Stefan Feuerriegel, Ce Zhang
HIGHLIGHT:	In this paper, we develop a novel framework for UDA of time series data, called CLUDA.
940, TITLE:	Cross-Level Distillation and Feature Denoising for Cross-Domain Few-Shot Classification
https://openreview.nd	et/forum?id=Kn-HA8DFik
AUTHORS:	Hao ZHENG, Runqi Wang, Jianzhuang Liu, Asako Kanezaki
HIGHLIGHT:	However, in practice, the base and the target datasets of few-shot classification are usually from different
domains, which is the	e problem of cross-domain few-shot classification. We tackle this problem by making a small proportion of
unlabeled images in t	the target domain accessible in the training stage.
941, TITLE: Approximation https://openreview.net AUTHORS: HIGHLIGHT:	Achieving Sub-linear Regret in Infinite Horizon Average Reward Constrained MDP with Linear Function et/forum?id=zZhX4eYNeeh Arnob Ghosh, Xingyu Zhou, Ness Shroff We study the infinite horizon average reward constrained Markov Decision Process (CMDP).
942, TITLE:	Exploring and Exploiting Decision Boundary Dynamics for Adversarial Robustness
https://openreview.nd	et/forum?id=aRTKuscKByJ
AUTHORS:	Yuancheng Xu, Yanchao Sun, Micah Goldblum, Tom Goldstein, Furong Huang
HIGHLIGHT:	However, it is unclear whether existing robust training methods effectively increase the margin for each
vulnerable point duri	ng training. To understand this, we propose a continuous-time framework for quantifying the relative speed of the
decision boundary w	ith respect to each individual point.
943, TITLE:	Benchmarking Constraint Inference in Inverse Reinforcement Learning
https://openreview.nd	et/forum?id=vINj_Hv9szL
AUTHORS:	Guiliang Liu, Yudong Luo, Ashish Gaurav, Kasra Rezaee, Pascal Poupart
HIGHLIGHT:	In this paper, we construct an ICRL benchmark in the context of two major application domains: robot control
and autonomous driv	ing.
944, TITLE:	SCoMoE: Efficient Mixtures of Experts with Structured Communication
https://openreview.nd	et/forum?id=s-c96mSU0u5
AUTHORS:	zhiyuan zeng, Deyi Xiong
HIGHLIGHT:	To reduce the communication cost, we propose SCoMoE, an MoE architecture with structured all-to-all
communication, insp	ired by the hierarchical architecture of the communication topology.
945, TITLE: https://openreview.nd AUTHORS: HIGHLIGHT: constraints induce be reward function is gi from the expert data.	Learning Soft Constraints From Constrained Expert Demonstrations et/forum?id=8sSnD78NqTN Ashish Gaurav, Kasra Rezaee, Guiliang Liu, Pascal Poupart However, in many settings, the agent may optimize a reward function subject to some constraints, where the haviors that may be otherwise difficult to express with just a reward function. We consider the setting where the ven, and the constraints are unknown, and propose a method that is able to recover these constraints satisfactorily
946, TITLE:	Priors, Hierarchy, and Information Asymmetry for Skill Transfer in Reinforcement Learning
https://openreview.nd	st/forum?id=0v4VkCSkHNm
AUTHORS:	Sasha Salter, Kristian Hartikainen, Walter Goodwin, Ingmar Posner
HIGHLIGHT:	In this paper, we theoretically and empirically show the crucial expressivity-transferability trade-off of skills
across sequential task	cs, controlled by information asymmetry.

947, TITLE: Programmatically Grounded, Compositionally Generalizable Robotic Manipulation https://openreview.net/forum?id=rZ-wylY5VI AUTHORS: Renhao Wang, Jiayuan Mao, Joy Hsu, Hang Zhao, Jiajun Wu, Yang Gao HIGHLIGHT: To this end, we propose \ours, a {\it modular} approach to better leverage pretrained VL models by exploiting the syntactic and semantic structures of an input language instruction. Behavior Proximal Policy Optimization 948, TITLE: https://openreview.net/forum?id=3c13LptpIph AUTHORS: Zifeng Zhuang, Kun LEI, Jinxin Liu, Donglin Wang, Yilang Guo HIGHLIGHT: In this work, starting from the analysis of offline monotonic policy improvement, we get a surprising finding that some online on-policy algorithms are naturally able to solve offline RL. TabCaps: A Capsule Neural Network for Tabular Data Classification with BoW Routing 949, TITLE: https://openreview.net/forum?id=OgbtSLESnI AUTHORS: Jintai Chen, KuanLun Liao, Yanwen Fang, Danny Chen, Jian Wu HIGHLIGHT: In this paper, we propose to encapsulate all tabular features of an instance into vectorial features and process them collectively rather than have to deal with individual ones, which directly captures the representations at the instance level and benefits robust performances. 950, TITLE: An Exact Poly-Time Membership-Queries Algorithm for Extracting a Three-Layer ReLU Network https://openreview.net/forum?id=-CoNloheTs AUTHORS: Amit Daniely, Elad Granot HIGHLIGHT: In this work, we present a polynomial-time algorithm that can learn a depth-two ReLU network from queries under mild general position assumptions. 951, TITLE: Protein Representation Learning via Knowledge Enhanced Primary Structure Reasoning https://openreview.net/forum?id=VbCMhg7MRmj Hong-Yu Zhou, Yunxiang Fu, Zhicheng Zhang, Bian Cheng, Yizhou Yu AUTHORS: HIGHLIGHT: However, it fails to consider the semantic gap between protein sequences and natural language, and the resulting feature misalignment may adversely affect representation learning. To mitigate this, we propose Knowledge-exploited Autoencoder for Proteins (KeAP), which performs implicit knowledge encoding by learning to exploit knowledge for protein primary structure reasoning. GOOD: Exploring geometric cues for detecting objects in an open world 952, TITLE: https://openreview.net/forum?id=W-nZDQyuv8D AUTHORS: Haiwen Huang, Andreas Geiger, Dan Zhang HIGHLIGHT: This is because RGB-based models primarily rely on appearance similarity to detect novel objects and are also prone to overfitting short-cut cues such as textures and discriminative parts. To address these shortcomings of RGB-based object detectors, we propose incorporating geometric cues such as depth and normals, predicted by general-purpose monocular estimators. 953, TITLE: Cycle to Clique (Cy2C) Graph Neural Network: A Sight to See beyond Neighborhood Aggregation https://openreview.net/forum?id=7d-g8KozkiE Yun Young Choi, Sun Woo Park, Youngho Woo, U Jin Choi AUTHORS: HIGHLIGHT: This paper mathematically identifies the caliber of graph neural networks in classifying isomorphism classes of graphs with continuous node attributes up to their local topological properties. 954, TITLE: Multitask Prompt Tuning Enables Parameter-Efficient Transfer Learning https://openreview.net/forum?id=Nk2pDtuhTq AUTHORS: Zhen Wang, Rameswar Panda, Leonid Karlinsky, Rogerio Feris, Huan Sun, Yoon Kim HIGHLIGHT: In this paper, we propose multitask prompt tuning (MPT), which first learns a single transferable prompt by decomposing and distilling knowledge from multiple task-specific source prompts. Contrastive Audio-Visual Masked Autoencoder 955, TITLE: https://openreview.net/forum?id=QPtMRyk5rb AUTHORS: Yuan Gong, Andrew Rouditchenko, Alexander H. Liu, David Harwath, Leonid Karlinsky, Hilde Kuehne, James R. Glass HIGHLIGHT: Subsequently, we propose the Contrastive Audio-Visual Masked Auto-Encoder (CAV-MAE) by combining contrastive learning and masked data modeling, two major self-supervised learning frameworks, to learn a joint and coordinated audio-visual representation. 956, TITLE: Eva: Practical Second-order Optimization with Kronecker-vectorized Approximation https://openreview.net/forum?id=_Mic8V96Voy AUTHORS: Lin Zhang, Shaohuai Shi, Bo Li In this work, we present a memory- and time-efficient second-order algorithm named Eva with two novel HIGHLIGHT: techniques: 1) we construct the second-order information with the Kronecker factorization of small stochastic vectors over a minibatch of training data to reduce memory consumption, and 2) we derive an efficient update formula without explicitly computing the inverse of matrices using the Sherman-Morrison formula.

957, TITLE: Better Teacher Better Student: Dynamic Prior Knowledge for Knowledge Distillation https://openreview.net/forum?id=M0_sUuEyHs

AUTHORS: Martin Zong, Zengyu Qiu, Xinzhu Ma, Kunlin Yang, Chunya Liu, Jun Hou, Shuai Yi, Wanli Ouyang HIGHLIGHT: Particularly, we propose the dynamic prior knowledge (DPK), which integrates part of teacher's features as the prior knowledge before the feature distillation. 958. TITLE: SketchKnitter: Vectorized Sketch Generation with Diffusion Models https://openreview.net/forum?id=4eJ43EN2g6l AUTHORS: Qiang Wang, Haoge Deng, Yonggang Qi, Da Li, Yi-Zhe Song HIGHLIGHT: We show vectorized sketch generation can be identified as a reversal of the stroke deformation process. 959, TITLE: Towards Better Selective Classification https://openreview.net/forum?id=5gDz yTcst AUTHORS: Leo Feng, Mohamed Osama Ahmed, Hossein Hajimirsadeghi, Amir H. Abdi HIGHLIGHT: In this paper, we challenge the aforementioned methods and confirm that the superior performance of state-ofthe-art methods is owed to training a more generalizable classifier rather than their proposed selection mechanisms. 960. TITLE: Achieve Near-Optimal Individual Regret & Low Communications in Multi-Agent Bandits https://openreview.net/forum?id=QTXKTXJKIh Xuchuang Wang, Lin Yang, Yu-Zhen Janice Chen, Xutong Liu, Mohammad Hajiesmaili, Don Towsley, John AUTHORS: C.S. Lui In this paper, we propose a near-optimal algorithm on both individual and group regrets, in addition, we also HIGHLIGHT: propose a novel communication module in the algorithm, which only needs O(log(logT)) communication times where T is the number of decision rounds. 961. TITLE: Learning Hyper Label Model for Programmatic Weak Supervision https://openreview.net/forum?id=aCQt BrkSjC AUTHORS: Renzhi Wu, Shen-En Chen, Jieyu Zhang, Xu Chu HIGHLIGHT: In this work, we present a hyper label model that (once learned) infers the ground-truth labels for each dataset in a single forward pass without dataset-specific parameter learning. 962, TITLE: A Unified Approach to Reinforcement Learning, Quantal Response Equilibria, and Two-Player Zero-Sum Games https://openreview.net/forum?id=DpE5UYUQzZH AUTHORS: Samuel Sokota, Ryan D'Orazio, J Zico Kolter, Nicolas Loizou, Marc Lanctot, Ioannis Mitliagkas, Noam Brown, Christian Kroer HIGHLIGHT: In this work, we show that a single algorithm --- a simple extension to mirror descent with proximal regularization that we call magnetic mirror descent (MMD)---can produce strong results in both settings, despite their fundamental differences. 963, TITLE: Anisotropic Message Passing: Graph Neural Networks with Directional and Long-Range Interactions https://openreview.net/forum?id=socffUzSIlx AUTHORS: Moritz Thürlemann, Sereina Riniker HIGHLIGHT: In this work, an anisotropic state based on Cartesian multipoles is proposed as an addition to the existing hidden features. SYNC: SAFETY-AWARE NEURAL CONTROL FOR STABILIZING STOCHASTIC DELAY-964. TITLE: DIFFERENTIAL EQUATIONS https://openreview.net/forum?id= 8mS2NE-HXN AUTHORS: Jingdong Zhang, Qunxi Zhu, Wei Yang, Wei Lin HIGHLIGHT: Stabilization of the systems described by stochastic delay-differential equations is a challenging task in control community. Here, to achieve this task, we leverage neural networks to learn control policies using the information of the controlled systems in some prescribed regions. That Label's got Style: Handling Label Style Bias for Uncertain Image Segmentation 965, TITLE: https://openreview.net/forum?id=wZ2SVhOTzBX AUTHORS: Kilian Zepf, Eike Petersen, Jes Frellsen, Aasa Feragen This results in datasets that contain both data variability and differing label styles. In this paper, we demonstrate HIGHLIGHT: that applying state-of-the-art segmentation uncertainty models on such datasets can lead to model bias caused by the different label styles. 966, TITLE: Tensor-Based Sketching Method for the Low-Rank Approximation of Data Streams. https://openreview.net/forum?id=rOFKmzNTbC AUTHORS: Cuiyu Liu, Xiao Chuanfu, Mingshuo Ding, Chao Yang HIGHLIGHT: In this paper, from a subspace perspective, we propose a tensor-based sketching method for low-rank approximation of data streams. Differentiable Mathematical Programming for Object-Centric Representation Learning 967, TITLE: https://openreview.net/forum?id=1J-ZTr7aypY AUTHORS: Adeel Pervez, Phillip Lippe, Efstratios Gavves We propose topology-aware feature partitioning into \$k\$ disjoint partitions for given scene features as a method HIGHLIGHT:

for object-centric representation learning.

968, TITLE:	Semi-Implicit Variational Inference via Score Matching
AUTHORS:	Longlin Yu. Cheng Zhang
HIGHLIGHT:	In this paper, we propose SIVI-SM, a new method for SIVI based on an alternative training objective via score
matching.	
969, TITLE:	Scalable Subset Sampling with Neural Conditional Poisson Networks
https://openreview.ne	et/forum?id=p8hMBcPtvju
AUTHORS:	Adeel Pervez, Phillip Lippe, Efstratios Gavves
HIGHLIGHT:	We present a simple alternative method for sampling subsets based on \emph{conditional Poisson sampling}.
970, TITLE:	POPGym: Benchmarking Partially Observable Reinforcement Learning
https://openreview.ne	et/forum?id=chDrutUTs0K
AUTHORS:	Steven Morad, Ryan Kortvelesy, Matteo Bettini, Stephan Liwicki, Amanda Prorok
HIGHLIGHT:	We introduce Partially Observable Process Gym (POPGym), a two-part library containing (1) a diverse
collection of 14 partia	ally observable environments, each with multiple difficulties and (2) implementations of 13 memory model
baselines the most	in a single RL library.
971, TITLE:	Write and Paint: Generative Vision-Language Models are Unified Modal Learners
https://openreview.ne	et/forum?id=HgQR0mXQ1 a
AUTHORS:	Shizhe Diao, Wangchunshu Zhou, Xinsong Zhang, Jiawei Wang
HIGHLIGHT:	In this work, we disclose the potential of symmetrical generative vision-language pre-training in learning to
write and paint concu	irrently, and propose a new unified modal model, named DaVinci, trained with prefix language modeling and
prefix image modelir	ng, a simple generative self-supervised objective on image-text pairs.
972, TITLE: https://openreview.ne	Unsupervised Meta-learning via Few-shot Pseudo-supervised Contrastive Learning et/forum?id=TdTGGj7fYYJ Huiwon Lang Hankock Log Jinwoo Shin
HIGHLIGHT: Pseudo-supervised C	To overcome the limitations, we propose a simple yet effective unsupervised meta-learning framework, coined ontrast (PsCo), for few-shot classification.
973, TITLE:	Policy-Based Self-Competition for Planning Problems
https://openreview.ne	et/forum?id=SmufNDN90G
AUTHORS:	Jonathan Pirnay, Quirin Göttl, Jakob Burger, Dominik Gerhard Grimm
HIGHLIGHT: instead of its scalar p to-Plan' (GAZ PTP),	We leverage the idea of self-competition and directly incorporate a historical policy into the planning process erformance. Based on the recently introduced Gumbel AlphaZero (GAZ), we propose our algorithm GAZ 'Play- in which the agent learns to find strong trajectories by planning against possible strategies of its past self.
974, TITLE:	Active Image Indexing
https://openreview.ne	et/forum?id=K9RHxPpjn2
AUTHORS:	Pierre Fernandez, Matthijs Douze, Herve Jegou, Teddy Furon
HIGHLIGHT:	This paper improves the robustness of image copy detection with active indexing, that optimizes the interplay of
these two component	is.
975, TITLE:	LDMIC: Learning-based Distributed Multi-view Image Coding
https://openreview.ne	et/forum?id=ILQVw4cA5F9
AUTHORS:	Xinjie Zhang, Jiawei Shao, Jun Zhang
HIGHLIGHT:	We design a multi-view image compression framework based on symmetric distributed source coding
paradigm, which ach	ieves higher compression performance than previous multi-view image compression methods.
976, TITLE:	Few-shot Cross-domain Image Generation via Inference-time Latent-code Learning
https://openreview.ne	et/forum?id=sCYXJr3OJM8
AUTHORS:	Arnab Kumar Mondal, Piyush Tiwary, Parag Singla, Prathosh AP
HIGHLIGHT:	In this work, our objective is to adapt a Deep generative model trained on a large-scale source dataset to
multiple target domai	ins with scarce data.
977, TITLE:	Global Explainability of GNNs via Logic Combination of Learned Concepts
https://openreview.ne	et/forum?id=OTbRTIY4YS
AUTHORS:	Steve Azzolin, Antonio Longa, Pietro Barbiero, Pietro Lio, Andrea Passerini
HIGHLIGHT:	In this work, we propose GLGExplainer (Global Logic-based GNN Explainer), the first Global Explainer
capable of generating	g explanations as arbitrary Boolean combinations of learned graphical concepts.
978, TITLE:	Differentially Private \$L_2\$-Heavy Hitters in the Sliding Window Model
https://openreview.ne	et/forum?id=3UHoYrgIYkG
AUTHORS:	Jeremiah Blocki, Seunghoon Lee, Tamalika Mukherjee, Samson Zhou
HIGHLIGHT:	In this paper, we consider the problem of privately releasing the \$L_2\$-heavy hitters in the sliding window
model, which include using polylogarithmi	$e \ p_p$ -neavy nitters for $p_{14} \ p_{3}$ and in some sense are the strongest possible guarantees that can be achieved c space, but cannot be handled by existing techniques due to the sub-additivity of the L_2 norm.

979, TITLE: Gradient Boosting Performs Gaussian Process Inference https://openreview.net/forum?id=3VKiaagxw1S AUTHORS: Aleksei Ustimenko, Artem Beliakov, Liudmila Prokhorenkova HIGHLIGHT: This paper shows that gradient boosting based on symmetric decision trees can be equivalently reformulated as a kernel method that converges to the solution of a certain Kernel Ridge Regression problem. Calibrating Sequence likelihood Improves Conditional Language Generation 980, TITLE: https://openreview.net/forum?id=0qSOodKmJaN AUTHORS: Yao Zhao, Mikhail Khalman, Rishabh Joshi, Shashi Narayan, Mohammad Saleh, Peter J Liu HIGHLIGHT: In this work, we introduce sequence likelihood calibration (SLiC) where the likelihood of model generated sequences are calibrated to better align with reference sequences in the model's latent space. The Continuous CNN: from Task-Specific to Unified CNN Architecture 981, TITLE: https://openreview.net/forum?id=ZW5aK4yCRqU AUTHORS: David M Knigge, David W. Romero, Albert Gu, Efstratios Gavves, Erik J Bekkers, Jakub Mikolaj Tomczak, Mark Hoogendoorn, Jan-jakob Sonke HIGHLIGHT: Performant Convolutional Neural Network (CNN) architectures must be tailored to specific tasks in order to incorporate considerations such as input length, resolution, and dimentionality of the data. To overcome the need for such problemspecific CNN architectures, and the fragmentation they represent to the field, we introduce the \textit {Continuous Convolutional Neural Network} (CCNN): a single CNN architecture that can be used for tasks on data of arbitrary resolution, dimensionality and length without structural changes. 982, TITLE: SMART: Sentences as Basic Units for Text Evaluation https://openreview.net/forum?id=OIe3kpwl40D AUTHORS: Reinald Kim Amplayo, Peter J Liu, Yao Zhao, Shashi Narayan HIGHLIGHT: Widely used evaluation metrics for text generation either do not work well with longer texts or fail to evaluate all aspects of text quality. In this paper, we introduce a new metric called SMART to mitigate such limitations. 983, TITLE: Planckian Jitter: countering the color-crippling effects of color jitter on self-supervised training https://openreview.net/forum?id=Pia70sP2Oi1 AUTHORS: Simone Zini, Alex Gomez-Villa, Marco Buzzelli, Bart?omiej Twardowski, Andrew D. Bagdanov, Joost van de weijer HIGHLIGHT: In this paper, we analyze how the color jitter traditionally used in data augmentation negatively impacts the quality of the color features in learned feature representations. 984, TITLE: GAMR: A Guided Attention Model for (visual) Reasoning https://openreview.net/forum?id=iLMgk2IGNyv AUTHORS: Mohit Vaishnav, Thomas Serre HIGHLIGHT: Here, we present a novel transformer-based module for visual reasoning, the Guided Attention Model for (visual) Reasoning (\$textit {GAMR}\$), which instantiates an active vision theory -- positing that the brain solves complex visual reasoning problems dynamically -- via sequences of attention shifts to select and route task-relevant visual information into memory. 985, TITLE: Disentangling Learning Representations with Density Estimation https://openreview.net/forum?id=EMvG1Jdhw 8 AUTHORS: Eric Yeats, Frank Y Liu, Hai Li HIGHLIGHT: We present Gaussian Channel Autoencoder (GCAE), a method which achieves reliable disentanglement via scalable non-parametric density estimation of the latent space. 986, TITLE: Reward Design with Language Models https://openreview.net/forum?id=10uNUgI5K1 Minae Kwon, Sang Michael Xie, Kalesha Bullard, Dorsa Sadigh AUTHORS: HIGHLIGHT: This paper explores how to simplify reward design by using a large language model (LLM) such as GPT-3 as a proxy reward function, where the user provides a textual prompt containing a few examples (few-shot) or a description (zero-shot) of desired behavior. Our approach leverages this proxy reward function in an RL framework. 987, TITLE: Bayes-MIL: A New Probabilistic Perspective on Attention-based Multiple Instance Learning for Whole Slide Images https://openreview.net/forum?id= geIwiOyUhZ AUTHORS: Yufei CUI, Ziquan Liu, Xiangyu Liu, Xue Liu, Cong Wang, Tei-Wei Kuo, Chun Jason Xue, Antoni B. Chan HIGHLIGHT: In this paper, we propose Bayes-MIL to address the problem from a probabilistic perspective. 988, TITLE: When to Make and Break Commitments? https://openreview.net/forum?id=q8vgHfPdoQP AUTHORS: Alihan Hüyük, Zhaozhi Qian, Mihaela van der Schaar

HIGHLIGHT: Given the time pressure created by the continual cost of keeping a commitment, we aim to answer: When should a decision-maker break a commitment that is likely to fail—either to make an alternative commitment or to make no further commitments at all?

989, TITLE: CASR: Generating Complex Sequences with Autoregressive Self-Boost Refinement https://openreview.net/forum?id=SV11w1u3InX

AUTHORS: Hongwei Han, Mengyu Zhou, Shi Han, Xiu Li, Dongmei Zhang

HIGHLIGHT: However, for complex sequences, the heuristic rules to break down them may hurt performance, and increase additional exposure bias. To tackle these challenges, we propose a PLM-friendly autoregressive self-boost refinement framework, CASR.

990, TITLE: Coverage-centric Coreset Selection for High Pruning Rates

https://openreview.net/forum?id=QwKvL6wC8Yi AUTHORS: Haizhong Zheng, Rui Liu, Fan Lai, J

AUTHORS: Haizhong Zheng, Rui Liu, Fan Lai, Atul Prakash HIGHLIGHT: These methods perform well at low pruning rates; but at high pruning rates, they have been found to suffer a

actuation accuracy drop, performing worse than even random coreset selection. In this paper, we explore the reasons for this accuracy drop both theoretically and empirically.

991, TITLE: Long-Tailed Partial Label Learning via Dynamic Rebalancing

https://openreview.net/forum?id=sXfWoK4KvSW

AUTHORS: Feng Hong, Jiangchao Yao, Zhihan Zhou, Yanfeng Wang, Ya Zhang

HIGHLIGHT: We show that even with the auxiliary of an oracle class prior, the state-of-the-art methods underperform due to an adverse fact that the constant rebalancing in LT is harsh to the label disambiguation in PLL. To overcome this challenge, we thus propose a dynamic rebalancing method, termed as RECORDS, without assuming any prior knowledge about the class distribution.

992, TITLE: Learning with Stochastic Orders

https://openreview.net/forum?id=P3PJokAqGW

AUTHORS: Carles Domingo-Enrich, Yair Schiff, Youssef Mroueh

HIGHLIGHT: Learning high-dimensional distributions is often done with explicit likelihood modeling or implicit modeling via minimizing integral probability metrics (IPMs). In this paper, we expand this learning paradigm to stochastic orders, namely, the \emph{convex} or \emph{Choquet order} between probability measures.

 993, TITLE:
 Certifiably Robust Policy Learning against Adversarial Multi-Agent Communication

 https://openreview.net/forum?id=dCOL0inGl3e

 AUTHORS:
 Yanchao Sun, Ruijie Zheng, Parisa Hassanzadeh, Yongyuan Liang, Soheil Feizi, Sumitra Ganesh, Furong

 Huang
 Yanchao Sun, Ruijie Zheng, Parisa Hassanzadeh, Yongyuan Liang, Soheil Feizi, Sumitra Ganesh, Furong

 $\label{eq:HIGHLIGHT:} In this work, we consider an environment with N agents, where the attacker may arbitrarily change the communication from any $C<\frac {N-1}{2}$ agents to a victim agent.}$

994, TITLE: Using Both Demonstrations and Language Instructions to Efficiently Learn Robotic Tasks https://openreview.net/forum?id=4u42KCQxCn8

AUTHORS: Albert Yu, Ray Mooney

HIGHLIGHT: In such cases, a combination of both a demonstration and an instruction more concisely and effectively conveys the task to the robot than either modality alone. To instantiate this problem setting, we train a single multi-task policy on a few hundred challenging robotic pick-and-place tasks and propose DeL-TaCo (Joint Demo-Language Task Conditioning), a method for conditioning a robotic policy on task embeddings comprised of two components: a visual demonstration and a language instruction.

995, TITLE: Strategic Classification on Graphs

https://openreview.net/forum?id=TuHkVOjSAR

AUTHORS: Itay Eilat, Ben Finkelshtein, Chaim Baskin, Nir Rosenfeld

HIGHLIGHT: As we show through analysis and simulation, this can work either against the system---or for it. Based on this, we propose a differentiable framework for strategically-robust learning of graph-based classifiers.

996, TITLE: Over-Training with Mixup May Hurt Generalization

https://openreview.net/forum?id=JmkjrlVE-DG

AUTHORS: Zixuan Liu, Ziqiao Wang, Hongyu Guo, Yongyi Mao

HIGHLIGHT: In this work, we report a previously unobserved phenomenon in Mixup raining: on a number of standard datasets, the performance of Mixup-trained models starts to decay after training for a large number of epochs, giving rise to a U-shaped generalization curve.

997, TITLE: Formal Mathematics Statement Curriculum Learning https://openreview.net/forum?id=-P7G-8dmSh4 AUTHORS: Stanislas Polu, Jesse Michael Han, Kunhao Zheng, Mantas Baksys, Igor Babuschkin, Ilya Sutskever HIGHLIGHT: We explore the use of expert iteration in the context of language modeling applied to formal mathematics. 998, TITLE: D4FT: A Deep Learning Approach to Kohn-Sham Density Functional Theory https://openreview.net/forum?id=aBWnqqsuot7 Tianbo Li, Min Lin, Zheyuan Hu, Kunhao Zheng, Giovanni Vignale, Kenji Kawaguchi, A.H. Castro Neto, AUTHORS: Kostya S. Novoselov, Shuicheng YAN HIGHLIGHT: In this work, we propose a deep learning approach to KS-DFT. 999, TITLE: The Trade-off between Universality and Label Efficiency of Representations from Contrastive Learning https://openreview.net/forum?id=rvsbw2YthH AUTHORS: Zhenmei Shi, Jiefeng Chen, Kunyang Li, Jayaram Raghuram, Xi Wu, Yingyu Liang, Somesh Jha

HIGHLIGHT: There are two key desiderata for the representation: label efficiency (the ability to learn an accurate classifier on top of the representation with a small amount of labeled data) and universality (usefulness across a wide range of downstream tasks).

In this paper, we focus on one of the most popular instantiations of this paradigm: contrastive learning with linear probing, i.e., learning a linear predictor on the representation pre-trained by contrastive learning. 1000, TITLE: Learning Heterogeneous Interaction Strengths by Trajectory Prediction with Graph Neural Network https://openreview.net/forum?id=qU6NIcpaSi-AUTHORS: Seungwoong Ha, Hawoong Jeong HIGHLIGHT: In this work, we propose the relational attentive inference network (RAIN) to infer continuously weighted interaction graphs without any ground-truth interaction strengths. 1001, TITLE: Conditional Antibody Design as 3D Equivariant Graph Translation https://openreview.net/forum?id=LFHFQbjxIiP AUTHORS: Xiangzhe Kong, Wenbing Huang, Yang Liu HIGHLIGHT: In this paper, we propose Multi-channel Equivariant Attention Network (MEAN), an end-to-end model that is able to co-design 1D sequences and 3D structures of CDRs. 1002, TITLE: Solving Continuous Control via Q-learning https://openreview.net/forum?id=U5XOGxAgccS AUTHORS: Tim Seyde, Peter Werner, Wilko Schwarting, Igor Gilitschenski, Martin Riedmiller, Daniela Rus, Markus Wulfmeier HIGHLIGHT: However, most actor-critic methods come at the cost of added complexity: heuristics for stabilisation, compute requirements as well as wider hyperparameter search spaces. We show that these issues can be largely alleviated via Q-learning by combining action discretization with value decomposition, framing single-agent control as cooperative multi-agent reinforcement learning (MARL). 1003, TITLE: Masked Visual-Textual Prediction for Document Image Representation Pretraining https://openreview.net/forum?id=HE 75XY5Ljh AUTHORS: Yuechen Yu, Yulin Li, Chengquan Zhang, Xiaoqiang Zhang, Zengyuan Guo, Xiameng Qin, Kun Yao, Junyu Han, Errui Ding, Jingdong Wang HIGHLIGHT: In this paper, we present Masked Visual-Textual Prediction for document image representation pretraining, called MaskDoc. 1004, TITLE: Bit-Pruning: A Sparse Multiplication-Less Dot-Product https://openreview.net/forum?id=YUDiZcZTI8 AUTHORS: Yusuke Sekikawa, Shingo Yashima HIGHLIGHT: In this study, we realize energy-efficient neural networks by exploiting a \$\text{{mult}}-ess, sparse dotproduct. 1005, TITLE: Robust Active Distillation https://openreview.net/forum?id=ALDM5SN2r7M AUTHORS: Cenk Baykal, Khoa Trinh, Fotis Iliopoulos, Gaurav Menghani, Erik Vee HIGHLIGHT: In this paper, we present a parameter-free approach with provable guarantees to query the soft-labels of points that are simultaneously informative and correctly labeled by the teacher. Approximate Nearest Neighbor Search through Modern Error-Correcting Codes 1006, TITLE: https://openreview.net/forum?id=-jP rDkyfpI AUTHORS: Noam Touitou, Nissim Halabi In this paper, we introduce a better way of using LSH functions for ANNS. HIGHLIGHT: 1007, TITLE: Self-supervised learning with rotation-invariant kernels https://openreview.net/forum?id=8uu6JStuYm AUTHORS: Léon Zheng, Gilles Puy, Elisa Riccietti, Patrick Perez, Rémi Gribonval HIGHLIGHT: We introduce a regularization loss based on kernel mean embeddings with rotation-invariant kernels on the hypersphere (also known as dot-product kernels) for self-supervised learning of image representations. 1008, TITLE: MocoSFL: enabling cross-client collaborative self-supervised learning https://openreview.net/forum?id=2QGJXyMNoPz AUTHORS: Jingtao Li, Lingjuan Lyu, Daisuke Iso, Chaitali Chakrabarti, Michael Spranger HIGHLIGHT: Existing collaborative self-supervised learning (SSL) schemes are not suitable for cross-client applications because of their expensive computation and large local data requirements. To address these issues, we propose MocoSFL, a collaborative SSL framework based on Split Federated Learning (SFL) and Momentum Contrast (MoCo). 1009, TITLE: Architectural optimization over subgroups of equivariant neural networks https://openreview.net/forum?id=a6rCdfABJXg AUTHORS: Kaitlin Maile, Dennis George Wilson, Patrick Forré We propose the equivariance relaxation morphism, which preserves functionality while reparameterizing a HIGHLIGHT: group equivariant layer to operate with equivariance constraints on a subgroup, as well as the \$[G]\$-mixed equivariant layer, which mixes layers constrained to different groups to enable within-layer equivariance optimization.

1010, TITLE: DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking https://openreview.net/forum?id=kKF8 K-mBbS

AUTHORS: Gabriele Corso, Hannes Stärk, Bowen Jing, Regina Barzilay, Tommi S. Jaakkola HIGHLIGHT: We instead frame molecular docking as a generative modeling problem and develop DiffDock, a diffusion generative model over the non-Euclidean manifold of ligand poses. 1011, TITLE: ViewCo: Discovering Text-Supervised Segmentation Masks via Multi-View Semantic Consistency https://openreview.net/forum?id=2XLRBjY46O6 AUTHORS: Pengzhen Ren, Changlin Li, Hang Xu, Yi Zhu, Guangrun Wang, Jianzhuang Liu, Xiaojun Chang, Xiaodan Liang HIGHLIGHT: However, existing works focus on pixel grouping and cross-modal semantic alignment, while ignoring the correspondence among multiple augmented views of the same image. To overcome such limitation, we propose multi-View Consistent learning (ViewCo) for text-supervised semantic segmentation. Momentum Stiefel Optimizer, with Applications to Suitably-Orthogonal Attention, and Optimal Transport 1012. TITLE: https://openreview.net/forum?id=vCJ9-Ri-6xU AUTHORS: Lingkai Kong, Yuqing Wang, Molei Tao HIGHLIGHT: Yet, a new approach is proposed based on, for the first time, an interplay between thoughtfully designed continuous and discrete dynamics. 1013, TITLE: Re-calibrating Feature Attributions for Model Interpretation https://openreview.net/forum?id=WUWJIV2Yxtp AUTHORS: Peiyu Yang, NAVEED AKHTAR, Zeyi Wen, Mubarak Shah, Ajmal Saeed Mian HIGHLIGHT: We propose a re-calibration technique to calibrate existing integral-based attribution methods with valid references for a consistent explanation. 1014, TITLE: Meta-Learning Black-Box Optimization via Black-Box Optimization https://openreview.net/forum?id=mFDU0fP3EQH AUTHORS: Robert Tjarko Lange, Tom Schaul, Yutian Chen, Tom Zahavy, Valentin Dalibard, Chris Lu, Satinder Singh, Sebastian Flennerhag HIGHLIGHT: Hence, we propose to discover effective update rules for evolution strategies via meta-learning. 1015, TITLE: Modeling Sequential Sentence Relation to Improve Cross-lingual Dense Retrieval https://openreview.net/forum?id=-bVsNeR56KS AUTHORS: Shunyu Zhang, Yaobo Liang, MING GONG, Daxin Jiang, Nan Duan Motivated by an observation that the sentences in parallel documents are approximately in the same order, HIGHLIGHT: which is universal across languages, we propose to model this sequential sentence relation to facilitate cross-lingual representation learning. Emergent world representations: Exploring a sequence model trained on a synthetic task 1016, TITLE: https://openreview.net/forum?id=DeG07 TcZvT AUTHORS: Kenneth Li, Aspen K Hopkins, David Bau, Fernanda Viégas, Hanspeter Pfister, Martin Wattenberg HIGHLIGHT: Do these networks just memorize a collection of surface statistics, or do they rely on internal representations of the process that generates the sequences they see? We investigate this question by applying a variant of the GPT model to the task of predicting legal moves in a simple board game, Othello. PASHA: Efficient HPO and NAS with Progressive Resource Allocation 1017. TITLE: https://openreview.net/forum?id=syfgJE6nFRW AUTHORS: Ondrej Bohdal, Lukas Balles, Martin Wistuba, Beyza Ermis, Cedric Archambeau, Giovanni Zappella HIGHLIGHT: We propose an approach to tackle the challenge of tuning machine learning models trained on large datasets with limited computational resources. Broken Neural Scaling Laws 1018, TITLE: https://openreview.net/forum?id=sckjveqlCZ AUTHORS: Ethan Caballero, Kshitij Gupta, Irina Rish, David Krueger HIGHLIGHT: We present a smoothly broken power law functional form that accurately models the scaling behaviors (of artificial neural networks) (i.e. how the evaluation metric of interest varies as the amount of compute used for training, number of model parameters, or training dataset size varies) for each task from a very large and diverse set of upstream and downstream (i.e. zero-shot, prompted, and fine-tuned) tasks. 1019, TITLE: Metadata Archaeology: Unearthing Data Subsets by Leveraging Training Dynamics https://openreview.net/forum?id=PvLnIaJbt9 AUTHORS: Shoaib Ahmed Siddiqui, Nitarshan Rajkumar, Tegan Maharaj, David Krueger, Sara Hooker Our work provides a unified and efficient framework for Metadata Archaeology -- uncovering and inferring HIGHLIGHT: metadata of examples in a dataset. 1020, TITLE: Learnable Graph Convolutional Attention Networks https://openreview.net/forum?id=WsUMeHPo-2 AUTHORS: Adrián Javalov, Pablo Sanchez Martin, Amit Levi, Isabel Valera HIGHLIGHT: In this work, we aim at exploiting the strengths of both approaches to their full extent.

1021, TITLE: Explicit Box Detection Unifies End-to-End Multi-Person Pose Estimation

https://openreview.net AUTHORS: HIGHLIGHT: estimation, called ED- information.	/forum?id=s4WVupnJjmX Jie Yang, Ailing Zeng, Shilong Liu, Feng Li, Ruimao Zhang, Lei Zhang This paper presents a novel end-to-end framework with Explicit box Detection for multi-person Pose Pose, where it unifies the contextual learning between human-level (global) and keypoint-level (local)
1022, TITLE:	Sample-Efficient Reinforcement Learning by Breaking the Replay Ratio Barrier
https://openreview.net	/forum?id=OpC-9aBBVJe
AUTHORS:	Pierluca D'Oro, Max Schwarzer, Evgenii Nikishin, Pierre-Luc Bacon, Marc G Bellemare, Aaron Courville
HIGHLIGHT:	In this work, we show that fully or partially resetting the parameters of deep reinforcement learning agents
causes better replay ra	tio scaling capabilities to emerge.
1023, TITLE:	Simplified State Space Layers for Sequence Modeling
https://openreview.net	/forum?id=Ai8Hw3AXqks
AUTHORS:	Jimmy T.H. Smith, Andrew Warrington, Scott Linderman
HIGHLIGHT:	We build on the design of the S4 layer and introduce a new state space layer, the S5 layer.
1024, TITLE:	Distributed Extra-gradient with Optimal Complexity and Communication Guarantees
https://openreview.net	/forum?id=b3itJyarLM0
AUTHORS:	Ali Ramezani-Kebrya, Kimon Antonakopoulos, Igor Krawczuk, Justin Deschenaux, Volkan Cevher
HIGHLIGHT:	To this end, we propose a quantized generalized extra-gradient (Q-GenX), which is an unbiased and adaptive
compression method t	ailored to solve VIs.
1025, TITLE:	DINO: DETR with Improved DeNoising Anchor Boxes for End-to-End Object Detection
https://openreview.net	/forum?id=3mRwyG5one
AUTHORS:	Hao Zhang, Feng Li, Shilong Liu, Lei Zhang, Hang Su, Jun Zhu, Lionel Ni, Harry Shum
HIGHLIGHT:	We present DINO (DETR with Improved deNoising anchOr boxes), a strong end-to-end object detector.
1026, TITLE:	Contrastive Meta-Learning for Partially Observable Few-Shot Learning
https://openreview.net	/forum?id=6iVJOtr2zL2
AUTHORS:	Adam Jelley, Amos Storkey, Antreas Antoniou, Sam Devlin
HIGHLIGHT:	We consider the problem of learning a unified representation from partial observations, where useful features
may be present in only	y some of the views.
1027, TITLE: Dynamics https://openreview.net AUTHORS: HIGHLIGHT: model by embedding t	Optimistic Exploration with Learned Features Provably Solves Markov Decision Processes with Neural /forum?id=9kBCMNb5mc Sirui Zheng, Lingxiao Wang, Shuang Qiu, Zuyue Fu, Zhuoran Yang, Csaba Szepesvari, Zhaoran Wang We propose a novel algorithm that designs exploration incentives via learnable representations of the dynamics the neural dynamics into a kernel space induced by the system noise.
1028, TITLE:	CodeBPE: Investigating Subtokenization Options for Large Language Model Pretraining on Source Code
https://openreview.net	/forum?id=htL4UZ344nF
AUTHORS:	Nadezhda Chirkova, Sergey Troshin
HIGHLIGHT:	We propose subtokenziation that reduces average length by 1740% without downstream performance drop,
and show that a carefu	illy chosen subtokenization may significantly improve quality by 0.5-2%, possibly with some length increase.
1029, TITLE:	TILP: Differentiable Learning of Temporal Logical Rules on Knowledge Graphs
https://openreview.net	/forum?id=_X12NmQKvX
AUTHORS:	Siheng Xiong, Yuan Yang, Faramarz Fekri, James Clayton Kerce
HIGHLIGHT:	In this paper, we propose TILP, a differentiable framework for temporal logical rules learning.
1030, TITLE: https://openreview.net AUTHORS: HIGHLIGHT: transformations by \te; to describe the general batch normalization.	Scale-invariant Bayesian Neural Networks with Connectivity Tangent Kernel /forum?id=VZ5EaTl6dqa SungYub Kim, Sihwan Park, Kyung-Su Kim, Eunho Yang As a more fundamental solution, we propose new prior and posterior distributions invariant to scaling xtit{decomposing} the scale and connectivity of parameters, thereby allowing the resulting generalization bound lizability of a broad class of networks with the more practical class of transformations such as weight decay with
1031, TITLE:	LogicDP: Creating Labels for Graph Data via Inductive Logic Programming
https://openreview.net	/forum?id=2b2s9vd7wYv
AUTHORS:	Yuan Yang, Faramarz Fekri, James Clayton Kerce, Ali Payani
HIGHLIGHT:	In this work, we propose LogicDP, a data programming framework for graph data.
1032, TITLE: https://openreview.net AUTHORS: HIGHLIGHT: relations.	Hierarchical Protein Representations via Complete 3D Graph Networks /forum?id=9X-hgLDLYkQ Limei Wang, Haoran Liu, Yi Liu, Jerry Kurtin, Shuiwang Ji In this work, we propose to develop a novel hierarchical graph network, known as ProNet, to capture the

1033. TITLE: A Statistical Framework for Personalized Federated Learning and Estimation: Theory, Algorithms, and Privacy https://openreview.net/forum?id=FUiDMCr W4o AUTHORS: Kaan Ozkara, Antonious M. Girgis, Deepesh Data, Suhas Diggavi HIGHLIGHT: In this work, we begin with a generative framework that could potentially unify several different algorithms as well as suggest new algorithms. 1034, TITLE: WiNeRT: Towards Neural Ray Tracing for Wireless Channel Modelling and Differentiable Simulations https://openreview.net/forum?id=tPKKXeW33YU AUTHORS: Tribhuvanesh Orekondy, Pratik Kumar, Shreya Kadambi, Hao Ye, Joseph Soriaga, Arash Behboodi HIGHLIGHT: In this paper, we work towards a neural surrogate to model wireless electro-magnetic propagation effects in indoor environments. Grounding Graph Network Simulators using Physical Sensor Observations 1035, TITLE: https://openreview.net/forum?id=jsZsEd8VEY AUTHORS: Jonas Linkerhägner, Niklas Freymuth, Paul Maria Scheikl, Franziska Mathis-Ullrich, Gerhard Neumann HIGHLIGHT: In this work, we integrate sensory information to ground Graph Network Simulators on real world observations. 1036, TITLE: Implicit Regularization for Group Sparsity https://openreview.net/forum?id=d7Q0vVfJ0wO AUTHORS: Jiangyuan Li, Thanh V Nguyen, Chinmay Hegde, Raymond K. W. Wong HIGHLIGHT: We study the implicit regularization of gradient descent towards structured sparsity via a novel neural reparameterization, which we call a diagonally grouped linear neural network. 1037, TITLE: Noise Injection Node Regularization for Robust Learning https://openreview.net/forum?id=gmSZ-GPNY6 Noam Itzhak Levi, Itay Mimouni Bloch, Marat Freytsis, Tomer Volansky AUTHORS: We introduce Noise Injection Node Regularization (NINR), a method of injecting structured noise into Deep HIGHLIGHT: Neural Networks (DNN) during the training stage, resulting in an emergent regularizing effect. 1038, TITLE: Compositional Semantic Parsing with Large Language Models https://openreview.net/forum?id=gJW8hSGBys8 AUTHORS: Andrew Drozdov, Nathanael Schärli, Ekin Akyürek, Nathan Scales, Xinying Song, Xinyun Chen, Olivier Bousquet, Denny Zhou HIGHLIGHT: Previous research shows that appropriate prompting techniques enable large language models (LLMs) to solve artificial compositional generalization tasks such as SCAN. In this work, we identify additional challenges in more realistic semantic parsing tasks with larger vocabulary and refine these prompting techniques to address them. 1039, TITLE: Guess the Instruction! Making Language Models Stronger Zero-Shot Learners https://openreview.net/forum?id=FtOxgKe Zg2 AUTHORS: Seonghyeon Ye, Doyoung Kim, Joel Jang, Joongbo Shin, Minjoon Seo HIGHLIGHT: In this paper, we propose Flipped Learning, an alternative method of meta-training which trains the LM to generate the task instruction given the input instance and label. 1040. TITLE: Binding Language Models in Symbolic Languages https://openreview.net/forum?id=lH1PV42cbF AUTHORS: Zhoujun Cheng, Tianbao Xie, Peng Shi, Chengzu Li, Rahul Nadkarni, Yushi Hu, Caiming Xiong, Dragomir Radev, Mari Ostendorf, Luke Zettlemoyer, Noah A. Smith, Tao Yu HIGHLIGHT: We propose Binder, a training-free neural-symbolic framework that maps the task input to a program, which (1) allows binding a unified API of language model (LM) functionalities to a programming language (e.g., SQL, Python) to extend its grammar coverage and thus tackle more diverse questions, (2) adopts an LM as both the program parser and the underlying model called by the API during execution, and (3) requires only a few in-context exemplar annotations. 1041, TITLE: Summarization Programs: Interpretable Abstractive Summarization with Neural Modular Trees https://openreview.net/forum?id=ooxDOe7ZtBe AUTHORS: Swarnadeep Saha, Shiyue Zhang, Peter Hase, Mohit Bansal HIGHLIGHT: To this end, we propose the Summarization Program (SP), an interpretable modular framework consisting of an (ordered) list of binary trees, each encoding the step-by-step generative process of an abstractive summary sentence from the source document. 1042, TITLE: Lower Bounds on the Depth of Integral ReLU Neural Networks via Lattice Polytopes https://openreview.net/forum?id=2mvALOAWaxY AUTHORS: Christian Alexander Haase, Christoph Hertrich, Georg Loho HIGHLIGHT: We prove that the set of functions representable by ReLU neural networks with integer weights strictly increases with the network depth while allowing arbitrary width. 1043, TITLE: Deep Declarative Dynamic Time Warping for End-to-End Learning of Alignment Paths https://openreview.net/forum?id=UClBPxIZqnY Ming Xu, Sourav Garg, Michael Milford, Stephen Gould AUTHORS:

HIGHLIGHT: We instead propose a DTW layer based around deep declarative networks.

1044. TITLE: Diffusion Adversarial Representation Learning for Self-supervised Vessel Segmentation https://openreview.net/forum?id=H0gdPxSwkPb AUTHORS: Boah Kim, Yujin Oh, Jong Chul Ye HIGHLIGHT: Although learning-based segmentation approaches have been extensively studied, a large amount of groundtruth labels are required in supervised methods and confusing background structures make neural networks hard to segment vessels in an unsupervised manner. To address this, here we introduce a novel diffusion adversarial representation learning (DARL) model that leverages a denoising diffusion probabilistic model with adversarial learning, and apply it for vessel segmentation. 1045, TITLE: The Union of Manifolds Hypothesis https://openreview.net/forum?id=Rvee9CAX4fi AUTHORS: Bradley CA Brown, Anthony L. Caterini, Brendan Leigh Ross, Jesse C Cresswell, Gabriel Loaiza-Ganem HIGHLIGHT: This success would be impossible if there was no hidden low-dimensional structure in data of interest; this existence is posited by the manifold hypothesis, which states that the data lies on an unknown manifold of low intrinsic dimension. In this paper, we argue that this hypothesis does not properly capture the low-dimensional structure typically present in data. 1046, TITLE: Disparate Impact in Differential Privacy from Gradient Misalignment https://openreview.net/forum?id=qLOaeRvteqbx AUTHORS: Maria S. Esipova, Atiyeh Ashari Ghomi, Yaqiao Luo, Jesse C Cresswell HIGHLIGHT: In this work we study the fine-grained causes of unfairness in DPSGD and identify gradient misalignment due to inequitable gradient clipping as the most significant source. 1047, TITLE: Multi-Rate VAE: Train Once, Get the Full Rate-Distortion Curve https://openreview.net/forum?id=OJ8aSiCaMNK AUTHORS: Juhan Bae, Michael R. Zhang, Michael Ruan, Eric Wang, So Hasegawa, Jimmy Ba, Roger Baker Grosse In this paper, we introduce Multi-Rate VAE (MR-VAE), a computationally efficient framework for learning HIGHLIGHT: optimal parameters corresponding to various \$\beta\$ in a single training run. 1048, TITLE: Closing the gap: Exact maximum likelihood training of generative autoencoders using invertible layers https://openreview.net/forum?id=g8wBdhnstYz AUTHORS: Gianluigi Silvestri, Daan Roos, Luca Ambrogioni HIGHLIGHT: In this work, we provide an exact likelihood alternative to the variational training of generative autoencoders. 1049, TITLE: Benchmarking Offline Reinforcement Learning on Real-Robot Hardware https://openreview.net/forum?id=3k5CUGDLNdd AUTHORS: Nico Gürtler, Sebastian Blaes, Pavel Kolev, Felix Widmaier, Manuel Wuthrich, Stefan Bauer, Bernhard Schölkopf, Georg Martius HIGHLIGHT: The combination of offline reinforcement learning with large diverse datasets, however, has the potential to lead to a breakthrough in this challenging domain analogously to the rapid progress made in supervised learning in recent years. To coordinate the efforts of the research community toward tackling this problem, we propose a benchmark including: i) a large collection of data for offline learning from a dexterous manipulation platform on two tasks, obtained with capable RL agents trained in simulation; ii) the option to execute learned policies on a real-world robotic system and a simulation for efficient debugging. 1050, TITLE: PAC-NeRF: Physics Augmented Continuum Neural Radiance Fields for Geometry-Agnostic System Identification https://openreview.net/forum?id=tVkrbkz42vc AUTHORS: Xuan Li, Yi-Ling Qiao, Peter Yichen Chen, Krishna Murthy Jatavallabhula, Ming Lin, Chenfanfu Jiang, Chuang Gan HIGHLIGHT: In this work, we aim to identify parameters characterizing a physical system from a set of multi-view videos without any assumption on object geometry or topology. 1051, TITLE: Sign and Basis Invariant Networks for Spectral Graph Representation Learning https://openreview.net/forum?id=Q-UHqMorzil AUTHORS: Derek Lim, Joshua David Robinson, Lingxiao Zhao, Tess Smidt, Suvrit Sra, Haggai Maron, Stefanie Jegelka HIGHLIGHT: We introduce SignNet and BasisNet---new neural architectures that are invariant to two key symmetries displayed by eigenvectors: (i) sign flips, since if v is an eigenvector then so is -v; and (ii) more general basis symmetries, which occur in higher dimensional eigenspaces with infinitely many choices of basis eigenvectors. 1052, TITLE: KNN-Diffusion: Image Generation via Large-Scale Retrieval https://openreview.net/forum?id=x5mtJD2ovc AUTHORS: Shelly Sheynin, Oron Ashual, Adam Polyak, Uriel Singer, Oran Gafni, Eliya Nachmani, Yaniv Taigman HIGHLIGHT: In this work, we propose using large-scale retrieval methods, in particular, efficient k-Nearest-Neighbors (kNN), which offers novel capabilities: (1) training a substantially small and efficient text-to-image diffusion model without any text, (2) generating out-of-distribution images by simply swapping the retrieval database at inference time, and (3) performing text-driven local semantic manipulations while preserving object identity. 1053, TITLE: AudioGen: Textually Guided Audio Generation https://openreview.net/forum?id=CYK7RfcOzQ4

AUTHORS: Felix Kreuk, Gabriel Synnaeve, Adam Polyak, Uriel Singer, Alexandre Défossez, Jade Copet, Devi Parikh, Yaniv Taigman, Yossi Adi

HIGHLIGHT: In this work, we tackle the problem of generating audio samples conditioned on descriptive text captions. 1054, TITLE: On the Importance and Applicability of Pre-Training for Federated Learning https://openreview.net/forum?id=fWWFv--P0xP AUTHORS: Hong-You Chen, Cheng-Hao Tu, Ziwei Li, Han Wei Shen, Wei-Lun Chao HIGHLIGHT: Across multiple visual recognition benchmarks, we found that pre-training can not only improve FL, but also close its accuracy gap to the counterpart centralized learning, especially in the challenging cases of non-IID clients' data. 1055, TITLE: Fisher-Legendre (FishLeg) optimization of deep neural networks https://openreview.net/forum?id=c9lAOPvOHS AUTHORS: Jezabel R Garcia, Federica Freddi, Stathi Fotiadis, Maolin Li, Sattar Vakili, Alberto Bernacchia, Guillaume Hennequin HIGHLIGHT: We introduce a new approach to estimate the natural gradient via Legendre-Fenchel duality, provide a convergence proof, and show competitive performance on a number of benchmarks. 1056, TITLE: Multi-Objective Reinforcement Learning: Convexity, Stationarity and Pareto Optimality https://openreview.net/forum?id=TjEzIsyEsQ6 AUTHORS: Haoye Lu, Daniel Herman, Yaoliang Yu HIGHLIGHT: In this paper, we perform a rigorous analysis of policy induced value functions and use the insights to distinguish three views of Pareto optimality. MapTR: Structured Modeling and Learning for Online Vectorized HD Map Construction 1057. TITLE: https://openreview.net/forum?id=k7p_YAO7yE AUTHORS: Bencheng Liao, Shaoyu Chen, Xinggang Wang, Tianheng Cheng, Qian Zhang, Wenyu Liu, Chang Huang HIGHLIGHT: We propose a unified permutation-equivalent modeling approach, i.e., modeling map element as a point set with a group of equivalent permutations, which accurately describes the shape of map element and stabilizes the learning process. Discrete Predictor-Corrector Diffusion Models for Image Synthesis 1058, TITLE: https://openreview.net/forum?id=VM8batVBWvg AUTHORS: Jose Lezama, Tim Salimans, Lu Jiang, Huiwen Chang, Jonathan Ho, Irfan Essa HIGHLIGHT: We introduce Discrete Predictor-Corrector diffusion models (DPC), extending predictor-corrector samplers in Gaussian diffusion models to the discrete case. On the Certification of Classifiers for Outperforming Human Annotators 1059, TITLE: https://openreview.net/forum?id=X5ZMzRYqUjB AUTHORS: Qiongkai Xu, Christian Walder, Chenchen Xu HIGHLIGHT: In this paper, we first raise the challenge of evaluating the performance of both humans and models with respect to an oracle which is \$\textit {unobserved}\$. Towards Stable Test-time Adaptation in Dynamic Wild World 1060, TITLE: https://openreview.net/forum?id=g2YraF75Tj AUTHORS: Shuaicheng Niu, Jiaxiang Wu, Yifan Zhang, Zhiquan Wen, Yaofo Chen, Peilin Zhao, Mingkui Tan HIGHLIGHT: In this paper, we investigate the unstable reasons and find that the batch norm layer is a crucial factor hindering TTA stability. 1061, TITLE: Does Decentralized Learning with Non-IID Unlabeled Data Benefit from Self Supervision? https://openreview.net/forum?id=2L9gzS80tA4 AUTHORS: Lirui Wang, Kaiqing Zhang, Yunzhu Li, Yonglong Tian, Russ Tedrake HIGHLIGHT: In this work, we carefully study decentralized learning with unlabeled data through the lens of self-supervised learning (SSL), specifically contrastive visual representation learning. 1062, TITLE: PEER: A Collaborative Language Model https://openreview.net/forum?id=KbYevcLjnc Timo Schick, Jane A. Yu, Zhengbao Jiang, Fabio Petroni, Patrick Lewis, Gautier Izacard, Qingfei You, AUTHORS: Christoforos Nalmpantis, Edouard Grave, Sebastian Riedel As a consequence, they lack several abilities crucial for collaborative writing: They are unable to update HIGHLIGHT: existing texts, difficult to control and incapable of verbally planning or explaining their actions. To address these shortcomings, we introduce PEER, a collaborative language model that is trained to imitate the entire writing process itself. 1063, TITLE: Embedding Fourier for Ultra-High-Definition Low-Light Image Enhancement https://openreview.net/forum?id=5N0wtJZ89r9 AUTHORS: Chongyi Li, Chun-Le Guo, man zhou, Zhexin Liang, Shangchen Zhou, Ruicheng Feng, Chen Change Loy HIGHLIGHT: Unlike existing methods that address the problem in the spatial domain, we propose a new solution, UHDFour, that embeds Fourier transform into a cascaded network. Implicit regularization in Heavy-ball momentum accelerated stochastic gradient descent 1064, TITLE: https://openreview.net/forum?id=ZzdBhtEH9yB AUTHORS: Avrajit Ghosh, He Lyu, Xitong Zhang, Rongrong Wang HIGHLIGHT: We explore the implicit regularization in (SGD+M) and (GD+M) through a series of experiments validating our theory.

1065. TITLE: Selective Annotation Makes Language Models Better Few-Shot Learners https://openreview.net/forum?id=qY1hlv7gwg AUTHORS: Hongjin SU, Jungo Kasai, Chen Henry Wu, Weijia Shi, Tianlu Wang, Jiayi Xin, Rui Zhang, Mari Ostendorf, Luke Zettlemoyer, Noah A. Smith, Tao Yu HIGHLIGHT: Based on this framework, we propose an unsupervised, graph-based selective annotation method, voke-k, to select diverse, representative examples to annotate. 1066, TITLE: Imitating Human Behaviour with Diffusion Models https://openreview.net/forum?id=Pv1GPQzRrC8 AUTHORS: Tim Pearce, Tabish Rashid, Anssi Kanervisto, Dave Bignell, Mingfei Sun, Raluca Georgescu, Sergio Valcarcel Macua, Shan Zheng Tan, Ida Momennejad, Katja Hofmann, Sam Devlin HIGHLIGHT: We introduce several innovations to make diffusion models suitable for sequential environments; designing suitable architectures, investigating the role of guidance, and developing reliable sampling strategies. 1067. TITLE: Free Lunch for Domain Adversarial Training: Environment Label Smoothing https://openreview.net/forum?id=GPTjnA57h 3 AUTHORS: YiFan Zhang, xue wang, Jian Liang, Zhang Zhang, Liang Wang, Rong Jin, Tieniu Tan HIGHLIGHT: Despite its success, we observe training instability from DAT, mostly due to over-confident domain discriminator and environment label noise. To address this issue, we proposed Environment Label Smoothing (ELS), which encourages the discriminator to output soft probability, which thus reduces the confidence of the discriminator and alleviates the impact of noisy environment labels. 1068, TITLE: Improved Sample Complexity for Reward-free Reinforcement Learning under Low-rank MDPs https://openreview.net/forum?id=jpsw-KuOi7r AUTHORS: Yuan Cheng, Ruiquan Huang, Yingbin Liang, Jing Yang HIGHLIGHT: In this paper we focus on reward-free RL under low-rank MDP models, which capture the representation learning in RL. 1069. TITLE: CircNet: Meshing 3D Point Clouds with Circumcenter Detection https://openreview.net/forum?id=zQWqV2tzDv AUTHORS: Huan Lei, Ruitao Leng, Liang Zheng, Hongdong Li HIGHLIGHT: Taking advantage of learning-based techniques in triangulation, existing methods enumerate the complete combinations of candidate triangles, which is both complex and inefficient. In this paper, we leverage the duality between a triangle and its circumcenter, and introduce a deep neural network that detects the circumcenters to achieve point cloud triangulation. Towards the Generalization of Contrastive Self-Supervised Learning 1070, TITLE: https://openreview.net/forum?id=XDJwuEYHhme AUTHORS: Weiran Huang, Mingyang Yi, Xuyang Zhao, Zihao Jiang HIGHLIGHT: To this end, we define a kind of \$(\sigma,\delta)\$-measure to mathematically quantify the data augmentation, and then provide an upper bound of the downstream classification error rate based on the measure. 1071, TITLE: Towards the Out-of-Distribution Generalization of Contrastive Self-Supervised Learning https://openreview.net/forum?id=n0Pb9T5kmb AUTHORS: Xuyang Zhao, Tianqi Du, Yisen Wang, Jun Yao, Weiran Huang HIGHLIGHT: In this paper, by focusing on the data augmentation used in SSL, we establish a theoretical framework for the OOD performance of contrastive-based self-supervised learning. 1072, TITLE: Representational Dissimilarity Metric Spaces for Stochastic Neural Networks https://openreview.net/forum?id=xjb563TH-GH AUTHORS: Lyndon Duong, Josue Nassar, Jingyang Zhou, Jules Berman, Jeroen Olieslagers, Alex H Williams HIGHLIGHT: However, these measures of deterministic representational similarity ignore the scale and geometric structure of noise, both of which play important roles in neural computation. To rectify this, we generalize previously proposed shape metrics (Williams et al. 2021) to quantify differences in stochastic representations. 1073, TITLE: On the Feasibility of Cross-Task Transfer with Model-Based Reinforcement Learning https://openreview.net/forum?id=KB1sc5pNKFv Yifan Xu, Nicklas Hansen, Zirui Wang, Yung-Chieh Chan, Hao Su, Zhuowen Tu AUTHORS: HIGHLIGHT: In this work, we investigate whether internal models learned by modern model-based RL algorithms can be leveraged to solve new, distinctly different tasks faster. 1074, TITLE: Multimodal Federated Learning via Contrastive Representation Ensemble https://openreview.net/forum?id=Hnk1WRMAYqg AUTHORS: Qiying Yu, Yimu Wang, Ke Xu, Yang Liu, Jingjing Liu In this work, we propose \textit{Contrastive Representation Ensemble and Aggregation for Multimodal FL HIGHLIGHT: (CreamFL)}, a multimodal federated learning framework that enables training larger server models from clients with heterogeneous model architectures and data modalities, while only communicating knowledge on public dataset.

1075, TITLE: Improved Learning-augmented Algorithms for k-means and k-medians Clustering https://openreview.net/forum?id=dCSFiAl_VO3

AUTHORS: Thy Dinh Nguyen, Anamay Chaturvedi, Huy Nguyen HIGHLIGHT: We consider the problem of clustering in the learning-augmented setting. 1076, TITLE: Open-Vocabulary Object Detection upon Frozen Vision and Language Models https://openreview.net/forum?id=MIMwy4kh9lf AUTHORS: Weicheng Kuo, Yin Cui, Xiuye Gu, AJ Piergiovanni, Anelia Angelova HIGHLIGHT: We present F-VLM, a simple open-vocabulary object detection method built uponFrozenVision andLanguageModels. 1077, TITLE: E-CRF: Embedded Conditional Random Field for Boundary-caused Class Weights Confusion in Semantic Segmentation https://openreview.net/forum?id=g1GnnCI1OrC AUTHORS: Jie Zhu, Huabin Huang, Banghuai Li, Leye Wang HIGHLIGHT: We call this issue Boundary-caused Class Weights Confusion (BCWC). We try to focus on this problem and propose a novel method named Embedded Conditional Random Field (E-CRF) to alleviate it. 1078, TITLE: Bag of Tricks for Unsupervised Text-to-Speech https://openreview.net/forum?id=SbR9mpTuBn AUTHORS: Yi Ren, Chen Zhang, Shuicheng YAN HIGHLIGHT: In this work, we introduce a bag of tricks to enable effective unsupervised TTS. MCAL: Minimum Cost Human-Machine Active Labeling 1079. TITLE: https://openreview.net/forum?id=1FxRPKrH8bw AUTHORS: Hang Oiu, Krishna Chintalapudi, Ramesh Govindan HIGHLIGHT: In this paper, we consider the problem of hybrid human-machine labeling, which trains a classifier to accurately auto-label part of the data set. 1080, TITLE: Relaxed Combinatorial Optimization Networks with Self-Supervision: Theoretical and Empirical Notes on the Cardinality-Constrained Case https://openreview.net/forum?id=h21vJhdzbwz AUTHORS: Runzhong Wang, Li Shen, Yiting Chen, Xiaokang Yang, Dacheng Tao, Junchi Yan HIGHLIGHT: In this paper, we aim to develop a new paradigm to solve the CO problem by incorporating the constraints into the network architecture and computational operators, which is a more natural learning pipeline and decouples the constraint violation penalty from the raw objective optimization. 1081, TITLE: ROCO: A General Framework for Evaluating Robustness of Combinatorial Optimization Solvers on Graphs https://openreview.net/forum?id=2r6YMqz4Mml AUTHORS: Han Lu, Zenan Li, Runzhong Wang, Qibing Ren, Xijun Li, Mingxuan Yuan, Jia Zeng, Xiaokang Yang, Junchi Yan HIGHLIGHT: In this paper, we study the robustness of a combinatorial solver as a blackbox regardless it is classic or learningbased though the latter can often be more interesting to the ML community. 1082, TITLE: DASHA: Distributed Nonconvex Optimization with Communication Compression and Optimal Oracle Complexity https://openreview.net/forum?id=VA1YpcNr7ul AUTHORS: Alexander Tyurin, Peter Richtárik HIGHLIGHT: We develop and analyze DASHA: a new family of methods for nonconvex distributed optimization problems. 1083, TITLE: Revocable Deep Reinforcement Learning with Affinity Regularization for Outlier-Robust Graph Matching https://openreview.net/forum?id=QjQibO3scV Chang Liu, Zetian Jiang, Runzhong Wang, Lingxiao Huang, Pinyan Lu, Junchi Yan AUTHORS: HIGHLIGHT: In this paper, we focus on learning the back-end solver under the most general form of GM: Lawler's QAP, whose input is the affinity matrix. 1084, TITLE: Liquid Structural State-Space Models https://openreview.net/forum?id=g4OTKRKfS7R AUTHORS: Ramin Hasani, Mathias Lechner, Tsun-Hsuan Wang, Makram Chahine, Alexander Amini, Daniela Rus HIGHLIGHT: We use the recently proposed parametrization and memorization techniques for training state-space models in a linearized version of liquid neural networks, and achieve SOTA on sequence modeling tasks. 1085, TITLE: 3D UX-Net: A Large Kernel Volumetric ConvNet Modernizing Hierarchical Transformer for Medical Image Segmentation https://openreview.net/forum?id=wsZsjOSytRA AUTHORS: Ho Hin Lee, Shunxing Bao, Yuankai Huo, Bennett A. Landman HIGHLIGHT: In this work, we propose a lightweight volumetric ConvNet, termed 3D UX-Net, which adapts the hierarchical transformer using ConvNet modules for robust volumetric segmentation. 1086, TITLE: SoftZoo: A Soft Robot Co-design Benchmark For Locomotion In Diverse Environments https://openreview.net/forum?id=Xyme9p1rpZw

AUTHORS: Tsun-Hsuan Wang, Pingchuan Ma, Andrew Everett Spielberg, Zhou Xian, Hao Zhang, Joshua B. Tenenbaum, Daniela Rus, Chuang Gan HIGHLIGHT: In this work, we introduce SoftZoo, a soft robot co-design platform for locomotion in diverse environments. 1087. TITLE: Words are all you need? Language as an approximation for representational similarity https://openreview.net/forum?id=O-G91-4cMdv AUTHORS: Raja Marjieh, Pol Van Rijn, Ilia Sucholutsky, Theodore Sumers, Harin Lee, Thomas L. Griffiths, Nori Jacoby HIGHLIGHT: We conducted an evaluation of 611 pre-trained models across three domains -- images, audio, video -- and found that there is a large gap in performance between human similarity judgments and pre-trained DNNs. To address this gap, we propose a new class of similarity approximation methods based on language. Hungry Hungry Hippos: Towards Language Modeling with State Space Models 1088. TITLE: https://openreview.net/forum?id=COZDy0WYGg AUTHORS: Tri Dao, Daniel Y Fu, Khaled Kamal Saab, Armin W Thomas, Atri Rudra, Christopher Re HIGHLIGHT: In this paper, we make progress on understanding the expressivity gap between SSMs and attention in language modeling, and on reducing the hardware barrier between SSMs and attention. 1089, TITLE: Anti-Symmetric DGN: a stable architecture for Deep Graph Networks https://openreview.net/forum?id=J3Y7cgZOOS AUTHORS: Alessio Gravina, Davide Bacciu, Claudio Gallicchio HIGHLIGHT: In this work, we present Anti-Symmetric Deep Graph Networks (A-DGNs), a framework for stable and nondissipative DGN design, conceived through the lens of ordinary differential equations. 1090, TITLE: Information-Theoretic Diffusion https://openreview.net/forum?id=UvmDCdSPDOW AUTHORS: Xianghao Kong, Rob Brekelmans, Greg Ver Steeg We introduce a new mathematical foundation for diffusion models inspired by classic results in information HIGHLIGHT: theory that connects Information with Minimum Mean Square Error estimators, the so-called I-MMSE relations. 1091. TITLE: FunkNN: Neural Interpolation for Functional Generation https://openreview.net/forum?id=BT4N v7CLrk AUTHORS: AmirEhsan Khorashadizadeh, Anadi Chaman, Valentin Debarnot, Ivan Dokmani? HIGHLIGHT: The question is then how to extrapolate the spectrum in a data-driven way while meeting the above design criteria. Our answer is FunkNN---a novel convolutional network which learns how to reconstruct continuous images at arbitrary coordinates and can be applied to any image dataset. How to Train your HIPPO: State Space Models with Generalized Orthogonal Basis Projections 1092, TITLE: https://openreview.net/forum?id=klK17OQ3KB AUTHORS: Albert Gu, Isys Johnson, Aman Timalsina, Atri Rudra, Christopher Re HIGHLIGHT: Consequently, the theoretical mechanism by which S4 models long-range dependencies actually remains unexplained. We derive a more general and intuitive formulation of the HiPPO framework, which provides a simple mathematical interpretation of S4 as a decomposition onto exponentially-warped Legendre polynomials, explaining its ability to capture long dependencies. 1093, TITLE: GNNSafe: Out-of-Distribution Detection for Graph Neural Networks with Energy Models https://openreview.net/forum?id=zoz7Ze4STUL AUTHORS: Qitian Wu, Yiting Chen, Chenxiao Yang, Junchi Yan HIGHLIGHT: In this paper, we identify a provably effective OOD discriminator based on an energy function directly extracted from a graph neural network trained with standard supervised classification loss. DIFFormer: Scalable (Graph) Transformers Induced by Energy Constrained Diffusion 1094, TITLE: https://openreview.net/forum?id=j6zUzrapY3L AUTHORS: Qitian Wu, Chenxiao Yang, Wentao Zhao, Yixuan He, David Wipf, Junchi Yan To this end, we introduce an energy constrained diffusion model which encodes a batch of instances from a HIGHLIGHT: dataset into evolutionary states that progressively incorporate other instances' information by their interactions. 1095, TITLE: Graph Neural Networks are Inherently Good Generalizers: Insights by Bridging GNNs and MLPs https://openreview.net/forum?id=dqnNW2omZL6 AUTHORS: Chenxiao Yang, Qitian Wu, Jiahua Wang, Junchi Yan HIGHLIGHT: This paper pinpoints the major source of GNNs' performance gain to their intrinsic generalization capabilities, by introducing an intermediate model class dubbed as P(ropagational)MLP, which is identical to standard MLP in training, and then adopt GNN's architecture in testing. \$\Lambda\$-DARTS: Mitigating Performance Collapse by Harmonizing Operation Selection among Cells 1096, TITLE: https://openreview.net/forum?id=oztkQizr3kk AUTHORS: Sajad Movahedi, Melika Adabinejad, Ayyoob Imani, Arezou Keshavarz, Mostafa Dehghani, Azadeh Shakery, Babak N Araabi

HIGHLIGHT: However, the weight-sharing framework used for cell-search in DARTS and the convergence of architecture parameters has not been analyzed yet. In this paper, we provide a thorough and novel theoretical and empirical analysis on DARTS and its point of convergence.

1097. TITLE: Promptagator: Few-shot Dense Retrieval From 8 Examples https://openreview.net/forum?id=gmL46YMpu2J AUTHORS: Zhuyun Dai, Vincent Y Zhao, Ji Ma, Yi Luan, Jianmo Ni, Jing Lu, Anton Bakalov, Kelvin Guu, Keith Hall, Ming-Wei Chang HIGHLIGHT: In this paper, we suggest to work on Few-shot Dense Retrieval, a setting where each task comes with a short description and a few examples. 1098, TITLE: VoGE: A Differentiable Volume Renderer using Gaussian Ellipsoids for Analysis-by-Synthesis https://openreview.net/forum?id=AdPJb9cud Y AUTHORS: Angtian Wang, Peng Wang, Jian Sun, Adam Kortylewski, Alan Yuille HIGHLIGHT: In this paper, we propose VoGE, which uses ray tracing to capture nearest components with their volume density distributions on the rays and aggregates via integral of the volume densities based on Gaussian ellipsoids, which brings more efficient and stable gradients. 1099, TITLE: CLIP-ViP: Adapting Pre-trained Image-Text Model to Video-Language Alignment https://openreview.net/forum?id=GNjzMAgawq AUTHORS: Hongwei Xue, Yuchong Sun, Bei Liu, Jianlong Fu, Ruihua Song, Hougiang Li, Jiebo Luo HIGHLIGHT: However, adapting image-text pre-trained models to video-text pre-training (i.e., post-pretraining) has not demonstrated a significant advantage yet. In this paper, we tackle this challenge by raising and addressing two questions: 1) what are the factors hindering post-pretraining CLIP from improving performance on video-text tasks, and 2) how to mitigate the impact of these factors. 1100. TITLE: Diminishing Return of Value Expansion Methods in Model-Based Reinforcement Learning https://openreview.net/forum?id=H4Ncs5jhTCu AUTHORS: Daniel Palenicek, Michael Lutter, Joao Carvalho, Jan Peters Specifically, this paper addresses the value expansion class of model-based approaches. HIGHLIGHT: Boosting the Cycle Counting Power of Graph Neural Networks with I\$^2\$-GNNs 1101, TITLE: https://openreview.net/forum?id=kDSmxOspsXO AUTHORS: Yinan Huang, Xingang Peng, Jianzhu Ma, Muhan Zhang Specifically, we prove that Subgraph MPNNs fail to count more-than-4-cycles at node level, implying that node HIGHLIGHT: representations cannot correctly encode the surrounding substructures like ring systems with more than four atoms. To overcome this limitation, we propose I\$^2\$-GNNs to extend Subgraph MPNNs by assigning different identifiers for the root node and its neighbors in each subgraph. 1102, TITLE: Graph-based Deterministic Policy Gradient for Repetitive Combinatorial Optimization Problems https://openreview.net/forum?id=yHIIM9BgOo AUTHORS: Zhongyuan Zhao, Ananthram Swami, Santiago Segarra HIGHLIGHT: We propose an actor-critic framework for graph-based machine learning pipelines with non-differentiable blocks, and apply it to repetitive combinatorial optimization problems (COPs) under hard constraints. 1103, TITLE: Estimating individual treatment effects under unobserved confounding using binary instruments https://openreview.net/forum?id=ULsuEVQbV-9 AUTHORS: Dennis Frauen, Stefan Feuerriegel HIGHLIGHT: In this paper, we propose a novel, multiply robust machine learning framework, called MRIV, for estimating ITEs using binary IVs and thus yield an unbiased ITE estimator. Safe Reinforcement Learning From Pixels Using a Stochastic Latent Representation 1104, TITLE: https://openreview.net/forum?id=b39dQt_uffW Yannick Hogewind, Thiago D. Simão, Tal Kachman, Nils Jansen AUTHORS: HIGHLIGHT: We address the problem of safe reinforcement learning from pixel observations. This Looks Like It Rather Than That: ProtoKNN For Similarity-Based Classifiers 1105, TITLE: https://openreview.net/forum?id=lh-HRYxuoRr AUTHORS: Yuki Ukai, Tsubasa Hirakawa, Takayoshi Yamashita, Hironobu Fujiyoshi HIGHLIGHT: Due to this difficulty, the effectiveness of similarity-based classifiers (e.g., k-nearest neighbor (KNN)) on the 'this looks like that' framework have not been sufficiently examined. To alleviate this problem, we propose ProtoKNN, an extension of ProtoPNet that adopts KNN classifiers. 1106, TITLE: The Role of ImageNet Classes in Fréchet Inception Distance https://openreview.net/forum?id=4oXTQ6m_ws8 AUTHORS: Tuomas Kynkäänniemi, Tero Karras, Miika Aittala, Timo Aila, Jaakko Lehtinen HIGHLIGHT: While remarkably successful, the metric is known to sometimes disagree with human judgement. We investigate a root cause of these discrepancies, and visualize what FID "looks at" in generated images. 1107, TITLE: Bidirectional Language Models Are Also Few-shot Learners https://openreview.net/forum?id=wCFB37bzud4

AUTHORS: Ajay Patel, Bryan Li, Mohammad Sadegh Rasooli, Noah Constant, Colin Raffel, Chris Callison-Burch

HIGHLIGHT: We present SAP (Sequential Autoregressive Prompting), a technique that enables the prompting of bidirectional models. 1108, TITLE: MICN: Multi-scale Local and Global Context Modeling for Long-term Series Forecasting https://openreview.net/forum?id=zt53IDUR1U AUTHORS: Huiqiang Wang, Jian Peng, Feihu Huang, Jince Wang, Junhui Chen, Yifei Xiao HIGHLIGHT: To solve the above problems, we propose to combine local features and global correlations to capture the overall view of time series (e.g., fluctuations, trends). 1109. TITLE: Revisit Finetuning strategy for Few-Shot Learning to Strengthen the Equivariance of Emdeddings https://openreview.net/forum?id=tXc-riXhmx Heng Wang, Tan Yue, Xiang Ye, Zihang He, Bohan Li, Yong Li AUTHORS: HIGHLIGHT: To extract the undistorted features, we designed Linear-Probing-Finetuning with Firth-Bias (LP-FT-FB) to yield an accurate bias on the limited samples for better finetuning the pre-trained feature extractor, imposing equivariance on the whole model. 1110, TITLE: Synthetic Data Generation of Many-to-Many Datasets via Random Graph Generation https://openreview.net/forum?id=Q120 4COf-K AUTHORS: Kai Xu, Georgi Ganev, Emile Joubert, Rees Davison, Olivier Van Acker, Luke Robinson HIGHLIGHT: We synthesise datasets with many-to-many relationships by first generating the relationships via random graph generation and then generating the data attributes. 1111, TITLE: Memorization Capacity of Neural Networks with Conditional Computation https://openreview.net/forum?id=rB3zRN0lBYr AUTHORS: Erdem Koyuncu HIGHLIGHT: We study the fundamental limits of neural conditional computation from the perspective of memorization capacity. 1112, TITLE: Exploring The Role of Mean Teachers in Self-supervised Masked Auto-Encoders https://openreview.net/forum?id=7sn6Vxp92xV Youngwan Lee, Jeffrey Ryan Willette, Jonghee Kim, Juho Lee, Sung Ju Hwang AUTHORS: From this analysis, we present a simple SSL method, The Reconstruction-Consistent Masked Auto-Encoder HIGHLIGHT: (RC-MAE) by adding an EMA teacher to MAE. 1113, TITLE: The Best of Both Worlds: Accurate Global and Personalized Models through Federated Learning with Data-Free Hyper-Knowledge Distillation https://openreview.net/forum?id=29V3AWjVAFi AUTHORS: Huancheng Chen, Chaining Wang, Haris Vikalo HIGHLIGHT: We propose FedHKD (Federated Hyper-Knowledge Distillation), a novel FL algorithm in which clients rely on knowledge distillation (KD) to train local models. Learning to Solve Constraint Satisfaction Problems with Recurrent Transformers 1114. TITLE: https://openreview.net/forum?id=udNhDCr2KQe AUTHORS: Zhun Yang, Adam Ishay, Joohyung Lee HIGHLIGHT: We show that the Transformer model extended with recurrence is a viable approach to learning to solve CSPs in an end-to-end manner, having clear advantages over the state-of-the-art methods such as Graph Neural Networks, SATNet, and some neuro-symbolic models. 1115, TITLE: Learning Simultaneous Navigation and Construction in Grid Worlds https://openreview.net/forum?id=NEtep2C7yD AUTHORS: Wenyu Han, Haoran Wu, Eisuke Hirota, Alexander Gao, Lerrel Pinto, Ludovic Righetti, Chen Feng HIGHLIGHT: We propose to study a new learning task, mobile construction, to enable an agent to build designed structures in 1/2/3D grid worlds while navigating in the same evolving environments. 1116, TITLE: Leveraging Future Relationship Reasoning for Vehicle Trajectory Prediction https://openreview.net/forum?id=CGBCTp2M61A AUTHORS: Daehee Park, Hobin Ryu, Yunseo Yang, Jegyeong Cho, Jiwon Kim, Kuk-Jin Yoon HIGHLIGHT: In this paper, we propose a new method to formulate a stochastic future relationship among agents using lane structure. 1117, TITLE: Neural Networks and the Chomsky Hierarchy https://openreview.net/forum?id=WbxHAzkeQcn AUTHORS: Gregoire Deletang, Anian Ruoss, Jordi Grau-Moya, Tim Genewein, Li Kevin Wenliang, Elliot Catt, Chris Cundy, Marcus Hutter, Shane Legg, Joel Veness, Pedro A Ortega In this work, we conduct an extensive empirical study (10250 models, 15 tasks) to investigate whether insights HIGHLIGHT: from the theory of computation can predict the limits of neural network generalization in practice. 1118, TITLE: Federated Nearest Neighbor Machine Translation https://openreview.net/forum?id=R1U5G2spbLd AUTHORS: Yichao Du, Zhirui Zhang, Bingzhe Wu, Lemao Liu, Tong Xu, Enhong Chen

HIGHLIGHT: In this paper, we propose a novel federated nearest neighbor (FedNN) machine translation framework that, instead of multi-round model-based interactions, leverages one-round memorization-based interaction to share knowledge across different clients to build low-overhead privacy-preserving systems.

1119, TITLE:	Unsupervised Manifold Alignment with Joint Multidimensional Scaling
https://openreview.net	/forum?id=IUpjsrK.ltz4 Deviens Chen Bewen Fen, Carles Oliver, Karsten Bergwardt
HIGHLIGHT	We introduce Joint Multidimensional Scaling, a novel approach for unsupervised manifold alignment, which
maps datasets from tw	o different domains, without any known correspondences between data instances across the datasets, to a
common low-dimensi	onal Euclidean space.
1120 TITLE:	I abel Propagation with Weak Supervision
https://openreview.net	t/forum?id=aCuFa-RRotI
AUTHORS:	Rattana Pukdee, Dylan Sam, Pradeep Kumar Ravikumar, Nina Balcan
HIGHLIGHT:	In this paper, we introduce a novel analysis of the classical label propagation algorithm (LPA) (Zhu &
Ghahramani, 2002) th	at moreover takes advantage of useful prior information, specifically probabilistic hypothesized labels on the
uniabeled data.	
1121, TITLE:	Out-of-distribution Detection with Implicit Outlier Transformation
https://openreview.net	t/forum?id=hdghx6wbGuD
AUTHORS:	Qizhou Wang, Junjie Ye, Feng Liu, Quanyu Dai, Marcus Kalander, Tongliang Liu, Jianye HAO, Bo Han
a novel OF-based ann	roach that makes the model perform well for unseen OOD data, can be weaken. To address this issue, we propose
a nover OE based app	rouch that makes the model perform wen for anseen 00D statations, even for anseen 00D cases.
1122, TITLE:	Pushing the Limits of Fewshot Anomaly Detection in Industry Vision: Graphcore
https://openreview.net	/forum?id=xzmqxHdZAwO
AUTHORS: HIGHLIGHT	Guoyang Ale, Jinbao Wang, Jiaqi Liu, Yaochu Jin, Feng Zheng To push the limits, we reveal that rotation-invariant feature property has a significant impact in industrial-based.
FSAD.	To push the minus, we reveal that rotation invariant reactive property has a significant impact in industrial based
1123, TITLE:	Reproducible Bandits
AUTHORS:	/Iorum/Id=gcD2UTUGMC2 Alkis Kalavasis, Grigoris Velegkas, Hossein Esfandiari, Vahah Mirrokni, Andreas Krause, Amin Karbasi
HIGHLIGHT:	In this paper, we introduce the notion of reproducible policies in the context of stochastic bandits, one of the
canonical problems in	interactive learning.
1124, 111LE: Networks	Which Layer is Learning Faster? A Systematic Exploration of Layer-wise Convergence Rate for Deep Neural
https://openreview.net	t/forum?id=wlMDF1jQF86
AUTHORS:	Yixiong Chen, Alan Yuille, Zongwei Zhou
HIGHLIGHT:	This work demonstrates that the shallower layers of DNNs tend to converge faster than the deeper layers.
1125 TITLE:	Improving Differentiable Neural Architecture Search by Encouraging Transferability
https://openreview.net	t/forum?id=Tl8OmiibP99
AUTHORS:	Parth Sheth, Pengtao Xie
HIGHLIGHT:	Existing methods for solving this problem have a variety of limitations, such as cannot prevent the happening of
architecture degenerat	ion, being excessively restrictive in setting the number of skip connections, etc. To address these limitations, we
encouraging tri-level of	optimization framework which improves the architecture of a main model by encouraging good transferability to
an auxiliary model.	
1126 TITLE.	False in the Indextine Direct of Crank Manual ODE for Madeline Demonical Sectors
https://openreview.net	Ennancing the inductive Blases of Graph Neural ODE for Modeling Dynamical Systems t/forum?id=ATLF1_izD87
AUTHORS:	Suresh Bishnoi, Ravinder Bhattoo, Jayadeva Jayadeva, Sayan Ranu, N M Anoop Krishnan
HIGHLIGHT:	In this paper, we present a graph-based neural ODE, Gnode, to learn the time evolution of dynamical systems.
1127 TITLE.	A Differential Coornerties View and Evaluinghility of CNN on Evaluing Cranks
https://openreview.net	A Differential Geometric view and Explainability of GNN on Evolving Graphs
AUTHORS:	Yazheng Liu, Xi Zhang, Sihong Xie
HIGHLIGHT:	We propose a smooth parameterization of the GNN predicted distributions using axiomatic attribution, where
the distributions are or	n a low- dimensional manifold within a high-dimensional embedding space.
1128. TITLE:	FastFill: Efficient Compatible Model Undate
https://openreview.net	t/forum?id=rnRiiHw8Vy
AUTHORS:	Florian Jaeckle, Fartash Faghri, Ali Farhadi, Oncel Tuzel, Hadi Pouransari
HIGHLIGHT:	In this work, we introduce \method: a compatible model update process using feature alignment and policy
based partial backfillin	ng to promptly elevate retrieval performance.

1129, TITLE: Versatile Neural Processes for Learning Implicit Neural Representations https://openreview.net/forum?id=2nLeOOfAjK

AUTHORS:	Zongyu Guo, Cuiling Lan, Zhizheng Zhang, Yan Lu, Zhibo Chen
HIGHLIGHT:	In this paper, we propose an efficient NP framework dubbed Versatile Neural Processes (VNP), which largely
increases the capabili	ty of approximating functions.
1130, TITLE:	BAYES RISK CTC: CONTROLLABLE CTC ALIGNMENT IN SEQUENCE-TO-SEQUENCE TASKS
https://openreview.ne	tt/forum?id=Bd7GueaTxUz
AUTHORS:	Jinchuan Tian, Brian Yan, Jianwei Yu, CHAO WENG, Dong Yu, Shinji Watanabe
HIGHLIGHT:	Thus, the motivation of this work is to make the CTC alignment prediction controllable and thus equip CTC
with extra functionali	ties.
1131, TITLE:	Dilated convolution with learnable spacings
https://openreview.ne	tt/forum?id=Q3-1vRh3HOA
AUTHORS:	Ismail Khalfaoui Hassani, Thomas Pellegrini, Timothée Masquelier
HIGHLIGHT:	This paper presents a new method to increase the RF size without increasing the number of parameters.
1132, TITLE:	Iterative Circuit Repair Against Formal Specifications
https://openreview.ne	tt/forum?id=SEcSahl0Ql
AUTHORS:	Matthias Cosler, Frederik Schmitt, Christopher Hahn, Bernd Finkbeiner
HIGHLIGHT:	We present a deep learning approach for repairing sequential circuits against formal specifications given in
linear-time temporal	logic (LTL).
1133, TITLE:	On Pre-training Language Model for Antibody
https://openreview.net	tt/forum?id=zaq4LV55xHl
AUTHORS:	Danqing Wang, Fei YE, Hao Zhou
HIGHLIGHT:	Here, to investigate the problem, we aim to answer the following key questions: (1) How do pre-trained
language models perf	form in antibody tasks with different specificity? (2) How many benefits will the model gain if we introduce the
specific biological me	echanism to the pretraining process? (3) Do the learned antibody pre-trained representations make sense in real-
world antibody proble	ems, like drug discovery and immune process understanding? Previously, no benchmark available largely
hindered the study to	answer these questions.
1134, TITLE: https://openreview.ne AUTHORS: HIGHLIGHT: paragraph explicitly.	TempCLR: Temporal Alignment Representation with Contrastive Learning tt/forum?id=CIFOsnhZvON Yuncong Yang, Jiawei Ma, Shiyuan Huang, Long Chen, Xudong Lin, Guangxing Han, Shih-Fu Chang In this paper, we propose a contrastive learning framework TempCLR to compare the full video and the
1135, TITLE: https://openreview.ne AUTHORS: HIGHLIGHT: decomposition.	Behind the Scenes of Gradient Descent: A Trajectory Analysis via Basis Function Decomposition tt/forum?id=TPiwkltUSu Jianhao Ma, Lingjun Guo, Salar Fattahi This work analyzes the solution trajectory of gradient-based algorithms via a novel basis function
1136, TITLE: https://openreview.net AUTHORS: Mishra HIGHLIGHT: Networks (GNNs).	Gradient Gating for Deep Multi-Rate Learning on Graphs tr/forum?id=JpRExTbl1- T. Konstantin Rusch, Benjamin Paul Chamberlain, Michael W. Mahoney, Michael M. Bronstein, Siddhartha We present Gradient Gating (G\$^2\$), a novel framework for improving the performance of Graph Neural
1137, TITLE:	Light Sampling Field and BRDF Representation for Physically-based Neural Rendering
https://openreview.ne	tt/forum?id=yYEb8v65X8
AUTHORS:	Jing Yang, Hanyuan Xiao, Wenbin Teng, Yunxuan Cai, Yajie Zhao
HIGHLIGHT:	Thus, this paper proposes a novel lighting representation that models direct and indirect light locally through
light sampling strateg	ty in a learned light sampling field.
1138, TITLE:	MoDem: Accelerating Visual Model-Based Reinforcement Learning with Demonstrations
https://openreview.ne	tt/forum?id=JdTnc9gjVfJ
AUTHORS:	Nicklas Hansen, Yixin Lin, Hao Su, Xiaolong Wang, Vikash Kumar, Aravind Rajeswaran
HIGHLIGHT:	In this work, we find that leveraging just a handful of demonstrations can dramatically improve the sample-
efficiency of model-b	based RL.
1139, TITLE:	Rethinking Graph Lottery Tickets: Graph Sparsity Matters
https://openreview.ne	tt/forum?id=fjh7UGQgOB
AUTHORS:	Bo Hui, Da Yan, Xiaolong Ma, Wei-Shinn Ku
HIGHLIGHT:	Specifically, we find that the performance of a sparsified GNN degrades significantly when the graph sparsity
goes beyond a certain	a extent. Therefore, we propose two techniques to improve GNN performance when the graph sparsity is high.
1140, TITLE:	Hyperbolic Deep Reinforcement Learning
https://openreview.ne	t/forum?id=TfBHFLgv77
AUTHORS:	Edoardo Cetin, Benjamin Paul Chamberlain, Michael M. Bronstein, Jonathan J Hunt

HIGHLIGHT: Consequently, encoding features that capture the hierarchical relationships between states into the model's latent representations is often conducive to recovering effective policies. In this work, we study a new class of deep RL algorithms that promote encoding such relationships by using hyperbolic space to model latent representations.

1141. TITLE: Diffusion Policies as an Expressive Policy Class for Offline Reinforcement Learning https://openreview.net/forum?id=AHvFDPi-FA AUTHORS: Zhendong Wang, Jonathan J Hunt, Mingyuan Zhou HIGHLIGHT: In this paper, we propose representing the policy as a diffusion model, a recent class of highly-expressive deep generative models. 1142, TITLE: Trainable Weight Averaging: Efficient Training by Optimizing Historical Solutions https://openreview.net/forum?id=8wbnpOJY-f AUTHORS: Tao Li, Zhehao Huang, Qinghua Tao, Yingwen Wu, Xiaolin Huang HIGHLIGHT: In this paper, we realize that the averaging coefficients could be determined in a trainable manner and propose Trainable Weight Averaging (TWA), a novel optimization method in the reduced subspace spanned by historical solutions. 1143, TITLE: Augmented Lagrangian is Enough for Optimal Offline RL with General Function Approximation and Partial Coverage https://openreview.net/forum?id=ZsvWb6mJnMv AUTHORS: Paria Rashidinejad, Hanlin Zhu, Kunhe Yang, Stuart Russell, Jiantao Jiao HIGHLIGHT: In this paper, we leverage the marginalized importance sampling (MIS) formulation of RL and present the first set of offline RL algorithms that are statistically optimal and practical under general function approximation and single-policy concentrability, bypassing the need for uncertainty quantification. Learning to CROSS exchange to solve min-max vehicle routing problems 1144, TITLE: https://openreview.net/forum?id=ZcnzsHC10Y AUTHORS: Minjun Kim, Junyoung Park, Jinkyoo Park Inspired by CE, we propose Neuro CE (NCE), a fundamental operator of \textit{learned} meta-heuristic, to HIGHLIGHT: solve various min-max VRPs while overcoming the limitations of CE, i.e., the expensive \$\mathcal {O} (n^4)\$ search cost. 1145, TITLE: Transformers are Sample-Efficient World Models https://openreview.net/forum?id=vhFu1Acb0xb AUTHORS: Vincent Micheli, Eloi Alonso, François Fleuret HIGHLIGHT: Motivated by the success of Transformers in sequence modeling tasks, we introduce IRIS, a data-efficient agent that learns in a world model composed of a discrete autoencoder and an autoregressive Transformer. 1146, TITLE: PGrad: Learning Principal Gradients For Domain Generalization https://openreview.net/forum?id=CgCmwcfgEdH AUTHORS: Zhe Wang, Jake Grigsby, Yanjun Qi HIGHLIGHT: In this work, we develop a novel DG training strategy, we call PGrad, to learn a robust gradient direction, improving models' generalization ability on unseen domains. 1147, TITLE: On the Robustness of Safe Reinforcement Learning under Observational Perturbations https://openreview.net/forum?id=jbIYfq4Tr-AUTHORS: Zuxin Liu, Zijian Guo, Zhepeng Cen, Huan Zhang, Jie Tan, Bo Li, Ding Zhao HIGHLIGHT: We show that baseline adversarial attack techniques for standard RL tasks are not always effective for safe RL and proposed two new approaches - one maximizes the cost and the other maximizes the reward. 1148, TITLE: Generate rather than Retrieve: Large Language Models are Strong Context Generators https://openreview.net/forum?id=fB0hRu9GZUS AUTHORS: Wenhao Yu, Dan Iter, Shuohang Wang, Yichong Xu, Mingxuan Ju, Soumya Sanyal, Chenguang Zhu, Michael Zeng, Meng Jiang HIGHLIGHT: In this paper, we present a novel perspective for solving knowledge-intensive tasks by replacing document retrievers with large language model generators. Protein Sequence and Structure Co-Design with Equivariant Translation 1149, TITLE: https://openreview.net/forum?id=pRCMXcfdihq Chence Shi, Chuanrui Wang, Jiarui Lu, Bozitao Zhong, Jian Tang AUTHORS: HIGHLIGHT: In this paper, we propose a new approach capable of protein sequence and structure co-design, which iteratively translates both protein sequence and structure into the desired state from random initialization, based on context features given a priori. Hybrid RL: Using both offline and online data can make RL efficient 1150, TITLE: https://openreview.net/forum?id=yyBis80iUuU Yuda Song, Yifei Zhou, Ayush Sekhari, Drew Bagnell, Akshay Krishnamurthy, Wen Sun AUTHORS: HIGHLIGHT: We consider a hybrid reinforcement learning setting (Hybrid RL), in which an agent has access to an offline dataset and the ability to collect experience via real-world online interaction.

1151, TITLE: IS SYNTHETIC DATA FROM GENERATIVE MODELS READY FOR IMAGE RECOGNITION? https://openreview.net/forum?id=nUmCcZ5RKF
AUTHORS: Ruifei He, Shuyang Sun, Xin Yu, Chuhui Xue, Wenqing Zhang, Philip Torr, Song Bai, XIAOJUAN QI HIGHLIGHT: In this work, we extensively study whether and how synthetic images generated from state-of-the-art text-toimage generation models can be used for image recognition tasks, and focus on two perspectives: synthetic data for improving classification models in the data-scare settings (i.e. zero-shot and few-shot), and synthetic data for large-scale model pre-training for transfer learning.

Concept Gradient: Concept-based Interpretation Without Linear Assumption 1152, TITLE: https://openreview.net/forum?id= 01dDd3f78 AUTHORS: Andrew Bai, Chih-Kuan Yeh, Pradeep Kumar Ravikumar, Neil Y.C. Lin, Cho-Jui Hsieh HIGHLIGHT: In this work we proposed Concept Gradient (CG), which extends concept-based, gradient interpretation methods to non-linear concept functions. Generalize Learned Heuristics to Solve Large-scale Vehicle Routing Problems in Real-time 1153. TITLE: https://openreview.net/forum?id=6ZajpxqTlQ AUTHORS: Qingchun Hou, Jingwei Yang, Yiqiang Su, Xiaoqing Wang, Yuming Deng We contribute in the three directions: We propose a Two-stage Divide Method (TAM) to generate sub-route HIGHLIGHT: sequence rather than node sequence for generalizing the heuristics learned on small-scale-VRPs to solve large-scale VRPs in realtime. 1154. TITLE: Re-parameterizing Your Optimizers rather than Architectures https://openreview.net/forum?id=B92TMCG 7rp AUTHORS: Xiaohan Ding, Honghao Chen, Xiangyu Zhang, Kaiqi Huang, Jungong Han, Guiguang Ding HIGHLIGHT: In this paper, we propose to incorporate model-specific prior knowledge into optimizers by modifying the gradients according to a set of model-specific hyper-parameters. 1155, TITLE: Provable Unsupervised Data Sharing for Offline Reinforcement Learning https://openreview.net/forum?id=MTTPLcwvqTt AUTHORS: Hao Hu, Yiqin Yang, Qianchuan Zhao, Chongjie Zhang In this paper, we examine the theoretical benefit of unlabeled data in the context of linear MDPs and propose a HIGHLIGHT: novel and Provable Data Sharing algorithm, which we refer to as PDS, to utilize such unlabeled data for offline RL. 1156, TITLE: Diffusion Posterior Sampling for General Noisy Inverse Problems https://openreview.net/forum?id=OnD9zGAGT0k AUTHORS: Hyungjin Chung, Jeongsol Kim, Michael Thompson Mccann, Marc Louis Klasky, Jong Chul Ye HIGHLIGHT: In this work, we extend diffusion solvers to efficiently handle general noisy (non)linear inverse problems via the Laplace approximation of the posterior sampling. 1157, TITLE: Recitation-Augmented Language Models https://openreview.net/forum?id=-cqvvvb-NkI AUTHORS: Zhiqing Sun, Xuezhi Wang, Yi Tay, Yiming Yang, Denny Zhou HIGHLIGHT: We propose a new paradigm to help Large Language Models (LLMs) generate more accurate factual knowledge without retrieving from an external corpus, called RECITation-augmented gEneration (RECITE). Masked Unsupervised Self-training for Label-free Image Classification 1158, TITLE: https://openreview.net/forum?id=ZAKkiVxiAM9 AUTHORS: Junnan Li, Silvio Savarese, Steven Hoi In this paper, we aim to leverage the abundant unlabeled data from a target domain to improve the performance HIGHLIGHT: of a pre-trained zero-shot classifier, by unsupervised finetuning of the pre-trained model. 1159, TITLE: Statistical Efficiency of Score Matching: The View from Isoperimetry https://openreview.net/forum?id=TD7AnQjNzR6 AUTHORS: Frederic Koehler, Alexander Heckett, Andrej Risteski HIGHLIGHT: Though this estimator is known to be consistent, its unclear whether (and when) its statistical efficiency is comparable to that of maximum likelihood --- which is known to be (asymptotically) optimal. We initiate this line of inquiry in this paper, and show a tight connection between statistical efficiency of score matching and the isoperimetric properties of the distribution being estimated --- i.e. the Poincarlye, log-Sobolev and isoperimetric constant --- quantities which govern the mixing time of Markov processes like Langevin dynamics. 1160, TITLE: Understanding weight-magnitude hyperparameters in training binary networks https://openreview.net/forum?id=uBKBoix9NXa AUTHORS: Joris Quist, Yunqiang Li, Jan van Gemert HIGHLIGHT: The magnitude is interpretable for real-valued weights, but loses its meaning for binary weights. In this paper we offer a new interpretation of these magnitude-based hyperparameters based on higher-order gradient filtering during network optimization. 1161, TITLE: Efficient Planning in a Compact Latent Action Space https://openreview.net/forum?id=cA77NrVEugn AUTHORS: zhengyao jiang, Tianjun Zhang, Michael Janner, Yueying Li, Tim Rocktäschel, Edward Grefenstette, Yuandong Tian

HIGHLIGHT: To advance efficient planning for high-dimensional continuous control, we propose Trajectory Autoencoding Planner (TAP), which learns low-dimensional latent action codes with a state-conditional VQ-VAE. 1162, TITLE: Leveraging Importance Weights in Subset Selection https://openreview.net/forum?id=9Nj gNdvqYf AUTHORS: Gui Citovsky, Giulia DeSalvo, Sanjiv Kumar, Srikumar Ramalingam, Afshin Rostamizadeh, Yunjuan Wang HIGHLIGHT: We present a subset selection algorithm designed to work with arbitrary model families in a practical batch setting. 1163, TITLE: Why adversarial training can hurt robust accuracy https://openreview.net/forum?id=-CA8yFkPc7O AUTHORS: Jacob Clarysse, Julia Hörrmann, Fanny Yang HIGHLIGHT: In this paper, we demonstrate that, surprisingly, the opposite can be true for a natural class of perceptible perturbations --- even though adversarial training helps when enough data is available, it may in fact hurt robust generalization in the small sample size regime. 1164, TITLE: Sequential Learning of Neural Networks for Prequential MDL https://openreview.net/forum?id=dMMPUvNSYJr AUTHORS: Jorg Bornschein, Yazhe Li, Marcus Hutter HIGHLIGHT: In this study, we evaluate approaches for computing prequential description lengths for image classification datasets with neural networks. 1165, TITLE: Learning Probabilistic Topological Representations Using Discrete Morse Theory https://openreview.net/forum?id=cXMHOD-xOas AUTHORS: Xiaoling Hu, Dimitris Samaras, Chao Chen HIGHLIGHT: In this paper, we propose the first deep learning based method to learn topological/structural representations. 1166, TITLE: Multi-objective optimization via equivariant deep hypervolume approximation https://openreview.net/forum?id=fSa5IjNMmmi Jim Boelrijk, Bernd Ensing, Patrick Forré AUTHORS: HIGHLIGHT: However, the computational complexity for the calculation of the hypervolume scales unfavorably with increasing number of objectives and data points, which restricts its use in those common multi-objective optimization frameworks. To overcome these restrictions we propose to approximate the hypervolume function with a deep neural network, which we call DeepHV. 1167, TITLE: DiffusER: Diffusion via Edit-based Reconstruction https://openreview.net/forum?id=nG9RF9z1yy3 AUTHORS: Machel Reid, Vincent Josua Hellendoorn, Graham Neubig We propose a generally applicable text generative model which takes inspiration from diffusion models and HIGHLIGHT: parameterises generation steps as text editing steps without compromising performance and adding flexibility. 1168, TITLE: Provable Sim-to-real Transfer in Continuous Domain with Partial Observations https://openreview.net/forum?id=S31oTB72m0G AUTHORS: Jiachen Hu, Han Zhong, Chi Jin, Liwei Wang HIGHLIGHT: In this paper, we study the sim-to-real transfer in continuous domain with partial observations, where the simulated environments and real-world environments are modeled by linear quadratic Gaussian (LQG) systems. 1169. TITLE: DynaMS: Dyanmic Margin Selection for Efficient Deep Learning https://openreview.net/forum?id=7oPAgqxNb20 AUTHORS: Jiaxing Wang, Yong Li, Jingwei Zhuo, Xupeng Shi, WEIZHONG ZHANG, Lixing Gong, Tong Tao, Pengzhang Liu, Yongjun Bao, Weipeng Yan HIGHLIGHT: In this paper, we propose dynamic margin selection (DynaMS). TANGOS: Regularizing Tabular Neural Networks through Gradient Orthogonalization and Specialization 1170. TITLE: https://openreview.net/forum?id=n6H86gW8u0d AUTHORS: Alan Jeffares, Tennison Liu, Jonathan Crabbé, Fergus Imrie, Mihaela van der Schaar HIGHLIGHT: In this work, we introduce Tabular Neural Gradient Orthogonalization and Specialization (TANGOS), a novel framework for regularization in the tabular setting built on latent unit attributions. Model-based Causal Bayesian Optimization 1171, TITLE: https://openreview.net/forum?id=Vk-34OQ7rFo AUTHORS: Scott Sussex, Anastasia Makarova, Andreas Krause HIGHLIGHT: We propose the {\em model-based causal Bayesian optimization algorithm (MCBO)} that learns a full system model instead of only modeling intervention-reward pairs. 1172, TITLE: On Explaining Neural Network Robustness with Activation Path https://openreview.net/forum?id=piIsx-G3Gux AUTHORS: Ziping Jiang HIGHLIGHT: This work investigates the robustness of neural networks from the activation pattern perspective. 1173, TITLE: Structure by Architecture: Structured Representations without Regularization

https://openreview.net/forum?id=O lFCPaF48t AUTHORS: Felix Leeb, Giulia Lanzillotta, Yashas Annadani, Michel Besserve, Stefan Bauer, Bernhard Schölkopf HIGHLIGHT: Unlike most methods which rely on matching an arbitrary, relatively unstructured, prior distribution for sampling, we propose a sampling technique that relies solely on the independence of latent variables, thereby avoiding the trade-off between reconstruction quality and generative performance typically observed in VAEs. Compressing multidimensional weather and climate data into neural networks 1174, TITLE: https://openreview.net/forum?id=Y5SEe3dfniJ AUTHORS: Langwen Huang, Torsten Hoefler HIGHLIGHT: Weather and climate simulations produce petabytes of high-resolution data that are later analyzed by researchers in order to understand climate change or severe weather. We propose a new method of compressing this multidimensional weather and climate data: a coordinate-based neural network is trained to overfit the data, and the resulting parameters are taken as a compact representation of the original grid-based data. 1175, TITLE: Timing is Everything: Learning to Act Selectively with Costly Actions and Budgetary Constraints https://openreview.net/forum?id= BoPed4tYww AUTHORS: David Henry Mguni, Aivar Sootla, Juliusz Krzysztof Ziomek, Oliver Slumbers, Zipeng Dai, Kun Shao, Jun Wang HIGHLIGHT: In this paper, we intro- duce a reinforcement learning (RL) framework named Learnable Impulse Control Reinforcement Algorithm (LICRA), for learning to optimally select both when to act and which actions to take when actions incur costs. DECAP: Decoding CLIP Latents for Zero-shot Captioning 1176, TITLE: https://openreview.net/forum?id=Lt8bMlhiwx2 AUTHORS: Wei Li, Linchao Zhu, Longyin Wen, Yi Yang HIGHLIGHT: In this work, we propose a simple framework, named DeCap, for zero-shot captioning. 1177, TITLE: Robust Explanation Constraints for Neural Networks https://openreview.net/forum?id= hHYaKu0jcj AUTHORS: Matthew Robert Wicker, Juveon Heo, Luca Costabello, Adrian Weller HIGHLIGHT: Relying on constraint relaxation techniques from non-convex optimization, we develop a method that upperbounds the largest change an adversary can make to a gradient-based explanation via bounded manipulation of either the input features or model parameters. 1178, TITLE: Offline Reinforcement Learning via High-Fidelity Generative Behavior Modeling https://openreview.net/forum?id=42zs3ga2kpy AUTHORS: Huayu Chen, Cheng Lu, Chengyang Ying, Hang Su, Jun Zhu HIGHLIGHT: In this work, we show that due to the limited distributional expressivity of policy models, previous methods might still select unseen actions during training, which deviates from their initial motivation. 1179, TITLE: Optimizing Spca-based Continual Learning: A Theoretical Approach https://openreview.net/forum?id=Vf6WcUDnY7c AUTHORS: Chunchun Yang, Malik Tiomoko, Zengfu Wang HIGHLIGHT: This paper proposes a theoretical analysis of a simple but efficient continual learning algorithm. 1180, TITLE: Value Memory Graph: A Graph-Structured World Model for Offline Reinforcement Learning https://openreview.net/forum?id=UYcIheNY9Pf AUTHORS: Deyao Zhu, Li Erran Li, Mohamed Elhoseiny HIGHLIGHT: Focusing on the offline RL setting, we aim to build a simple and discrete world model that abstracts the original environment. 1181, TITLE: Domain Generalisation via Domain Adaptation: An Adversarial Fourier Amplitude Approach https://openreview.net/forum?id=7IG0wsTND7w AUTHORS: Minyoung Kim, Da Li, Timothy Hospedales HIGHLIGHT: We tackle the domain generalisation (DG) problem by posing it as a domain adaptation (DA) task where we adversarially synthesise the worst-case 'target' domain and adapt a model to that worst-case domain, thereby improving the model's robustness. 1182, TITLE: Statistical Theory of Differentially Private Marginal-based Data Synthesis Algorithms https://openreview.net/forum?id=hxUwnEGxW87 AUTHORS: Ximing Li, Chendi Wang, Guang Cheng HIGHLIGHT: In this paper, we study DP data synthesis algorithms based on Bayesian networks (BN) from a statistical perspective. Online Low Rank Matrix Completion 1183, TITLE: https://openreview.net/forum?id=47KG AvNqeZ AUTHORS: Soumyabrata Pal, Prateek Jain HIGHLIGHT: Here, we propose OCTAL (Online Collaborative filTering using iterAtive user cLustering) that guarantees nearly optimal regret of $O(\text{mathsf}(N)) \times \{T_{1/2})\$.

1184, TITLE: A Primal-Dual Framework for Transformers and Neural Networks https://openreview.net/forum?id=U T8-5hClV AUTHORS: Tan Minh Nguyen, Tam Minh Nguyen, Nhat Ho, Andrea L. Bertozzi, Richard Baraniuk, Stanley Osher HIGHLIGHT: We show that the self-attention corresponds to the support vector expansion derived from a support vector regression problem and provide a principled framework for constructing new attention mechanisms from popular neural network layers. 1185, TITLE: Scaffolding a Student to Instill Knowledge https://openreview.net/forum?id=N4K5ck-BTT AUTHORS: Anil Kag, Durmus Alp Emre Acar, Aditya Gangrade, Venkatesh Saligrama HIGHLIGHT: We propose a novel knowledge distillation (KD) method to selectively instill teacher knowledge into a student model motivated by situations where the student's capacity is significantly smaller than that of the teachers. Understanding the Generalization of Adam in Learning Neural Networks with Proper Regularization 1186, TITLE: https://openreview.net/forum?id=iUYpN14qjTF AUTHORS: Difan Zou, Yuan Cao, Yuanzhi Li, Quanquan Gu HIGHLIGHT: However, it has been observed in many deep learning applications such as image classification, Adam can converge to a different solution with a worse test error compared to (stochastic) gradient descent, even with a fine-tuned regularization. In this paper, we provide a theoretical explanation for this phenomenon: we show that in the nonconvex setting of learning over-parameterized two-layer convolutional neural networks starting from the same random initialization, for a class of data distributions (inspired from image data), Adam and gradient descent (GD) can converge to different global solutions of the training objective with provably different generalization errors, even with weight decay regularization. 1187. TITLE: Learning ReLU networks to high uniform accuracy is intractable https://openreview.net/forum?id=nchvKfvNeX0 AUTHORS: Julius Berner, Philipp Grohs, Felix Voigtlaender In this paper we precisely quantify the number of training samples needed for any conceivable training HIGHLIGHT: algorithm to guarantee a given uniform accuracy on any learning problem formulated over target classes containing (or consisting of) ReLU neural networks of a prescribed architecture. 1188, TITLE: How Sharpness-Aware Minimization Minimizes Sharpness? https://openreview.net/forum?id=5spDgWmpY6x AUTHORS: Kaiyue Wen, Tengyu Ma, Zhiyuan Li HIGHLIGHT: We prove the implicit bias of Sharpness-Aware Minimization (SAM) is minimizing the top eigenvalue of Hessian in the full-batch setting or minimizing the trace of Hessian when batch size is 1. The Implicit Bias of Minima Stability in Multivariate Shallow ReLU Networks 1189, TITLE: https://openreview.net/forum?id=xtbog7cfsr AUTHORS: Mor Shpigel Nacson, Rotem Mulayoff, Greg Ongie, Tomer Michaeli, Daniel Soudry HIGHLIGHT: We study the type of solutions to which stochastic gradient descent converges when used to train a single hidden-layer multivariate ReLU network with the quadratic loss. 1190, TITLE: MAESTRO: Open-Ended Environment Design for Multi-Agent Reinforcement Learning https://openreview.net/forum?id=sKWlRDzPfd7 Mikayel Samvelyan, Akbir Khan, Michael D Dennis, Minqi Jiang, Jack Parker-Holder, Jakob Nicolaus AUTHORS: Foerster, Roberta Raileanu, Tim Rocktäschel It is thus crucial to consider the dependency between the environment and co-player when shaping a curriculum HIGHLIGHT: in multi-agent domains. In this work, we use this insight and extend Unsupervised Environment Design (UED) to multi-agent environments. Leveraging Unlabeled Data to Track Memorization 1191, TITLE: https://openreview.net/forum?id=ORp91sAbzI AUTHORS: Mahsa Forouzesh, Hanie Sedghi, Patrick Thiran HIGHLIGHT: It is therefore important to track and evaluate the robustness of models against noisy label memorization. We propose a metric, called \$\textit {susceptibility}\$, to gauge such memorization for neural networks. 1192. TITLE: Learning Vortex Dynamics for Fluid Inference and Prediction https://openreview.net/forum?id=nYWqxUwFc3x AUTHORS: Yitong Deng, Hong-Xing Yu, Jiajun Wu, Bo Zhu We propose a novel machine learning method based on differentiable vortex particles to infer and predict fluid HIGHLIGHT: dynamics from a single video. 1193, TITLE: Quality-Similar Diversity via Population Based Reinforcement Learning https://openreview.net/forum?id=bLmSMXbqXr Shuang Wu, Jian Yao, Haobo Fu, Ye Tian, Chao Qian, Yaodong Yang, QIANG FU, Yang Wei AUTHORS: In this work, we present the quality-similar diversity problem that features diversity among policies of similar HIGHLIGHT: qualities. To derive the gradient of the user-specified diversity with respect to a policy, which is not trivially available, we introduce a set of BD estimators and connect it with the classical policy gradient theorem.

1194, TITLE: Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow

https://openreview.net/forum?id=XVjTT1nw5z

AUTHORS: Xingchao Liu, Chengyue Gong, qiang liu

HIGHLIGHT: We present rectified flow, a simple approach to learning (neural) ordinary differential equation (ODE) models to transport between two empirically observed distributions \$\pi_0\$ and \$\pi_1\$, hence providing a unified solution to generative modeling and domain transfer, among various other tasks involving distribution transport.

1195, TITLE: Learning Diffusion Bridges on Constrained Domains

https://openreview.net/forum?id=WH1yCa0TbB

AUTHORS: Xingchao Liu, Lemeng Wu, Mao Ye, qiang liu

HIGHLIGHT:However, because diffusion processes are most naturally applied on the unconstrained Euclidean space
 \mathbf{R}^{0} , key challenges arise for developing diffusion based models for learning data on constrained and structured domains.
We present a simple and unified framework to achieve this that can be easily adopted to various types of domains, including product
spaces of any type (be it bounded/unbounded, continuous/discrete, categorical/ordinal, or their mix).

1196, TITLE: Indiscriminate Poisoning Attacks on Unsupervised Contrastive Learning

https://openreview.net/forum?id=f0a_dWEYg-Td

AUTHORS: Hao He, Kaiwen Zha, Dina Katabi

HIGHLIGHT: This paper is the first to consider indiscriminate poisoning attacks of contrastive learning. We propose contrastive poisoning, the first effective such attack on CL.

1197, TITLE: Decompositional Generation Process for Instance-Dependent Partial Label Learning

https://openreview.net/forum?id=lKOfilXucGB

AUTHORS: Congyu Qiao, Ning Xu, Xin Geng

HIGHLIGHT: In this paper, we consider instance-dependent PLL and assume that the generation process of the candidate labels could decompose into two sequential parts, where the correct label emerges first in the mind of the annotator but then the incorrect labels related to the feature are also selected with the correct label as candidate labels due to uncertainty of labeling.

1198, TITLE: FLIP: A Provable Defense Framework for Backdoor Mitigation in Federated Learning https://openreview.net/forum?id=Xo2E217 M4n

AUTHORS: Kaiyuan Zhang, Guanhong Tao, Qiuling Xu, Siyuan Cheng, Shengwei An, Yingqi Liu, Shiwei Feng, Guangyu Shen, Pin-Yu Chen, Shiqing Ma, Xiangyu Zhang

HIGHLIGHT: Existing work on robust aggregation and certified FL robustness does not study how hardening benign clients can affect the global model (and the malicious clients). In this work, we theoretically analyze the connection among cross-entropy loss, attack success rate, and clean accuracy in this setting.

1199, TITLE: Minimum Variance Unbiased N:M Sparsity for the Neural Gradients

https://openreview.net/forum?id=vuD2xEtxZcj

AUTHORS: Brian Chmiel, Itay Hubara, Ron Banner, Daniel Soudry

HIGHLIGHT: To this end, we first establish a tensor-level optimality criteria. Previous works aimed to minimize the meansquare-error (MSE) of each pruned block. We show that while minimization of the MSE works fine for pruning the weights and activations, it catastrophically fails for the neural gradients. Instead, we show that accurate pruning of the neural gradients requires an unbiased minimum-variance pruning mask.

1200, TITLE: Incremental Learning of Structured Memory via Closed-Loop Transcription

https://openreview.net/forum?id=XrgjF5-M3xi

AUTHORS: Shengbang Tong, Xili Dai, Ziyang Wu, Mingyang Li, Brent Yi, Yi Ma

HIGHLIGHT: This work proposes a minimal computational model for learning structured memories of multiple object classes in an incremental setting.

1201, TITLE: Knowledge-in-Context: Towards Knowledgeable Semi-Parametric Language Models https://openreview.net/forum?id=a2jNdqE2102

 AUTHORS:
 Xiaoman Pan, Wenlin Yao, Hongming Zhang, Dian Yu, Dong Yu, Jianshu Chen

 HIGHLIGHT:
 In this paper, we develop a novel semi-parametric language model architecture, Knowledge-in-Context (KiC),

which empowers a parametric text-to-text language model with a knowledge-rich external memory.

1202, TITLE: NewModel: Improving DeBERTa using ELECTRA-Style Pre-Training with Gradient-Disentangled Embedding Sharing

https://openreview.net/forum?id=sE7-XhLxHA

AUTHORS: Pengcheng He, Jianfeng Gao, Weizhu Chen

HIGHLIGHT: This paper presents a new pre-trained language model, NewModel, which improves the original DeBERTa model by replacing mask language modeling (MLM) with replaced token detection (RTD), a more sample-efficient pre-training task.

1203, TITLE: Pushing the Accuracy-Fairness Tradeoff Frontier with Introspective Self-play

https://openreview.net/forum?id=MofT9KEF0kw

AUTHORS: Jeremiah Zhe Liu, Krishnamurthy Dj Dvijotham, Jihyeon Lee, Quan Yuan, Balaji Lakshminarayanan, Deepak Ramachandran

HIGHLIGHT: In this work, we propose \$Introspective Self-play\$ (ISP), a simple approach to improve the uncertainty estimation of a deep neural network under dataset bias, by adding an auxiliary \$introspection\$ task requiring a model to predict the bias for each data point in addition to the label.

1204, TITLE: Asymptotic Instance-Optimal Algorithms for Interactive Decision Making https://openreview.net/forum?id=oGVu9spZaJJ AUTHORS: Kefan Dong, Tengyu Ma HIGHLIGHT: In this paper, we design the first asymptotic instance-optimal algorithm for general interactive decision making problems with finite number of decisions under mild conditions. The hidden uniform cluster prior in self-supervised learning 1205, TITLE: https://openreview.net/forum?id=04K3PMtMckp AUTHORS: Mido Assran, Randall Balestriero, Quentin Duval, Florian Bordes, Ishan Misra, Piotr Bojanowski, Pascal Vincent, Michael Rabbat, Nicolas Ballas HIGHLIGHT: By moving away from conventional uniformity priors and instead preferring power-law distributed feature clusters, we show that one can improve the quality of the learned representations on real-world class-imbalanced datasets. Globally Optimal Training of Neural Networks with Threshold Activation Functions 1206, TITLE: https://openreview.net/forum?id= 9k5kTgyHT AUTHORS: Tolga Ergen, Halil Ibrahim Gulluk, Jonathan Lacotte, Mert Pilanci HIGHLIGHT: To this end, we study weight decay regularized training problems of deep neural networks with threshold activations. 1207. TITLE: Mosaic Representation Learning for Self-supervised Visual Pre-training https://openreview.net/forum?id=JAezPMehaUu AUTHORS: Zhaoqing Wang, Ziyu Chen, Yaqian Li, Yandong Guo, Jun Yu, Mingming Gong, Tongliang Liu HIGHLIGHT: However, it overlooks the diverse contextual backgrounds, which reduces the variance of the input views and degenerates the performance. To address this problem, we propose a mosaic representation learning framework (MosRep), consisting of a new data augmentation strategy that enriches the backgrounds of each small crop and improves the quality of visual representations. 1208, TITLE: FluidLab: A Differentiable Environment for Benchmarking Complex Fluid Manipulation https://openreview.net/forum?id=Cp-io BoFaE AUTHORS: Zhou Xian, Bo Zhu, Zhenjia Xu, Hsiao-Yu Tung, Antonio Torralba, Katerina Fragkiadaki, Chuang Gan HIGHLIGHT: In this work, we introduce FluidLab, a simulation environment with a diverse set of manipulation tasks involving complex fluid dynamics. Task Ambiguity in Humans and Language Models 1209, TITLE: https://openreview.net/forum?id=QrnDe 9ZFd8 AUTHORS: Alex Tamkin, Kunal Handa, Avash Shrestha, Noah Goodman HIGHLIGHT We motivate the direction of studying task ambiguity in humans and language models, evaluating them on a new benchmark of ambiguously-specified tasks and develop methods for improving performance. 1210, TITLE: Preference Transformer: Modeling Human Preferences using Transformers for RL https://openreview.net/forum?id=Peot1SFDX0 AUTHORS: Changyeon Kim, Jongjin Park, Jinwoo Shin, Honglak Lee, Pieter Abbeel, Kimin Lee HIGHLIGHT: In this paper, we present Preference Transformer, a neural architecture that models human preferences using transformers. 1211, TITLE: Molecule Generation For Target Protein Binding with Structural Motifs https://openreview.net/forum?id=Rq13idF0F73 AUTHORS: ZAIXI ZHANG, Qi Liu, Shuxin Zheng, Yaosen Min HIGHLIGHT: Although deep generative models and geometric deep learning have made great progress in drug design, existing works either sample in the 2D graph space or fail to generate valid molecules with realistic substructures. To tackle these problems, we propose a Fragment-based LigAnd Generation framework (FLAG), to generate 3D molecules with valid and realistic substructures fragment-by-fragment. 1212, TITLE: Deep Generative Modeling on Limited Data with Regularization by Nontransferable Pre-trained Models https://openreview.net/forum?id=M9u ctqFUlg AUTHORS: Yong Zhong, Hong Tao Liu, Xiaodong Liu, Fan Bao, Weiran Shen, Chongxuan Li HIGHLIGHT: Inspired by the classical perspective of the bias-variance tradeoff, we propose regularized deep generative model (Reg-DGM), which leverages a nontransferable pre-trained model to reduce the variance of generative modeling with limited data. 1213, TITLE: Can CNNs Be More Robust Than Transformers? https://openreview.net/forum?id=TKIFuQHHECj AUTHORS: Zeyu Wang, Yutong Bai, Yuyin Zhou, Cihang Xie HIGHLIGHT: In this paper, we question that belief by closely examining the design of Transformers. Risk-Aware Reinforcement Learning with Coherent Risk Measures and Non-linear Function Approximation 1214, TITLE: https://openreview.net/forum?id=-RwZOVybbj AUTHORS: Thanh Lam, Arun Verma, Bryan Kian Hsiang Low, Patrick Jaillet

HIGHLIGHT: We study the risk-aware reinforcement learning (RL) problem in the episodic finite-horizon Markov decision process with unknown transition and reward functions.

1215, TITLE: Bi-level Physics-Informed Neural Networks for PDE Constrained Optimization using Broyden's Hypergradients https://openreview.net/forum?id=kkpL4zUXtiw AUTHORS: Zhongkai Hao, Chengyang Ying, Hang Su, Jun Zhu, Jian Song, Ze Cheng HIGHLIGHT: In this paper, we present a novel bi-level optimization framework to resolve the challenge by decoupling the optimization of the targets and constraints. Analog Bits: Generating Discrete Data using Diffusion Models with Self-Conditioning 1216, TITLE: https://openreview.net/forum?id=3itjR9QxFw AUTHORS: Ting Chen, Ruixiang ZHANG, Geoffrey Hinton HIGHLIGHT: We present Bit Diffusion: a simple and generic approach for generating discrete data with continuous diffusion models. 1217, TITLE: Understanding Edge-of-Stability Training Dynamics with a Minimalist Example https://openreview.net/forum?id=p7EagBsMAEO AUTHORS: Xingyu Zhu, Zixuan Wang, Xiang Wang, Mo Zhou, Rong Ge HIGHLIGHT: In this paper, we study EoS phenomenon by constructing a simple function that has the same behavior. 1218, TITLE: Learning Proximal Operators to Discover Multiple Optima https://openreview.net/forum?id=PzBGIu-llo7 AUTHORS: Lingxiao Li, Noam Aigerman, Vladimir Kim, Jiajin Li, Kristjan Greenewald, Mikhail Yurochkin, Justin Solomon HIGHLIGHT: We present an end-to-end method to learn the proximal operator of a family of training problems so that multiple local minima can be quickly obtained from initial guesses by iterating the learned operator, emulating the proximal-point algorithm that has fast convergence. We further present an exhaustive benchmark for multi-solution optimization to demonstrate the effectiveness of our method. 1219, TITLE: Neural Radiance Field Codebooks https://openreview.net/forum?id=mX56bKDybu5 AUTHORS: Matthew Wallingford, Aditya Kusupati, Alex Fang, Vivek Ramanujan, Aniruddha Kembhavi, Roozbeh Mottaghi, Ali Farhadi HIGHLIGHT: Learning such representations for complex scenes and tasks remains an open challenge. Towards this goal, we introduce Neural Radiance Field Codebooks (NRC), a scalable method for learning object-centric representations through novel view reconstruction. FiT: Parameter Efficient Few-shot Transfer Learning for Personalized and Federated Image Classification 1220, TITLE: https://openreview.net/forum?id=9aokcgBVIj1 AUTHORS: Aliaksandra Shysheya, John F Bronskill, Massimiliano Patacchiola, Sebastian Nowozin, Richard E Turner HIGHLIGHT: Modern deep learning systems are increasingly deployed in situations such as personalization and federated learning where it is necessary to support i) learning on small amounts of data, and ii) communication efficient distributed training protocols. In this work, we develop FiLM Transfer (FiT) which fulfills these requirements in the image classification setting by combining ideas from transfer learning (fixed pretrained backbones and fine-tuned FiLM adapter layers) and meta-learning (automatically configured Naive Bayes classifiers and episodic training) to yield parameter efficient models with superior classification accuracy at low-shot. 1221, TITLE: AANG : Automating Auxiliary Learning https://openreview.net/forum?id=vtVDI3w BLL AUTHORS: Lucio M. Dery, Paul Michel, Mikhail Khodak, Graham Neubig, Ameet Talwalkar HIGHLIGHT: In this work, we present an approach for automatically generating a suite of auxiliary objectives. 1222, TITLE: NeRF-SOS: Any-View Self-supervised Object Segmentation on Complex Scenes https://openreview.net/forum?id=kfOtMqYJlUU AUTHORS: Zhiwen Fan, Peihao Wang, Yifan Jiang, Xinyu Gong, Dejia Xu, Zhangyang Wang HIGHLIGHT: We propose a novel collaborative contrastive loss for NeRF to segment objects in complex real-world scenes, without any annotation. 1223, TITLE: Private Federated Learning Without a Trusted Server: Optimal Algorithms for Convex Losses https://openreview.net/forum?id=TVY6GoURrw AUTHORS: Andrew Lowy, Meisam Razaviyayn HIGHLIGHT: In this work, we provide tight (up to logarithms) upper and lower bounds for ISRL-DP FL with convex/strongly convex loss functions and homogeneous (i.i.d.) silo data. 1224, TITLE: Cheap Talk Discovery and Utilization in Multi-Agent Reinforcement Learning https://openreview.net/forum?id=cddbeL1HWaD AUTHORS: Yat Long Lo, Christian Schroeder de Witt, Samuel Sokota, Jakob Nicolaus Foerster, Shimon Whiteson HIGHLIGHT: Current methods require these channels to be constantly accessible and known to the agents a priori. In this work, we lift these requirements such that the agents must discover the cheap talk channels and learn how to use them.

1225, TITLE: Analyzing Tree Architectures in Ensembles via Neural Tangent Kernel https://openreview.net/forum?id=V_06QV-kZX

AUTHORS: Ryuichi Kanoh, Mahito Sugiyama HIGHLIGHT: In this paper, we formulate and analyze the Neural Tangent Kernel (NTK) induced by soft tree ensembles for arbitrary tree architectures. 1226, TITLE: Evaluating Long-Term Memory in 3D Mazes https://openreview.net/forum?id=yHLvIIE9RGN AUTHORS: Jurgis Pašukonis, Timothy P Lillicrap, Danijar Hafner HIGHLIGHT: In this paper, we introduce the Memory Maze, a 3D domain of randomized mazes specifically designed for evaluating long-term memory in agents. With Memory Maze, we propose an online reinforcement learning benchmark, a diverse offline dataset, and an offline probing evaluation. 1227, TITLE: Proactive Multi-Camera Collaboration for 3D Human Pose Estimation https://openreview.net/forum?id=CPIy9TWFYBG Hai Ci, Mickel Liu, Xuehai Pan, fangwei zhong, Yizhou Wang AUTHORS: HIGHLIGHT: This work introduces a multi-agent reinforcement learning (MARL) scheme to proactive Multi-Camera Collaboration for 3D Human Pose Estimation (MCC-HPE) in dynamic human crowds. 1228. TITLE: Become a Proficient Player with Limited Data through Watching Pure Videos https://openreview.net/forum?id=Sy-o2N0hF4f AUTHORS: Weirui Ye, Yunsheng Zhang, Pieter Abbeel, Yang Gao HIGHLIGHT: We introduce a two-phase training pipeline as follows: for the pre-training phase, we implicitly extract the hidden action embedding from videos and pre-train the visual representation and the environment dynamics network through a novel cycle consistency objective based on vector quantization; for down-stream tasks, we finetune with small amount of task data based on the learned models. 1229, TITLE: Hidden Markov Transformer for Simultaneous Machine Translation https://openreview.net/forum?id=9y0HFvaAYD6 Shaolei Zhang, Yang Feng AUTHORS: HIGHLIGHT: In this paper, we propose Hidden Markov Transformer (HMT), which treats the moments of starting translating as hidden events and the target sequence as the corresponding observed events, thereby organizing them as a hidden Markov model. 1230, TITLE: Rank Preserving Framework for Asymmetric Image Retrieval https://openreview.net/forum?id=dYHYXZ3uGdQ AUTHORS: Hui Wu, Min Wang, Wengang Zhou, Houqiang Li HIGHLIGHT: Considering that the primary concern of the users is the rank of the returned images, we propose a generic rank preserving framework, which achieves feature compatibility and the order consistency between query and gallery models simultaneously. Mega: Moving Average Equipped Gated Attention 1231, TITLE: https://openreview.net/forum?id=qNLe3iq2E1 AUTHORS: Xuezhe Ma, Chunting Zhou, Xiang Kong, Junxian He, Liangke Gui, Graham Neubig, Jonathan May, Luke Zettlemoyer HIGHLIGHT: In this paper, we introduce Mega, a simple, theoretically grounded, single-head gated attention mechanism equipped with (exponential) moving average to incorporate inductive bias of position-aware local dependencies into the positionagnostic attention mechanism. 1232, TITLE: Parallel Deep Neural Networks Have Zero Duality Gap https://openreview.net/forum?id=6zrOr Rdhjs AUTHORS: Yifei Wang, Tolga Ergen, Mert Pilanci HIGHLIGHT: In this paper, we particularly prove that the duality gap for deeper linear networks with vector outputs is nonzero 1233, TITLE: Multi-domain image generation and translation with identifiability guarantees https://openreview.net/forum?id=U2g8OGONA V AUTHORS: Shaoan Xie, Lingjing Kong, Mingming Gong, Kun Zhang HIGHLIGHT: Inspired by the recent advances in nonlinear Independent Component Analysis (ICA) theory, we propose a new method to learn the joint distribution from the marginals by enforcing a specific type of minimal change across domains. 1234, TITLE: Pessimism in the Face of Confounders: Provably Efficient Offline Reinforcement Learning in Partially Observable Markov Decision Processes https://openreview.net/forum?id=PbkBDQ5 UbV AUTHORS: Miao Lu, Yifei Min, Zhaoran Wang, Zhuoran Yang HIGHLIGHT: In particular, we aim to learn an optimal policy from a dataset collected by a behavior policy which possibly depends on the latent state. Understanding Zero-shot Adversarial Robustness for Large-Scale Models 1235, TITLE: https://openreview.net/forum?id=P4bXCawRi5J AUTHORS: Chengzhi Mao, Scott Geng, Junfeng Yang, Xin Wang, Carl Vondrick HIGHLIGHT: In this work, we identify and explore the problem of adapting large-scale models for zero-shot adversarial robustness.

1236, TITLE:Continual evaluation for lifelong learning: Identifying the stability gaphttps://openreview.net/forum?id=Zy350cRstc6AUTHORS:Matthias De Lange, Gido M van de Ven, Tinne TuytelaarsHIGHLIGHT:Instead, we establish a framework for continual evaluation that uses per-iteration evaluation and define a new
set of metrics that enables quantifying the stability gap.
1237, TITLE: Transformer-based model for symbolic regression via joint supervised learning https://openreview.net/forum?id=ULzyv9M1j5 AUTHORS: Wenqiang Li, Weijun Li, Linjun Sun, Min Wu, Lina Yu, Jingyi Liu, Yanjie Li, Songsong Tian HIGHLIGHT: However, recent transformer-based methods for SR focus more attention on large scale training data and ignore the ill-posed problem: the lack of sufficient supervision, i.e. expressions that may be completely different have the same supervision because of their same skeleton, which makes it challenging to deal with data that may be from the same expression skeleton but with different coefficients. Therefore, we present a transformer-based model for SR with the ability to alleviate this problem.
1238, TITLE: Robust and Controllable Object-Centric Learning through Energy-based Models https://openreview.net/forum?id=wcNtbEtcGIC
AUTHORS: Ruixiang ZHANG, Tong Che, Boris Ivanovic, Renhao Wang, Marco Pavone, Yoshua Bengio, Liam Paull HIGHLIGHT: In this work, we present EGO, a conceptually simple and general approach to learning object-centric representation through energy-based model. Pavone
1239, TITLE:Robust Fair Clustering: A Novel Fairness Attack and Defense Frameworkhttps://openreview.net/forum?id=4LMIZY7gt7hAUTHORS:Anshuman Chhabra, Peizhao Li, Prasant Mohapatra, Hongfu Liu
HIGHLIGHT: We propose a highly effective & novel fairness attack against state-of-the-art fair clustering models, & for self- completeness, we propose a defense framework based on consensus clustering & graph representation learning that is robust to our attack.
1240, TITLE: Learning to Jointly Share and Prune Weights for Grounding Based Vision and Language Models https://openreview.net/forum?id=UMERaIHMwB3 AUTHORS: Shangqian Gao, Burak Uzkent, Yilin Shen, Heng Huang, Hongxia Jin HIGHLIGHT: Leveraging this feature of transformers, we propose weight sharing across two transformer backbones and within the same transformer backbone and pruning across two backbones in a unified framework
1241 TITLE. Grand Domain Adaptation via Theory-Grounded Spectral Regularization
https://openreview.net/forum?id=OysfLgrk8mk AUTHORS: Yuning You, Tianlong Chen, Zhangyang Wang, Yang Shen HIGHLIGHT: This paper targets at designing theory-grounded algorithms for graph domain adaptation (GDA).
1242, TITLE: Understanding Why Generalized Reweighting Does Not Improve Over ERM https://openreview.net/forum?id=ashPce_W8F-
AUTHORS: Runtian Zhai, Chen Dan, J Zico Kolter, Pradeep Kumar Ravikumar HIGHLIGHT: But a line of recent work has empirically shown that these approaches do not significantly improve over ERM in real applications with distribution shift. The goal of this work is to obtain a comprehensive theoretical understanding of this intriguing phenomenon.
1243, TITLE: Particle-based Variational Inference with Preconditioned Functional Gradient Flow https://openreview.net/forum?id=60phWWAE3cS
AUTHORS: Hanze Dong, Xi Wang, LIN Yong, Tong Zhang HIGHLIGHT: This paper remedies the problem by proposing a general framework to obtain tractable functional gradient flow estimates. This paper remedies the problem by proposing a general framework to obtain tractable functional gradient flow
1244, TITLE: Combating Exacerbated Heterogeneity for Robust Decentralized Models https://openreview.net/forum?id=eKllxpLOOm AUTHORS: Jianing Zhu, Jiangchao Yao, Tongliang Liu, guanming yao, Jianliang Xu, Bo Han
HIGHLIGHT: We discover that the attribution behind this phenomenon is the generated adversarial data could exacerbate the data heterogeneity among local clients, making the wrapped federated learning perform poorly. To deal with this problem, we propose a novel framework termed as Slack Federated Adversarial Training (SFAT), assigning the client-wise slack during aggregation to combat the intensified heterogeneity.
1245, TITLE: Bidirectional Propagation for Cross-Modal 3D Object Detection https://openreview.net/forum?id=gYs_cRuK7V
AUTHORS:Yifan Zhang, Qijian Zhang, Junhui Hou, Yixuan Yuan, Guoliang XingHIGHLIGHT:In this paper, in contrast to existing pixel-to-point feature propagation, we investigate an opposite point-to-pixeldirection, allowing point-wise features to flow inversely into the 2D image branch.
1246, TITLE: TimesNet: Temporal 2D-Variation Modeling for General Time Series Analysis https://openreview.net/forum?id=ju_Uqw384Oq AUTHORS: Haixu Wu Tengge Hu, Yong Liu, Hang Zhou, Jianmin Wang, Mingsheng Long

1247. TITLE: Learning without Prejudices: Continual Unbiased Learning via Benign and Malignant Forgetting https://openreview.net/forum?id=gfPUokHsW-AUTHORS: Myeongho Jeon, Hyoje Lee, Yedarm Seong, Myungjoo Kang HIGHLIGHT: We term such type of forgetting benign forgetting, and categorize detrimental forgetting as malignant forgetting. Based on this finding, our objective in this study is twofold: (a) to discourage malignant forgetting by generating previous representations, and (b) encourage benign forgetting by employing contrastive learning in conjunction with feature-level augmentation. 1248, TITLE: FINDE: Neural Differential Equations for Finding and Preserving Invariant Quantities https://openreview.net/forum?id=tLScKVhcCR AUTHORS: Takashi Matsubara, Takaharu Yaguchi HIGHLIGHT: To this end, we propose first integral-preserving neural differential equation (FINDE). 1249, TITLE: A Holistic View of Noise Transition Matrix in Deep Learning and Beyond https://openreview.net/forum?id=aFzaXRImWE AUTHORS: LIN Yong, Renjie Pi, WEIZHONG ZHANG, Xiaobo Xia, Jiahui Gao, Xiao Zhou, Tongliang Liu, Bo Han HIGHLIGHT: In this paper, we explore learning statistically consistent classifiers under label noise by estimating the noise transition matrix T. 1250, TITLE: NORM: Knowledge Distillation via N-to-One Representation Matching https://openreview.net/forum?id=CRNwGauQpb6 AUTHORS: Xiaolong Liu, Lujun Li, Chao Li, Anbang Yao HIGHLIGHT: In this paper, we present \$N\$-to-\$O\$ne \$R\$epresentation \$M\$atching (NORM), a new two-stage knowledge distillation method, which relies on a linear Feature Transform (FT) module. 1251, TITLE: GPViT: A High Resolution Non-Hierarchical Vision Transformer with Group Propagation https://openreview.net/forum?id=IowKt5rYWsK AUTHORS: Chenhongyi Yang, Jiarui Xu, Shalini De Mello, Elliot J. Crowley, Xiaolong Wang HIGHLIGHT: We present the Group Propagation Vision Transformer (GPViT): a novel non- hierarchical (i.e. non-pyramidal) transformer model designed for general visual recognition with high-resolution features. 1252, TITLE: Deep Learning meets Nonparametric Regression: Are Weight-Decayed DNNs Locally Adaptive? https://openreview.net/forum?id=0Q9H Pgx132 AUTHORS: Kaiqi Zhang, Yu-Xiang Wang We study the theory of neural network (NN) from the lens of classical nonparametric regression problems with HIGHLIGHT: a focus on NN's ability to adaptively estimate functions with heterogeneous smoothness — a property of functions in Besov or Bounded Variation (BV) classes. 1253, TITLE: Sparse Token Transformer with Attention Back Tracking https://openreview.net/forum?id=VV0hSE8AxCw AUTHORS: Heejun Lee, Minki Kang, Youngwan Lee, Sung Ju Hwang However, previous token pruning approaches often remove tokens during the feed-forward stage without HIGHLIGHT: consideration of their impact on later layers' attentions, which has a potential risk of dropping out important tokens for the given task. To tackle this issue, we propose an attention back-tracking method that tracks the importance of each attention in a Transformer architecture from the outputs to the inputs, to preserve the tokens that have a large impact on the final predictions. 1254, TITLE: Maximizing Communication Efficiency for Large-scale Training via 0/1 Adam https://openreview.net/forum?id=-CefY2EOupj AUTHORS: Yucheng Lu, Conglong Li, Minjia Zhang, Christopher De Sa, Yuxiong He HIGHLIGHT: In this paper, we demonstrate the non-linearity in Adam causes slow convergence even when 1-bit compression or local steps are individually applied. PD-MORL: Preference-Driven Multi-Objective Reinforcement Learning Algorithm 1255, TITLE: https://openreview.net/forum?id=zS9sRyaPFlJ AUTHORS: Toygun Basaklar, Suat Gumussoy, Umit Ogras HIGHLIGHT: To this end, we propose a novel MORL algorithm that trains a single universal network to cover the entire preference space scalable to continuous robotic tasks. 1256, TITLE: Learning Symbolic Models for Graph-structured Physical Mechanism https://openreview.net/forum?id=f2wN4v_2__W AUTHORS: Hongzhi Shi, Jingtao Ding, Yufan Cao, quanming yao, Li Liu, Yong Li HIGHLIGHT: In this paper, we propose a new approach that generalizes symbolic regression to graph-structured physical mechanisms. Linear Connectivity Reveals Generalization Strategies 1257, TITLE: https://openreview.net/forum?id=hY6M0JHl3uL AUTHORS: Jeevesh Juneja, Rachit Bansal, Kyunghyun Cho, João Sedoc, Naomi Saphra HIGHLIGHT: In contrast to existing results from image classification, we find that among text classifiers (trained on MNLI, QQP, and CoLA), some pairs of finetuned models have large barriers of increasing loss on the linear paths between them.

1258, TITLE: Dirichlet-based Uncertainty Calibration for Active Domain Adaptation https://openreview.net/forum?id=4WM4cy42B81 AUTHORS: Mixue Xie, Shuang Li, Rui Zhang, Chi Harold Liu HIGHLIGHT: Despite active DA methods address this by further proposing targetness to measure the representativeness of target domain characteristics, their predictive uncertainty is usually based on the prediction of deterministic models, which can easily be miscalibrated on data with distribution shift. Considering this, we propose a Dirichlet-based Uncertainty Calibration (DUC) approach for active DA, which simultaneously achieves the mitigation of miscalibration and the selection of informative target samples. 1259, TITLE: Accurate Image Restoration with Attention Retractable Transformer https://openreview.net/forum?id=IloMJ5rqfnt AUTHORS: Jiale Zhang, Yulun Zhang, Jinjin Gu, Yongbing Zhang, Linghe Kong, Xin Yuan Obviously, this strategy could result in restricted receptive fields. To address this issue, we propose HIGHLIGHT: \textbf{A}ttention \textbf{R}etractable \textbf{T}ransformer (ART) for image restoration, which presents both dense and sparse attention modules in the network. 1260, TITLE: Causal Representation Learning for Instantaneous and Temporal Effects https://openreview.net/forum?id=itZ6ggvMnzS AUTHORS: Phillip Lippe, Sara Magliacane, Sindy Löwe, Yuki M Asano, Taco Cohen, Efstratios Gavves HIGHLIGHT: This effectively creates "instantaneous" effects and invalidates previous identifiability results. To address this issue, we propose iCITRIS, a causal representation learning method that allows for instantaneous effects in temporal sequences with known intervention targets. 1261, TITLE: Visual Imitation Learning with Patch Rewards https://openreview.net/forum?id=OnM3R47KIiU Minghuan Liu, Tairan He, Weinan Zhang, Shuicheng YAN, Zhongwen Xu AUTHORS: HIGHLIGHT: In this work, we propose to measure the expertise of various local regions of image samples, or called patches, and recover multi-dimensional patch rewards accordingly. 1262, TITLE: Diffusion Models Already Have A Semantic Latent Space https://openreview.net/forum?id=pd1P2eUBVfq AUTHORS: Mingi Kwon, Jaeseok Jeong, Youngjung Uh HIGHLIGHT: To address the problem, we propose asymmetric reverse process (Asyrp) which discovers the semantic latent space in frozen pretrained diffusion models. Gradient-Guided Importance Sampling for Learning Binary Energy-Based Models 1263, TITLE: https://openreview.net/forum?id=9DZKk85Z4zA AUTHORS: Meng Liu, Haoran Liu, Shuiwang Ji HIGHLIGHT: Although ratio matching is a sound method to learn discrete EBMs, it suffers from expensive computation and excessive memory requirement, thereby resulting in difficulties for learning EBMs on high-dimensional data. Motivated from these limitations, in this study, we propose ratio matching with gradient-guided importance sampling (RMwGGIS). Dataset Pruning: Reducing Training Data by Examining Generalization Influence 1264, TITLE: https://openreview.net/forum?id=4wZiAXD29TQ AUTHORS: Shuo Yang, Zeke Xie, Hanyu Peng, Min Xu, Mingming Sun, Ping Li HIGHLIGHT: To answer these, we propose dataset pruning, an optimization-based sample selection method that can (1) examine the influence of removing a particular set of training samples on model's generalization ability with theoretical guarantee, and (2) construct a smallest subset of training data that yields strictly constrained generalization gap. 1265, TITLE: Plateau in Monotonic Linear Interpolation --- A "Biased" View of Loss Landscape for Deep Networks https://openreview.net/forum?id=z289SIQOQna AUTHORS: Xiang Wang, Annie N. Wang, Mo Zhou, Rong Ge In this paper, we show that the MLI property is not necessarily related to the hardness of optimization problems, HIGHLIGHT: and empirical observations on MLI for deep neural networks depend heavily on the biases. 1266, TITLE: Crossformer: Transformer Utilizing Cross-Dimension Dependency for Multivariate Time Series Forecasting https://openreview.net/forum?id=vSVLM2j9eie AUTHORS: Yunhao Zhang, Junchi Yan HIGHLIGHT: To fill the gap, we propose Crossformer, a Transformer-based model utilizing cross-dimension dependency for MTS forecasting. 1267, TITLE: BrainBERT: Self-supervised representation learning for Intracranial Electrodes https://openreview.net/forum?id=xmcYx reUn6 AUTHORS: Christopher Wang, Vighnesh Subramaniam, Adam Uri Yaari, Gabriel Kreiman, Boris Katz, Ignacio Cases, Andrei Barbu HIGHLIGHT: We create a reusable Transformer, BrainBERT, for intracranial recordings bringing modern representation learning approaches to neuroscience.

1268, TITLE: Nonlinear Reconstruction for Operator Learning of PDEs with Discontinuities

https://openreview.net/forum?id=CrfhZAsJDsZ AUTHORS: Samuel Lanthaler, Roberto Molinaro, Patrik Hadorn, Siddhartha Mishra HIGHLIGHT: This paper investigates, both theoretically and empirically, the operator learning of PDEs with discontinuous solutions. 1269, TITLE: Discovering Informative and Robust Positives for Video Domain Adaptation https://openreview.net/forum?id=vk-j5pQY3Gv Chang Liu, Kunpeng Li, Michael Stopa, Jun Amano, Yun Fu AUTHORS: HIGHLIGHT: This paper presents a non-contrastive learning framework without relying on negative samples for unsupervised video domain adaptation. Composing Ensembles of Pre-trained Models via Iterative Consensus 1270, TITLE: https://openreview.net/forum?id=gmwDKo-4cY AUTHORS: Shuang Li, Yilun Du, Joshua B. Tenenbaum, Antonio Torralba, Igor Mordatch HIGHLIGHT: In this work, we propose a unified framework for composing ensembles of different pre-trained models -combining the strengths of each individual model to solve various multimodal problems in a zero-shot manner. 1271, TITLE: Automated Data Augmentations for Graph Classification https://openreview.net/forum?id=vTb1JI0Gps AUTHORS: Youzhi Luo, Michael Curtis McThrow, Wing Yee Au, Tao Komikado, Kanji Uchino, Koji Maruhashi, Shuiwang Ji HIGHLIGHT: In this work, we propose GraphAug, a novel automated data augmentation method aiming at computing labelinvariant augmentations for graph classification. 1272, TITLE: Learning Label Encodings for Deep Regression https://openreview.net/forum?id=k60XE b0Ix6 AUTHORS: Deval Shah, Tor M. Aamodt Lacking heretofore have been automated approaches to find a good label encoding for a given application. This HIGHLIGHT: paper introduces \emph {Regularized Label Encoding Learning (RLEL)} for end-to-end training of an entire network and its label encodings. 1273, TITLE: Fair Attribute Completion on Graph with Missing Attributes https://openreview.net/forum?id=9vcXCMp9VEp AUTHORS: Dongliang Guo, Zhixuan Chu, Sheng Li HIGHLIGHT: In this paper, we propose FairAC, a fair attribute completion method, to complement missing information and learn fair node embeddings for graphs with missing attributes. 1274, TITLE: Robustness to corruption in pre-trained Bayesian neural networks https://openreview.net/forum?id=kUI41mY8bH1 AUTHORS: Xi Wang, Laurence Aitchison HIGHLIGHT: We develop ShiftMatch, a new training-data-dependent likelihood for robustness to corruption in Bayesian neural networks (BNNs). NTFields: Neural Time Fields for Physics-Informed Robot Motion Planning 1275, TITLE: https://openreview.net/forum?id=ApF0dmi1 9K AUTHORS: Ruiqi Ni, Ahmed H Qureshi HIGHLIGHT: Recent developments have also led to physics-informed deep neural models capable of representing complex dynamical Partial Differential Equations (PDEs). Inspired by these developments, we propose Neural Time Fields (NTFields) for robot motion planning in cluttered scenarios. 1276, TITLE: Meta-learning Adaptive Deep Kernel Gaussian Processes for Molecular Property Prediction https://openreview.net/forum?id=KXRSh0sdVTP AUTHORS: Wenlin Chen, Austin Tripp, José Miguel Hernández-Lobato HIGHLIGHT: We propose Adaptive Deep Kernel Fitting with Implicit Function Theorem (ADKF-IFT), a novel framework for learning deep kernel Gaussian processes (GPs) by interpolating between meta-learning and conventional deep kernel learning. 1277. TITLE: ERL-Re\$^2\$: Efficient Evolutionary Reinforcement Learning with Shared State Representation and Individual Policy Representation https://openreview.net/forum?id=FYZCHEtt6H0 AUTHORS: Jianye HAO, Pengyi Li, Hongyao Tang, YAN ZHENG, Xian Fu, Zhaopeng Meng HIGHLIGHT: In this paper, we propose Evolutionary Reinforcement Learning with Two-scale State Representation and Policy Representation (ERL-Re§^2\$), a novel solution to the aforementioned two drawbacks. 1278, TITLE: Deep Ensembles for Graphs with Higher-order Dependencies https://openreview.net/forum?id=hZftxQGJ4Re AUTHORS: Steven Krieg, William Burgis, Patrick Soga, Nitesh Chawla In the presence of higher-order sequential dependencies, we show that the tendency of traditional graph HIGHLIGHT: representations to underfit each node's neighborhood causes existing GNNs to generalize poorly. To address this, we propose a novel Deep Graph Ensemble (DGE), which captures neighborhood variance by training an ensemble of GNNs on different neighborhood subspaces of the same node within a higher-order network structure.

1279. TITLE: Denoising Masked Autoencoders are Certifiable Robust Vision Learners https://openreview.net/forum?id=zDjtZZBZtqK AUTHORS: QuanLin Wu, Hang Ye, Yuntian Gu, Huishuai Zhang, Liwei Wang, Di He HIGHLIGHT: In this paper, we propose a new self-supervised method, which is called denoising masked autoencoders (DMAE), for learning certified robust classifiers of images. 1280, TITLE: Sound Randomized Smoothing in Floating-Point Arithmetic https://openreview.net/forum?id=HaHCoGcpV9 AUTHORS: Vaclav Voracek, Matthias Hein HIGHLIGHT: We discuss the implicit assumptions of randomized smoothing and show that they do not apply to generic image classification models whose smoothed versions are commonly certified. In order to overcome this problem, we propose a sound approach to randomized smoothing when using floating-point precision with essentially equal speed for quantized input. 1281. TITLE: Provably Efficient Risk-Sensitive Reinforcement Learning: Iterated CVaR and Worst Path https://openreview.net/forum?id=Yn0xg-kHNW-AUTHORS: Yihan Du, Siwei Wang, Longbo Huang In this paper, we study a novel episodic risk-sensitive Reinforcement Learning (RL) problem, named Iterated HIGHLIGHT: CVaR RL, which aims to maximize the tail of the reward-to-go at each step, and focuses on tightly controlling the risk of getting into catastrophic situations at each stage. Test-Time Robust Personalization for Federated Learning 1282. TITLE: https://openreview.net/forum?id=3aBuJEza5sq AUTHORS: Liangze Jiang, Tao Lin In this work, we identify the pitfalls of existing works under test-time distribution shifts and propose Federated HIGHLIGHT: Test-time Head Ensemble plus tuning (FedTHE+), which personalizes FL models with robustness to various test-time distribution shifts. Along with this, we build a benchmark for assessing performance and robustness of personalized FL methods during deployment. 1283, TITLE: Meta-prediction Model for Distillation-Aware NAS on Unseen Datasets https://openreview.net/forum?id=SEh5SfEQtqB AUTHORS: Hayeon Lee, Sohyun An, Minseon Kim, Sung Ju Hwang To eliminate excessive computational cost of DaNAS methods and the sub-optimality of rapid NAS methods, HIGHLIGHT: we propose a distillation-aware meta accuracy prediction model which can predict a given architecture's final performances on a dataset when performing KD with a given teacher, without having to actually train it on the target task. 1284, TITLE: Denoising Diffusion Error Correction Codes https://openreview.net/forum?id=rLwC0 MG-4w AUTHORS: Yoni Choukroun, Lior Wolf HIGHLIGHT: In this work, we propose to employ denoising diffusion models for the soft decoding of linear codes at arbitrary block lengths. 1285, TITLE: Dynamic Prompt Learning via Policy Gradient for Semi-structured Mathematical Reasoning https://openreview.net/forum?id=DHyHRBwJUTN AUTHORS: Pan Lu, Liang Qiu, Kai-Wei Chang, Ying Nian Wu, Song-Chun Zhu, Tanmay Rajpurohit, Peter Clark, Ashwin Kalyan HIGHLIGHT: To fill the gap, we present Tabular Math Word Problems (TabMWP), a new dataset containing 38,431 opendomain grade-level problems that require mathematical reasoning on both textual and tabular data. 1286, TITLE: Phase transition for detecting a small community in a large network https://openreview.net/forum?id=iN3Lh-Vy2TH AUTHORS: Jiashun Jin, Tracy Ke, Paxton Turner, Anru Zhang HIGHLIGHT: Using Sinkhorn's theorem, we show that the signal captured by the X^2-test may be a modeling artifact, and it may disappear once we replace the Erdös-Renyi model by a broader network model. 1287, TITLE: The Power of Regularization in Solving Extensive-Form Games https://openreview.net/forum?id=bPiHuNUNv R Mingyang Liu, Asuman E. Ozdaglar, Tiancheng Yu, Kaiqing Zhang AUTHORS: HIGHLIGHT: In this paper, we investigate the power of {\it regularization}, a common technique in reinforcement learning and optimization, in solving extensive-form games (EFGs). 1288, TITLE: Progressive Compressed Auto-Encoder for Self-supervised Representation Learning https://openreview.net/forum?id=8T4gmZbTkW7 AUTHORS: Jin Li, Yaoming Wang, XIAOPENG ZHANG, Yabo Chen, Dongsheng Jiang, Wenrui Dai, Chenglin Li, Hongkai Xiong, Qi Tian This redundancy is neglected by existing methods and causes non-negligible overheads in computation and HIGHLIGHT: storage that do not necessarily benefit self-supervised learning. In this paper, we present a novel approach named Progressive Compressed AutoEncoder (PCAE) to address this problem by progressively compacting tokens and retaining the least necessary

121

information for representation.

1289, TITLE: CFlowNets: Continuous control with Generative Flow Networks https://openreview.net/forum?id=yAYHho4fATa AUTHORS: Yinchuan Li, Shuang Luo, Haozhi Wang, Jianye HAO HIGHLIGHT: In this paper, we propose generative continuous flow networks (CFlowNets) that can be applied to continuous
1290, TITLE: Causal Balancing for Domain Generalization https://openreview.net/forum?id=F91SROvVJ_6 AUTHORS: Xinyi Wang, Michael Saxon, Jiachen Li, Hongyang Zhang, Kun Zhang, William Yang Wang HIGHLIGHT: We propose a balanced mini-batch sampling strategy to transform a biased data distribution into a spurious-free balanced distribution, based on the invariance of the underlying causal mechanisms for the data generation process.
1291, TITLE:Breaking Correlation Shift via Conditional Invariant Regularizerhttps://openreview.net/forum?id=-jTaz3CMk72AUTHORS:Mingyang Yi, Ruoyu Wang, Jiacheng Sun, Zhenguo Li, Zhi-Ming MaHIGHLIGHT:This paper proposes an algorithm to make the model to generalize on data with spurious correlation, the methodcan be implemented without information on spurious feature.
1292, TITLE: GRACE-C: Generalized Rate Agnostic Causal Estimation via Constraints https://openreview.net/forum?id=B_pCIsX8KL_ AUTHORS: Mohammadsajad Abavisani, David Danks, Sergey Plis HIGHLIGHT: We present an algorithm that combines constraint programming with both theoretical insights into the problem structure and prior information about admissible causal interactions to achieve speed up of multiple orders of magnitude.
1293, TITLE:Equiformer: Equivariant Graph Attention Transformer for 3D Atomistic Graphshttps://openreview.net/forum?id=KwmPfARgOTDAUTHORS:Yi-Lun Liao, Tess SmidtHIGHLIGHT:In this paper, we present Equiformer, a graph neural network leveraging the strength of Transformerarchitectures and incorporating SE(3)/E(3)-equivariant features based on irreducible representations (irreps).
1294, TITLE:Automating Nearest Neighbor Search Configuration with Constrained Optimizationhttps://openreview.net/forum?id=KfptQCEKVW4AUTHORS:Philip Sun, Ruiqi Guo, Sanjiv KumarHIGHLIGHT:This is becoming an increasingly unrealistic demand as ANN search grows in popularity. To tackle this obstacleto ANN adoption, this work proposes a constrained optimization-based approach to tuning quantization-based ANN algorithms.
1295, TITLE:Truncated Diffusion Probabilistic Models and Diffusion-based Adversarial Auto-Encodershttps://openreview.net/forum?id=HDxgaKk956lAUTHORS:Huangjie Zheng, Pengcheng He, Weizhu Chen, Mingyuan ZhouHIGHLIGHT:However, this approach is slow and costly because it needs many forward and reverse steps. We propose afaster and cheaper approach that adds noise not until the data become pure random noise, but until they reach a hidden noisy-datadistribution that we can confidently learn.
1296, TITLE:NTK-SAP: Improving neural network pruning by aligning training dynamicshttps://openreview.net/forum?id=-5EWhW_4qWPAUTHORS:Yite Wang, Dawei Li, Ruoyu SunHIGHLIGHT:Recent advances in neural tangent kernel (NTK) theory suggest that the training dynamics of large enoughneural networks is closely related to the spectrum of the NTK. Motivated by this finding, we propose to prune the connections thathave the least influence on the spectrum of the NTK.
1297, TITLE:TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Secondhttps://openreview.net/forum?id=cp5PvcI6w8_AUTHORS:Noah Hollmann, Samuel Müller, Katharina Eggensperger, Frank HutterHIGHLIGHT:We present TabPFN, a trained Transformer that can do supervised classification for small tabular datasets inless than a second, needs no hyperparameter tuning and is competitive with state-of-the-art classification methods.
1298, TITLE:Filter-Recovery Network for Multi-Speaker Audio-Visual Speech Separationhttps://openreview.net/forum?id=fiB2RjmgwQ6AUTHORS:Haoyue Cheng, Zhaoyang Liu, Wayne Wu, Limin WangHIGHLIGHT:To deal with these, we propose a valid method BFRNet, including a basic audio-visual speech separator and aFilter-Recovery Network (FRNet).
1299, TITLE:Sparse Q-Learning: Offline Reinforcement Learning with Implicit Value Regularizationhttps://openreview.net/forum?id=ueYYgo2pSSUAUTHORS:Haoran Xu, Li Jiang, Jianxiong Li, Zhuoran Yang, Zhaoran Wang, Victor Wai Kin Chan, Xianyuan ZhanHIGHLIGHT:In this work, we make a key finding that the in-sample learning paradigm arises under the \textit {Implicit ValueRegularization}(IVR) framework.

1300, TITLE: Topology-aware robust optimization https://openreview.net/forum?id=ylMq8MBnAp

AUTHORS: Fengchun Qiao, Xi Peng To this end, we propose topology-aware robust optimization (TRO) that seamlessly integrates distributional HIGHLIGHT: topology in a principled optimization framework. 1301. TITLE: Limitless Stability for Graph Convolutional Networks https://openreview.net/forum?id=XqcQhVUr2h0 AUTHORS: Christian Koke HIGHLIGHT: This work establishes rigorous, novel and widely applicable stability guarantees and transferability bounds for general graph convolutional networks -- without reference to any underlying limit object or statistical distribution. 1302, TITLE: Token Merging: Your ViT But Faster https://openreview.net/forum?id=JroZRaRw7Eu AUTHORS: Daniel Bolya, Cheng-Yang Fu, Xiaoliang Dai, Peizhao Zhang, Christoph Feichtenhofer, Judy Hoffman HIGHLIGHT: We introduce Token Merging (ToMe), a simple method to increase the throughput of existing ViT models without needing to train. 1303, TITLE: Spatial Attention Kinetic Networks with E(n)-Equivariance https://openreview.net/forum?id=3DIpIf3wQMC AUTHORS: Yuanqing Wang, John Chodera HIGHLIGHT: In this paper, we propose a simple alternative functional form that uses neurally parametrized linear combinations of edge vectors to achieve equivariance while still universally approximating node environments. 1304, TITLE: Revisiting the Entropy Semiring for Neural Speech Recognition https://openreview.net/forum?id=SNgLnzFOeiD AUTHORS: Oscar Chang, Dongseong Hwang, Olivier Siohan HIGHLIGHT: In this work, we revisit the entropy semiring for neural speech recognition models, and show how alignment entropy can be used to supervise models through regularization or distillation. 1305, TITLE: Neural Groundplans: Persistent Neural Scene Representations from a Single Image https://openreview.net/forum?id=Pza24zf9FpS AUTHORS: Prafull Sharma, Ayush Tewari, Yilun Du, Sergey Zakharov, Rares Andrei Ambrus, Adrien Gaidon, William T. Freeman, Fredo Durand, Joshua B. Tenenbaum, Vincent Sitzmann HIGHLIGHT: We present a method to map 2D image observations of a scene to a persistent 3D scene representation, enabling novel view synthesis and disentangled representation of the movable and immovable components of the scene. 1306, TITLE: Stochastic Differentially Private and Fair Learning https://openreview.net/forum?id=3nM5uhPlfv6 AUTHORS: Andrew Lowy, Devansh Gupta, Meisam Razaviyayn HIGHLIGHT: In this paper, we provide the first stochastic differentially private algorithm for fair learning that is guaranteed to converge. 1307, TITLE: Volumetric Optimal Transportation by Fast Fourier Transform https://openreview.net/forum?id=EVrz7UM-ZDm AUTHORS: Na Lei, DONGSHENG An, Min Zhang, Xiaoyin Xu, David Gu HIGHLIGHT: In this work, we propose a novel and powerful method, the FFT-OT (fast Fourier transform-optimal transport), to compute the 3-dimensional OT problems. 1308, TITLE: Function-Consistent Feature Distillation https://openreview.net/forum?id=pgHNOcxEdRI AUTHORS: Dongyang Liu, Meina Kan, Shiguang Shan, Xilin CHEN Considering this, we argue that the similarity between teacher and student features should \textit {not} be HIGHLIGHT: measured merely based on their appearance (i.e. L2 distance), but should, more importantly, be measured by their difference in function, namely how the lateral parts of the network will read, decode, and process them. Therefore, we propose Function-Consistent Feature Distillation (FCFD), which explicitly optimizes the functional similarity between teacher and student features. 1309, TITLE: Learning to Linearize Deep Neural Networks for Secure and Efficient Private Inference https://openreview.net/forum?id=BGF9IeDfmlH Souvik Kundu, Shunlin Lu, Yuke Zhang, Jacqueline Tiffany Liu, Peter Anthony Beerel AUTHORS: HIGHLIGHT: We present an automated linearization method to train a DNN with limited ReLU budget for inference in yielding models able to perform significantly better than exiting private inference SOTA both in terms of potentially improved latency and accuracy. 1310, TITLE: Decompose to Generalize: Species-Generalized Animal Pose Estimation https://openreview.net/forum?id=nQai B1Zrt AUTHORS: Guangrui Li, Yifan Sun, Zongxin Yang, Yi Yang HIGHLIGHT: This paper challenges the cross-species generalization problem for animal pose estimation, aiming to learn a pose estimator that can be well generalized to novel species.

1311, TITLE: Image as Set of Points

https://openreview.net/forum?id=awnvqZja69

AUTHORS: HIGHLIGHT: Context Clusters.	Xu Ma, Yuqian Zhou, Huan Wang, Can Qin, Bin Sun, Chang Liu, Yun Fu In this work, we introduce a straightforward and promising paradigm for visual representation, which is called
1312, TITLE: https://openreview.ne	Trainability Preserving Neural Pruning et/forum?id=AZFvpnnewr Huan Wang, Yun Fu
HIGHLIGHT: trainability against pr	In this paper, we present trainability preserving pruning (TPP), a scalable method to preserve network runing, aiming for improved pruning performance.
1313, TITLE:	DrML: Diagnosing and Rectifying Vision Models using Language
https://openreview.ne	et/forum?id=D-zfUK7BR6c
AUTHORS:	Yuhui Zhang, Jeff Z. HaoChen, Shih-Cheng Huang, Kuan-Chieh Wang, James Zou, Serena Yeung
HIGHLIGHT:	Through a combination of theoretical explanation and empirical verification, we present conditions under which
classifiers trained on	embeddings from one modality can be equivalently applied to embeddings from another modality.
1314, TITLE:	Adversarial Attacks on Adversarial Bandits
https://openreview.ne	et/forum?id=bBpT6dEjeRG
AUTHORS:	Yuzhe Ma, Zhijin Zhou
HIGHLIGHT:	We study a security threat to adversarial multi-armed bandit, in which an attacker perturbs the loss or reward
signal to control the b	behavior of the victim bandit player.
1315, TITLE: https://openreview.ne	Harnessing Out-Of-Distribution Examples via Augmenting Content and Style
AUTHORS:	Zhuo Huang, Xiaobo Xia, Li Shen, Bo Han, Mingming Gong, Chen Gong, Tongliang Liu
HIGHLIGHT:	To Harness OOD data, this paper proposes HOOD method that can leverage the content and style from each
image instance to ide	entify benign and malign OOD data.
1316, TITLE: https://openreview.ne	TaskPrompter: Spatial-Channel Multi-Task Prompting for Dense Scene Understanding et/forum?id=-CwPopPJda Hanrong Ve, Dan Xu
HIGHLIGHT: manner, which can n learning objectives, b capacity will be used task prompting transf	It is barely explored in the literature to model these three perspectives in each network layer in an end-to-end ot only minimize the effort of carefully designing empirical structures for the three multi-task representation but also greatly improve the representation learning capability of the multi-task network since all the model to optimize the three objectives together. In this paper, we propose TaskPrompter, a novel spatial-channel multi-former framework to achieve this target.
1317, TITLE:	Learning Domain-Agnostic Representation for Disease Diagnosis
https://openreview.ne	et/forum?id=-HHJZIRpGb
AUTHORS:	Churan Wang, Jing Li, Xinwei Sun, Fandong Zhang, Yizhou Yu, Yizhou Wang
HIGHLIGHT:	To disentangle disease-related features, we first leverage structural causal modeling to explicitly model disease-
related and center-eff	fects that are provable to be disentangled from each other. Guided by this, we propose a novel Domain Agnostic
Representation Mode	el (DarMo) based on variational Auto-Encoder.
1318, TITLE:	BEVDistill: Cross-Modal BEV Distillation for Multi-View 3D Object Detection
https://openreview.ne	et/forum?id=-2zfgNS917
HIGHLIGHT:	In this paper, we explore the incorporation of LiDAR-based detectors for multi-view 3D object detection.
1319, TITLE: https://openreview.ne	Suppressing the Heterogeneity: A Strong Feature Extractor for Few-shot Segmentation et/forum?id=CGuvK3U09LH Zhengdong Hu, Yifan Sun, Yi Yang
HIGHLIGHT: (MuHS).	Motivated by these observations, we propose a feature extractor with Multi-level Heterogeneity Suppressing
1320, TITLE:	Sparse and Hierarchical Masked Modeling for Convolutional Representation Learning
https://openreview.ne	et/forum?id=NRxydtWup1S
HIGHLIGHT:	Keyu Tian, Yi Jiang, qishuai diao, Chen Lin, Liwei Wang, Zenuan Yuan
masKed modeling.	This paper presents a simple yet powerful framework to pre-train convolutional network (convnet) with Sparse
1321, TITLE:	On amortizing convex conjugates for optimal transport
https://openreview.ne	et/forum?id=TQ5WUwS_4ai
HIGHLIGHT: Wasserstein-2 optima	This paper focuses on computing the convex conjugate operation that arises when solving Euclidean al transport problems.
1322, TITLE:	Molecular Geometry Pretraining with SE(3)-Invariant Denoising Distance Matching
https://openreview.ne	et/forum?id=CjTHVo1dvR
AUTHORS:	Shengchao Liu, Hongyu Guo, Jian Tang

HIGHLIGHT: Motivated by the dynamic nature of 3D molecules, where the continuous motion of a molecule in the 3D Euclidean space forms a smooth potential energy surface, we propose a 3D coordinate denoising pretraining framework to model such an energy landscape.

1323, TITLE: DAG Matters! GFlowNets Enhanced Explainer for Graph Neural Networks https://openreview.net/forum?id=jgmuRzM-sb6 AUTHORS: Wenqian Li, Yinchuan Li, Zhigang Li, Jianye HAO, Yan Pang HIGHLIGHT: However, the exponential size of candidate subgraphs limits the applicability of state-of-the-art methods to large-scale GNNs. We enhance on this through a different approach: by proposing a generative structure - GFlowNets-based GNN Explainer (GFlowExplainer), we turn the optimization problem into a step-by-step generative problem. 1324, TITLE: Continuous-Discrete Convolution for (3+1)D Geometry-Sequence Modeling in Proteins https://openreview.net/forum?id=P5Z-Z19XJ7 AUTHORS: Hehe Fan, Zhangyang Wang, Yi Yang, Mohan Kankanhalli HIGHLIGHT: This paper proposes a Continuous-Discrete Convolution (CDConv) that uses irregular and regular approaches to model the geometry and sequence structures, respectively. 1325, TITLE: Solving stochastic weak Minty variational inequalities without increasing batch size https://openreview.net/forum?id=ejR4E1jaH9k AUTHORS: Thomas Pethick, Olivier Fercoq, Puya Latafat, Panagiotis Patrinos, Volkan Cevher HIGHLIGHT: This paper introduces a family of stochastic extragradient-type algorithms for a class of nonconvex-nonconcave problems characterized by the weak Minty variational inequality (MVI). 1326. TITLE: A General Framework For Proving The Equivariant Strong Lottery Ticket Hypothesis https://openreview.net/forum?id=vVJZtlZB9D AUTHORS: Damien Ferbach, Christos Tsirigotis, Gauthier Gidel, Joey Bose In this paper, we generalize the SLTH to functions that preserve the action of the group \$G\$---i.e. \$G\$-HIGHLIGHT: equivariant network---and prove, with high probability, that one can approximate any \$G\$-equivariant network of fixed width and depth by pruning a randomly initialized overparametrized \$G\$-equivariant network to a \$G\$-equivariant subnetwork. 1327, TITLE: CO3: Cooperative Unsupervised 3D Representation Learning for Autonomous Driving https://openreview.net/forum?id=QUaDoIdgo0 AUTHORS: Runjian Chen, Yao Mu, Runsen Xu, Wenqi Shao, Chenhan Jiang, Hang Xu, Yu Qiao, Zhenguo Li, Ping Luo HIGHLIGHT: In this paper, we propose CO3, namely {Co}operative {Co}ntrastive Learning and {Co}ntextual Shape Prediction, to learn 3D representation for outdoor-scene point clouds in an unsupervised manner. 1328, TITLE: FedSpeed: Larger Local Interval, Less Communication Round, and Higher Generalization Accuracy https://openreview.net/forum?id=bZjxxYURKT AUTHORS: Yan Sun, Li Shen, Tiansheng Huang, Liang Ding, Dacheng Tao HIGHLIGHT: Its performance suffers from the non-vanishing biases introduced by the local inconsistent optimal and the rugged client-drifts by the local over-fitting. In this paper, we propose a novel and practical method, FedSpeed, to alleviate the negative impacts posed by these problems. 1329, TITLE: Share Your Representation Only: Guaranteed Improvement of the Privacy-Utility Tradeoff in Federated Learning https://openreview.net/forum?id=oJpVVGXu9i AUTHORS: Zebang Shen, Jiayuan Ye, Anmin Kang, Hamed Hassani, Reza Shokri HIGHLIGHT: In this paper, we consider a representation federated learning objective that encourages various parties to collaboratively refine the consensus part of the model, with differential privacy guarantees, while separately allowing sufficient freedom for local personalization (without releasing it).

1330, TITLE: EquiMod: An Equivariance Module to Improve Self-Supervised Learning

https://openreview.net/forum?id=eDLwjKmtYFt

AUTHORS: Alexandre DEVILLERS, Mathieu Lefort

HIGHLIGHT: In this work, we introduce a generic equivariance module that structures the learned latent space, in the sense that our module learns to predict the displacement in the embedding space caused by the augmentations.

1331, TITLE: KwikBucks: Correlation Clustering with Cheap-Weak and Expensive-Strong Signals

https://openreview.net/forum?id=p0JSSa1AuV

AUTHORS: Sandeep Silwal, Sara Ahmadian, Andrew Nystrom, Andrew McCallum, Deepak Ramachandran, Seyed Mehran Kazemi

HIGHLIGHT: Inspired by text clustering, we study correlation clustering where similarities must be queried via an expensive model (e.g. a large language model) with additional help from a cheap but noisy model (e.g. an embedding based model).

1332, TITLE: Few-Shot Domain Adaptation For End-to-End Communication

https://openreview.net/forum?id=4F1gvduDeL

AUTHORS:Jayaram Raghuram, Yijing Zeng, Dolores Garcia, Rafael Ruiz, Somesh Jha, Joerg Widmer, Suman BanerjeeHIGHLIGHT:We focus on a generative channel model based on the Gaussian mixture density network (MDN), and propose aregularized, parameter-efficient adaptation of the MDN using a set of affine transformations.

1333, TITLE: Online Bias Correction for Task-Free Continual Learning https://openreview.net/forum?id=18XzeuYZh AUTHORS: Aristotelis Chrysakis, Marie-Francine Moens HIGHLIGHT: In this work, we show both theoretically and empirically how experience replay biases the outputs of the model towards recent stream observations. Don't fear the unlabelled: safe semi-supervised learning via debiasing 1334, TITLE: https://openreview.net/forum?id=TN9gO4x0Ep3 AUTHORS: Hugo Schmutz, Olivier HUMBERT, Pierre-Alexandre Mattei HIGHLIGHT: We propose a slight modification of most common semi-supervised learning methods to make them safe by debiasing their risk estimate. In particular, we apply it successfully to Fixmatch. Learning a Data-Driven Policy Network for Pre-Training Automated Feature Engineering 1335. TITLE: https://openreview.net/forum?id=688hNNMigVX AUTHORS: Liyao Li, Haobo Wang, Liangyu Zha, Qingyi Huang, Sai Wu, Gang Chen, Junbo Zhao HIGHLIGHT: This paper proposes a novel AutoFE framework Feature Set Data-Driven Search (FETCH), a pipeline mainly for feature generation and selection. 1336, TITLE: Actionable Neural Representations: Grid Cells from Minimal Constraints https://openreview.net/forum?id=xfqDe72zh41 AUTHORS: Will Dorrell, Peter E. Latham, Timothy E. J. Behrens, James C. R. Whittington HIGHLIGHT: We study a novel definition of an optimal representation of structured spaces, and show that it can be used to derive the brain's grid cells and their perturbations normatively. 1337, TITLE: A Message Passing Perspective on Learning Dynamics of Contrastive Learning https://openreview.net/forum?id=VBTJqqWjxMv AUTHORS: Yifei Wang, Qi Zhang, Tianqi Du, Jiansheng Yang, Zhouchen Lin, Yisen Wang HIGHLIGHT: In this paper, we show that if we cast a contrastive objective equivalently into the function space, then its learning dynamics admits an interpretable form. 1338, TITLE: Zeroth-Order Optimization with Trajectory-Informed Derivative Estimation https://openreview.net/forum?id=n1bLgxHW6jW AUTHORS: Yao Shu, Zhongxiang Dai, Weicong Sng, Arun Verma, Patrick Jaillet, Bryan Kian Hsiang Low HIGHLIGHT: To this end, we propose a trajectory-informed derivative estimation method which only uses the optimization trajectory (i.e., the history of function queries during optimization) and hence eliminates the need for additional function queries to estimate a derivative. 1339, TITLE: Neuroevolution is a Competitive Alternative to Reinforcement Learning for Skill Discovery https://openreview.net/forum?id=6BHlZgyPOZY AUTHORS: Felix Chalumeau, Raphael Boige, Bryan Lim, Valentin Macé, Maxime Allard, Arthur Flajolet, Antoine Cully, Thomas PIERROT HIGHLIGHT: In this work, we demonstrate that less widely-used neuroevolution methods, specifically Quality Diversity (QD), are a competitive alternative to information-theory-augmented RL for skill discovery. 1340. TITLE: Uniform-in-time propagation of chaos for the mean field gradient Langevin dynamics https://openreview.net/forum?id= JScUk9TBUn AUTHORS: Taiji Suzuki, Atsushi Nitanda, Denny Wu HIGHLIGHT: In this work, we establish a quantitative weak propagation of chaos result for the system, with a finite-particle discretization error of \$\mathcal{O}(1/N)\$ \textit{uniformly over time}, where \$N\$ is the width of the neural network. 1341. TITLE: Individual Privacy Accounting with Gaussian Differential Privacy https://openreview.net/forum?id=JmC Tld3v-f AUTHORS: Antti Koskela, Marlon Tobaben, Antti Honkela This kind of analysis has been carried out for the R\'envi differential privacy by Feldman and Zrnic (2021), HIGHLIGHT: however not yet for the so-called optimal privacy accountants. We make first steps in this direction by providing a careful analysis using the Gaussian differential privacy which gives optimal bounds for the Gaussian mechanism, one of the most versatile DP mechanisms. 1342, TITLE: Evolving Populations of Diverse RL Agents with MAP-Elites https://openreview.net/forum?id=CBfYffLqWqb AUTHORS: Thomas PIERROT, Arthur Flajolet HIGHLIGHT: Furthermore, existing approaches mixing ME with RL tend to be tied to a specific RL algorithm, which effectively prevents their use on problems where the corresponding RL algorithm fails. To address these shortcomings, we introduce a flexible framework that allows the use of any RL algorithm within a population update and alleviates the aforementioned limitations by evolving populations of agents (whose definition include hyperparameters and all learnable parameters) instead of just policies. 1343, TITLE: RandProx: Primal-Dual Optimization Algorithms with Randomized Proximal Updates https://openreview.net/forum?id=cB4N3G5udUS

AUTHORS: Laurent Condat, Peter Richtárik

HIGHLIGHT: We propose a new primal-dual algorithm, in which the dual update is randomized; equivalently, the proximity operator of one of the function in the problem is replaced by a stochastic oracle. 1344, TITLE: Fast Nonlinear Vector Quantile Regression https://openreview.net/forum?id=UxqUgchwXkK AUTHORS: Aviv A. Rosenberg, Sanketh Vedula, Yaniv Romano, Alexander Bronstein HIGHLIGHT: We extend Vector Quantile Regression to support non-linear specification, while ensuring monotonicity and scaling to millions of samples. 1345, TITLE: Hierarchical Abstraction for Combinatorial Generalization in Object Rearrangement https://openreview.net/forum?id=fGG6vHp3W9W AUTHORS: Michael Chang, Alyssa Li Dayan, Franziska Meier, Thomas L. Griffiths, Sergey Levine, Amy Zhang HIGHLIGHT: Worse, these entities are often unknown and must be inferred from sensory percepts. We present a hierarchical abstraction approach to uncover these underlying entities and achieve combinatorial generalization from unstructured inputs. 1346, TITLE: Transfer NAS with Meta-learned Bayesian Surrogates https://openreview.net/forum?id=paGvsrl4Ntr AUTHORS: Gresa Shala, Thomas Elsken, Frank Hutter, Josif Grabocka HIGHLIGHT: This is in contrast to the manual design process by researchers and engineers that leverage previous deep learning experiences by, e.g., transferring architectures from previously solved, related problems. We propose to adopt this human design strategy and introduce a novel surrogate for NAS, that is meta-learned across prior architecture evaluations across different datasets. 1347. TITLE: Scaling Up Probabilistic Circuits by Latent Variable Distillation https://openreview.net/forum?id=067CGykiZTS AUTHORS: Anji Liu, Honghua Zhang, Guy Van den Broeck This phenomenon suggests that the existing optimizers fail to exploit the full expressive power of large PCs. We HIGHLIGHT: propose to overcome such bottleneck by latent variable distillation: we leverage the less tractable but more expressive deep generative models to provide extra supervision over the latent variables of PCs. 1348, TITLE: UL2: Unifying Language Learning Paradigms https://openreview.net/forum?id=6ruVLB727MC AUTHORS: Yi Tay, Mostafa Dehghani, Vinh Q. Tran, Xavier Garcia, Jason Wei, Xuezhi Wang, Hyung Won Chung, Dara Bahri, Tal Schuster, Steven Zheng, Denny Zhou, Neil Houlsby, Donald Metzler HIGHLIGHT: This paper presents a unified framework for pre-training models that are universally effective across datasets and setups. Bitrate-Constrained DRO: Beyond Worst Case Robustness To Unknown Group Shifts 1349, TITLE: https://openreview.net/forum?id=2QzNuaRHn4Z AUTHORS: Amrith Setlur, Don Dennis, Benjamin Eysenbach, Aditi Raghunathan, Chelsea Finn, Virginia Smith, Sergey Levine HIGHLIGHT: In this work, we address limitations in prior approaches by assuming a more nuanced form of group shift: conditioned on the label, we assume that the true group function is simple. 1350, TITLE: Feature selection and low test error in shallow low-rotation ReLU networks https://openreview.net/forum?id=swEskiem99 AUTHORS: Matus Telgarsky HIGHLIGHT: This work establishes low test error of gradient flow (GF) and stochastic gradient descent (SGD) on two-layer ReLU networks with standard initialization scale, in three regimes where key sets of weights rotate little (either naturally due to GF and SGD, or due to an artificial constraint), and making use of margins as the core analysis technique. 1351, TITLE: The Asymmetric Maximum Margin Bias of Quasi-Homogeneous Neural Networks https://openreview.net/forum?id=IM4xp7kGI5V AUTHORS: Daniel Kunin, Atsushi Yamamura, Chao Ma, Surya Ganguli In this work, we explore the maximum-margin bias of quasi-homogeneous neural networks trained with HIGHLIGHT: gradient flow on an exponential loss and past a point of separability. Coupled Multiwavelet Operator Learning for Coupled Differential Equations 1352. TITLE: https://openreview.net/forum?id=kIo C6QmMOM AUTHORS: Xiongye Xiao, Defu Cao, Ruochen Yang, Gaurav Gupta, Chenzhong Yin, Gengshuo Liu, Radu Balan, Paul Bogdan HIGHLIGHT: Towards this end, we propose a \textit{coupled multiwavelets neural operator} (CMWNO) learning scheme by decoupling the coupled integral kernels during the multiwavelet decomposition and reconstruction procedures in the Wavelet space. 1353, TITLE: Multi-Objective Online Learning https://openreview.net/forum?id=dKkMnCWfVmm Jiyan Jiang, Wenpeng Zhang, Shiji Zhou, Lihong Gu, Xiaodong Zeng, Wenwu Zhu AUTHORS: HIGHLIGHT: This paper presents a systematic study of multi-objective online learning.

1354, TITLE: Sparse Distributed Memory is a Continual Learner

https://openreview.net/forum?id=JknGeelZJpHP AUTHORS: Trenton Bricken, Xander Davies, Deepak Singh, Dmitry Krotov, Gabriel Kreiman HIGHLIGHT: Building on work using Sparse Distributed Memory (SDM) to connect a core neural circuit with the powerful Transformer model, we create a modified Multi-Layered Perceptron (MLP) that is a strong continual learner. 1355, TITLE: UniMax: Fairer and More Effective Language Sampling for Large-Scale Multilingual Pretraining https://openreview.net/forum?id=kXwdL1cWOAi AUTHORS: Hyung Won Chung, Xavier Garcia, Adam Roberts, Yi Tay, Orhan Firat, Sharan Narang, Noah Constant HIGHLIGHT: In this paper, we propose a new sampling method, UniMax, that delivers more uniform coverage of head languages while mitigating overfitting on tail languages by explicitly capping the number of repeats over each languages corpus. 1356, TITLE: GNNInterpreter: A Probabilistic Generative Model-Level Explanation for Graph Neural Networks https://openreview.net/forum?id=rqq6Dh8t4d AUTHORS: Xiaoqi Wang, Han Wei Shen HIGHLIGHT: In this paper, we propose a model-agnostic model-level explanation method for different GNNs that follow the message passing scheme, GNNInterpreter, to explain the high-level decision-making process of the GNN model. 1357, TITLE: Learning to reason over visual objects https://openreview.net/forum?id=uR6x8Be7o_M Shanka Subhra Mondal. Taylor Whittington Webb, Jonathan Cohen AUTHORS: HIGHLIGHT: Motivated by the goal of designing AI systems with this capacity, recent work has focused on evaluating whether neural networks can learn to solve RPM-like problems. This work has generally found that strong performance on these problems requires the incorporation of inductive biases that are specific to the RPM problem format, raising the question of whether such models might be more broadly useful. 1358, TITLE: Contextual Convolutional Networks https://openreview.net/forum?id=PldynS56bN AUTHORS: Shuxian Liang, Xu Shen, Tongliang Liu, Xian-Sheng Hua This paper presents a new Convolutional Neural Network, named Contextual Convolutional Network, that HIGHLIGHT: capably serves as a general-purpose backbone for visual recognition. 1359, TITLE: Imitating Graph-Based Planning with Goal-Conditioned Policies https://openreview.net/forum?id=6lUEy1J5R7p AUTHORS: Junsu Kim, Younggyo Seo, Sungsoo Ahn, Kyunghwan Son, Jinwoo Shin HIGHLIGHT: However, the sample-efficiency of such RL schemes still remains a challenge, particularly for long-horizon tasks. To address this issue, we present a simple vet effective self-imitation scheme which distills a subgoal-conditioned policy into the target-goal-conditioned policy. 1360, TITLE: Statistical Inference for Fisher Market Equilibrium https://openreview.net/forum?id=KemSBwOYJC AUTHORS: Luofeng Liao, Yuan Gao, Christian Kroer HIGHLIGHT: In this paper we focus on the specific case of linear Fisher markets. A theoretical study of inductive biases in contrastive learning 1361, TITLE: https://openreview.net/forum?id=AuEgNlEAmed AUTHORS: Jeff Z. HaoChen, Tengyu Ma HIGHLIGHT: In this work, we provide the first theoretical analysis of self-supervised learning that incorporates the effect of inductive biases originating from the model class. 1362, TITLE: Easy Differentially Private Linear Regression https://openreview.net/forum?id=rSUCajhLsQ AUTHORS: Kareem Amin, Matthew Joseph, Mónica Ribero, Sergei Vassilvitskii HIGHLIGHT: In this paper, we study an algorithm which uses the exponential mechanism to select a model with high Tukey depth from a collection of non-private regression models. 1363, TITLE: From Play to Policy: Conditional Behavior Generation from Uncurated Robot Data https://openreview.net/forum?id=c7rM7F7iOiN Zichen Jeff Cui, Yibin Wang, Nur Muhammad Mahi Shafiullah, Lerrel Pinto AUTHORS: HIGHLIGHT: In this work, we present Conditional Behavior Transformers (C-BeT), a method that combines the multi-modal generation ability of Behavior Transformer with future-conditioned goal specification. 1364, TITLE: A Closer Look at Model Adaptation using Feature Distortion and Simplicity Bias https://openreview.net/forum?id=wkg b4-IwTZ AUTHORS: Puja Trivedi, Danai Koutra, Jayaraman J. Thiagarajan HIGHLIGHT: Given the strong effectiveness of LP+FT, we propose incorporating hardness-promoting perturbations during LP to obtain initializations for FT that further decrease SB. 1365, TITLE: Digging into Backbone Design on Face Detection https://openreview.net/forum?id=NkJOhtNKX91 AUTHORS: Yang Liu, Fei Wang, Lei Shang, Jiankang Deng, Baigui Sun, Xuansong Xie

HIGHLIGHT: Considering the intrinsic design property and the virtual importance role of the face detection backbone, we thus ask a critical question: How to employ NAS to search FD-friendly backbone architecture? To cope with this question, we propose a distribution-dependent stage-aware ranking score (DDSAR-Score) to explicitly characterize the stage-level expressivity and identify the individual importance of each stage, thus satisfying the aforementioned design criterion of the FD backbone. 1366, TITLE: Understanding and Adopting Rational Behavior by Bellman Score Estimation https://openreview.net/forum?id=WzGdBqcBicl

AUTHORS: Kuno Kim, Stefano Ermon

HIGHLIGHT: In this work, we make a key observation that knowing how changes in the underlying rewards affect the optimal behavior allows one to solve a variety of aforementioned problems.

1367, TITLE: Towards Smooth Video Composition

https://openreview.net/forum?id=W918Ora75q

AUTHORS: Qihang Zhang, Ceyuan Yang, Yujun Shen, Yinghao Xu, Bolei Zhou

HIGHLIGHT: This work investigates how to model the temporal relations for composing a video with arbitrary number of frames, from a few to even infinite, using generative adversarial networks (GANs).

 1368, TITLE:
 Can Agents Run Relay Race with Strangers? Generalization of RL to Out-of-Distribution Trajectories

 https://openreview.net/forum?id=ipflrGaf7ry

 AUTHORS:
 Li-Cheng Lan, Huan Zhang, Cho-Jui Hsieh

HIGHLIGHT: In this paper, we evaluate and improve the generalization performance for rein- forcement learning (RL) agents on the set of "controllable" states, where good policies exist in these states to achieve high rewards.

1369, TITLE:Continuous PDE Dynamics Forecasting with Implicit Neural Representationshttps://openreview.net/forum?id=B73niNjbPsAUTHORS:Yuan Yin, Matthieu Kirchmeyer, Jean-Yves Franceschi, Alain Rakotomamonjy, patrick gallinariHIGHLIGHT:We address this problem by introducing a new data-driven approach, DINo, that models a PDE's flow with

continuous-time dynamics of spatially continuous functions.

 1370, TITLE:
 STUNT: Few-shot Tabular Learning with Self-generated Tasks from Unlabeled Tables

 https://openreview.net/forum?id=_xlsjehDvlY

 AUTHORS:
 Jaehyun Nam, Jihoon Tack, Kyungmin Lee, Hankook Lee, Jinwoo Shin

HIGHLIGHT: In this paper, we propose a simple yet effective framework for few-shot tabular learning, coined Self-generated Tasks from UNlabeled Tables (STUNT).

1371, TITLE: Language Models Can (kind of) Reason: A Systematic Formal Analysis of Chain-of-Thought https://openreview.net/forum?id=qFVVBzXxR2V

AUTHORS: Abulhair Saparov, He He

HIGHLIGHT: To enable systematic exploration of the reasoning ability of LLMs, we present a new synthetic questionanswering dataset called PrOntoQA, where each example is generated from a synthetic world model represented in first-order logic.

1372, TITLE: Understanding the Covariance Structure of Convolutional Filters https://openreview.net/forum?id=WGApODQvwRg

AUTHORS: Asher Trockman, Devin Willmott, J Zico Kolter

HIGHLIGHT: In this work, we first observe that such learned filters have highly-structured covariance matrices, and moreover, we find that covariances calculated from small networks may be used to effectively initialize a variety of larger networks of different depths, widths, patch sizes, and kernel sizes, indicating a degree of model-independence to the covariance structure.

1373, TITLE: Masked Distillation with Receptive Tokens

https://openreview.net/forum?id=mWRngkvIki3

AUTHORS: Tao Huang, Yuan Zhang, Shan You, Fei Wang, Chen Qian, Jian Cao, Chang Xu HIGHLIGHT: In this paper, we introduce a learnable embedding dubbed receptive token to locate the pixels of interests (PoIs)

in the feature map, with a distillation mask generated via pixel-wise attention.

1374, TITLE: Robust Multivariate Time-Series Forecasting: Adversarial Attacks and Defense Mechanisms https://openreview.net/forum?id=ctmLBs8IITa

AUTHORS: Linbo Liu, Youngsuk Park, Trong Nghia Hoang, Hilaf Hasson, Luke Huan

HIGHLIGHT: This work studies the threats of adversarial attack on multivariate probabilistic forecasting models and viable defense mechanisms.

 1375, TITLE:
 Efficient Deep Reinforcement Learning Requires Regulating Statistical Overfitting

 https://openreview.net/forum?id=14-kr46GvP

 AUTHORS:
 Qiyang Li, Aviral Kumar, Ilya Kostrikov, Sergey Levine

 UICHI LCUT:
 In attionance was attempt to understand the primary bettlengels in sergels officient deep PL by examining

HIGHLIGHT: In this paper, we attempt to understand the primary bottleneck in sample-efficient deep RL by examining several potential hypotheses such as non-stationarity, excessive action distribution shift, and overfitting.

1376, TITLE: Offline Reinforcement Learning with Differentiable Function Approximation is Provably Efficient https://openreview.net/forum?id=6jfbOWzWTcE AUTHORS: Ming Yin, Mengdi Wang, Yu-Xiang Wang

HIGHLIGHT: Towards bridging the gap, we take a step by considering offline reinforcement learning with \emph{differentiable function class approximation} (DFA). 1377, TITLE: MECTA: Memory-Economic Continual Test-Time Model Adaptation https://openreview.net/forum?id=N92hjSf5NNh AUTHORS: Junyuan Hong, Lingjuan Lyu, Jiayu Zhou, Michael Spranger HIGHLIGHT: In this paper, we provide a novel solution, dubbed MECTA, to drastically improve the memory efficiency of gradient-based CTA. Robust Algorithms on Adaptive Inputs from Bounded Adversaries 1378, TITLE: https://openreview.net/forum?id=I29Kt0RwChs Yeshwanth Cherapanamjeri, Sandeep Silwal, David Woodruff, Fred Zhang, Qiuyi Zhang, Samson Zhou AUTHORS: HIGHLIGHT: We study dynamic algorithms robust to adaptive inputs generated from sources with bounded capabilities, such as sparsity or limited interaction. 1379, TITLE: Interpretability with full complexity by constraining feature information https://openreview.net/forum?id=R OL5mLhsv AUTHORS: Kieran A Murphy, Danielle Bassett HIGHLIGHT: The optimal compression of each feature---at every stage of approximation---allows fine-grained inspection of how feature values are similar or distinct with regards to the prediction. We develop a framework for extracting insight from the spectrum of approximate models and demonstrate its utility on a range of tabular datasets. 1380, TITLE: Chasing All-Round Graph Representation Robustness: Model, Training, and Optimization https://openreview.net/forum?id=7ik5gWiC18M AUTHORS: Chunhui Zhang, Yijun Tian, Mingxuan Ju, Zheyuan Liu, Yanfang Ye, Nitesh Chawla, Chuxu Zhang HIGHLIGHT: We identify a fundamental issue in graph adversarial learning and then propose a novel method to enlarge the model capacity and enrich the representation diversity of adversarial samples. 1381, TITLE: What shapes the loss landscape of self supervised learning? https://openreview.net/forum?id=3zSn48RUO8M AUTHORS: Liu Ziyin, Ekdeep Singh Lubana, Masahito Ueda, Hidenori Tanaka We derive an analytically tractable theory of SSL landscape and show that it accurately captures an array of HIGHLIGHT: collapse phenomena and identifies their causes. 1382, TITLE: No Reason for No Supervision: Improved Generalization in Supervised Models https://openreview.net/forum?id=3Y5Uhf5KgGK AUTHORS: Mert Bülent Sar?y?ld?z, Yannis Kalantidis, Karteek Alahari, Diane Larlus HIGHLIGHT: In this paper, we propose a supervised learning setup that leverages the best of both worlds. 1383, TITLE: Simple initialization and parametrization of sinusoidal networks via their kernel bandwidth https://openreview.net/forum?id=yVqC6gCNf4d AUTHORS: Filipe de Avila Belbute-Peres, J Zico Kolter HIGHLIGHT: We perform a theoretical analysis of a simplified sinusoidal network and use this to propose an informed initialization scheme. 1384, TITLE: Linear Convergence of Natural Policy Gradient Methods with Log-Linear Policies https://openreview.net/forum?id=-z9hdsyUwVQ AUTHORS: Rui Yuan, Simon Shaolei Du, Robert M. Gower, Alessandro Lazaric, Lin Xiao HIGHLIGHT: We consider infinite-horizon discounted Markov decision processes and study the convergence rates of the natural policy gradient (NPG) and the Q-NPG methods with the log-linear policy class. 1385, TITLE: EVA3D: Compositional 3D Human Generation from 2D Image Collections https://openreview.net/forum?id=g7U9jD 2CUr AUTHORS: Fangzhou Hong, Zhaoxi Chen, Yushi LAN, Liang Pan, Ziwei Liu HIGHLIGHT: In this work, we propose, EVA3D, an unconditional 3D human generative model learned from 2D image collections only. Nearly Minimax Optimal Offline Reinforcement Learning with Linear Function Approximation: Single-Agent 1386, TITLE: MDP and Markov Game https://openreview.net/forum?id=UP GHHPw7rP AUTHORS: Wei Xiong, Han Zhong, Chengshuai Shi, Cong Shen, Liwei Wang, Tong Zhang HIGHLIGHT: In this paper, we focus on offline RL with linear function approximation and propose two new algorithms, LinPEVI-ADV+ and LinPMVI-ADV+, for single-agent MDPs and two-player zero-sum Markov games (MGs), respectively. 1387, TITLE: Quantile Risk Control: A Flexible Framework for Bounding the Probability of High-Loss Predictions https://openreview.net/forum?id=p6jsTidUkPx Jake Snell, Thomas P Zollo, Zhun Deng, Toniann Pitassi, Richard Zemel AUTHORS: HIGHLIGHT: In this work, we propose a flexible framework to produce a variety of bounds on quantiles of the loss

130

distribution incurred by a predictor.

1388, TITLE: Git Re-Basin: Merging Models modulo Permutation Symmetries

https://openreview.net/forum?id=CQsmMYmlP5T

AUTHORS: Samuel Ainsworth, Jonathan Hayase, Siddhartha Srinivasa

HIGHLIGHT: We argue that neural network loss landscapes contain (nearly) a single basin after accounting for all possible permutation symmetries of hidden units a la Entezari et al. 2021. We introduce three algorithms to permute the units of one model to bring them into alignment with a reference model in order to merge the two models in weight space.

1389. TITLE: The Role of Coverage in Online Reinforcement Learning

https://openreview.net/forum?id=LQIjzPdDt3q

AUTHORS: Tengyang Xie, Dylan J Foster, Yu Bai, Nan Jiang, Sham M. Kakade

HIGHLIGHT: Concretely, we show that coverability---that is, existence of a data distribution that satisfies a ubiquitous coverage condition called concentrability---can be viewed as a structural property of the underlying MDP, and can be exploited by standard algorithms for sample-efficient exploration, even when the agent does not know said distribution.

1390, TITLE: Stateful Active Facilitator: Coordination and Environmental Heterogeneity in Cooperative Multi-Agent Reinforcement Learning

https://openreview.net/forum?id=B4maZQLLW0

AUTHORS: Dianbo Liu, Vedant Shah, Oussama Boussif, Cristian Meo, Anirudh Goyal, Tianmin Shu, Michael Curtis Mozer, Nicolas Heess, Yoshua Bengio

HIGHLIGHT: We formalize the notions of coordination level and heterogeneity level of an environment and present HECOGrid, a suite of multi-agent RL environments that facilitates empirical evaluation of different MARL approaches across different levels of coordination and environmental heterogeneity by providing a quantitative control over coordination and heterogeneity levels of the environment.

1391, TITLE: Learning Achievement Structure for Structured Exploration in Domains with Sparse Reward https://openreview.net/forum?id=NDW19qcUpvy

AUTHORS:

Zihan Zhou, Animesh Garg

We propose Structured Exploration with Achievements (SEA), a multi-stage reinforcement learning algorithm HIGHLIGHT: that learns the environment structure with offline data and uses the learned structure to learn different skills and improve overall exploration with online environment interactions in a particular type of environment that has an internal achievement system.

1392, TITLE: PINTO: Faithful Language Reasoning Using Prompted-Generated Rationales

https://openreview.net/forum?id=WBXbRs63oVu

AUTHORS: PeiFeng Wang, Aaron Chan, Filip Ilievski, Muhao Chen, Xiang Ren

HIGHLIGHT: In this paper, we propose PINTO, an LM pipeline that rationalizes via prompt-based learning, and learns to faithfully reason over rationales via counterfactual regularization.

Excess Risk of Two-Layer ReLU Neural Networks in Teacher-Student Settings and its Superiority to Kernel 1393, TITLE: Methods

https://openreview.net/forum?id=6doXHqwMayf

AUTHORS: Shunta Akiyama, Taiji Suzuki

HIGHLIGHT: While deep learning has outperformed other methods for various tasks, theoretical frameworks that explain its reason have not been fully established. We investigate the excess risk of two-layer ReLU neural networks in a teacher-student regression model, in which a student network learns an unknown teacher network through its outputs.

1394, TITLE: GEASS: Neural causal feature selection for high-dimensional biological data https://openreview.net/forum?id=aKcS3xojnwY

AUTHORS: Mingze Dong, Yuval Kluger

HIGHLIGHT: Here we present GEASS (Granger fEAture Selection of Spatiotemporal data), which identifies sparse Granger causality mechanisms of high dimensional spatiotemporal data by a single neural network.

1395, TITLE: Conditional Positional Encodings for Vision Transformers

https://openreview.net/forum?id=3KWnuT-R1bh

Xiangxiang Chu, Zhi Tian, Bo Zhang, Xinlong Wang, Chunhua Shen AUTHORS:

HIGHLIGHT: We propose a conditional positional encoding (CPE) scheme for vision Transformers.

1396. TITLE: Characterizing intrinsic compositionality in transformers with Tree Projections https://openreview.net/forum?id=sAOOeI878Ns

AUTHORS:

Shikhar Murty, Pratyusha Sharma, Jacob Andreas, Christopher D Manning HIGHLIGHT: We provide a method to functionally project a transformer into the space of tree structured models and use it to uncover intrinsic compositionality of transformers trained on language data.

1397, TITLE: Augmentation Component Analysis: Modeling Similarity via the Augmentation Overlaps https://openreview.net/forum?id=5vM51iamNeL AUTHORS: Lu Han, Han-Jia Ye, De-Chuan Zhan So the augmentation feature, composed of the distribution of augmentations, can act as the ideal embedding, HIGHLIGHT:

and similarity over them reveals how much the augmentations of two samples overlap. Without computational burdens to explicitly estimate its value, we propose Augmentation Component Analysis (ACA) with a contrastive-like loss to learn principal components and an on-the-fly projection loss to embed data.

1398, TITLE: ModelAngelo: Automated Model Building for Cryo-EM Maps https://openreview.net/forum?id=65XDF nwI61 AUTHORS: Kiarash Jamali, Dari Kimanius, Sjors HW Scheres HIGHLIGHT: Recent advances in machine learning applications to protein structure prediction show potential for automating this process. Taking inspiration from these techniques, we have built ModelAngelo for automated model building of proteins in cryo-EM maps. 1399, TITLE: Distilling Cognitive Backdoor within an Image https://openreview.net/forum?id=S3D9NLzjnQ5 AUTHORS: Hanxun Huang, Xingjun Ma, Sarah Monazam Erfani, James Bailey HIGHLIGHT: This paper proposes a simple method to distill and detect backdoor patterns within an image: \emph{Cognitive Distillation} (CD). 3D generation on ImageNet 1400, TITLE: https://openreview.net/forum?id=U2WjB9xxZ9q AUTHORS: Ivan Skorokhodov, Aliaksandr Siarohin, Yinghao Xu, Jian Ren, Hsin-Ying Lee, Peter Wonka, Sergey Tulyakov HIGHLIGHT: In this work, for the first time, we propose a 3D generator which works on non-aligned datasets. 1401, TITLE: Faster Last-iterate Convergence of Policy Optimization in Zero-Sum Markov Games https://openreview.net/forum?id=bRwBpKrNzF7 AUTHORS: Shicong Cen, Yuejie Chi, Simon Shaolei Du, Lin Xiao HIGHLIGHT: In this paper, we focus on the most basic setting of competitive multi-agent RL, namely two-player zero-sum Markov games, and study equilibrium finding algorithms in both the infinite-horizon discounted setting and the finite-horizon episodic setting. 1402, TITLE: Rethinking the Expressive Power of GNNs via Graph Biconnectivity https://openreview.net/forum?id=r9hNv76KoT3 AUTHORS: Bohang Zhang, Shengjie Luo, Liwei Wang, Di He HIGHLIGHT: In this paper, we take a fundamentally different perspective to study the expressive power of GNNs beyond the WL test. 1403, TITLE: Interaction-Based Disentanglement of Entities for Object-Centric World Models https://openreview.net/forum?id=JQc2VowqCzz AUTHORS: Akihiro Nakano, Masahiro Suzuki, Yutaka Matsuo HIGHLIGHT: This paper proposes STEDIE, a new model that disentangles object representations based on interactions, into interaction-relevant relational features and interaction-irrelevant global features without supervision. One Transformer Can Understand Both 2D & 3D Molecular Data 1404, TITLE: https://openreview.net/forum?id=vZTp1oPV3PC AUTHORS: Shengjie Luo, Tianlang Chen, Yixian Xu, Shuxin Zheng, Tie-Yan Liu, Liwei Wang, Di He HIGHLIGHT: We believe a general-purpose neural network model for chemistry should be able to handle molecular tasks across data modalities. To achieve this goal, in this work, we develop a novel Transformer-based Molecular model called Transformer-M, which can take molecular data of 2D or 3D formats as input and generate meaningful semantic representations. 1405, TITLE: Linear Mode Connectivity of Deep Neural Networks via Permutation Invariance and Renormalization https://openreview.net/forum?id=gU5sJ6ZggcX AUTHORS: Keller Jordan, Hanie Sedghi, Olga Saukh, Rahim Entezari, Behnam Neyshabur HIGHLIGHT: In this paper we empirically investigate the conjecture from Entezari et al. (2021) which states that if permutation invariance is taken into account, then there should be no loss barrier to the linear interpolation between SGD solutions. 1406. TITLE: Learning to Compose Soft Prompts for Compositional Zero-Shot Learning https://openreview.net/forum?id=S8-A2FXnIh AUTHORS: Nihal V. Nayak, Peilin Yu, Stephen Bach We introduce compositional soft prompting (CSP), a parameter-efficient learning technique to improve the HIGHLIGHT: zero-shot compositionality of large-scale pretrained vision-language models (VLMs) like CLIP. 1407. TITLE: Diffusion-GAN: Training GANs with Diffusion https://openreview.net/forum?id=HZf7UbpWHuA AUTHORS: Zhendong Wang, Huangjie Zheng, Pengcheng He, Weizhu Chen, Mingyuan Zhou HIGHLIGHT: In this paper, we propose Diffusion-GAN, a novel GAN framework that leverages a forward diffusion chain to generate Gaussian-mixture distributed instance noise. 1408, TITLE: Unsupervised Learning for Combinatorial Optimization Needs Meta Learning https://openreview.net/forum?id=-ENYHCE8zBp AUTHORS: Haoyu Peter Wang, Pan Li Albeit with some advantages over traditional solvers, the current framework optimizes an averaged performance HIGHLIGHT: over the distribution of historical problem instances, which misaligns with the actual goal of CO that looks for a good solution to every future encountered instance. With this observation, we propose a new objective of unsupervised learning for CO where the goal of learning is to search good initialization for future problem instances rather than give direct solutions.

1409, TITLE: Decepticons: Corrupted Transformers Breach Privacy in Federated Learning for Language Models https://openreview.net/forum?id=r0BrY4BiEXO AUTHORS: Liam H Fowl, Jonas Geiping, Steven Reich, Yuxin Wen, Wojciech Czaja, Micah Goldblum, Tom Goldstein HIGHLIGHT: We propose a novel attack that reveals private user text by deploying malicious parameter vectors, and which succeeds even with mini-batches, multiple users, and long sequences. Adaptive Optimization in the \$\infty\$-Width Limit 1410, TITLE: https://openreview.net/forum?id=zgVDqw9ZUES AUTHORS: Etai Littwin, Greg Yang HIGHLIGHT: We derive the infinite width limits of neural networks trained with adaptive optimizers. 1411. TITLE: Sparse Mixture-of-Experts are Domain Generalizable Learners https://openreview.net/forum?id=RecZ9nB9Q4 AUTHORS: Bo Li, Yifei Shen, Jingkang Yang, Yezhen Wang, Jiawei Ren, Tong Che, Jun Zhang, Ziwei Liu HIGHLIGHT: In this paper, we propose to explore an orthogonal direction, i.e., the design of the backbone architecture. 1412, TITLE: Offline Q-learning on Diverse Multi-Task Data Both Scales And Generalizes https://openreview.net/forum?id=4-k7kUavAj AUTHORS: Aviral Kumar, Rishabh Agarwal, Xinyang Geng, George Tucker, Sergey Levine HIGHLIGHT: Drawing on the learnings from these works, we re-examine previous design choices and find that with appropriate choices: ResNets, cross-entropy based distributional backups, and feature normalization, offline Q-learning algorithms exhibit strong performance that scales with model capacity. 1413, TITLE: Diffusion-based Image Translation using disentangled style and content representation https://openreview.net/forum?id=Nayau9fwXU AUTHORS: Gihyun Kwon, Jong Chul Ye Unfortunately, due to the stochastic nature of diffusion models, it is often difficult to maintain the original HIGHLIGHT: content of the image during the reverse diffusion. To address this, here we present a novel diffusion-based image translation method using disentangled style and content representation. 1414, TITLE: Empowering Networks With Scale and Rotation Equivariance Using A Similarity Convolution https://openreview.net/forum?id=NJENsJ37sQ AUTHORS: Zikai Sun, Thierry Blu HIGHLIGHT: In this paper, we devise a method that provides networks with equivariance with respect to translation, rotation, and scaling simultaneously. 1415, TITLE: What learning algorithm is in-context learning? Investigations with linear models https://openreview.net/forum?id=0g0X4H8yN4I Ekin Akyürek, Jacob Andreas, Dale Schuurmans, Tengyu Ma, Denny Zhou AUTHORS: HIGHLIGHT: We prove that the transformers can implement learning algorithms for linear models based e.g gradient descent, then observe they closely match the predictors of known algorithms, transitioning between different predictors as transformer depth vary. 1416, TITLE: Enhancing Meta Learning via Multi-Objective Soft Improvement Functions https://openreview.net/forum?id=hCmjBJeGXcu AUTHORS: Runsheng Yu, Weiyu Chen, Xinrun Wang, James Kwok HIGHLIGHT: However, existing MOO solvers need to access all the objectives' gradients in each iteration, and cannot scale to the huge number of tasks in typical meta-learning settings. To alleviate this problem, we propose a scalable gradient-based solver with the use of mini-batch. 1417 TITLE: \$\mathrm{SE}(3)\$-Equivariant Attention Networks for Shape Reconstruction in Function Space https://openreview.net/forum?id=RDy3IbvjMqT AUTHORS: Evangelos Chatzipantazis, Stefanos Pertigkiozoglou, Edgar Dobriban, Kostas Daniilidis HIGHLIGHT: We propose a method for 3D shape reconstruction from unoriented point clouds. 1418, TITLE: Understanding Ensemble, Knowledge Distillation and Self-Distillation in Deep Learning https://openreview.net/forum?id=Uuf2q9TfXGA Zeyuan Allen-Zhu, Yuanzhi Li AUTHORS: We show that ensemble/knowledge distillation in \emph{deep learning} works very differently from traditional HIGHLIGHT: learning theory (such as boosting or NTKs). We develop a theory showing that when data has a structure we refer to as "multi-view", then ensemble of independently trained neural networks can provably improve test accuracy, and such superior test accuracy can also be provably distilled into a single model. 1419, TITLE: How Can GANs Learn Hierarchical Generative Models for Real-World Distributions https://openreview.net/forum?id=7h5KSs2PCRi Zeyuan Allen-Zhu, Yuanzhi Li AUTHORS: In this work, we formally study how GANs can efficiently learn certain hierarchically generated distributions HIGHLIGHT: that are close to the distribution of real-life images.

1420, TITLE: Spotlight: Mobile UI Understanding using Vision-Language Models with a Focus

https://openreview.ne	t/forum?id=9yE2xEj0BH7
AUTHORS:	Gang Li, Yang Li
HIGHLIGHT:	In this paper, we propose Spotlight, a vision-only approach for mobile UI understanding.
1421, TITLE:	A Control-Centric Benchmark for Video Prediction
https://openreview.net	et/forum?id=rimcq1oIFeR
AUTHORS:	Stephen Tian, Chelsea Finn, Jiajun Wu
HIGHLIGHT:	We find empirically that for planning robotic manipulation, existing metrics can be unreliable at predicting
execution success. To	o address this, we propose a benchmark for action-conditioned video prediction in the form of a control
benchmark that evalue	nates a given model for simulated robotic manipulation through sampling-based planning.
1422, TITLE: https://openreview.ne AUTHORS: HIGHLIGHT: SGD and Adam.	Heavy-tailed Noise Does Not Explain the Gap Between SGD and Adam, but Sign Descent Might et/forum?id=a65YK0cqH8g Frederik Kunstner, Jacques Chen, Jonathan Wilder Lavington, Mark Schmidt We find evidence that stochasticity and heavy-tailed noise are not major factors in the performance gap between
1423, TITLE:	Building Normalizing Flows with Stochastic Interpolants
https://openreview.ne	et/forum?id=li7qeBbCR1t
AUTHORS:	Michael Samuel Albergo, Eric Vanden-Eijnden
HIGHLIGHT:	A simple generative model based on a continuous-time normalizing flow between any pair of base and target
distributions is propo	sed.
1424, TITLE:	Dual Student Networks for Data-Free Model Stealing
https://openreview.net	tt/forum?id=VE1s3e5xriA
AUTHORS:	James Beetham, Navid Kardan, Ajmal Saeed Mian, Mubarak Shah
HIGHLIGHT:	We propose a Dual Student method where two students are symmetrically trained in order to provide the
generator a criterion	to generate samples that the two students disagree on.
1425, TITLE:	Competitive Physics Informed Networks
https://openreview.net	et/forum?id=z9SIj-IM7tn
AUTHORS:	Qi Zeng, Yash Kothari, Spencer H Bryngelson, Florian Tobias Schaefer
HIGHLIGHT:	This strategy is called "physics-informed neural networks" (PINNs), but it currently cannot produce high-
accuracy solutions, ty	vpically attaining about \$0.1\%\$ relative error. We present an adversarial approach that overcomes this limitation,
which we call compe	titive PINNs (CPINNs).
1426, TITLE:	Energy-Inspired Self-Supervised Pretraining for Vision Models
https://openreview.net	et/forum?id=ZMz-sW6gCLF
AUTHORS:	Ze Wang, Jiang Wang, Zicheng Liu, Qiang Qiu
HIGHLIGHT:	Motivated by the fact that forward and backward passes of a deep network naturally form symmetric mappings
between input and ou	tiput representations, we introduce a simple yet effective self-supervised vision model pretraining framework
inspired by energy-ba	ased models (EBMs).
1427, TITLE:	Effectively Modeling Time Series with Simple Discrete State Spaces
https://openreview.net	et/forum?id=2EpjkjzdCAa
AUTHORS:	Michael Zhang, Khaled Kamal Saab, Michael Poli, Tri Dao, Karan Goel, Christopher Re
HIGHLIGHT:	For efficient training and inference, we introduce an algorithm that reduces the memory and compute of a
forward pass with the	e companion matrix.
1428, TITLE: Systems https://openreview.net AUTHORS: HIGHLIGHT: learning-to-rank liter:	Fantastic Rewards and How to Tame Them: A Case Study on Reward Learning for Task-Oriented Dialogue et/forum?id=086pmarAris Yihao Feng, Shentao Yang, Shujian Zhang, Jianguo Zhang, Caiming Xiong, Mingyuan Zhou, Huan Wang Specifically, we introduce two generalized objectives for reward-function learning, inspired from the classical ature.
1429, TITLE: https://openreview.net AUTHORS: HIGHLIGHT: leverages supervision kernel.	Supervision Complexity and its Role in Knowledge Distillation t/forum?id=8jU7wy7N7mA Hrayr Harutyunyan, Ankit Singh Rawat, Aditya Krishna Menon, Seungyeon Kim, Sanjiv Kumar In order to study the generalization behavior of a distilled student, we propose a new theoretical framework that a complexity: a measure of alignment between teacher-provided supervision and the student's neural tangent
1430, TITLE:	Transferable Unlearnable Examples
https://openreview.ne	st/forum?id=-htnolWDLvP
AUTHORS:	Jie Ren, Han Xu, Yuxuan Wan, Xingjun Ma, Lichao Sun, Jiliang Tang
HIGHLIGHT:	However, their unlearnable effects significantly decrease when used in other training settings and datasets. To

tackle this issue, we propose a novel unlearnable strategy based on Clustering Separability Discriminant (CSD), which aims to better transfer the unlearnable effects to other training settings and datasets by enhancing the linear separability.

1431, TITLE: Random Laplacian Features for Learning with Hyperbolic Space https://openreview.net/forum?id=3pfNb4pZBNp AUTHORS: Tao Yu, Christopher De Sa HIGHLIGHT: In this paper, we propose a simpler approach: learn a hyperbolic embedding of the input, then map once from it to Euclidean space using a mapping that encodes geometric priors by respecting the isometries of hyperbolic space, and finish with a standard Euclidean network. 1432, TITLE: Neural Causal Models for Counterfactual Identification and Estimation https://openreview.net/forum?id=vouQcZS8KfW AUTHORS: Kevin Muyuan Xia, Yushu Pan, Elias Bareinboim HIGHLIGHT: In this paper, we study the evaluation of counterfactual statements through neural models. 1433, TITLE: SIMPLE: A Gradient Estimator for k-Subset Sampling https://openreview.net/forum?id=GPJVuyX4p h AUTHORS: kareem ahmed, Zhe Zeng, Mathias Niepert, Guy Van den Broeck In this work, we fall back to discrete \$k\$-subset sampling on the forward pass. HIGHLIGHT: 1434, TITLE: Learning Iterative Neural Optimizers for Image Steganography https://openreview.net/forum?id=gLPkzWjdhBN Varsha Kishore, Xiangyu Chen, Kilian Q Weinberger AUTHORS: HIGHLIGHT: In this paper, we argue that image steganography is inherently performed on the (elusive) manifold of natural images, and propose to train an iterative neural network to perform the optimization steps. 1435, TITLE: Confidence-Conditioned Value Functions for Offline Reinforcement Learning https://openreview.net/forum?id=Zeb5mTuqT5 AUTHORS: Joey Hong, Aviral Kumar, Sergey Levine To do so, in this work, we propose learning value functions that additionally condition on the degree of HIGHLIGHT: conservatism, which we dub confidence-conditioned value functions. 1436, TITLE: On the Sensitivity of Reward Inference to Misspecified Human Models https://openreview.net/forum?id=hJqGbUpDGV AUTHORS: Joey Hong, Kush Bhatia, Anca Dragan HIGHLIGHT: We investigate the impact of assuming wrong human models on reward learning. 1437, TITLE: How Much Data Are Augmentations Worth? An Investigation into Scaling Laws, Invariance, and Implicit Regularization https://openreview.net/forum?id=3aQs3MCSexD AUTHORS: Jonas Geiping, Micah Goldblum, Gowthami Somepalli, Ravid Shwartz-Ziv, Tom Goldstein, Andrew Gordon Wilson HIGHLIGHT: In this paper, we disentangle several key mechanisms through which data augmentations operate. 1438, TITLE: Fundamental limits on the robustness of image classifiers https://openreview.net/forum?id=gpmL0D4VjN4 AUTHORS: Zheng Dai, David Gifford HIGHLIGHT: We prove that image classifiers are fundamentally sensitive to small perturbations in their inputs. 1439, TITLE: Evolve Smoothly, Fit Consistently: Learning Smooth Latent Dynamics For Advection-Dominated Systems https://openreview.net/forum?id=Z4s73sJYQM AUTHORS: Zhong Yi Wan, Leonardo Zepeda-Nunez, Anudhyan Boral, Fei Sha HIGHLIGHT: We present a data-driven, space-time continuous framework to learn surrogate models for complex physical systems described by advection-dominated partial differential equations. 1440, TITLE: Understanding Influence Functions and Datamodels via Harmonic Analysis https://openreview.net/forum?id=cxCEOSF99f AUTHORS: Nikunj Saunshi, Arushi Gupta, Mark Braverman, Sanjeev Arora HIGHLIGHT: This paper establishes connections between datamodels, influence functions and Fourier coefficients using theoretical tools from harmonic analysis of Boolean functions. 1441, TITLE: BC-IRL: Learning Generalizable Reward Functions from Demonstrations https://openreview.net/forum?id=Ovnwe sDQW AUTHORS: Andrew Szot, Amy Zhang, Dhruv Batra, Zsolt Kira, Franziska Meier We introduce BC-IRL a new inverse reinforcement learning method that learns reward functions that generalize HIGHLIGHT: better when compared to maximum-entropy IRL approaches. TextGrad: Advancing Robustness Evaluation in NLP by Gradient-Driven Optimization 1442, TITLE: https://openreview.net/forum?id=5tKXUZil3X

AUTHORS:Bairu Hou, Jinghan Jia, Yihua Zhang, Guanhua Zhang, Yang Zhang, Sijia Liu, Shiyu ChangHIGHLIGHT:To bridge the gap, we propose TextGrad, a new attack generator using gradient-driven optimization, supportinghigh-accuracy and high-quality assessment of adversarial robustness in NLP.

1443, TITLE: Information Plane Analysis for Dropout Neural Networks https://openreview.net/forum?id=bQB6qozaBw AUTHORS: Linara Adilova, Bernhard C Geiger, Asja Fischer HIGHLIGHT: In this work we show how the stochasticity induced by dropout layers can be utilized to estimate MI in a theoretically sound manner. Greedy Actor-Critic: A New Conditional Cross-Entropy Method for Policy Improvement 1444, TITLE: https://openreview.net/forum?id=eSQh8rG8Oa AUTHORS: Samuel Neumann, Sungsu Lim, Ajin George Joseph, Yangchen Pan, Adam White, Martha White HIGHLIGHT: In this work, we explore an alternative update for the actor, based on an extension of the cross entropy method (CEM) to condition on inputs (states). Characteristic Neural Ordinary Differential Equation 1445. TITLE: https://openreview.net/forum?id=loIfC8WHevK AUTHORS: Xingzi Xu, Ali Hasan, Khalil Elkhalil, Jie Ding, Vahid Tarokh HIGHLIGHT: We propose Characteristic-Neural Ordinary Differential Equations (C-NODEs), a framework for extending Neural Ordinary Differential Equations (NODEs) beyond ODEs. 1446, TITLE: Fast Sampling of Diffusion Models with Exponential Integrator https://openreview.net/forum?id=Loek7hfb46P AUTHORS: Qinsheng Zhang, Yongxin Chen HIGHLIGHT: Our goal is to develop a fast sampling method for DMs with a much less number of steps while retaining high sample quality. gDDIM: Generalized denoising diffusion implicit models 1447, TITLE: https://openreview.net/forum?id=1hKE9gjvz-Qinsheng Zhang, Molei Tao, Yongxin Chen AUTHORS: HIGHLIGHT: Our goal is to extend the denoising diffusion implicit model (DDIM) to general diffusion models~(DMs) besides isotropic diffusions. 1448, TITLE: Panning for Gold in Federated Learning: Targeted Text Extraction under Arbitrarily Large-Scale Aggregation https://openreview.net/forum?id=A9WQaxYsfx AUTHORS: Hong-Min Chu, Jonas Geiping, Liam H Fowl, Micah Goldblum, Tom Goldstein HIGHLIGHT: In this work, we propose the first attack on FL that achieves targeted extraction of sequences that contain privacy-critical phrases, whereby we employ maliciously modified parameters to allow the transformer itself to filter relevant sequences from aggregated user data and encode them in the gradient update. Artificial Neuronal Ensembles with Learned Context Dependent Gating 1449, TITLE: https://openreview.net/forum?id=dBk3hsg-n6 AUTHORS: Matthew James Tilley, Michelle Miller, David Freedman HIGHLIGHT: We introduce Learned Context Dependent Gating (LXDG), a method to flexibly allocate and recall 'artificial neuronal ensembles', using a particular network structure and a new set of regularization terms. Learning Language Representations with Logical Inductive Bias 1450, TITLE: https://openreview.net/forum?id=rGeZuBRahju AUTHORS: Jianshu Chen HIGHLIGHT: In this paper, we seek to go further beyond and explore a new logical inductive bias for better language representation learning. 1451, TITLE: How Does Self-supervised Learning Work? A Representation Learning Perspective https://openreview.net/forum?id=Dzmd-Cc8OI AUTHORS: Yiwen Kou, Zixiang Chen, Yuan Cao, Quanquan Gu HIGHLIGHT: In this paper, we aim to theoretically understand a special kind of SSL approaches based on pre-training and fine-tuning. 1452, TITLE: Provable Robustness against Wasserstein Distribution Shifts via Input Randomization https://openreview.net/forum?id=HJFVrpCaGE AUTHORS: Aounon Kumar, Alexander Levine, Tom Goldstein, Soheil Feizi HIGHLIGHT: In this work, we present provable robustness guarantees on the accuracy of a model under bounded Wasserstein shifts of the data distribution. Denoising Diffusion Samplers 1453, TITLE: https://openreview.net/forum?id=8pvnfTAbu1f AUTHORS: Francisco Vargas, Will Sussman Grathwohl, Arnaud Doucet HIGHLIGHT: We explore here a similar idea to sample approximately from unnormalized probability density functions and estimate their normalizing constants. 1454, TITLE: How I Learned to Stop Worrying and Love Retraining https://openreview.net/forum?id= nF5imFKQI AUTHORS: Max Zimmer, Christoph Spiegel, Sebastian Pokutta

HIGHLIGHT: Going a step further, we propose similarly imposing a budget on the initial dense training phase and show that the resulting simple and efficient method is capable of outperforming significantly more complex or heavily parameterized state-of-the-art approaches that attempt to sparsify the network during training.

1455. TITLE: GOGGLE: Generative Modelling for Tabular Data by Learning Relational Structure https://openreview.net/forum?id=fPVRcJqspu AUTHORS: Tennison Liu, Zhaozhi Qian, Jeroen Berrevoets, Mihaela van der Schaar HIGHLIGHT: In this work, we learn and exploit relational structure underlying tabular data to better model variable dependence, and as a natural means to introduce regularization on relationships and include prior knowledge. 1456, TITLE: Progressive Prompts: Continual Learning for Language Models without Forgetting https://openreview.net/forum?id=UJTgQBc91 AUTHORS: Anastasia Razdaibiedina, Yuning Mao, Rui Hou, Madian Khabsa, Mike Lewis, Amjad Almahairi HIGHLIGHT: We introduce Progressive Prompts - a simple and efficient approach for continual learning in language models. 1457, TITLE: Deep Learning From Crowdsourced Labels: Coupled Cross-Entropy Minimization, Identifiability, and Regularization https://openreview.net/forum?id= qVhsWyWB9 AUTHORS: Shahana Ibrahim, Tri Nguyen, Xiao Fu HIGHLIGHT: Our analysis reveals for the first time that the CCEM can indeed correctly identify the annotators' confusion characteristics and the desired "ground-truth" neural classifier under realistic conditions, e.g., when only incomplete annotator labeling and finite samples are available. 1458, TITLE: Projective Proximal Gradient Descent for Nonconvex Nonsmooth Optimization: Fast Convergence Without Kurdyka-Lojasiewicz (KL) Property https://openreview.net/forum?id=yEsj8pGN11 AUTHORS: Yingzhen Yang, Ping Li In this paper, we propose Projected Proximal Gradient Descent (PPGD) which solves a class of nonconvex and HIGHLIGHT: nonsmooth optimization problems, where the nonconvexity nd nonsmoothness come from a nonsmooth regularization term which is nonconvex but piecewise convex. 1459, TITLE: First Steps Toward Understanding the Extrapolation of Nonlinear Models to Unseen Domains https://openreview.net/forum?id=7wrq3vHcMM AUTHORS: Kefan Dong, Tengyu Ma HIGHLIGHT: We prove that the family of nonlinear models of the form $f(x) = \int \frac{1}{x^2} \int \frac{1}{x^2} dx$, where f_i is an \emph{arbitrary} function on the subset of features \$x i\$, can extrapolate to unseen distributions, if the covariance of the features is well-conditioned. 1460, TITLE: Variable Compositionality Reliably Emerges in Neural Networks https://openreview.net/forum?id=-Yzz6v1X7V-AUTHORS: Henry Conklin, Kenny Smith HIGHLIGHT: We introduce a variation-based framework for interpreting the mappings produced by neural networks in emergent communication games and find that they reliably exhibit straight-forward compositional structure, with a degree of natural language-like variation that obscures their compositionality under measures used in previous work. 1461, TITLE: Multiple sequence alignment as a sequence-to-sequence learning problem https://openreview.net/forum?id=8efJYMBrNb AUTHORS: Edo Dotan, Yonatan Belinkov, Oren Avram, Elya Wygoda, Noa Ecker, Michael Alburquerque, Omri Keren, Gil Loewenthal, Tal Pupko HIGHLIGHT: Here we introduce BetaAlign, a novel methodology for aligning sequences using a natural language processing (NLP) approach. 1462, TITLE: A Mixture-of-Expert Approach to RL-based Dialogue Management https://openreview.net/forum?id=4FBUihxz5nm AUTHORS: Yinlam Chow, Azamat Tulepbergenov, Ofir Nachum, Dhawal Gupta, Moonkyung Ryu, Mohammad Ghavamzadeh, Craig Boutilier HIGHLIGHT: We use reinforcement learning (RL) to develop a dialogue agent that avoids being short-sighted (outputting generic utterances) and maximizes overall user satisfaction. 1463, TITLE: f-DM: A Multi-stage Diffusion Model via Progressive Signal Transformation https://openreview.net/forum?id=iBdwKIsg4m Jiatao Gu, Shuangfei Zhai, Yizhe Zhang, Miguel Ángel Bautista, Joshua M. Susskind AUTHORS: HIGHLIGHT: In this work, we propose f-DM, a generalized family of DMs which allows progressive signal transformation. 1464, TITLE: Progressive Mix-Up for Few-Shot Supervised Multi-Source Domain Transfer https://openreview.net/forum?id=H7M 5K5qKJV AUTHORS: Ronghang Zhu, Ronghang Zhu, Xiang Yu, Sheng Li HIGHLIGHT: The multi-source setting further prevents the transfer task as excessive domain gap introduced from all the

source domains. To tackle this problem, we newly propose a progressive mix-up (P-Mixup) mechanism to introduce an intermediate mix-up domain, pushing both the source domains and the few-shot target domain aligned to this mix-up domain.

1465. TITLE: Efficient approximation of neural population structure and correlations with probabilistic circuits https://openreview.net/forum?id=XC yGI-0j9 AUTHORS: Koosha Khalvati, Samantha Johnson, Stefan Mihalas, Michael A Buice HIGHLIGHT: We present a computationally efficient framework to model a wide range of population structures with high order correlations and a large number of neurons. 1466, TITLE: Exploring perceptual straightness in learned visual representations https://openreview.net/forum?id=4cOfD2qL6T AUTHORS: Anne Harrington, Vasha DuTell, Ayush Tewari, Mark Hamilton, Simon Stent, Ruth Rosenholtz, William T. Freeman HIGHLIGHT: In this paper, we explore the relationship between network architecture, robustness, biologically-inspired filtering mechanisms, and representational straightness in response to time-varying input; we identify curvature as a useful way of evaluating neural network representations. 1467. TITLE: Is Forgetting Less a Good Inductive Bias for Forward Transfer? https://openreview.net/forum?id=dL351x-mTEs AUTHORS: Jiefeng Chen, Arslan Chaudhry, Timothy Nguyen, Dilan Gorur HIGHLIGHT: We argue that the measure of forward transfer to a task should not be affected by the restricted updates on the task by the continual learner to preserve previous tasks. Learning Structured Representations by Embedding Class Hierarchy 1468. TITLE: https://openreview.net/forum?id=7J-30ilaUZM AUTHORS: Sigi Zeng, Remi Tachet des Combes, Han Zhao HIGHLIGHT: We propose to learn structured representations that preserve the hierarchy between label classes by using CPCC as a regularizer. TypeT5: Seq2seq Type Inference using Static Analysis 1469, TITLE: https://openreview.net/forum?id=4TyNEhI2GdN AUTHORS: Jiavi Wei, Greg Durrett, Isil Dillig In this paper, we present a new type inference method that treats type prediction as a code completion task by HIGHLIGHT: leveraging CodeT5, a state-of-the-art seq2seq pre-trained language model for code. Dichotomy of Control: Separating What You Can Control from What You Cannot 1470, TITLE: https://openreview.net/forum?id=DEGjDDV22pI AUTHORS: Sherry Yang, Dale Schuurmans, Pieter Abbeel, Ofir Nachum In this work, we propose the dichotomy of control (DoC), a future-conditioned supervised learning framework HIGHLIGHT: that separates mechanisms within a policy's control (actions) from those outside of a policy's control (environment stochasticity). Revisiting Curiosity for Exploration in Procedurally Generated Environments 1471, TITLE: https://openreview.net/forum?id=j3GK3_xZydY AUTHORS: Kaixin Wang, Kuangqi Zhou, Bingyi Kang, Jiashi Feng, Shuicheng YAN HIGHLIGHT: We consider lifelong and episodic curiosities used in prior works, and compare the performance of all lifelongepisodic combinations on the commonly used MiniGrid benchmark. 1472, TITLE: Can Neural Networks Learn Implicit Logic from Physical Reasoning? https://openreview.net/forum?id=HVoJCRLBvVk AUTHORS: Aaron Traylor, Roman Feiman, Ellie Pavlick HIGHLIGHT: We test the hypothesis that neural networks without inherent inductive biases for logical reasoning can acquire an implicit representation of negation and disjunction. 1473, TITLE: ESCHER: Eschewing Importance Sampling in Games by Computing a History Value Function to Estimate Regret https://openreview.net/forum?id=35QyoZv8cKO AUTHORS: Stephen Marcus McAleer, Gabriele Farina, Marc Lanctot, Tuomas Sandholm HIGHLIGHT: In this paper we propose an unbiased model-free method that does not require any importance sampling. 1474, TITLE: Serving Graph Compression for Graph Neural Networks https://openreview.net/forum?id=T-qVtA3pAxG AUTHORS: Si Si, Felix Yu, Ankit Singh Rawat, Cho-Jui Hsieh, Sanjiv Kumar HIGHLIGHT: In this paper, we study graph compression to reduce the storage requirement for GNN in serving. 1475, TITLE: On The Specialization of Neural Modules https://openreview.net/forum?id=Fh97BDaR6I Devon Jarvis, Richard Klein, Benjamin Rosman, Andrew M Saxe AUTHORS: HIGHLIGHT: To this end we introduce a minimal space of datasets motivated by practical systematic generalization benchmarks.

1476, TITLE: HomoDistil: Homotopic Task-Agnostic Distillation of Pre-trained Transformers https://openreview.net/forum?id=D7srTrGhAs AUTHORS: Chen Liang, Haoming Jiang, Zheng Li, Xianfeng Tang, Bing Yin, Tuo Zhao HIGHLIGHT: Such a large prediction discrepancy often diminishes the benefits of knowledge distillation. To address this challenge, we propose Homotopic Distillation (HomoDistil), a novel task-agnostic distillation approach equipped with iterative pruning. 1477. TITLE: FIGARO: Controllable Music Generation using Learned and Expert Features https://openreview.net/forum?id=NyR8OZFHw6i Dimitri von Rütte, Luca Biggio, Yannic Kilcher, Thomas Hofmann AUTHORS: HIGHLIGHT: In this work, we release FIGARO, a Transformer-based conditional model trained to generate symbolic music based on a sequence of high-level control codes. 1478, TITLE: Language models are multilingual chain-of-thought reasoners https://openreview.net/forum?id=fR3wGCk-IXp AUTHORS: Freda Shi, Mirac Suzgun, Markus Freitag, Xuezhi Wang, Suraj Srivats, Soroush Vosoughi, Hyung Won Chung, Yi Tay, Sebastian Ruder, Denny Zhou, Dipanjan Das, Jason Wei HIGHLIGHT: We introduce the Multilingual Grade School Math (MGSM) benchmark, by manually translating 250 gradeschool math problems from the GSM8K dataset (Cobbe et al., 2021) into ten typologically diverse languages. 1479, TITLE: Is Reinforcement Learning (Not) for Natural Language Processing?: Benchmarks, Baselines, and Building Blocks for Natural Language Policy Optimization https://openreview.net/forum?id=8aHzds2uUyB AUTHORS: Rajkumar Ramamurthy, Prithviraj Ammanabrolu, Kianté Brantley, Jack Hessel, Rafet Sifa, Christian Bauckhage, Hannaneh Hajishirzi, Yejin Choi HIGHLIGHT: We tackle the problem of aligning pre-trained large language models (LMs) with human preferences. 1480, TITLE: Learning multi-scale local conditional probability models of images https://openreview.net/forum?id=VZX2I VVJKH Zahra Kadkhodaie, Florentin Guth, Stéphane Mallat, Eero P Simoncelli AUTHORS: But these models are implicit, and the means by which these networks capture complex global statistical HIGHLIGHT: structure, apparently without suffering from the curse of dimensionality, remain a mystery. To study this, we generalize a multi-scale model class motivated by the renormalization group of theoretical physics. 1481, TITLE: Subsampling in Large Graphs Using Ricci Curvature https://openreview.net/forum?id=w9WUQkBvpI AUTHORS: Shushan Wu, Huimin Cheng, Jiazhang Cai, Ping Ma, Wenxuan Zhong HIGHLIGHT: Based on the asymptotic results about the within-community edge and between-community edge's OR curvature, we propose a subsampling algorithm based on our theoretical results, the Ollivier-Ricci curvature Gradient-based subsampling (ORG-sub) algorithm. 1482, TITLE: Universal Few-shot Learning of Dense Prediction Tasks with Visual Token Matching https://openreview.net/forum?id=88nT0j5jAn AUTHORS: Donggyun Kim, Jinwoo Kim, Seongwoong Cho, Chong Luo, Seunghoon Hong We propose Visual Token Matching (VTM), a universal few-shot learner for arbitrary dense prediction tasks. HIGHLIGHT: Scaling up and Stabilizing Differentiable Planning with Implicit Differentiation 1483. TITLE: https://openreview.net/forum?id=PYbe4MoHf32 AUTHORS: Linfeng Zhao, Huazhe Xu, Lawson L.S. Wong HIGHLIGHT: However, an issue prevents it from scaling up to larger-scale problems: they need to differentiate through forward iteration layers to compute gradients, which couples forward computation and backpropagation and needs to balance forward planner performance and computational cost of the backward pass. To alleviate this issue, we propose to differentiate through the Bellman fixed-point equation to decouple forward and backward passes for Value Iteration Network and its variants, which enables constant backward cost (in planning horizon) and flexible forward budget and helps scale up to large tasks. 1484, TITLE: Score-based Continuous-time Discrete Diffusion Models https://openreview.net/forum?id=BYWWwSY2G5s AUTHORS: Haoran Sun, Lijun Yu, Bo Dai, Dale Schuurmans, Hanjun Dai HIGHLIGHT: In this paper, we extend diffusion models to discrete variables by introducing a stochastic jump process where the reverse process denoises via a continuous-time Markov chain. 1485, TITLE: Is Model Ensemble Necessary? Model-based RL via a Single Model with Lipschitz Regularized Value Function https://openreview.net/forum?id=hNyJBk3CwR Ruijie Zheng, Xiyao Wang, Huazhe Xu, Furong Huang AUTHORS: HIGHLIGHT: In this paper, we provide both practical and theoretical insights on the empirical success of the probabilistic dynamics model ensemble through the lens of Lipschitz continuity. Disentangling with Biological Constraints: A Theory of Functional Cell Types 1486, TITLE: https://openreview.net/forum?id=9Z GfhZnGH AUTHORS: James C. R. Whittington, Will Dorrell, Surya Ganguli, Timothy Behrens HIGHLIGHT: Moreover, such disentangled representations are highly sought after in machine learning. Here we mathematically prove that simple biological constraints on neurons, namely nonnegativity and energy efficiency in both activity and

weights, promote such sought after disentangled representations by enforcing neurons to become selective for single factors of task variation.

1487, TITLE: Learning rigid dynamics with face interaction graph networks

https://openreview.net/forum?id=J7Uh781A05p

AUTHORS: Kelsey R Allen, Yulia Rubanova, Tatiana Lopez-Guevara, William F Whitney, Alvaro Sanchez-Gonzalez, Peter Battaglia, Tobias Pfaff

HIGHLIGHT: Here we introduce the ``Face Interaction Graph Network" (FIGNet) which extends beyond GNN-based methods, and computes interactions between mesh faces, rather than nodes.

1488, TITLE: Images as Weight Matrices: Sequential Image Generation Through Synaptic Learning Rules

https://openreview.net/forum?id=ddad0PNUvV

AUTHORS: Kazuki Irie, Jürgen Schmidhuber

HIGHLIGHT: We train neural nets to execute sequences of synaptic learning rules to sequentially generate natural images (instead of weight matrices).

1489, TITLE: Why (and When) does Local SGD Generalize Better than SGD?

https://openreview.net/forum?id=svCcui6Drl

AUTHORS: Xinran Gu, Kaifeng Lyu, Longbo Huang, Sanjeev Arora

HIGHLIGHT: The main contributions of this paper include (i) the derivation of an SDE that captures the long-term behavior of Local SGD with a small learning rate, after approaching the manifold of minima, (ii) a comparison between the SDEs of Local SGD and SGD, showing that Local SGD induces a stronger drift term that can result in a stronger effect of regularization, e.g., a faster reduction of sharpness, and (iii) empirical evidence validating that having small learning rate and long enough training time enables the generalization improvement over SGD but removing either of the two conditions leads to no improvement.

1490, TITLE: Depth Separation with Multilayer Mean-Field Networks

https://openreview.net/forum?id=uzFQpkEzOo

AUTHORS: Yunwei Ren, Mo Zhou, Rong Ge

HIGHLIGHT: Previous results often focus on representation power, for example, Safran et al. (2019) constructed a function that is easy to approximate using a 3-layer network but not approximable by any 2-layer network. In this paper, we show that this separation is in fact algorithmic: one can learn the function constructed by Safran et al. (2019) using an overparametrized network with polynomially many neurons ef?ciently.

1491, TITLE: Analogical Networks for Memory-Modulated 3D Parsing

https://openreview.net/forum?id=SRIQZTh0IK

AUTHORS:Nikolaos Gkanatsios, Mayank Singh, Zhaoyuan Fang, Shubham Tulsiani, Katerina FragkiadakiHIGHLIGHT:We present Analogical Networks, a model that casts fine-grained visual parsing into analogical inference:instead of mapping input scenes to part labels, which is hard to adapt in a few-shot manner to novel inputs, our model retrieves relatedscenes from memory and their corresponding part structures, and predicts analogous part structures in the input scene, via an end-to-end learnable modulation mechanism.

 1492, TITLE:
 Beyond Lipschitz: Sharp Generalization and Excess Risk Bounds for Full-Batch GD

 https://openreview.net/forum?id=pOyi9KqE56b

 AUTHORS:
 Konstantinos Nikolakakis, Farzin Haddadpour, Amin Karbasi, Dionysios Kalogerias

 HIGHLIGHT:
 We provide sharp path-dependent generalization and excess risk guarantees for the full-batch Gradient Descent

 (GD) algorithm on smooth losses (possibly non-Lipschitz, possibly nonconvex).

 1493, TITLE:
 GNNDelete: A General Unlearning Strategy for Graph Neural Networks

 https://openreview.net/forum?id=X9yCkmT5Qrl

 AUTHORS:
 Jiali Cheng, George Dasoulas, Huan He, Chirag Agarwal, Marinka Zitnik

 HIGHLIGHT:
 We consider the problem of graph unlearning, wherein graph neural network (GNN) model is trained to

 specified accuracy and then deployed while a sequence of requests arrives to delete graph elements (nodes, edges) from the model.

1494, TITLE: ReAct: Synergizing Reasoning and Acting in Language Models

https://openreview.net/forum?id=WE vluYUL-X

AUTHORS: Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik R Narasimhan, Yuan Cao HIGHLIGHT: In this paper, we explore the use of LLMs to generate both reasoning traces and task-specific actions in an interleaved manner, allowing for greater synergy between the two: reasoning traces help the model induce, track, and update action plans as well as handle exceptions, while actions allow it to interface with external sources, such as knowledge bases or environments, to gather additional information.

 1495, TITLE:
 Towards convergence to Nash equilibria in two-team zero-sum games

 https://openreview.net/forum?id=4BPFwvKOvo5

 AUTHORS:
 Fivos Kalogiannis, Ioannis Panageas, Emmanouil-Vasileios Vlatakis-Gkaragkounis

 HIGHLIGHT:
 On a brighter note, we propose a first-order method that leverages control theory techniques and under some conditions enjoys last-iterate local convergence to a Nash equilibrium.

1496, TITLE: DensePure: Understanding Diffusion Models towards Adversarial Robustness https://openreview.net/forum?id=p7hvOJ6Gq0i

AUTHORS: Zhongzhu Chen, Kun Jin, Chaowei Xiao, Jiongxiao Wang, Weili Nie, Mingyan Liu, Anima Anandkumar, Bo Li, Dawn Song HIGHLIGHT: However, the theoretical understanding of why diffusion models are able to improve the certified robustness is still lacking, preventing from further improvement. In this study, we close this gap by analyzing the fundamental properties of diffusion models and establishing the conditions under which they can enhance certified robustness. Where to Diffuse, How to Diffuse and How to get back: Learning in Multivariate Diffusions 1497, TITLE: https://openreview.net/forum?id=osei3IzUia AUTHORS: Raghav Singhal, Mark Goldstein, Rajesh Ranganath HIGHLIGHT: In this work, we study linear Multivariate Diffusion Models (MDMs). 1498, TITLE: Spatio-temporal point processes with deep non-stationary kernels https://openreview.net/forum?id=PsIk0kO3hKd AUTHORS: Zheng Dong, Xiuyuan Cheng, Yao Xie HIGHLIGHT: This paper develops a deep non-stationary influence kernel for spatio-temporal point processes with a novel parameterization that enables us to well approximate complicated kernels in a low-rank form. 1499, TITLE: Scalable Batch-Mode Deep Bayesian Active Learning via Equivalence Class Annealing https://openreview.net/forum?id=GRZtigJljLY AUTHORS: Renyu Zhang, Aly A Khan, Robert L. Grossman, Yuxin Chen HIGHLIGHT: In this paper, we propose Batch-BALanCe, a scalable batch-mode active learning algorithm, which combines insights from decision-theoretic active learning, combinatorial information measure, and diversity sampling. 1500, TITLE: Explicitly Minimizing the Blur Error of Variational Autoencoders https://openreview.net/forum?id=9krnQ-ue9M AUTHORS: Gustav Bredell, Kyriakos Flouris, Krishna Chaitanya, Ertunc Erdil, Ender Konukoglu Here we propose a new formulation of the reconstruction term for the VAE that specifically penalizes the HIGHLIGHT: generation of blurry images while at the same time still maximizing the ELBO under the modeled distribution. 1501. TITLE: Is Conditional Generative Modeling all you need for Decision Making? https://openreview.net/forum?id=sP1fo2K9DFG AUTHORS: Anurag Ajay, Yilun Du, Abhi Gupta, Joshua B. Tenenbaum, Tommi S. Jaakkola, Pulkit Agrawal HIGHLIGHT: Recent improvements in conditional generative modeling have made it possible to generate high-quality images from language descriptions alone. We investigate whether these methods can directly address the problem of sequential decisionmaking. 1502, TITLE: TEMPERA: Test-Time Prompt Editing via Reinforcement Learning https://openreview.net/forum?id=gSHyqBijPFO AUTHORS: Tianjun Zhang, Xuezhi Wang, Denny Zhou, Dale Schuurmans, Joseph E. Gonzalez HIGHLIGHT: In this work, we propose Test-time Prompt Editing using Reinforcement learning (TEMPERA). 1503, TITLE: Evaluating Representations with Readout Model Switching https://openreview.net/forum?id=Fsd-6ax4T1m AUTHORS: Yazhe Li, Jorg Bornschein, Marcus Hutter In this paper, we treat the evaluation of representations as a model selection problem and propose to use the HIGHLIGHT: Minimum Description Length (MDL) principle to devise an evaluation metric. 1504, TITLE: Pseudoinverse-Guided Diffusion Models for Inverse Problems https://openreview.net/forum?id=9 gsMA8MRKQ AUTHORS: Jiaming Song, Arash Vahdat, Morteza Mardani, Jan Kautz HIGHLIGHT: Models trained for specific inverse problems work well but are limited to their particular use cases, whereas methods that use problem-agnostic models are general but often perform worse empirically. To address this dilemma, we introduce Pseudoinverse-guided Diffusion Models (\$\Pi\$GDM), an approach that uses problem-agnostic models to close the gap in performance. 1505, TITLE: Planning with Language Models through Iterative Energy Minimization https://openreview.net/forum?id=cVFD6qE8gnY Hongyi Chen, Yilun Du, Yiye Chen, Patricio A. Vela, Joshua B. Tenenbaum AUTHORS: HIGHLIGHT: The typical autoregressive generation procedures of language models preclude sequential refinement of earlier steps, which limits the effectiveness of a predicted plan. In this paper, we suggest an approach towards integrating planning with language models based on the idea of iterative energy minimization, and illustrate how such a procedure leads to improved RL performance across different tasks. 1506, TITLE: Last Layer Re-Training is Sufficient for Robustness to Spurious Correlations https://openreview.net/forum?id=Zb6c8A-Fghk Polina Kirichenko, Pavel Izmailov, Andrew Gordon Wilson AUTHORS: We propose a simple method based on retraining the last layer of a neural network which achieves strong results HIGHLIGHT: on spurious correlation benchmarks. 1507, TITLE: Don't forget the nullspace! Nullspace occupancy as a mechanism for out of distribution failure

141

https://openreview.net/forum?id=39z0zPZ0AvB AUTHORS: Daksh Idnani, Vivek Madan, Naman Goyal, David J. Schwab, Shanmukha Ramakrishna Vedantam HIGHLIGHT: In this work, we identify a particular failure mode of OoD generalization for discriminative classifiers that is based on test data (from a new domain) lying in the nullspace of features learnt from source data. ContraNorm: A Contrastive Learning Perspective on Oversmoothing and Beyond 1508, TITLE: https://openreview.net/forum?id=SM7XkJouWHm AUTHORS: Xiaojun Guo, Yifei Wang, Tianqi Du, Yisen Wang HIGHLIGHT: Accordingly, inspired by the power of contrastive learning in preventing dimensional collapse, we propose a novel normalization layer ContraNorm. 1509, TITLE: Accelerated Single-Call Methods for Constrained Min-Max Optimization https://openreview.net/forum?id=HRwN7IQLUKA Yang Cai, Weiqiang Zheng AUTHORS: HIGHLIGHT: We study first-order methods for constrained min-max optimization. 1510, TITLE: Performance Bounds for Model and Policy Transfer in Hidden-parameter MDPs https://openreview.net/forum?id=20gBzEzgtiI AUTHORS: Haotian Fu, Jiayu Yao, Omer Gottesman, Finale Doshi-Velez, George Konidaris HIGHLIGHT: We first show that the value function of HiP-MDPs is Lipschitz continuous under certain conditions. We then derive regret bounds for both settings through the lens of Lipschitz continuity. 1511, TITLE: The Lie Derivative for Measuring Learned Equivariance https://openreview.net/forum?id=JL7Va5Vv15J AUTHORS: Nate Gruver, Marc Anton Finzi, Micah Goldblum, Andrew Gordon Wilson HIGHLIGHT: In order to better understand the role of equivariance in recent vision models, we introduce the Lie derivative, a method for measuring equivariance with strong mathematical foundations and minimal hyperparameters. Agree to Disagree: Diversity through Disagreement for Better Transferability 1512, TITLE: https://openreview.net/forum?id=K7CbYObvYhY AUTHORS: Matteo Pagliardini, Martin Jaggi, François Fleuret, Sai Praneeth Karimireddy Instead, we advocate for learning an ensemble of models which capture a diverse set of predictive features. HIGHLIGHT: Towards this, we propose a new algorithm D-BAT (Diversity-By-disAgreement Training), which enforces agreement among the models on the training data, but disagreement on the OOD data. 1513, TITLE: Taking a Step Back with KCal: Multi-Class Kernel-Based Calibration for Deep Neural Networks https://openreview.net/forum?id=p_jIy5QFB7 AUTHORS: Zhen Lin, Shubhendu Trivedi, Jimeng Sun This paper proposes a new Kernel-based calibration method called KCal. HIGHLIGHT: 1514, TITLE: SemPPL: Predicting Pseudo-Labels for Better Contrastive Representations https://openreview.net/forum?id=TAVBJ4aHsWt AUTHORS: Matko Bošnjak, Pierre Harvey Richemond, Nenad Tomasev, Florian Strub, Jacob C Walker, Felix Hill, Lars Holger Buesing, Razvan Pascanu, Charles Blundell, Jovana Mitrovic HIGHLIGHT: We propose a new semi-supervised learning method, Semantic Positives via Pseudo-Labels (SEMPPL), that combines labelled and unlabelled data to learn informative representations. 1515, TITLE: Transfer Learning with Deep Tabular Models https://openreview.net/forum?id=b0RuGUYo8pA AUTHORS: Roman Levin, Valeriia Cherepanova, Avi Schwarzschild, Arpit Bansal, C. Bayan Bruss, Tom Goldstein, Andrew Gordon Wilson, Micah Goldblum HIGHLIGHT: In this work, we explore the benefits that representation learning provides for knowledge transfer in the tabular domain. 1516, TITLE: Understanding Train-Validation Split in Meta-Learning with Neural Networks https://openreview.net/forum?id=JVlyfHEEm0k AUTHORS: Xinzhe Zuo, Zixiang Chen, Huaxiu Yao, Yuan Cao, Quanquan Gu HIGHLIGHT: In this paper, we study the benefit of train-validation split for classification problems with neural network models trained by gradient descent. 1517, TITLE: Revisiting Robustness in Graph Machine Learning https://openreview.net/forum?id=h1o7Ry9Zctm AUTHORS: Lukas Gosch, Daniel Sturm, Simon Geisler, Stephan Günnemann HIGHLIGHT: However, because manual inspection of a graph is difficult, it is unclear if the studied perturbations always preserve a core assumption of adversarial examples: that of unchanged semantic content. To address this problem, we introduce a more principled notion of an adversarial graph, which is aware of semantic content change. 1518, TITLE: Few-shot Backdoor Attacks via Neural Tangent Kernels https://openreview.net/forum?id=a701GJ-rwy

AUTHORS: Jonathan Hayase, Sewoong Oh

HIGHLIGHT: Central to these attacks is the trade-off between the success rate of the attack and the number of corrupted training examples injected. We pose this attack as a novel bilevel optimization problem: construct strong poison examples that maximize the attack success rate of the trained model.

1519, TITLE:	Hyperparameter Optimization through Neural Network Partitioning
https://openreview.ne	t/forum?id=nAgdXgfmqj
AUTHORS:	Bruno Kacper Mlodozeniec, Matthias Reisser, Christos Louizos
HIGHLIGHT: likelihood, an optimiz	In this work, we propose a simple and efficient way for optimizing hyperparameters inspired by the marginal zation objective that requires no validation data.
1520, TITLE:	Symmetries, Flat Minima and the Conserved Quantities of Gradient Flow
https://openreview.ne	t/forum?id=92pciCOunFb
AUTHORS:	Bo Zhao, Iordan Ganev, Kobin Walters, Kose Yu, Nima Denmamy
the low-loss valleys.In	mportantly, we introduce a novel set of nonlinear, data-dependent symmetries for neural networks.
1521, TITLE:	Hebbian Deep Learning Without Feedback
AUTHOPS:	Viorum/Id=8ga4MKJI Adrian Jaura'a Haatar Caraia Padriguaz, Oinghai Gua, Timalaan Maraitia
HIGHLIGHT.	Here grounded on recent theory for Hebbian learning in soft winner-take-all networks, we present multilayer
SoftHebb, i.e. an algo	rithm that trains deep neural networks, without any feedback, target, or error signals.
1522, TITLE:	Making Better Decision by Directly Planning in Continuous Control
AUTHOPS:	Viorum/id=romu/idxyr
HIGHLIGHT: continuous action spa	We propose a novel POMP algorithm with a D3P planner module to achieve the efficient planning in the ce control problem.
1523, TITLE:	(Certified!!) Adversarial Robustness for Free!
https://openreview.ne	t/forum?id=JLg5aHHv7j
AUTHORS:	Nicholas Carlini, J Zico Kolter, Florian Tramer, Krishnamurthy Dj Dvijotham, Leslie Rice, Mingjie Sun
HIGHLIGHT:	In this paper we show how to achieve state-of-the-art certified adversarial robustness to 2-norm bounded
perturbations by relying	ng exclusively on off-the-shelf pretrained models.
1524, TITLE:	MMVAE+: Enhancing the Generative Quality of Multimodal VAEs without Compromises
https://openreview.ne	t/forum?id=sdQGxouELX
AUTHORS:	Emanuele Palumbo, Imant Daunhawer, Julia E Vogt
HIGHLIGHT:	In particular mixture-based models achieve good coherence only at the expense of sample diversity and a
resulting lack of gene	rative quality. We present a novel variant of the mixture-of-experts multimodal variational autoencoder that
improves its generativ	e quality, while maintaining high semantic coherence.
1525, TITLE:	Scaling Laws For Deep Learning Based Image Reconstruction
https://openreview.ne	t/forum?id=op-ceGueqc4
AUTHORS:	Tobit Klug, Reinhard Heckel
HIGHLIGHT:	In this work, we study whether major performance gains are expected from scaling up the training set size.
1526, TITLE:	Canary in a Coalmine: Better Membership Inference with Ensembled Adversarial Queries
https://openreview.ne	VIORUM/10=D/SBTEBERU Vuvin Wan, Arnit Bansal, Hamid Kazami, Fitan Barania, Micah Galdhlum, Janas Gaining, Tam Galdstain
HIGHLIGHT:	In this work, we use adversarial tools to directly optimize for queries that are discriminative and diverse.
1527, TITLE:	Revisiting Populations in multi-agent Communication
https://openreview.ne	Vforum?id=n-UHRIdPju Dayl Michal Mathiay Dita Kary Walloon Mathawaan Olivian Tialaman Angoliki Lagariday
HIGHLIGHT	Faul Michel, Matineu Kita, Kory wanace Mathewson, Onvier Theining paradigm for referential games based on the
idea of "partitioning"	the agents into sender-receiver pairs and limiting co-adaptation across pairs.
1528, TITLE:	Sequential Gradient Coding For Straggler Mitigation
https://openreview.ne	t/forum?id=-lGvSmht7a
AUTHORS:	Nikhil Krishnan Muralee Krishnan, MohammadReza Ebrahimi, Ashish J Khisti
HIGHLIGHT: where processing of e	In this paper, we consider the distributed computation of a sequence of gradients $\left(g(1),g(2),dots,g(J)\right)$, each gradient $g(t)$ starts in round- t and finishes by round- $t+T$.
1529. TITLE:	Learning MLPs on Graphs: A Unified View of Effectiveness. Robustness, and Efficiency
https://openreview.ne	t/forum?id=Cs3r5KLdoj
AUTHORS:	Yijun Tian, Chuxu Zhang, Zhichun Guo, Xiangliang Zhang, Nitesh Chawla
HIGHLIGHT:	In this paper, we ascribe the lack of effectiveness and robustness to three significant challenges: 1) the
misalignment between	n content feature and label spaces, 2) the strict hard matching to teacher's output, and 3) the sensitivity to node
teature noises.	

1530, TITLE: Generating Sequences by Learning to Self-Correct https://openreview.net/forum?id=hH36JeQZDaO AUTHORS: Sean Welleck, Ximing Lu, Peter West, Faeze Brahman, Tianxiao Shen, Daniel Khashabi, Yejin Choi HIGHLIGHT: We present Self-Correction, an approach that decouples an imperfect base generator (an off-the-shelf language model or supervised sequence-to-sequence model) from a separate corrector that learns to iteratively correct imperfect generations. Learning About Progress From Experts 1531, TITLE: https://openreview.net/forum?id=sKc6fgce1zs AUTHORS: Jake Bruce, Ankit Anand, Bogdan Mazoure, Rob Fergus HIGHLIGHT: In this work, we explore the use of expert demonstrations in long-horizon tasks to learn a monotonically increasing function that summarizes progress. Learning Fair Graph Representations via Automated Data Augmentations 1532, TITLE: https://openreview.net/forum?id=1 OGWcP1s9w AUTHORS: Hongyi Ling, Zhimeng Jiang, Youzhi Luo, Shuiwang Ji, Na Zou In this work, we propose a method, known as Graphair, to learn fair representations based on automated graph HIGHLIGHT: data augmentations. 1533, TITLE: Emergence of Maps in the Memories of Blind Navigation Agents https://openreview.net/forum?id=ITt4KjHSsyl AUTHORS: Erik Wijmans, Manolis Savva, Irfan Essa, Stefan Lee, Ari S. Morcos, Dhruv Batra HIGHLIGHT: Overall, this paper presents no new techniques for the AI audience, but a surprising finding, an insight, and an explanation. 1534, TITLE: Latent Neural ODEs with Sparse Bayesian Multiple Shooting https://openreview.net/forum?id=moIIFZfi 1b Valerii Iakovlev, Cagatay Yildiz, Markus Heinonen, Harri Lähdesmäki AUTHORS: HIGHLIGHT: We propose a principled multiple shooting technique for neural ODEs that splits the trajectories into manageable short segments, which are optimized in parallel, while ensuring probabilistic control on continuity over consecutive segments. 1535, TITLE: \$\mathcal{O}\$-GNN: incorporating ring priors into molecular modeling https://openreview.net/forum?id=5cFfz6yMVPU AUTHORS: Jinhua Zhu, Kehan Wu, Bohan Wang, Yingce Xia, Shufang Xie, Qi Meng, Lijun Wu, Tao Qin, Wengang Zhou, Houqiang Li, Tie-Yan Liu In this work, we design a new variant of GNN, ring-enhanced GNN (\$\mathcal{O}\$-GNN), that explicitly HIGHLIGHT: models rings in addition to atoms and bonds in compounds. 1536, TITLE: Transformers Learn Shortcuts to Automata https://openreview.net/forum?id=De4FYqjFueZ AUTHORS: Bingbin Liu, Jordan T. Ash, Surbhi Goel, Akshay Krishnamurthy, Cyril Zhang HIGHLIGHT: This raises the question: what solutions are these shallow and non-recurrent models finding? We investigate this question in the setting of learning automata, discrete dynamical systems naturally suited to recurrent modeling and expressing algorithmic tasks. 1537, TITLE: Obtaining More Generalizable Fair Classifiers on Imbalanced Datasets https://openreview.net/forum?id=zVrw4OH1Lch AUTHORS: Zhun Deng, Jiayao Zhang, Linjun Zhang, Ting Ye, Yates Coley, Weijie J Su, James Zou HIGHLIGHT: In this paper, we propose a theoretically principled, yet flexible approach that encourages both classification and fairness generalization and can be flexibly combined with many existing fair learning methods with logits-based losses. 1538, TITLE: Understanding The Robustness of Self-supervised Learning Through Topic Modeling https://openreview.net/forum?id=7Cb7Faxa1OB Zeping Luo, Shiyou Wu, Cindy Weng, Mo Zhou, Rong Ge AUTHORS: HIGHLIGHT: In this paper, we focus on the context of topic modeling and highlight a key advantage of self-supervised learning - when applied to data generated by topic models, self-supervised learning can be oblivious to the specific model, and hence is less susceptible to model misspecification. 1539, TITLE: Provably Efficient Neural Offline Reinforcement Learning via Perturbed Rewards https://openreview.net/forum?id=WOquZTLCBO1 AUTHORS: Thanh Nguyen-Tang, Raman Arora We propose a novel offline reinforcement learning (RL) algorithm, namely PEturbed-Reward Value Iteration HIGHLIGHT: (PERVI) which amalgamates the randomized value function idea with the pessimism principle. 1540, TITLE: Neuromechanical Autoencoders: Learning to Couple Elastic and Neural Network Nonlinearity https://openreview.net/forum?id=QubsmJT A0 AUTHORS: Deniz Oktay, Mehran Mirramezani, Eder Medina, Ryan P Adams HIGHLIGHT: Just as deep neural networks provide flexible and massively-parametric function approximators for perceptual

and control tasks, cellular solid metamaterials are promising as a rich and learnable space for approximating a variety of actuation
tasks. In this work we take advantage of these complementary computational concepts to co-design materials and neural network controls to achieve nonintuitive mechanical behavior.

1541, TITLE: Towards Universal Visual Reward and Representation via Value-Implicit Pre-Training https://openreview.net/forum?id=YJ7o2wetJ2 AUTHORS: Yecheng Jason Ma, Shagun Sodhani, Dinesh Jayaraman, Osbert Bastani, Vikash Kumar, Amy Zhang We introduce $\int \frac{V}{\delta} = \frac{1}{2} \frac{V}{\delta}$ HIGHLIGHT: visual representation capable of generating dense and smooth reward functions for unseen robotic tasks. Bridging the Gap to Real-World Object-Centric Learning 1542, TITLE: https://openreview.net/forum?id=b9tUk-f aG AUTHORS: Maximilian Seitzer, Max Horn, Andrii Zadaianchuk, Dominik Zietlow, Tianjun Xiao, Carl-Johann Simon-Gabriel, Tong He, Zheng Zhang, Bernhard Schölkopf, Thomas Brox, Francesco Locatello However, current methods are restricted to simulated data or require additional information in the form of HIGHLIGHT: motion or depth in order to successfully discover objects. In this work, we overcome this limitation by showing that reconstructing features from models trained in a self-supervised manner is a sufficient training signal for object-centric representations to arise in a fully unsupervised way. 1543, TITLE: Towards a Unified Theoretical Understanding of Non-contrastive Learning via Rank Differential Mechanism https://openreview.net/forum?id=cIbjyd2Vcy AUTHORS: Zhijian Zhuo, Yifei Wang, Jinwen Ma, Yisen Wang HIGHLIGHT: In this work, we propose a new understanding for non-contrastive learning, named the Rank Differential Mechanism (RDM). 1544, TITLE: Stay Moral and Explore: Learn to Behave Morally in Text-based Games https://openreview.net/forum?id=CtS2Rs aYk AUTHORS: Zijing Shi, Meng Fang, Yunqiu Xu, Ling Chen, Yali Du In this paper, we propose a general framework named Moral Awareness Adaptive Learning (MorAL) that HIGHLIGHT: enhances the morality capacity of an agent using a plugin moral-aware learning model. 1545, TITLE: Learning to Induce Causal Structure https://openreview.net/forum?id=hp_RwhKDJ5 AUTHORS: Nan Rosemary Ke, Silvia Chiappa, Jane X Wang, Jorg Bornschein, Anirudh Goyal, Melanie Rey, Theophane Weber, Matthew Botvinick, Michael Curtis Mozer, Danilo Jimenez Rezende HIGHLIGHT: In our work, we instead treat the inference process as a black box and design a neural network architecture that learns the mapping from both observational and interventional data to graph structures via supervised training on synthetic graphs. 1546, TITLE: Geometrically regularized autoencoders for non-Euclidean data https://openreview.net/forum?id= q7A0m3vXH0 AUTHORS: Cheongjae Jang, Yonghyeon Lee, Yung-Kyun Noh, Frank C. Park HIGHLIGHT: Given the recent surge of interest in machine learning problems involving non-Euclidean data, in this paper we address the regularization of autoencoders on curved spaces. Online Boundary-Free Continual Learning by Scheduled Data Prior 1547, TITLE: https://openreview.net/forum?id=qco4ekz2Epm AUTHORS: Hyunseo Koh, Minhyuk Seo, Jihwan Bang, Hwanjun Song, Deokki Hong, Seulki Park, Jung-Woo Ha, Jonghyun Choi HIGHLIGHT: To this end, we propose a scheduled transfer of previously learned knowledge. 1548, TITLE: Efficient Learning of Rationalizable Equilibria in General-Sum Games https://openreview.net/forum?id=HjOo2k8lhFl AUTHORS: Yuanhao Wang, Dingwen Kong, Yu Bai, Chi Jin HIGHLIGHT: We develop provably efficient algorithms for finding approximate CE and CCE that are also rationalizable. 1549, TITLE: Energy-Based Test Sample Adaptation for Domain Generalization https://openreview.net/forum?id=3dnrKbeVatv AUTHORS: Zehao Xiao, Xiantong Zhen, Shengcai Liao, Cees G. M. Snoek HIGHLIGHT: In this paper, we propose energy-based sample adaptation at test time for domain generalization. 1550, TITLE: Revisiting adapters with adversarial training https://openreview.net/forum?id=HPdxC1THU8T Sylvestre-Alvise Rebuffi, Francesco Croce, Sven Gowal AUTHORS: HIGHLIGHT: We establish that using the classification token of a Vision Transformer (ViT) as an adapter is enough to match the classification performance of dual normalization layers, while using significantly less additional parameters. EPISODE: Episodic Gradient Clipping with Periodic Resampled Corrections for Federated Learning with 1551, TITLE: Heterogeneous Data https://openreview.net/forum?id=ytZIYmztET AUTHORS: Michael Crawshaw, Yajie Bao, Mingrui Liu

HIGHLIGHT: In this paper, we design EPISODE, the very first algorithm to solve FL problems with heterogeneous data in the nonconvex and relaxed smoothness setting.

1552, TITLE: On the Trade-Off between Actionable Explanations and the Right to be Forgotten https://openreview.net/forum?id=HWt4BBZjVW AUTHORS: Martin Pawelczyk, Tobias Leemann, Asia Biega, Gjergji Kasneci HIGHLIGHT: To date it is unknown whether these two principles can be operationalized simultaneously. Therefore, we introduce and study the problem of recourse invalidation in the context of data deletion requests. Computing all Optimal Partial Transports 1553. TITLE: https://openreview.net/forum?id=gwcQajoXNF Abhijeet Phatak, Sharath Raghvendra, CHITTARANJAN TRIPATHY, Kaiyi Zhang AUTHORS: HIGHLIGHT: In this paper, we consider the question of computing the OT-profile. 1554, TITLE: Probabilistically Robust Recourse: Navigating the Trade-offs between Costs and Robustness in Algorithmic Recourse https://openreview.net/forum?id=sC-PmTsiTB AUTHORS: Martin Pawelczyk, Teresa Datta, Johan Van den Heuvel, Gjergji Kasneci, Himabindu Lakkaraju HIGHLIGHT: Furthermore, prior approaches do not provide end users with any agency over navigating the aforementioned trade-offs. In this work, we address the above challenges by proposing the first algorithmic framework which enables users to effectively manage the recourse cost vs. robustness trade-offs. 1555, TITLE: Pseudo-label Training and Model Inertia in Neural Machine Translation https://openreview.net/forum?id=eXkhH12DTD9 AUTHORS: Benjamin Hsu, Anna Currey, Xing Niu, Maria Nadejde, Georgiana Dinu HIGHLIGHT: We study inertia effects under different training settings and we identify distribution simplification as a mechanism behind the observed results. 1556, TITLE: HyperDeepONet: learning operator with complex target function space using the limited resources via hypernetwork https://openreview.net/forum?id=OAw6V3ZAhSd AUTHORS: Jae Yong Lee, SungWoong CHO, Hyung Ju Hwang HIGHLIGHT: In this study, we propose HyperDeepONet, which uses the expressive power of the hypernetwork to enable learning of a complex operator with smaller set of parameters. 1557, TITLE: Latent Graph Inference using Product Manifolds https://openreview.net/forum?id=JLR_B7n_Wqr AUTHORS: Haitz Sáez de Ocáriz Borde, Anees Kazi, Federico Barbero, Pietro Lio In this work, we generalize the discrete Differentiable Graph Module (dDGM) for latent graph learning. HIGHLIGHT: Meta Temporal Point Processes 1558, TITLE: https://openreview.net/forum?id=QZfdDpTX1uM AUTHORS: Wonho Bae, Mohamed Osama Ahmed, Frederick Tung, Gabriel L. Oliveira HIGHLIGHT: In this work, we propose to train TPPs in a meta learning framework, where each sequence is treated as a different task, via a novel framing of TPPs as neural processes (NPs). We introduce context sets to model TPPs as an instantiation of NPs. 1559, TITLE: Implicit Bias of Large Depth Networks: a Notion of Rank for Nonlinear Functions https://openreview.net/forum?id=6iDHce-0B-a AUTHORS: Arthur Jacot HIGHLIGHT: We show that the representation cost of fully connected neural networks with homogeneous nonlinearities which describes the implicit bias in function space of networks with L2-regularization or with losses such as the cross-entropy converges as the depth of the network goes to infinity to a notion of rank over nonlinear functions. 1560, TITLE: Neural Compositional Rule Learning for Knowledge Graph Reasoning https://openreview.net/forum?id=F8VKQyDgRVj AUTHORS: Kewei Cheng, Nesreen Ahmed, Yizhou Sun HIGHLIGHT: In this paper, we propose an end-to-end neural model for learning compositional logic rules called NCRL. 1561, TITLE: MetaGL: Evaluation-Free Selection of Graph Learning Models via Meta-Learning https://openreview.net/forum?id=C1ns08q9jZ AUTHORS: Namyong Park, Ryan A. Rossi, Nesreen Ahmed, Christos Faloutsos HIGHLIGHT: In this work, we develop the first meta-learning approach for evaluation-free graph learning model selection, called MetaGL, which utilizes the prior performances of existing methods on various benchmark graph datasets to automatically select an effective model for the new graph, without any model training or evaluations. 1562, TITLE: A Stable and Scalable Method for Solving Initial Value PDEs with Neural Networks https://openreview.net/forum?id=vsMyHUq_C1c

AUTHORS: Marc Anton Finzi, Andres Potapczynski, Matthew Choptuik, Andrew Gordon Wilson

HIGHLIGHT: We develop Neural-IVP, a new method for solving initial value PDEs with Neural Networks that is both stable and scalable.

1563, TITLE:Quantifying and Mitigating the Impact of Label Errors on Model Disparity Metricshttps://openreview.net/forum?id=RUzSobdYy0VAUTHORS:Julius Adebayo, Melissa Hall, Bowen Yu, Bobbie Chern

HIGHLIGHT: Towards mitigating the impact of training-time label error, we present an approach to estimate how changing a single training input's label affects a model's group disparity metric on a test set.