1, TITLE: What Matters for On-Policy Deep Actor-Critic Methods? A Large-Scale Study
https://openreview.net/forum?id=nIAxjsniDzg
AUTHORS: Marcin Andrychowicz, Anton Raichuk, Piotr Stanczyk, Manu Orsini, Sertan Girgin, Rapha?l Marinier, Leonard Hussenot, Matthieu Geist, Olivier Pietquin, Marcin Michalski, Sylvain Gelly, Olivier Bachem
HIGHLIGHT: We conduct a large-scale empirical study that provides insights and practical recommendations for the training of on-policy deep actor-critic RL agents.

2, TITLE: Theoretical Analysis of Self-Training with Deep Networks on Unlabeled Data
https://openreview.net/forum?id=rC8sJ4i6kaH
AUTHORS: Colin Wei, Kendrick Shen, Yining Chen, Tengyu Ma
HIGHLIGHT: This paper provides accuracy guarantees for self-training with deep networks on polynomial unlabeled samples for semi-supervised learning, unsupervised domain adaptation, and unsupervised learning.

3, TITLE: Learning to Reach Goals via Iterated Supervised Learning
https://openreview.net/forum?id=rALA0Xo6yNJ
AUTHORS: Dibya Ghosh, Abhishek Gupta, Ashwin Reddy, Coline Manon Devin, Benjamin Eysenbach, Sergey Levine
HIGHLIGHT: We propose and formally analyze a simple RL method for learning goal-reaching policies via iterated supervised learning.

4, TITLE: Deep symbolic regression: Recovering mathematical expressions from data via risk-seeking policy gradients
https://openreview.net/forum?id=m5Qsh0kBQG
AUTHORS: Brenden K Petersen, Mikel Landajuela Larma, Terrell N. Mundhenk, Claudio Prata Santiago, Soo Kyung Kim, Joanne Taery Kim
HIGHLIGHT: A deep learning approach to symbolic regression, in which an autoregressive RNN emits a distribution over expressions that is optimized using a risk-seeking policy gradient.

5, TITLE: Optimal Rates for Averaged Stochastic Gradient Descent under Neural Tangent Kernel Regime
https://openreview.net/forum?id=PULSD5qI2N1
AUTHORS: Atsushi Nitanda, Taiji Suzuki
HIGHLIGHT: This is the first paper to overcome technical challenges of achieving the optimal convergence rate under the NTK regime.

6, TITLE: Free Lunch for Few-shot Learning: Distribution Calibration
https://openreview.net/forum?id=JWOiYxMG92s
AUTHORS: Shuo Yang, Lu Liu, Min Xu
HIGHLIGHT: We proposed a simple but effective few-shot classification method, which can estimate an unseen class distribution from one sample. Code is available at: https://github.com/ShuoYang-1998/ICLR2021-Oral_Distribution_Calibration

7, TITLE: Scalable Learning and MAP Inference for Nonsymmetric Determinantal Point Processes
https://openreview.net/forum?id=HajQFbx_yB
AUTHORS: Mike Gartrell, Insu Han, Elvis Dohmatob, Jennifer Gillenwater, Victor-Emmanuel Brunel
HIGHLIGHT: We propose scalable learning and maximum a posteriori (MAP) inference algorithms for nonsymmetric determinantal point processes (DPPs).

8, TITLE: Randomized Automatic Differentiation
https://openreview.net/forum?id=xpx9zj7CUlY
AUTHORS: Deniz Oktay, Nick McGreivy, Joshua Aduol, Alex Beatson, Ryan P Adams
HIGHLIGHT: We develop a general framework and approach for randomized automatic differentiation (RAD), which can allow unbiased gradient estimates to be computed with reduced memory in return for variance.

9, TITLE: Learning Generalizable Visual Representations via Interactive Gameplay
https://openreview.net/forum?id=UuchYL8wSZo
AUTHORS: Luca Weihs, Aniruddha Kembhavi, Kiana Ehsani, Sarah M Pratt, Winson Han, Alvaro Herrasti, Eric Kolve, Dustin Schwenk, Roozbeh Mottaghi, Ali Farhadi
HIGHLIGHT: We show the representation learned through interaction and gameplay generalizes better compared to passive and static representation learning methods.

10, TITLE: Global Convergence of Three-layer Neural Networks in the Mean Field Regime
https://openreview.net/forum?id=KvyxFqZS_D
AUTHORS: Huy Tuan Pham, Phan-Minh Nguyen

HIGHLIGHT: We propose a rigorous framework for three-layer neural networks in the mean field regime and prove a global convergence guarantee.

11, TITLE: Rao-Blackwellizing the Straight-Through Gumbel-Softmax Gradient Estimator
https://openreview.net/forum?id=Mk6PZtgAgfq
AUTHORS: Max B Paulus, Chris J. Maddison, Andreas Krause
HIGHLIGHT: We reduce the variance of the straight-through Gumbel-Softmax estimator to improve its performance.

12, TITLE: Rethinking Attention with Performers
https://openreview.net/forum?id=Ua6zuk0WRH
AUTHORS: Krzysztof Marcin Choromanski, Valerii Likhosherstov, David Dohan, Xingyou Song, Andreea Gane, Tamas Sarlos, Peter Hawkins, Jared Quincy Davis, Afroz Mohiuddin, Lukasz Kaiser, David Benjamin Belanger, Lucy J Colwell, Adrian Weller
HIGHLIGHT: We introduce Performers, linear full-rank-attention Transformers via provable random feature approximation methods, without relying on sparsity or low-rankness.

13, TITLE: Getting a CLUE: A Method for Explaining Uncertainty Estimates
https://openreview.net/forum?id=XSLF1XFq5h
AUTHORS: Javier Antoran, Umang Bhatt, Tameem Adel, Adrian Weller, Jos? Miguel Hern?ndez-Lobato
HIGHLIGHT: We introduce a method to help explain uncertainties of any differentiable probabilistic model by perturbing input features.

14, TITLE: When Do Curricula Work?
https://openreview.net/forum?id=tW4QEInpni
AUTHORS: Xiaoxia Wu, Ethan Dyer, Behnam Neyshabur
HIGHLIGHT: We conduct extensive experiments over thousands of orderings to investigate the effectiveness of three kinds of learning: curriculum, anti-curriculum, and random-curriculum.

15, TITLE: Federated Learning Based on Dynamic Regularization
https://openreview.net/forum?id=B7v4QMR6Z9w
AUTHORS: Durmus Alp Emre Acar, Yue Zhao, Ramon Matas, Matthew Mattina, Paul Whatmough, Venkatesh Saligrama
HIGHLIGHT: We present, FedDyn, a novel dynamic regularization method for Federated Learning where the risk objective for each device is dynamically updated to ensure the device optima is asymptotically consistent with stationary points of the global loss.

16, TITLE: Geometry-aware Instance-reweighted Adversarial Training
https://openreview.net/forum?id=iAX0l6Cz8ub
AUTHORS: Jingfeng Zhang, Jianing Zhu, Gang Niu, Bo Han, Masashi Sugiyama, Mohan Kankanhalli
HIGHLIGHT: This paper has proposed a novel adversarial training method, i.e., geometry-aware instance-reweighted adversarial training (GAIRAT), which sheds new lights on improving the adversarial training.

17, TITLE: Co-Mixup: Saliency Guided Joint Mixup with Supermodular Diversity
https://openreview.net/forum?id=gvxJzw8kW4b
AUTHORS: JangHyun Kim, Wonho Choo, Hosan Jeong, Hyun Oh Song
HIGHLIGHT: We propose a new perspective on joint mixup augmentation and formulate the optimal construction of a batch of mixup data.

18, TITLE: SenSeI: Sensitive Set Invariance for Enforcing Individual Fairness
https://openreview.net/forum?id=DktZb97_Fx
AUTHORS: Mikhail Yurochkin, Yuekai Sun
HIGHLIGHT: We propose a new invariance-enforcing regularizer for training individually fair ML systems.

19, TITLE: End-to-end Adversarial Text-to-Speech
https://openreview.net/forum?id=rsf1z-JSj87
AUTHORS: Jeff Donahue, Sander Dieleman, Mikolaj Binkowski, Erich Elsen, Karen Simonyan
HIGHLIGHT: Efficient, adversarially trained feed-forward text-to-speech model producing high-quality speech, learnt end-to-end in a single stage.

20, TITLE: Dataset Condensation with Gradient Matching
https://openreview.net/forum?id=mSAKhLYLSsl

AUTHORS: Bo Zhao, Konda Reddy Mopuri, Hakan Bilen
HIGHLIGHT: This paper proposes a training set synthesis technique that learns to produce a small set of informative samples for training deep neural networks from scratch in a small fraction of computational cost while achieving as close results as possible.

21, TITLE: Rethinking Architecture Selection in Differentiable NAS
https://openreview.net/forum?id=PKubaeJkw3
AUTHORS: Ruochen Wang, Minhao Cheng, Xiangning Chen, Xiaocheng Tang, Cho-Jui Hsieh
HIGHLIGHT: We propose an alternative perturbation-based architecture selection that directly measures each operation's influence on the supernet.

22, TITLE: A Distributional Approach to Controlled Text Generation
https://openreview.net/forum?id=jWkw45-9AbL
AUTHORS: Muhammad Khalifa, Hady Elsahar, Marc Dymetman
HIGHLIGHT: We propose a novel approach to Controlled Text Generation, relying on Constraints over Distributions, Information Geometry, and Sampling from Energy-Based Models.

23, TITLE: Learning Cross-Domain Correspondence for Control with Dynamics Cycle-Consistency
https://openreview.net/forum?id=QIRlze3I6hX
AUTHORS: Qiang Zhang, Tete Xiao, Alexei A Efros, Lerrel Pinto, Xiaolong Wang
HIGHLIGHT: We learn correspondence across domains with different modalities, physics parameters, and morphologies for control tasks (in both simulation and real robot) with Dynamics Cycle-Consistency.

24, TITLE: Human-Level Performance in No-Press Diplomacy via Equilibrium Search
https://openreview.net/forum?id=0-uUGPbIjD
AUTHORS: Jonathan Gray, Adam Lerer, Anton Bakhtin, Noam Brown
HIGHLIGHT: We present an agent that approximates a one-step equilibrium in no-press Diplomacy using no-regret learning and show that it exceeds human-level performance

25, TITLE: Parrot: Data-Driven Behavioral Priors for Reinforcement Learning
https://openreview.net/forum?id=Ysuv-WOFeKR
AUTHORS: Avi Singh, Huihan Liu, Gaoyue Zhou, Albert Yu, Nicholas Rhinehart, Sergey Levine
HIGHLIGHT: We propose a method for pre-training a prior for reinforcement learning using data from a diverse range of tasks, and use this prior to speed up learning of new tasks.

26, TITLE: Invariant Representations for Reinforcement Learning without Reconstruction
https://openreview.net/forum?id=-2FCwDKRREu
AUTHORS: Amy Zhang, Rowan Thomas McAllister, Roberto Calandra, Yarin Gal, Sergey Levine
HIGHLIGHT: Our goal is to learn representations that provide for effective downstream control and invariance to task-irrelevant details.

27, TITLE: Do 2D GANs Know 3D Shape? Unsupervised 3D Shape Reconstruction from 2D Image GANs
https://openreview.net/forum?id=FGqiDsBUKL0
AUTHORS: Xingang Pan, Bo Dai, Ziwei Liu, Chen Change Loy, Ping Luo
HIGHLIGHT: Unsupervised 3D Shape Reconstruction from 2D Image GANs

28, TITLE: Varying Coefficient Neural Network with Functional Targeted Regularization for Estimating Continuous Treatment Effects
https://openreview.net/forum?id=RmB-88r9dL
AUTHORS: Lizhen Nie, Mao Ye, qiang liu, Dan Nicolae
HIGHLIGHT: We propose a varying coefficient network and a functional targeted regularization for estimating continuous treatment.

29, TITLE: Rethinking the Role of Gradient-based Attribution Methods for Model Interpretability
https://openreview.net/forum?id=dYeAHXnpWJ4
AUTHORS: Suraj Srinivas, Francois Fleuret
HIGHLIGHT: Input-gradients in discriminative neural net models capture information regarding an implicit density model, rather than that of the underlying discriminative model which it is intended to explain.

30, TITLE: Neural Synthesis of Binaural Speech
https://openreview.net/forum?id=uAX8q61EVRu

AUTHORS: Alexander Richard, Dejan Markovic, Israel D. Gebru, Steven Krenn, Gladstone Alexander Butler, Fernando Torre, Yaser Sheikh
HIGHLIGHT: We propose an end-to-end approach to neural binaural sound synthesis that for the first time outperforms DSP-based methods in a qualitative evaluation and in a perceptual study.
We will release a first-of-its-kind binaural audio dataset as a benchmark for future research.

31, TITLE: DiffWave: A Versatile Diffusion Model for Audio Synthesis
https://openreview.net/forum?id=a-xFK8Ymz5J
AUTHORS: Zhifeng Kong, Wei Ping, Jiaji Huang, Kexin Zhao, Bryan Catanzaro
HIGHLIGHT: DiffWave is a versatile diffusion probabilistic model for waveform generation, which matches the state-of-the-art neural vocoder in terms of quality and can generate abundant realistic voices in time-domain without any conditional information.

32, TITLE: An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale
https://openreview.net/forum?id=YicbFdNTTy
AUTHORS: Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, Neil Houlsby
HIGHLIGHT: Transformers applied directly to image patches and pre-trained on large datasets work really well on image classification.

33, TITLE: On the mapping between Hopfield networks and Restricted Boltzmann Machines
https://openreview.net/forum?id=RGJbergVIoO
AUTHORS: Matthew Smart, Anton Zilman
HIGHLIGHT: Hopfield networks with correlated patterns can be mapped to Restricted Boltzmann Machines with orthogonal weights.

34, TITLE: SMiRL: Surprise Minimizing Reinforcement Learning in Unstable Environments
https://openreview.net/forum?id=cPZOyoDloxl
AUTHORS: Glen Berseth, Daniel Geng, Coline Manon Devin, Nicholas Rhinehart, Chelsea Finn, Dinesh Jayaraman, Sergey Levine
HIGHLIGHT: Using Bayesian surprise as an unsupervised intrinsic reward function to learn complex behaviors in unstable environments.

35, TITLE: Evolving Reinforcement Learning Algorithms
https://openreview.net/forum?id=0XXpJ4OtjW
AUTHORS: John D Co-Reyes, Yingjie Miao, Daiyi Peng, Quoc V Le, Sergey Levine, Honglak Lee, Aleksandra Faust
HIGHLIGHT: We meta-learn RL algorithms by evolving computational graphs which compute the loss function for a value-based model-free RL agent to optimize.

36, TITLE: Growing Efficient Deep Networks by Structured Continuous Sparsification
https://openreview.net/forum?id=wb3wxCObbRT
AUTHORS: Xin Yuan, Pedro Henrique Pamplona Savarese, Michael Maire
HIGHLIGHT: We propose an efficient training method that dynamically grows and prunes neural network architectures.

37, TITLE: Deformable DETR: Deformable Transformers for End-to-End Object Detection
https://openreview.net/forum?id=gZ9hCDWe6ke
AUTHORS: Xizhou Zhu, Weijie Su, Lewei Lu, Bin Li, Xiaogang Wang, Jifeng Dai
HIGHLIGHT: Deformable DETR is an efficient and fast-converging end-to-end object detector. It mitigates the high complexity and slow convergence issues of DETR via a novel sampling-based efficient attention mechanism.

38, TITLE: EigenGame: PCA as a Nash Equilibrium
https://openreview.net/forum?id=NzTU59SYbNq
AUTHORS: Ian Gemp, Brian McWilliams, Claire Vernade, Thore Graepel
HIGHLIGHT: We formulate the solution to PCA as the Nash of a suitable game with accompanying algorithm that we demonstrate on a 200TB dataset.

39, TITLE: Augmenting Physical Models with Deep Networks for Complex Dynamics Forecasting
https://openreview.net/forum?id=kmG8vRXTFv
AUTHORS: Vincent LE GUEN, Yuan Yin, J?r?mie DONA, Ibrahim Ayed, Emmanuel de Bezenac, Nicolas THOME, patrick gallinari
HIGHLIGHT: We propose a new principled framework for combining physical models with deep data-driven networks, for which we provide theoretical decomposition guarantees.

40, TITLE: Complex Query Answering with Neural Link Predictors
https://openreview.net/forum?id=Mos9F9kDwkz
AUTHORS: Erik Arakelyan, Daniel Daza, Pasquale Minervini, Michael Cochez
HIGHLIGHT: We show how to answer complex queries by answering their sub-queries via neural link predictors, aggregating results via t-norms and t-conorms, and identifying the optimal variable substitutions by solving an optimisation problem.

41, TITLE: Towards Nonlinear Disentanglement in Natural Data with Temporal Sparse Coding
https://openreview.net/forum?id=EbIDjBynYJ8
AUTHORS: David A. Klindt, Lukas Schott, Yash Sharma, Ivan Ustyuzhaninov, Wieland Brendel, Matthias Bethge, Dylan Paiton
HIGHLIGHT: Our work addresses key issues in disentanglement research for moving towards more natural settings.

42, TITLE: Self-training For Few-shot Transfer Across Extreme Task Differences
https://openreview.net/forum?id=O3Y56aqpChA
AUTHORS: Cheng Perng Phoo, Bharath Hariharan
HIGHLIGHT: Self-training a source domain classifier on unlabeled data from the target domain improves cross-domain few-shot transfer.

43, TITLE: Score-Based Generative Modeling through Stochastic Differential Equations
https://openreview.net/forum?id=PxTIG12RRHS
AUTHORS: Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, Ben Poole
HIGHLIGHT: A general framework for training and sampling from score-based models that unifies and generalizes previous methods, allows likelihood computation, and enables controllable generation.

44, TITLE: Share or Not? Learning to Schedule Language-Specific Capacity for Multilingual Translation
https://openreview.net/forum?id=Wj4ODo0uyCF
AUTHORS: Biao Zhang, Ankur Bapna, Rico Sennrich, Orhan Firat
HIGHLIGHT: We investigate and improve parameter-sharing strategies in multilingual Transformers by utilizing conditional computation.

45, TITLE: Image GANs meet Differentiable Rendering for Inverse Graphics and Interpretable 3D Neural Rendering
https://openreview.net/forum?id=yWkP7JuHX1
AUTHORS: Yuxuan Zhang, Wenzheng Chen, Huan Ling, Jun Gao, Yinan Zhang, Antonio Torralba, Sanja Fidler
HIGHLIGHT: We marry generative models with differentiable rendering to extract and disentangle 3D knowledge learned implicitly by generative image synthesis models

46, TITLE: How Neural Networks Extrapolate: From Feedforward to Graph Neural Networks
https://openreview.net/forum?id=UH-cmocLJC
AUTHORS: Keyulu Xu, Mozhi Zhang, Jingling Li, Simon Shaolei Du, Ken-Ichi Kawarabayashi, Stefanie Jegelka
HIGHLIGHT: We study how neural networks trained by gradient descent extrapolate.

47, TITLE: Contrastive Explanations for Reinforcement Learning via Embedded Self Predictions
https://openreview.net/forum?id=Ud3DSz72nYR
AUTHORS: Zhengxian Lin, Kin-Ho Lam, Alan Fern
HIGHLIGHT: We introduced the embedded self-prediction (ESP) model for producing meaningful and sound contrastive explanations for RL agents.

48, TITLE: Improved Autoregressive Modeling with Distribution Smoothing
https://openreview.net/forum?id=rJA5Pz7lHKb
AUTHORS: Chenlin Meng, Jiaming Song, Yang Song, Shengjia Zhao, Stefano Ermon
HIGHLIGHT: Inspired by randomized smoothing for adversarial defense, we incorporate randomized smoothing techniques into autoregressive generative modeling.

49, TITLE: MONGOOSE: A Learnable LSH Framework for Efficient Neural Network Training
https://openreview.net/forum?id=wWK7yXkULyh
AUTHORS: Beidi Chen, Zichang Liu, Binghui Peng, Zhaozhuo Xu, Jonathan Lingjie Li, Tri Dao, Zhao Song, Anshumali Shrivastava, Christopher Re
HIGHLIGHT: We propose MONGOOSE, a learnable LSH framework for efficient neural network training.

50, TITLE: Gradient Projection Memory for Continual Learning
https://openreview.net/forum?id=3AOj0RCNC2
AUTHORS: Gobinda Saha, Kaushik Roy
HIGHLIGHT: To avoid catastrophic forgetting in continual learning, we propose a novel approach where a neural network learns new tasks by taking gradient steps in the orthogonal direction to the gradient subspaces deemed important for past tasks.

51, TITLE: Why Are Convolutional Nets More Sample-Efficient than Fully-Connected Nets?
https://openreview.net/forum?id=uCY5MuAxcxU
AUTHORS: Zhiyuan Li, Yi Zhang, Sanjeev Arora
HIGHLIGHT: We construct a single natural distribution on which any fully-connected networks trained with SGD requires \Omega(d^2) samples to generalize while O (1) samples suffices for convolutional architectures.

52, TITLE: Iterated learning for emergent systematicity in VQA
https://openreview.net/forum?id=Pd_oMxH8IlF
AUTHORS: Ankit Vani, Max Schwarzer, Yuchen Lu, Eeshan Dhekane, Aaron Courville
HIGHLIGHT: We use iterated learning to encourage the emergence of structure in the generated programs for neural module networks.

53, TITLE: Coupled Oscillatory Recurrent Neural Network (coRNN): An accurate and (gradient) stable architecture for learning long time dependencies
https://openreview.net/forum?id=F3s69XzWOia
AUTHORS: T. Konstantin Rusch, Siddhartha Mishra
HIGHLIGHT: A biologically motivated and discretized ODE based RNN for learning long-term dependencies, with rigorous bounds mitigating the exploding and vanishing gradient problem.

54, TITLE: Sparse Quantized Spectral Clustering
https://openreview.net/forum?id=pBqLS-7KYAF
AUTHORS: Zhenyu Liao, Romain Couillet, Michael W. Mahoney
HIGHLIGHT: Here, we exploit tools from random matrix theory to make precise statements about how the eigenspectrum of a matrix changes under such nonlinear transformations.

55, TITLE: Graph-Based Continual Learning
https://openreview.net/forum?id=HHSEKOnPvaO
AUTHORS: Binh Tang, David S. Matteson
HIGHLIGHT: In this work, we propose to augment such an array with a learnable random graph that captures pairwise similarities between its samples, and use it not only to learn new tasks but also to guard against forgetting.

56, TITLE: Dynamic Tensor Rematerialization
https://openreview.net/forum?id=Vfs_2RnOD0H
AUTHORS: Marisa Kirisame, Steven Lyubomirsky, Altan Haan, Jennifer Brennan, Mike He, Jared Roesch, Tianqi Chen, Zachary Tatlock
HIGHLIGHT: We present an online algorithm for rematerialization (recomputing intermediate activations during backpropagation instead of storing them), which enables training under low memory, finding that it is competitive with offline techniques.

57, TITLE: Gradient Vaccine: Investigating and Improving Multi-task Optimization in Massively Multilingual Models
https://openreview.net/forum?id=F1vEjWK-lH_
AUTHORS: Zirui Wang, Yulia Tsvetkov, Orhan Firat, Yuan Cao
HIGHLIGHT: In this paper, we attempt to peek into the black-box of multilingual optimization through the lens of loss function geometry.

58, TITLE: CPT: Efficient Deep Neural Network Training via Cyclic Precision
https://openreview.net/forum?id=87ZwsaQNHPZ
AUTHORS: Yonggan Fu, Han Guo, Meng Li, Xin Yang, Yining Ding, Vikas Chandra, Yingyan Lin
HIGHLIGHT: We propose Cyclic Precision Training towards better accuracy-efficiency trade-offs in DNN training.

59, TITLE: Learning a Latent Simplex in Input Sparsity Time
https://openreview.net/forum?id=04LZCAxMSco
AUTHORS: Ainesh Bakshi, Chiranjib Bhattacharyya, Ravi Kannan, David Woodruff, Samson Zhou
HIGHLIGHT: We obtain the first input sparsity runtime algorithm for the problem of learning a latent simplex.

60, TITLE: Expressive Power of Invariant and Equivariant Graph Neural Networks
https://openreview.net/forum?id=lxHgXYN4bwl
AUTHORS: Waiss Azizian, marc lelarge
HIGHLIGHT: In this paper, we propose a theoretical framework able to compare the expressive power of these GNN architectures.

61, TITLE: Discovering a set of policies for the worst case reward
https://openreview.net/forum?id=PUkhWz65dy5
AUTHORS: Tom Zahavy, Andre Barreto, Daniel J Mankowitz, Shaobo Hou, Brendan O'Donoghue, Iurii Kemaev, Satinder Singh
HIGHLIGHT: We study the problem of how to construct a set of policies that can be composed together to solve a collection of reinforcement learning tasks.

62, TITLE: Model-Based Visual Planning with Self-Supervised Functional Distances
https://openreview.net/forum?id=UcoXdfrORC
AUTHORS: Stephen Tian, Suraj Nair, Frederik Ebert, Sudeep Dasari, Benjamin Eysenbach, Chelsea Finn, Sergey Levine
HIGHLIGHT: We combine model-based planning with dynamical distance learning to solve visual goal-reaching tasks, using random, unlabeled, experience.

63, TITLE: Noise against noise: stochastic label noise helps combat inherent label noise
https://openreview.net/forum?id=80FMcTSZ6J0
AUTHORS: Pengfei Chen, Guangyong Chen, Junjie Ye, jingwei zhao, Pheng-Ann Heng
HIGHLIGHT: SGD noise induced by stochastic label noise helps escape sharp minima and prevents overconfidence, hence can mitigate the effects of inherent label noise and improve generalization.

64, TITLE: Contrastive Behavioral Similarity Embeddings for Generalization in Reinforcement Learning
https://openreview.net/forum?id=qda7-sVg84
AUTHORS: Rishabh Agarwal, Marlos C. Machado, Pablo Samuel Castro, Marc G Bellemare
HIGHLIGHT: A contrastive representation learning method which encodes behavioural similarity in RL for improving generalization.

65, TITLE: VAEBM: A Symbiosis between Variational Autoencoders and Energy-based Models
https://openreview.net/forum?id=5m3SEczOV8L
AUTHORS: Zhisheng Xiao, Karsten Kreis, Jan Kautz, Arash Vahdat
HIGHLIGHT: We introduce an energy-based generative model where the data distribution is defined jointly by a VAE and an energy network.

66, TITLE: Geometry-Aware Gradient Algorithms for Neural Architecture Search
https://openreview.net/forum?id=MuSYkd1hxRP
AUTHORS: Liam Li, Mikhail Khodak, Nina Balcan, Ameet Talwalkar
HIGHLIGHT: Studying the right single-level optimization geometry yields state-of-the-art methods for NAS.

67, TITLE: Learning-based Support Estimation in Sublinear Time
https://openreview.net/forum?id=tilovEHA3YS
AUTHORS: Talya Eden, Piotr Indyk, Shyam Narayanan, Ronitt Rubinfeld, Sandeep Silwal, Tal Wagner
HIGHLIGHT: A learning-based algorithm for support size estimation, which given a sufficiently accurate predictor, improves both provably and empirically over state-of-the-art algorithms that do not use a predictor.

68, TITLE: Deciphering and Optimizing Multi-Task Learning: a Random Matrix Approach
https://openreview.net/forum?id=Cri3xz59ga
AUTHORS: Malik Tiomoko, Hafiz Tiomoko Ali, Romain Couillet
HIGHLIGHT: This paper provides a theoretical analysis of Multi Task Learning schemes for large dimensional data

69, TITLE: Autoregressive Entity Retrieval
https://openreview.net/forum?id=5k8F6UU39V
AUTHORS: Nicola De Cao, Gautier Izacard, Sebastian Riedel, Fabio Petroni
HIGHLIGHT: We address entity retrieval by generating their unique name identifiers, left to right, in an autoregressive fashion, and conditioned on the context showing SOTA results in more than 20 datasets with a tiny fraction of the memory of recent systems.

70, TITLE: Systematic generalisation with group invariant predictions
https://openreview.net/forum?id=b9PoimzZFJ
AUTHORS: Faruk Ahmed, Yoshua Bengio, Harm van Seijen, Aaron Courville
HIGHLIGHT: Invariance penalties across splits of a biased dataset can improve systematic generalisation

71, TITLE: Iterative Empirical Game Solving via Single Policy Best Response
https://openreview.net/forum?id=R4aWTjmrEKM
AUTHORS: Max Smith, Thomas Anthony, Michael Wellman
HIGHLIGHT: We introduce two variations of PSRO designed to reduce the amount of simulation required during DRL training.

72, TITLE: Understanding the role of importance weighting for deep learning
https://openreview.net/forum?id=_WnwtieRHxM
AUTHORS: Da Xu, Yuting Ye, Chuanwei Ruan
HIGHLIGHT: We study the theoretical properties of importance weighting for deep learning.

73, TITLE: Long-tail learning via logit adjustment
https://openreview.net/forum?id=37nvvqkCo5
AUTHORS: Aditya Krishna Menon, Sadeep Jayasumana, Ankit Singh Rawat, Himanshu Jain, Andreas Veit, Sanjiv Kumar
HIGHLIGHT: Adjusting classifier logits based on class priors, either post-hoc or during training, can improve performance on rare classes.

74, TITLE: DDPNOpt: Differential Dynamic Programming Neural Optimizer
https://openreview.net/forum?id=6s7ME_X5_Un
AUTHORS: Guan-Horng Liu, Tianrong Chen, Evangelos Theodorou
HIGHLIGHT: We introduce a new class of optimal-control-theoretic training methods, DDPNOpt, that performs a distinct backward pass inherited with Bellman optimality and generates layer-wise feedback policies to robustify training over existing training methods

75, TITLE: Learning with feature dependent label noise: a progressive approach
https://openreview.net/forum?id=ZPa2SyGcbwh
AUTHORS: Yikai Zhang, Songzhu Zheng, Pengxiang Wu, Mayank Goswami, Chao Chen
HIGHLIGHT: We propose a progressive label correction approach for noisy label learning task

76, TITLE: Information Laundering for Model Privacy
https://openreview.net/forum?id=dyaIRud1zXg
AUTHORS: Xinran Wang, Yu Xiang, Jun Gao, Jie Ding
HIGHLIGHT: We propose information laundering, a novel framework for enhancing model privacy.

77, TITLE: Mutual Information State Intrinsic Control
https://openreview.net/forum?id=OthEq8I5v1
AUTHORS: Rui Zhao, Yang Gao, Pieter Abbeel, Volker Tresp, Wei Xu
HIGHLIGHT: Motivated by the self-consciousness concept in psychology, we propose a new intrinsic objective that encourages the agent to have maximum control on the environment.

78, TITLE: Benefit of deep learning with non-convex noisy gradient descent: Provable excess risk bound and superiority to kernel methods
https://openreview.net/forum?id=2m0g1wEafh
AUTHORS: Taiji Suzuki, Shunta Akiyama
HIGHLIGHT: We consider a teacher-student regression model, and eventually show that {\it any} linear estimator can be outperformed by deep learning in a sense of the minimax optimal rate especially for a high dimension setting.

79, TITLE: How Does Mixup Help With Robustness and Generalization?
https://openreview.net/forum?id=8yKEo06dKNo
AUTHORS: Linjun Zhang, Zhun Deng, Kenji Kawaguchi, Amirata Ghorbani, James Zou
HIGHLIGHT: A theoretical point of view for Mixup training

80, TITLE: Dataset Inference: Ownership Resolution in Machine Learning

https://openreview.net/forum?id=hvdKKV2yt7T
AUTHORS: Pratyush Maini, Mohammad Yaghini, Nicolas Papernot
HIGHLIGHT: We introduce 'Dataset Inference' as a new method towards resolving model ownership.

81, TITLE: Individually Fair Gradient Boosting
https://openreview.net/forum?id=JBAa9we1AL
AUTHORS: Alexander Vargo, Fan Zhang, Mikhail Yurochkin, Yuekai Sun
HIGHLIGHT: We propose an algorithm for training individually fair gradient boosted decision trees classifiers.

82, TITLE: Large Scale Image Completion via Co-Modulated Generative Adversarial Networks
https://openreview.net/forum?id=sSjqmfsk95O
AUTHORS: Shengyu Zhao, Jonathan Cui, Yilun Sheng, Yue Dong, Xiao Liang, Eric I-Chao Chang, Yan Xu
HIGHLIGHT: Bridging the gap between between image-conditional and unconditional GAN architectures via co-modulation

83, TITLE: Self-Supervised Policy Adaptation during Deployment
https://openreview.net/forum?id=o_V-MjyyGV_
AUTHORS: Nicklas Hansen, Rishabh Jangir, Yu Sun, Guillem Aleny?, Pieter Abbeel, Alexei A Efros, Lerrel Pinto, Xiaolong Wang
HIGHLIGHT: Generalization across enviroments is known to be hard. We propose a self-supervised method for policy adaptation during deployment that assumes no prior knowledge of the test environment, yet still obtain significant improvements.

84, TITLE: Sharpness-aware Minimization for Efficiently Improving Generalization
https://openreview.net/forum?id=6Tm1mposlrM
AUTHORS: Pierre Foret, Ariel Kleiner, Hossein Mobahi, Behnam Neyshabur
HIGHLIGHT: Motivated by the connection between geometry of the loss landscape and generalization, we introduce a procedure for simultaneously minimizing loss value and loss sharpness.

85, TITLE: PMI-Masking: Principled masking of correlated spans
https://openreview.net/forum?id=3Aoft6NWFej
AUTHORS: Yoav Levine, Barak Lenz, Opher Lieber, Omri Abend, Kevin Leyton-Brown, Moshe Tennenholtz, Yoav Shoham
HIGHLIGHT: Joint masking of correlated tokens significantly speeds up and improves BERT's pretraining

86, TITLE: Very Deep VAEs Generalize Autoregressive Models and Can Outperform Them on Images
https://openreview.net/forum?id=RLRXCV6DbEJ
AUTHORS: Rewon Child
HIGHLIGHT: We argue deeper VAEs should perform better, implement one, and show it outperforms all PixelCNN-based autoregressive models in likelihood, while being substantially more efficient.

87, TITLE: Data-Efficient Reinforcement Learning with Self-Predictive Representations
https://openreview.net/forum?id=uCQfPZwRaUu
AUTHORS: Max Schwarzer, Ankesh Anand, Rishab Goel, R Devon Hjelm, Aaron Courville, Philip Bachman
HIGHLIGHT: We propose a temporal, self-supervised objective for RL agents and show that it significantly improves data efficiency in a setting limited to just 2h of gameplay on Atari.

88, TITLE: Watch-And-Help: A Challenge for Social Perception and Human-AI Collaboration
https://openreview.net/forum?id=w_7JMpGZRh0
AUTHORS: Xavier Puig, Tianmin Shu, Shuang Li, Zilin Wang, Yuan-Hong Liao, Joshua B. Tenenbaum, Sanja Fidler, Antonio Torralba
HIGHLIGHT: We introduce Watch-And-Help (WAH), a challenge for testing social intelligence in agents.
For this challenge, we build VirtualHome-Social, a multi-agent household environment, and provide a benchmark including both planning and learning based baselines.

89, TITLE: A Good Image Generator Is What You Need for High-Resolution Video Synthesis
https://openreview.net/forum?id=6puCSjH3hwA
AUTHORS: Yu Tian, Jian Ren, Menglei Chai, Kyle Olszewski, Xi Peng, Dimitris N. Metaxas, Sergey Tulyakov
HIGHLIGHT: We present a framework that leverages contemporary image generators to render high-resolution videos.

90, TITLE: UPDeT: Universal Multi-agent RL via Policy Decoupling with Transformers
https://openreview.net/forum?id=v9c7hr9ADKx

AUTHORS: Siyi Hu, Fengda Zhu, Xiaojun Chang, Xiaodan Liang
HIGHLIGHT: In this paper, we make the first attempt to explore a universal multi-agent reinforcement learning pipeline, designing one single architecture to fit tasks with the requirement of different observation and action configurations.

91, TITLE: BUSTLE: Bottom-Up program Synthesis Through Learning-guided Exploration
https://openreview.net/forum?id=yHeg4PbFHh
AUTHORS: Augustus Odena, Kensen Shi, David Bieber, Rishabh Singh, Charles Sutton, Hanjun Dai
HIGHLIGHT: We use a learned model to guide a bottom-up program synthesis search to efficiently synthesize spreadsheet programs.

92, TITLE: Improving Adversarial Robustness via Channel-wise Activation Suppressing
https://openreview.net/forum?id=zQTezqCCtNx
AUTHORS: Yang Bai, Yuyuan Zeng, Yong Jiang, Shu-Tao Xia, Xingjun Ma, Yisen Wang
HIGHLIGHT: Training with Channel-wise Activation Suppressing (CAS) can help imrove the robustness of adversarial training.

93, TITLE: What are the Statistical Limits of Offline RL with Linear Function Approximation?
https://openreview.net/forum?id=30EvkP2aQLD
AUTHORS: Ruosong Wang, Dean Foster, Sham M. Kakade
HIGHLIGHT: Exponential lower bounds for batch RL with linear function approximation.

94, TITLE: Unlearnable Examples: Making Personal Data Unexploitable
https://openreview.net/forum?id=iAmZUo0DxC0
AUTHORS: Hanxun Huang, Xingjun Ma, Sarah Monazam Erfani, James Bailey, Yisen Wang
HIGHLIGHT: We present a type of error-minimizing noise that can make training examples unlearnable to deep learning.

95, TITLE: Learning Mesh-Based Simulation with Graph Networks
https://openreview.net/forum?id=roNqYL0_XP
AUTHORS: Tobias Pfaff, Meire Fortunato, Alvaro Sanchez-Gonzalez, Peter Battaglia
HIGHLIGHT: We introduce a general method for learning the dynamics of complex physics systems accurately and efficiently on meshes

96, TITLE: Locally Free Weight sharing for Network Width Search
https://openreview.net/forum?id=S0UdquAnr9k
AUTHORS: Xiu Su, Shan You, Tao Huang, Fei Wang, Chen Qian, Changshui Zhang, Chang Xu
HIGHLIGHT: One-shot locally free weight sharing supernet for searching optimal network width

97, TITLE: Graph Convolution with Low-rank Learnable Local Filters
https://openreview.net/forum?id=9OHFhefeB86
AUTHORS: Xiuyuan Cheng, Zichen Miao, Qiang Qiu
HIGHLIGHT: Graph convolution model for landmark data with local graph regularization and provable graph signal representation expressiveness and stability.

98, TITLE: Regularized Inverse Reinforcement Learning
https://openreview.net/forum?id=HgLO8yalfwc
AUTHORS: Wonseok Jeon, Chen-Yang Su, Paul Barde, Thang Doan, Derek Nowrouzezahrai, Joelle Pineau
HIGHLIGHT: We propose tractable solutions of regularized IRL and algorithms to acquire those solutions for both discrete and continuous control problems.

99, TITLE: Interpreting Graph Neural Networks for NLP With Differentiable Edge Masking
https://openreview.net/forum?id=WznmQa42ZAx
AUTHORS: Michael Sejr Schlichtkrull, Nicola De Cao, Ivan Titov
HIGHLIGHT: We present a novel post-hoc interpretation method for graph neural networks, and apply it to analyse two models from the NLP literature.

100, TITLE: Deep Neural Network Fingerprinting by Conferrable Adversarial Examples
https://openreview.net/forum?id=VqzVhqxkjH1
AUTHORS: Nils Lukas, Yuxuan Zhang, Florian Kerschbaum
HIGHLIGHT: Proposal of a new property called "conferrability" for adversarial examples that we use as a method for DNN fingerprinting robust to model extraction.

101, TITLE: Tent: Fully Test-Time Adaptation by Entropy Minimization
https://openreview.net/forum?id=uXl3bZLkr3c
AUTHORS: Dequan Wang, Evan Shelhamer, Shaoteng Liu, Bruno Olshausen, Trevor Darrell
HIGHLIGHT: Deep networks can generalize better during testing by adapting to feedback from their own predictions.

102, TITLE: GAN "Steerability" without optimization
https://openreview.net/forum?id=zDy_nQCXiIj
AUTHORS: Nurit Spingarn, Ron Banner, Tomer Michaeli
HIGHLIGHT: Extracting linear & nonlinear semantic directions in GAN latent space without any required optimization

103, TITLE: Contrastive Divergence Learning is a Time Reversal Adversarial Game
https://openreview.net/forum?id=MLSvqIHRidA
AUTHORS: Omer Yair, Tomer Michaeli
HIGHLIGHT: We present an alternative derivation of the classical Contrastive divergence method, which reveals that it is in fact an adversarial learning procedure.

104, TITLE: Topology-Aware Segmentation Using Discrete Morse Theory
https://openreview.net/forum?id=LGgdb4TS4Z
AUTHORS: Xiaoling Hu, Yusu Wang, Li Fuxin, Dimitris Samaras, Chao Chen
HIGHLIGHT: In this paper, we propose a new approach to train deep image segmentation networks for better topological accuracy.

105, TITLE: Are Neural Rankers still Outperformed by Gradient Boosted Decision Trees?
https://openreview.net/forum?id=Ut1vF_q_vC
AUTHORS: Zhen Qin, Le Yan, Honglei Zhuang, Yi Tay, Rama Kumar Pasumarthi, Xuanhui Wang, Michael Bendersky, Marc Najork
HIGHLIGHT: Are Neural Rankers still Outperformed by Gradient Boosted Decision Trees?

106, TITLE: Predicting Infectiousness for Proactive Contact Tracing
https://openreview.net/forum?id=lVgB2FUbzuQ
AUTHORS: Yoshua Bengio, Prateek Gupta, Tegan Maharaj, Nasim Rahaman, Martin Weiss, Tristan Deleu, Eilif Benjamin Muller, Meng Qu, victor schmidt, Pierre-luc St-charles, hannah alsdurf, Olexa Bilaniuk, david buckeridge, gaetan caron, pierre luc carrier, Joumana Ghosn, satya ortiz gagne, Christopher Pal, Irina Rish, Bernhard Sch?lkopf, abhinav sharma, Jian Tang, andrew williams
HIGHLIGHT: Proposes a framework called Proactive Contact Tracing which uses distributed inference of expected Covid-19 infectiousness to provide individualized, private recommendations.

107, TITLE: Regularization Matters in Policy Optimization - An Empirical Study on Continuous Control
https://openreview.net/forum?id=yr1mzrH3IC
AUTHORS: Zhuang Liu, Xuanlin Li, Bingyi Kang, Trevor Darrell
HIGHLIGHT: We show that conventional regularization methods (e.g., l2), which have been largely ignored in RL methods, can be very effective in policy optimization on continuous control tasks; we also analyze why they can help from several perspectives.

108, TITLE: Minimum Width for Universal Approximation
https://openreview.net/forum?id=O-XJwyoIF-k
AUTHORS: Sejun Park, Chulhee Yun, Jaeho Lee, Jinwoo Shin
HIGHLIGHT: We establish the tight bound on width for the universal approximability of neural network.

109, TITLE: Towards Robustness Against Natural Language Word Substitutions
https://openreview.net/forum?id=ks5nebunVn_
AUTHORS: Xinshuai Dong, Hong Liu, Rongrong Ji, Anh Tuan Luu
HIGHLIGHT: Model word substitutions at the vector space using convex hull towards robustness.

110, TITLE: Dynamics of Deep Equilibrium Linear Models
https://openreview.net/forum?id=p-NZIuwqhl4
AUTHORS: Kenji Kawaguchi
HIGHLIGHT: We analyze gradient dynamics of deep equilibrium linear models and mathematically prove its theoretical properties.

111, TITLE: Structured Prediction as Translation between Augmented Natural Languages
https://openreview.net/forum?id=US-TP-xnXI
AUTHORS: Giovanni Paolini, Ben Athiwaratkun, Jason Krone, Jie Ma, Alessandro Achille, RISHITA ANUBHAI, Cicero Nogueira dos Santos, Bing Xiang, Stefano Soatto
HIGHLIGHT: We propose a unified text-to-text approach to handle a variety of structured prediction tasks in a single model, allowing seamless multi-task training and providing extra benefits on low-resource scenarios.

112, TITLE: How Benign is Benign Overfitting ?
https://openreview.net/forum?id=g-wu9TMPODo
AUTHORS: Amartya Sanyal, Puneet K. Dokania, Varun Kanade, Philip Torr
HIGHLIGHT: Interpolating label noise hurts adversarial robustness

113, TITLE: Correcting experience replay for multi-agent communication
https://openreview.net/forum?id=xvxPuCkCNPO
AUTHORS: Sanjeevan Ahilan, Peter Dayan
HIGHLIGHT: We improve multi-agent learning by relabelling past experience to better reflect current communication policies

114, TITLE: Emergent Symbols through Binding in External Memory
https://openreview.net/forum?id=LSFCEb3GYU7
AUTHORS: Taylor Whittington Webb, Ishan Sinha, Jonathan Cohen
HIGHLIGHT: We introduce a new architecture, the Emergent Symbol Binding Network, that enables rapid learning of abstract rules and strong generalization of those rules to novel entities.

115, TITLE: Influence Estimation for Generative Adversarial Networks
https://openreview.net/forum?id=opHLcXxYTC_
AUTHORS: Naoyuki Terashita, Hiroki Ohashi, Yuichi Nonaka, Takashi Kanemaru
HIGHLIGHT: We propose an influence estimation method which predicts how GAN's output changes if an training instance is abesent, and propose to evaluate harmfulness of the instance by estimating how its absence improves GAN evaluation metric.

116, TITLE: PlasticineLab: A Soft-Body Manipulation Benchmark with Differentiable Physics
https://openreview.net/forum?id=xCcdBRQEDW
AUTHORS: Zhiao Huang, Yuanming Hu, Tao Du, Siyuan Zhou, Hao Su, Joshua B. Tenenbaum, Chuang Gan
HIGHLIGHT: We propose a soft-body manipulation benchmark with differentiable physics support.

117, TITLE: Implicit Normalizing Flows
https://openreview.net/forum?id=8PS8m9oYtNy
AUTHORS: Cheng Lu, Jianfei Chen, Chongxuan Li, Qiuhao Wang, Jun Zhu
HIGHLIGHT: We generalize normalizing flows, allowing the mapping to be implicitly defined by the roots of an equation and enlarging the expressiveness power while retaining the tractability.

118, TITLE: Support-set bottlenecks for video-text representation learning
https://openreview.net/forum?id=EqoXe2zmhrh
AUTHORS: Mandela Patrick, Po-Yao Huang, Yuki Asano, Florian Metze, Alexander G Hauptmann, Joao F. Henriques, Andrea Vedaldi
HIGHLIGHT: We use a generative objective to improve the instance discrimination limitations of contrastive learning to set new state-of-the-art results in text-to-video retrieval

119, TITLE: Winning the L2RPN Challenge: Power Grid Management via Semi-Markov Afterstate Actor-Critic
https://openreview.net/forum?id=LmUJqB1Cz8
AUTHORS: Deunsol Yoon, Sunghoon Hong, Byung-Jun Lee, Kee-Eung Kim
HIGHLIGHT: We present an off-policy actor-critic approach that effectively tackles the unique challenges in power grid management by reinforcement learning, adopting the hierarchical policy together with the afterstate representation.

120, TITLE: Learning Incompressible Fluid Dynamics from Scratch - Towards Fast, Differentiable Fluid Models that Generalize
https://openreview.net/forum?id=KUDUoRsEphu
AUTHORS: Nils Wandel, Michael Weinmann, Reinhard Klein
HIGHLIGHT: We present an unsupervised training framework for incompressible fluid dynamics that allows neural networks to perform fast, accurate, differentiable fluid simulations and generalize to new domain geometries.

121, TITLE: The Traveling Observer Model: Multi-task Learning Through Spatial Variable Embeddings
https://openreview.net/forum?id=qYda4oLEc1
AUTHORS: Elliot Meyerson, Risto Miikkulainen
HIGHLIGHT: Learn a single model across "unrelated" tasks by embedding their input and output variables in a shared space.

122, TITLE: Grounded Language Learning Fast and Slow
https://openreview.net/forum?id=wpSWuz_hyqA
AUTHORS: Felix Hill, Olivier Tieleman, Tamara von Glehn, Nathaniel Wong, Hamza Merzic, Stephen Clark
HIGHLIGHT: A language-learning agent with dual-coding external memory meta-learns to combine fast-mapped and semantic lexical knowledge to execute instructions in one-shot..

123, TITLE: Long-tailed Recognition by Routing Diverse Distribution-Aware Experts
https://openreview.net/forum?id=D9I3drBz4UC
AUTHORS: Xudong Wang, Long Lian, Zhongqi Miao, Ziwei Liu, Stella Yu
HIGHLIGHT: We aim to reduce both the bias and the variance of a long-tailed classifier by RoutIng Diverse Experts (RIDE).

124, TITLE: Differentially Private Learning Needs Better Features (or Much More Data)
https://openreview.net/forum?id=YTWGvpFOQD-
AUTHORS: Florian Tramer, Dan Boneh
HIGHLIGHT: Linear models with handcrafted features outperform end-to-end CNNs for differentially private learning

125, TITLE: Unsupervised Object Keypoint Learning using Local Spatial Predictability
https://openreview.net/forum?id=GJwMHetHc73
AUTHORS: Anand Gopalakrishnan, Sjoerd van Steenkiste, J?rgen Schmidhuber
HIGHLIGHT: We propose PermaKey, a novel method for learning object keypoint representations that leverages local predictability as a measure of objectness.

126, TITLE: On Statistical Bias In Active Learning: How and When to Fix It
https://openreview.net/forum?id=JiYq3eqTKY
AUTHORS: Sebastian Farquhar, Yarin Gal, Tom Rainforth
HIGHLIGHT: We formalize the bias introduced by active learning and investigate the situations in which it can be harmful and sometimes even helpful, further introducing novel corrective weights to remove it when doing so is beneficial.

127, TITLE: Implicit Convex Regularizers of CNN Architectures: Convex Optimization of Two- and Three-Layer Networks in Polynomial Time
https://openreview.net/forum?id=0N8jUH4JMv6
AUTHORS: Tolga Ergen, Mert Pilanci
HIGHLIGHT: We study the training problem for various CNN architectures with ReLU activations and introduce equivalent finite dimensional convex formulations that can be used to globally optimize these architectures.

128, TITLE: Generalization in data-driven models of primary visual cortex
https://openreview.net/forum?id=Tp7kI90Htd
AUTHORS: Konstantin-Klemens Lurz, Mohammad Bashiri, Konstantin Willeke, Akshay Jagadish, Eric Wang, Edgar Y. Walker, Santiago A Cadena, Taliah Muhammad, Erick Cobos, Andreas S. Tolias, Alexander S Ecker, Fabian H. Sinz
HIGHLIGHT: We introduce a novel network architecture which sets a new state of the art at predicting neural responses to visual input and successfully learns generalizing features of mouse visual cortex (V1).

129, TITLE: Mathematical Reasoning via Self-supervised Skip-tree Training
https://openreview.net/forum?id=YmqAnY0CMEy
AUTHORS: Markus Norman Rabe, Dennis Lee, Kshitij Bansal, Christian Szegedy
HIGHLIGHT: We demonstrate that self-supervised language modeling applied to mathematical formulas enables logical reasoning.

130, TITLE: Parameterization of Hypercomplex Multiplications
https://openreview.net/forum?id=rcQdycl0zyk
AUTHORS: Aston Zhang, Yi Tay, SHUAI Zhang, Alvin Chan, Anh Tuan Luu, Siu Hui, Jie Fu
HIGHLIGHT: We propose a new parameterization of hypercomplex multiplications for architectural flexibility and effectiveness.

131, TITLE: Distributional Sliced-Wasserstein and Applications to Generative Modeling
https://openreview.net/forum?id=QYjO70ACDK
AUTHORS: Khai Nguyen, Nhat Ho, Tung Pham, Hung Bui
HIGHLIGHT: A new optimal transport distance based on slicing approach and its applications to generative modeling.

132, TITLE: Async-RED: A Provably Convergent Asynchronous Block Parallel Stochastic Method using Deep Denoising Priors
https://openreview.net/forum?id=9EsrXMzlFQY
AUTHORS: Yu Sun, Jiaming Liu, Yiran Sun, Brendt Wohlberg, Ulugbek Kamilov
HIGHLIGHT: Our work develops a novel deep-regularized asynchronous parallel method with provable convergence guarantees for solving large-scale inverse problems.

133, TITLE: DeepAveragers: Offline Reinforcement Learning By Solving Derived Non-Parametric MDPs
https://openreview.net/forum?id=eMP1j9efXtX
AUTHORS: Aayam Kumar Shrestha, Stefan Lee, Prasad Tadepalli, Alan Fern
HIGHLIGHT: The paper introduces and investigates an offline RL approach based on optimally solving a finite-state MDP that is derived from the experience dataset using any latent state representation.

134, TITLE: Learning from Protein Structure with Geometric Vector Perceptrons
https://openreview.net/forum?id=1YLJDvSx6J4
AUTHORS: Bowen Jing, Stephan Eismann, Patricia Suriana, Raphael John Lamarre Townshend, Ron Dror
HIGHLIGHT: We introduce a novel graph neural network layer to learn from the structure of macromolecules.

135, TITLE: Behavioral Cloning from Noisy Demonstrations
https://openreview.net/forum?id=zrT3HcsWSAt
AUTHORS: Fumihiro Sasaki, Ryota Yamashina
HIGHLIGHT: We propose an imitation learning algorithm to learn from non-optimal (noisy) demonstrations without any environment interactions and annotations associated with the demonstrations.

136, TITLE: Undistillable: Making A Nasty Teacher That CANNOT teach students
https://openreview.net/forum?id=0zvfm-nZqQs
AUTHORS: Haoyu Ma, Tianlong Chen, Ting-Kuei Hu, Chenyu You, Xiaohui Xie, Zhangyang Wang
HIGHLIGHT: We propose the Nasty Teacher, a defensive approach to prevent unauthorized cloning from a teacher model through knowledge distillation.

137, TITLE: Multivariate Probabilistic Time Series Forecasting via Conditioned Normalizing Flows
https://openreview.net/forum?id=WiGQBFuVRv
AUTHORS: Kashif Rasul, Abdul-Saboor Sheikh, Ingmar Schuster, Urs M Bergmann, Roland Vollgraf
HIGHLIGHT: SOTA Multivariate probabilistic time series forecasting using RNNs or Attention to model the dynamics and normalizing flows for the emission model.

138, TITLE: Image Augmentation Is All You Need: Regularizing Deep Reinforcement Learning from Pixels
https://openreview.net/forum?id=GY6-6sTvGaf
AUTHORS: Denis Yarats, Ilya Kostrikov, Rob Fergus
HIGHLIGHT: The first successful demonstration that image augmentation can be applied to image-based Deep RL to achieve SOTA performance.

139, TITLE: HW-NAS-Bench: Hardware-Aware Neural Architecture Search Benchmark
https://openreview.net/forum?id=_0kaDkv3dVf
AUTHORS: Chaojian Li, Zhongzhi Yu, Yonggan Fu, Yongan Zhang, Yang Zhao, Haoran You, Qixuan Yu, Yue Wang, Cong Hao, Yingyan Lin
HIGHLIGHT: A Hardware-Aware Neural Architecture Search Benchmark

140, TITLE: Practical Real Time Recurrent Learning with a Sparse Approximation
https://openreview.net/forum?id=q3KSThy2GwB
AUTHORS: Jacob Menick, Erich Elsen, Utku Evci, Simon Osindero, Karen Simonyan, Alex Graves
HIGHLIGHT: We show how to make RTRL efficient with sparse RNNs, a sparse approximation, or both.

141, TITLE: Random Feature Attention
https://openreview.net/forum?id=QtTKTdVrFBB

AUTHORS: Hao Peng, Nikolaos Pappas, Dani Yogatama, Roy Schwartz, Noah Smith, Lingpeng Kong
HIGHLIGHT: We propose a random-feature-based attention that scales linearly in sequence length, and performs on par with strong transformer baselines on language modeling and machine translation.

142, TITLE: A Gradient Flow Framework For Analyzing Network Pruning
https://openreview.net/forum?id=rumv7QmLUue
AUTHORS: Ekdeep Singh Lubana, Robert Dick
HIGHLIGHT: This paper establishes the relationship between regularly used importance measures for network pruning and evolution of model parameters under gradient flow, thus providing useful insights into pruning early-on in training.

143, TITLE: Recurrent Independent Mechanisms
https://openreview.net/forum?id=mLcmdlEUxy-
AUTHORS: Anirudh Goyal, Alex Lamb, Jordan Hoffmann, Shagun Sodhani, Sergey Levine, Yoshua Bengio, Bernhard Sch?lkopf
HIGHLIGHT: Learning recurrent mechanisms which operate independently, and sparingly interact can lead to better generalization to out of distribution samples.

144, TITLE: The Intrinsic Dimension of Images and Its Impact on Learning
https://openreview.net/forum?id=XJk19XzGq2J
AUTHORS: Phil Pope, Chen Zhu, Ahmed Abdelkader, Micah Goldblum, Tom Goldstein
HIGHLIGHT: We measure the dimensionality of common used datasets, and experimentally investigate whether the links between dimensionality and learning that have been identified in the manifold learning literature describe the behaviors of deep neural networks.

145, TITLE: Uncertainty Sets for Image Classifiers using Conformal Prediction
https://openreview.net/forum?id=eNdiU_DbM9
AUTHORS: Anastasios Nikolas Angelopoulos, Stephen Bates, Michael Jordan, Jitendra Malik
HIGHLIGHT: We quantify uncertainty for image classifiers using prediction sets, with detailed experiments on Imagenet Val and V2.

146, TITLE: Sequential Density Ratio Estimation for Simultaneous Optimization of Speed and Accuracy
https://openreview.net/forum?id=Rhsu5qD36cL
AUTHORS: Akinori F Ebihara, Taiki Miyagawa, Kazuyuki Sakurai, Hitoshi Imaoka
HIGHLIGHT: With a novel sequential density estimation algorithm, we relax critical assumptions of the classical Sequential Probability Ratio Test to be applicable in various real-world scenarios.

147, TITLE: Disentangled Recurrent Wasserstein Autoencoder
https://openreview.net/forum?id=O7ms4LFdsX
AUTHORS: Jun Han, Martin Renqiang Min, Ligong Han, Xuan Zhang, Li Erran Li
HIGHLIGHT: We propose the first recurrent Wasserstein Autoencoder for learning disentangled representations of sequential data with theoretical analysis.

148, TITLE: Generalization bounds via distillation
https://openreview.net/forum?id=EGdFhBzmAwB
AUTHORS: Daniel Hsu, Ziwei Ji, Matus Telgarsky, Lan Wang
HIGHLIGHT: This paper provides a suite of mathematical tools to bound the generalization error of networks which possess low-complexity distillations.

149, TITLE: Neural Approximate Sufficient Statistics for Implicit Models
https://openreview.net/forum?id=SRDuJssQud
AUTHORS: Yanzhi Chen, Dinghuai Zhang, Michael U. Gutmann, Aaron Courville, Zhanxing Zhu
HIGHLIGHT: We learn low-dimensional near-sufficient statistics by infomax principle to improve likelihood-free inference methods.

150, TITLE: A Panda? No, It's a Sloth: Slowdown Attacks on Adaptive Multi-Exit Neural Network Inference
https://openreview.net/forum?id=9xC2tWEwBD
AUTHORS: Sanghyun Hong, Yigitcan Kaya, Ionu?-Vlad Modoranu, Tudor Dumitras
HIGHLIGHT: Is the computational savings provided by the input-adaptive 'multi-exit architectures' robust against adversarial perturbations? No.

151, TITLE: Orthogonalizing Convolutional Layers with the Cayley Transform
https://openreview.net/forum?id=Pbj8H_jEHYv
AUTHORS: Asher Trockman, J Zico Kolter
HIGHLIGHT: In this work, we propose and evaluate an alternative approach to directly parameterize convolutional layers that are constrained to be orthogonal.

152, TITLE: LambdaNetworks: Modeling long-range Interactions without Attention
https://openreview.net/forum?id=xTJEN-ggl1b
AUTHORS: Irwan Bello
HIGHLIGHT: Scalable framework for capturing long-range interactions between input and structured contextual information, which leads to strong improvements in vision tasks.

153, TITLE: Mind the Pad -- CNNs Can Develop Blind Spots
https://openreview.net/forum?id=m1CD7tPubNy
AUTHORS: Bilal Alsallakh, Narine Kokhlikyan, Vivek Miglani, Jun Yuan, Orion Reblitz-Richardson
HIGHLIGHT: The padding mechanism in CNNs can induce harmful spatial bias in the learned weights and in the feature maps, which can be mitigated with careful architectural choices.

154, TITLE: Meta-GMVAE: Mixture of Gaussian VAE for Unsupervised Meta-Learning
https://openreview.net/forum?id=wS0UFjsNYjn
AUTHORS: Dong Bok Lee, Dongchan Min, Seanie Lee, Sung Ju Hwang
HIGHLIGHT: we propose a principled unsupervised meta-learning model which meta-learns a set-level variational posterior, by matching it with multi-modal prior distribution obtained by EM.

155, TITLE: Fast Geometric Projections for Local Robustness Certification
https://openreview.net/forum?id=zWy1uxjDdZJ
AUTHORS: Aymeric Fromherz, Klas Leino, Matt Fredrikson, Bryan Parno, Corina Pasareanu
HIGHLIGHT: We present a fast, scalable procedure for checking local robustness in neural networks

156, TITLE: Fidelity-based Deep Adiabatic Scheduling
https://openreview.net/forum?id=NECTfffOvn1
AUTHORS: Eli Ovits, Lior Wolf
HIGHLIGHT: A new loss for applying supervised deep learning to the problem of scheduling adiabatic quantum computations

157, TITLE: On Self-Supervised Image Representations for GAN Evaluation
https://openreview.net/forum?id=NeRdBeTionN
AUTHORS: Stanislav Morozov, Andrey Voynov, Artem Babenko
HIGHLIGHT: We show that the state-of-the-art self-supervised representations should be used when comparing GANs on the non-Imagenet datasets

158, TITLE: Retrieval-Augmented Generation for Code Summarization via Hybrid GNN
https://openreview.net/forum?id=zv-typ1gPxA
AUTHORS: Shangqing Liu, Yu Chen, Xiaofei Xie, Jing Kai Siow, Yang Liu
HIGHLIGHT: This paper proposes a novel retrieval-augmented mechanism to augment the code semantics with hybrid graph neural network for source code summarization.
To evaluate the proposed approach, we release a new challenging benchmark, crawled from diversified large-scale open-source C projects (total 95k+ unique functions in the dataset).

159, TITLE: Self-supervised Visual Reinforcement Learning with Object-centric Representations
https://openreview.net/forum?id=xppLmXCbOw1
AUTHORS: Andrii Zadaianchuk, Maximilian Seitzer, Georg Martius
HIGHLIGHT: The combination of object-centric representations and goal-conditioned attention policies helps autonomous agents to learn useful multi-task policies in visual multi-object environments

160, TITLE: Identifying nonlinear dynamical systems with multiple time scales and long-range dependencies
https://openreview.net/forum?id=_XYzwxPIQu6
AUTHORS: Dominik Schmidt, Georgia Koppe, Zahra Monfared, Max Beutelspacher, Daniel Durstewitz
HIGHLIGHT: We introduce a novel regularization for ReLU-based vanilla RNN that mitigates the exploding vs. vanishing gradient problem while retaining a simple mathematical structure that makes the RNN's dynamical systems properties partly analytically tractable

161, TITLE: Neural Topic Model via Optimal Transport
https://openreview.net/forum?id=Oos98K9Lv-k
AUTHORS: He Zhao, Dinh Phung, Viet Huynh, Trung Le, Wray Buntine
HIGHLIGHT: This paper presents a neural topic model via optimal transport, which can discover more coherent and diverse topics and derive better document representations for both regular and short texts.

162, TITLE: Memory Optimization for Deep Networks
https://openreview.net/forum?id=bnY0jm4l59
AUTHORS: Aashaka Shah, Chao-Yuan Wu, Jayashree Mohan, Vijay Chidambaram, Philipp Kraehenbuehl
HIGHLIGHT: MONeT reduces the memory footprint of training while minimizing compute overhead by jointly optimizing checkpointing with operator optimizations.

163, TITLE: Stabilized Medical Image Attacks
https://openreview.net/forum?id=QfTXQiGYudJ
AUTHORS: Gege Qi, Lijun GONG, Yibing Song, Kai Ma, Yefeng Zheng
HIGHLIGHT: We propose a stabilized adversarial attack method for medical image analysis.

164, TITLE: Quantifying Differences in Reward Functions
https://openreview.net/forum?id=LwEQnp6CYev
AUTHORS: Adam Gleave, Michael D Dennis, Shane Legg, Stuart Russell, Jan Leike
HIGHLIGHT: A theoretically principled distance measure on reward functions that is quick to compute and predicts policy training performance.

165, TITLE: MARS: Markov Molecular Sampling for Multi-objective Drug Discovery
https://openreview.net/forum?id=kHSu4ebxFXY
AUTHORS: Yutong Xie, Chence Shi, Hao Zhou, Yuwei Yang, Weinan Zhang, Yong Yu, Lei Li
HIGHLIGHT: In this paper, we propose MARS, a method for multi-objective drug molecule discovery.

166, TITLE: Gauge Equivariant Mesh CNNs: Anisotropic convolutions on geometric graphs
https://openreview.net/forum?id=Jnspzp-oIZE
AUTHORS: Pim De Haan, Maurice Weiler, Taco Cohen, Max Welling
HIGHLIGHT: Expressive anisotropic mesh convolution without having to pick arbitrary kernel orientation by using gauge equivariance

167, TITLE: RMSprop can converge with proper hyper-parameter
https://openreview.net/forum?id=3UDSdyIcBDA
AUTHORS: Naichen Shi, Dawei Li, Mingyi Hong, Ruoyu Sun
HIGHLIGHT: Towards closing the gap between theory and practice, we prove that RMSprop can converge with proper choice of hyper-parameters under certain conditions.

168, TITLE: Revisiting Dynamic Convolution via Matrix Decomposition
https://openreview.net/forum?id=YwpZmcAehZ
AUTHORS: Yunsheng Li, Yinpeng Chen, Xiyang Dai, mengchen liu, Dongdong Chen, Ye Yu, Lu Yuan, Zicheng Liu, Mei Chen, Nuno Vasconcelos
HIGHLIGHT: Efficient network with dynamic matrix decomposition

169, TITLE: Explainable Deep One-Class Classification
https://openreview.net/forum?id=A5VV3UyIQz
AUTHORS: Philipp Liznerski, Lukas Ruff, Robert A. Vandermeulen, Billy Joe Franks, Marius Kloft, Klaus Robert Muller
HIGHLIGHT: We introduce an approach to explainable deep anomaly detection based on fully convolutional neural networks.

170, TITLE: Taking Notes on the Fly Helps Language Pre-Training
https://openreview.net/forum?id=lU5Rs_wCweN
AUTHORS: Qiyu Wu, Chen Xing, Yatao Li, Guolin Ke, Di He, Tie-Yan Liu
HIGHLIGHT: We improve the efficiency of language pre-training methods through providing better data utilization.

171, TITLE: Mixed-Features Vectors and Subspace Splitting
https://openreview.net/forum?id=l-LGlk4Yl6G
AUTHORS: Alejandro Pimentel-Alarc?n, Daniel L. Pimentel-Alarc?n

HIGHLIGHT: Motivated by metagenomics, recommender systems, dictionary learning, and related problems, this paper introduces subspace splitting(SS): the task of clustering the entries of what we call amixed-features vector, that is, a vector whose subsets of coordinates agree with a collection of subspaces.

172, TITLE: Neural Pruning via Growing Regularization
https://openreview.net/forum?id=o966_Is_nPA
AUTHORS: Huan Wang, Can Qin, Yulun Zhang, Yun Fu
HIGHLIGHT: We propose two new deep network pruning algorithms based a growing regularization paradigm.

173, TITLE: Practical Massively Parallel Monte-Carlo Tree Search Applied to Molecular Design
https://openreview.net/forum?id=6k7VdojAIK
AUTHORS: Xiufeng Yang, Tanuj Aasawat, Kazuki Yoshizoe
HIGHLIGHT: Novel massively parallel MCTS achieves state-of-the-art score in molecular design benchmark.

174, TITLE: Empirical Analysis of Unlabeled Entity Problem in Named Entity Recognition
https://openreview.net/forum?id=5jRVa89sZk
AUTHORS: Yangming Li, lemao liu, Shuming Shi
HIGHLIGHT: This work studys what are the impacts of unlabeled entity problem on NER models and how to effectively eliminate them by a general method.

175, TITLE: Deep Networks and the Multiple Manifold Problem
https://openreview.net/forum?id=O-6Pm_d_Q-
AUTHORS: Sam Buchanan, Dar Gilboa, John Wright
HIGHLIGHT: We prove a finite-time generalization result for deep fully-connected neural networks trained by gradient descent to classify structured data, where the required width, depth, and sample complexity depend only on intrinsic properties of the data.

176, TITLE: Knowledge distillation via softmax regression representation learning
https://openreview.net/forum?id=ZzwDy_wiWv
AUTHORS: Jing Yang, Brais Martinez, Adrian Bulat, Georgios Tzimiropoulos
HIGHLIGHT: We advocate for a method that optimizes the output feature of the penultimate layer of the student network and hence is directly related to representation learning.

177, TITLE: Nearest Neighbor Machine Translation
https://openreview.net/forum?id=7wCBOfJ8hJM
AUTHORS: Urvashi Khandelwal, Angela Fan, Dan Jurafsky, Luke Zettlemoyer, Mike Lewis
HIGHLIGHT: We augment the decoder of a pre-trained machine translation model with a nearest neighbor classifier, substantially improving performance in the single language-pair, multilingual and domain adaptation settings, without any additional training.

178, TITLE: WrapNet: Neural Net Inference with Ultra-Low-Precision Arithmetic
https://openreview.net/forum?id=3SqrRe8FWQ-
AUTHORS: Renkun Ni, Hong-min Chu, Oscar Castaneda, Ping-yeh Chiang, Christoph Studer, Tom Goldstein
HIGHLIGHT: We adapt neural networks to integer overflow and extreme low-bit accumulator, and show the efficacy on both software and hardware platforms.

179, TITLE: Wandering within a world: Online contextualized few-shot learning
https://openreview.net/forum?id=oZIvHV04XgC
AUTHORS: Mengye Ren, Michael Louis Iuzzolino, Michael Curtis Mozer, Richard Zemel
HIGHLIGHT: We propose a new continual few-shot learning paradigm and a new model.
Building upon this setting, we propose a new few-shot learning dataset based on large scale indoor imagery that mimics the visual experience of an agent wandering within a world.

180, TITLE: Few-Shot Learning via Learning the Representation, Provably
https://openreview.net/forum?id=pW2Q2xLwIMD
AUTHORS: Simon Shaolei Du, Wei Hu, Sham M. Kakade, Jason D. Lee, Qi Lei
HIGHLIGHT: We study when and how much representation learning can help few-shot learning by drastically reducing sample complexity on the target task.

181, TITLE: AdaGCN: Adaboosting Graph Convolutional Networks into Deep Models

https://openreview.net/forum?id=QkRbdiiEjM
AUTHORS: Ke Sun, Zhouchen Lin, Zhanxing Zhu
HIGHLIGHT: In this paper, we propose a novel RNN-like deep graph neural network architecture by incorporating AdaBoost into the computation of network; and the proposed graph convolutional network called AdaGCN~(Adaboosting Graph Convolutional Network) has the ability to efficiently extract knowledge from high-order neighbors of current nodes and then integrates knowledge from different hops of neighbors into the network in an Adaboost way.

182, TITLE: MultiModalQA: complex question answering over text, tables and images
https://openreview.net/forum?id=ee6W5UgQLa
AUTHORS: Alon Talmor, Ori Yoran, Amnon Catav, Dan Lahav, Yizhong Wang, Akari Asai, Gabriel Ilharco, Hannaneh Hajishirzi, Jonathan Berant
HIGHLIGHT: MultiModalQA: A question answering dataset that requires multi-modal multi-hop reasoning over wikipedia text, tables and images, accompanied by a new multi-hop model for tackling the task.

183, TITLE: Net-DNF: Effective Deep Modeling of Tabular Data
https://openreview.net/forum?id=73WTGs96kho
AUTHORS: Gal Elidan, Liran Katzir, Ran El-Yaniv
HIGHLIGHT: Neural network architecture for tabular data

184, TITLE: Optimal Regularization can Mitigate Double Descent
https://openreview.net/forum?id=7R7fAoUygoa
AUTHORS: Preetum Nakkiran, Prayaag Venkat, Sham M. Kakade, Tengyu Ma
HIGHLIGHT: Optimal regularization can provably avoid double-descent in certain settings.

185, TITLE: Meta Back-Translation
https://openreview.net/forum?id=3jjmdp7Hha
AUTHORS: Hieu Pham, Xinyi Wang, Yiming Yang, Graham Neubig
HIGHLIGHT: Use meta learning to teach the back-translation model to generate better back-translated sentences.

186, TITLE: Learning A Minimax Optimizer: A Pilot Study
https://openreview.net/forum?id=nkIDwI6oO4_
AUTHORS: Jiayi Shen, Xiaohan Chen, Howard Heaton, Tianlong Chen, Jialin Liu, Wotao Yin, Zhangyang Wang
HIGHLIGHT: This paper introduces the learning to optimize (L2O) methodology, called Twin L2O, for minimax optimization consisting of two LSTMs.

187, TITLE: A Wigner-Eckart Theorem for Group Equivariant Convolution Kernels
https://openreview.net/forum?id=ajOrOhQOsYx
AUTHORS: Leon Lang, Maurice Weiler
HIGHLIGHT: We parameterize equivariant convolution kernels by proving a generalization of the Wigner-Eckart theorem for spherical tensor operators.

188, TITLE: Viewmaker Networks: Learning Views for Unsupervised Representation Learning
https://openreview.net/forum?id=enoVQWLsfyL
AUTHORS: Alex Tamkin, Mike Wu, Noah Goodman
HIGHLIGHT: We present a new generative model that produces views for contrastive learning, matching or outperforming hand-crafted views on image, speech, and wearable sensor datasets

189, TITLE: Scalable Transfer Learning with Expert Models
https://openreview.net/forum?id=23ZjUGpjcc
AUTHORS: Joan Puigcerver, Carlos Riquelme Ruiz, Basil Mustafa, Cedric Renggli, Andr? Susano Pinto, Sylvain Gelly, Daniel Keysers, Neil Houlsby
HIGHLIGHT: We explore the use of expert representations for transfer with a simple, yet effective, strategy.

190, TITLE: Negative Data Augmentation
https://openreview.net/forum?id=Ovp8dvB8IBH
AUTHORS: Abhishek Sinha, Kumar Ayush, Jiaming Song, Burak Uzkent, Hongxia Jin, Stefano Ermon
HIGHLIGHT: We propose a framework to do Negative Data Augmentation for generative models and self-supervised learning

191, TITLE: Fantastic Four: Differentiable and Efficient Bounds on Singular Values of Convolution Layers
https://openreview.net/forum?id=JCRblSgs34Z

AUTHORS: Sahil Singla, Soheil Feizi
HIGHLIGHT: We derive four provable upper bounds on the largest singular value of convolution layers that are differentiable, independent of size of input image and can be computed efficiently during training with negligible overhead.

192, TITLE: CoDA: Contrast-enhanced and Diversity-promoting Data Augmentation for Natural Language Understanding
https://openreview.net/forum?id=Ozk9MrX1hvA
AUTHORS: Yanru Qu, Dinghan Shen, Yelong Shen, Sandra Sajeev, Weizhu Chen, Jiawei Han
HIGHLIGHT: In this paper, we propose a novel data augmentation frame-work dubbed CoDA, which synthesizes diverse and informative augmented examples by integrating multiple transformations organically.

193, TITLE: Teaching with Commentaries
https://openreview.net/forum?id=4RbdgBh9gE
AUTHORS: Aniruddh Raghu, Maithra Raghu, Simon Kornblith, David Duvenaud, Geoffrey Hinton
HIGHLIGHT: We propose a flexible framework for neural network teaching, demonstrate it in various settings, and find that it can improve performance and yield insights about datasets and the training process.

194, TITLE: MixKD: Towards Efficient Distillation of Large-scale Language Models
https://openreview.net/forum?id=UFGEelJkLu5
AUTHORS: Kevin J Liang, Weituo Hao, Dinghan Shen, Yufan Zhou, Weizhu Chen, Changyou Chen, Lawrence Carin
HIGHLIGHT: We propose MixKD, a distillation framework leveraging mixup for large-scale language models.

195, TITLE: FairFil: Contrastive Neural Debiasing Method for Pretrained Text Encoders
https://openreview.net/forum?id=N6JECD-PI5w
AUTHORS: Pengyu Cheng, Weituo Hao, Siyang Yuan, Shijing Si, Lawrence Carin
HIGHLIGHT: A debiasing method for large-scale pretrained text encoders via contrastive learning.

196, TITLE: Probabilistic Numeric Convolutional Neural Networks
https://openreview.net/forum?id=T1XmO8ScKim
AUTHORS: Marc Anton Finzi, Roberto Bondesan, Max Welling
HIGHLIGHT: We build a neural network which integrates internal discretization error and missing values probabilistically with GPs

197, TITLE: Computational Separation Between Convolutional and Fully-Connected Networks
https://openreview.net/forum?id=hkMoYYEkBoI
AUTHORS: eran malach, Shai Shalev-Shwartz
HIGHLIGHT: We show a computational separation between convolutional and fully-connected networks, proving that the former can leverage strong local structure in the data.

198, TITLE: On the Stability of Fine-tuning BERT: Misconceptions, Explanations, and Strong Baselines
https://openreview.net/forum?id=nzpLWnVAyah
AUTHORS: Marius Mosbach, Maksym Andriushchenko, Dietrich Klakow
HIGHLIGHT: We provide an analysis of the fine-tuning instability of BERT-based models and present a simple method to fix it

199, TITLE: Variational Information Bottleneck for Effective Low-Resource Fine-Tuning
https://openreview.net/forum?id=kvhzKz-_DMF
AUTHORS: Rabeeh Karimi mahabadi, Yonatan Belinkov, James Henderson
HIGHLIGHT: We propose to use Variational Information Bottleneck to suppress irrelevant features for an effective fine-tuning of large-scale language models in low-resource scenarios.

200, TITLE: Witches' Brew: Industrial Scale Data Poisoning via Gradient Matching
https://openreview.net/forum?id=01olnfLIbD
AUTHORS: Jonas Geiping, Liam H Fowl, W. Ronny Huang, Wojciech Czaja, Gavin Taylor, Michael Moeller, Tom Goldstein
HIGHLIGHT: Data poisoning attacks that successfully poison neural networks trained from scratch, even on large-scale datasets like ImageNet.

201, TITLE: DEBERTA: DECODING-ENHANCED BERT WITH DISENTANGLED ATTENTION
https://openreview.net/forum?id=XPZIaotutsD
AUTHORS: Pengcheng He, Xiaodong Liu, Jianfeng Gao, Weizhu Chen

HIGHLIGHT: A new model architecture DeBERTa is proposed that improves the BERT and RoBERTa models using disentangled attention and enhanced mask decoder.

202, TITLE: Optimism in Reinforcement Learning with Generalized Linear Function Approximation
https://openreview.net/forum?id=CBmJwzneppz
AUTHORS: Yining Wang, Ruosong Wang, Simon Shaolei Du, Akshay Krishnamurthy
HIGHLIGHT: A provably efficient (statistically and computationally) algorithm for reinforcement learning with generalized linear function approximation and no explicit dynamics assumptions.

203, TITLE: Graph Traversal with Tensor Functionals: A Meta-Algorithm for Scalable Learning
https://openreview.net/forum?id=6DOZ8XNNfGN
AUTHORS: Elan Sopher Markowitz, Keshav Balasubramanian, Mehrnoosh Mirtaheri, Sami Abu-El-Haija, Bryan Perozzi, Greg Ver Steeg, Aram Galstyan
HIGHLIGHT: GTTF is a meta-algorithm, upon which, many algorithms for graph learning can be implemented, automatically giving them efficiency and scale, yet unbiased learning.

204, TITLE: Diverse Video Generation using a Gaussian Process Trigger
https://openreview.net/forum?id=Qm7R_SdqTpT
AUTHORS: Gaurav Shrivastava, Abhinav Shrivastava
HIGHLIGHT: Diverse future frame synthesis by modeling the diversity of future states using a Gaussian Process, and using Bayesian inference to sample diverse future states.

205, TITLE: Signatory: differentiable computations of the signature and logsignature transforms, on both CPU and GPU
https://openreview.net/forum?id=lqU2cs3Zca
AUTHORS: Patrick Kidger, Terry Lyons
HIGHLIGHT: Differentiable, GPU-capable implementations of the (log)signature transform via novel algorithms.

206, TITLE: MoPro: Webly Supervised Learning with Momentum Prototypes
https://openreview.net/forum?id=0-EYBhgw80y
AUTHORS: Junnan Li, Caiming Xiong, Steven Hoi
HIGHLIGHT: MoPro is a new webly-supervised learning framework which advances representation learning using freely-available Web images.

207, TITLE: A Universal Representation Transformer Layer for Few-Shot Image Classification
https://openreview.net/forum?id=04cII6MumYV
AUTHORS: Lu Liu, William L. Hamilton, Guodong Long, Jing Jiang, Hugo Larochelle
HIGHLIGHT: code at: https://github.com/liulu112601/URT

208, TITLE: Primal Wasserstein Imitation Learning
https://openreview.net/forum?id=TtYSU29zgR
AUTHORS: Robert Dadashi, Leonard Hussenot, Matthieu Geist, Olivier Pietquin
HIGHLIGHT: A new Imitation Learning method based on optimal transport.

209, TITLE: Learning perturbation sets for robust machine learning
https://openreview.net/forum?id=MIDckA56aD
AUTHORS: Eric Wong, J Zico Kolter
HIGHLIGHT: We learn to characterize real-world changes in well-defined perturbation sets, which allow us train models which are empirically and certifiably robust to real-world adversarial changes.

210, TITLE: CopulaGNN: Towards Integrating Representational and Correlational Roles of Graphs in Graph Neural Networks
https://openreview.net/forum?id=XI-OJ5yyse
AUTHORS: Jiaqi Ma, Bo Chang, Xuefei Zhang, Qiaozhu Mei
HIGHLIGHT: We distinguish the representational and the correlational information encoded by the graphs in node-level prediction tasks, and propose a novel Copula Graph Neural Network to effectively leverage both information.

211, TITLE: On the Critical Role of Conventions in Adaptive Human-AI Collaboration
https://openreview.net/forum?id=8Ln-Bq0mZcy
AUTHORS: Andy Shih, Arjun Sawhney, Jovana Kondic, Stefano Ermon, Dorsa Sadigh
HIGHLIGHT: Training agents that can adapt to new settings in multi-agent games.

212, TITLE: On the Bottleneck of Graph Neural Networks and its Practical Implications
https://openreview.net/forum?id=i80OPhOCVH2
AUTHORS: Uri Alon, Eran Yahav
HIGHLIGHT: A novel observation on GNN limitations: in long-range problems, a computational bottleneck causes over-squashing of information.

213, TITLE: The geometry of integration in text classification RNNs
https://openreview.net/forum?id=42kiJ7n_8xO
AUTHORS: Kyle Aitken, Vinay Venkatesh Ramasesh, Ankush Garg, Yuan Cao, David Sussillo, Niru Maheswaranathan
HIGHLIGHT: We study text classification RNNs using tools from dynamical systems analysis, finding and explaining the geometry of low-dimensional attractor manifolds.

214, TITLE: Gradient Descent on Neural Networks Typically Occurs at the Edge of Stability
https://openreview.net/forum?id=jh-rTtvkGeM
AUTHORS: Jeremy Cohen, Simran Kaur, Yuanzhi Li, J Zico Kolter, Ameet Talwalkar
HIGHLIGHT: We trained neural networks using full-batch gradient descent -- you won't believe what happens next!

215, TITLE: CausalWorld: A Robotic Manipulation Benchmark for Causal Structure and Transfer Learning
https://openreview.net/forum?id=SK7A5pdrgov
AUTHORS: Ossama Ahmed, Frederik Tr?uble, Anirudh Goyal, Alexander Neitz, Manuel Wuthrich, Yoshua Bengio, Bernhard Sch?lkopf, Stefan Bauer
HIGHLIGHT: A benchmark to address the challenge of agents transferring their learned skills to related environments; primarily for causal structure and transfer learning.

216, TITLE: Empirical or Invariant Risk Minimization? A Sample Complexity Perspective
https://openreview.net/forum?id=jrA5GAccy_
AUTHORS: Kartik Ahuja, Jun Wang, Amit Dhurandhar, Karthikeyan Shanmugam, Kush R. Varshney
HIGHLIGHT: In this work, we provide a sample complexity comparison of the recent invariant risk minimization (IRM) framework with the classic empirical risk minimization (ERM) to answer when is IRM better than ERM in terms of out-of-distribution generalization?

217, TITLE: Scaling Symbolic Methods using Gradients for Neural Model Explanation
https://openreview.net/forum?id=V5j-jdoDDP
AUTHORS: Subham Sekhar Sahoo, Subhashini Venugopalan, Li Li, Rishabh Singh, Patrick Riley
HIGHLIGHT: In this work, we propose a technique for combining gradient-based methods with symbolic techniques to scale such analyses and demonstrate its application for model explanation.

218, TITLE: Control-Aware Representations for Model-based Reinforcement Learning
https://openreview.net/forum?id=dgd4EJqsbW5
AUTHORS: Brandon Cui, Yinlam Chow, Mohammad Ghavamzadeh
HIGHLIGHT: Two important questions in this area are how to learn a representation that is amenable to the control problem at hand, and how to achieve an end-to-end framework for representation learning and control. In this paper, we take a few steps towards addressing these questions.

219, TITLE: C-Learning: Learning to Achieve Goals via Recursive Classification
https://openreview.net/forum?id=tc5qisoB-C
AUTHORS: Benjamin Eysenbach, Ruslan Salakhutdinov, Sergey Levine
HIGHLIGHT: We reframe the goal-conditioned RL problem as one of predicting and controlling the future state of the world, and derive a principled algorithm to solve this problem.

220, TITLE: The Bootstrap Framework: Generalization Through the Lens of Online Optimization
https://openreview.net/forum?id=guetrIHLFGI
AUTHORS: Preetum Nakkiran, Behnam Neyshabur, Hanie Sedghi
HIGHLIGHT: We show empirical evidence that the performance gap between offline generalization and online optimization is small and propose an alternative framework for studying generalization.

221, TITLE: Improving VAEs' Robustness to Adversarial Attack
https://openreview.net/forum?id=-Hs_otp2RB
AUTHORS: Matthew JF Willetts, Alexander Camuto, Tom Rainforth, S Roberts, Christopher C Holmes

HIGHLIGHT: We show that regularisation methods first developed to obtain 'disentangled' VAEs increase the robustness of VAEs to adversarial attack; leveraging this insight we propose an even-more-robust hierarchical VAE.

222, TITLE: Learning Visual Representation from Human Interactions
https://openreview.net/forum?id=Qm8UNVCFdh
AUTHORS: Kiana Ehsani, Daniel Gordon, Thomas Hai Dang Nguyen, Roozbeh Mottaghi, Ali Farhadi
HIGHLIGHT: We learn a self-supervised visual representation from human's interactions with the visual world. For this study, we collect a dataset of human interactions capturing body part movements and gaze in their daily lives.

223, TITLE: EEC: Learning to Encode and Regenerate Images for Continual Learning
https://openreview.net/forum?id=lWaz5a9lcFU
AUTHORS: Ali Ayub, Alan Wagner
HIGHLIGHT: We train autoencoders with Neural Style Transfer to replay old tasks data for continual learning. The encoded features are converted into centroids and covariances to keep memory footprint from growing while keeping classifier performance stable.

224, TITLE: Impact of Representation Learning in Linear Bandits
https://openreview.net/forum?id=edJ_HipawCa
AUTHORS: Jiaqi Yang, Wei Hu, Jason D. Lee, Simon Shaolei Du
HIGHLIGHT: We show representation learning provably improves multi-task linear bandits.

225, TITLE: MODALS: Modality-agnostic Automated Data Augmentation in the Latent Space
https://openreview.net/forum?id=XjYgR6gbCEc
AUTHORS: Tsz-Him Cheung, Dit-Yan Yeung
HIGHLIGHT: In this work, we propose an automated data augmentation approach called MODALS (Modality-agnostic Automated Data Augmentation in the Latent Space) to augment data for any modality in a generic way.

226, TITLE: The Recurrent Neural Tangent Kernel
https://openreview.net/forum?id=3T9iFICe0Y9
AUTHORS: Sina Alemohammad, Zichao Wang, Randall Balestriero, Richard Baraniuk
HIGHLIGHT: In this paper we introduce and study the Recurrent Neural Tangent Kernel (RNTK), which provides new insights into the behavior of overparametrized RNNs, including how different time steps are weighted by the RNTK to form the output under different initialization parameters and nonlinearity choices, and how inputs of different lengths are treated.

227, TITLE: Projected Latent Markov Chain Monte Carlo: Conditional Sampling of Normalizing Flows
https://openreview.net/forum?id=MBpHUFrcG2x
AUTHORS: Chris Cannella, Mohammadreza Soltani, Vahid Tarokh
HIGHLIGHT: We introduce and demonstrate a novel MCMC technique for sampling from the exact conditional distributions known by normalizing flows.

228, TITLE: Learning the Pareto Front with Hypernetworks
https://openreview.net/forum?id=NjF772F4ZZR
AUTHORS: Aviv Navon, Aviv Shamsian, Gal Chechik, Ethan Fetaya
HIGHLIGHT: A novel approach for learning the entire Pareto front using hypernetworks

229, TITLE: Estimating and Evaluating Regression Predictive Uncertainty in Deep Object Detectors
https://openreview.net/forum?id=YLewtnvKgR7
AUTHORS: Ali Harakeh, Steven L. Waslander
HIGHLIGHT: In this work, we focus on estimating predictive distributions for bounding box regression output with variance networks.

230, TITLE: Predicting Classification Accuracy when Adding New Unobserved Classes
https://openreview.net/forum?id=Y9McSeEaqUh
AUTHORS: Yuli Slavutsky, Yuval Benjamini
HIGHLIGHT: A new prediction method of multiclass classification accuracy for an increased number of classes.

231, TITLE: BRECQ: Pushing the Limit of Post-Training Quantization by Block Reconstruction
https://openreview.net/forum?id=POWv6hDd9XH
AUTHORS: Yuhang Li, Ruihao Gong, Xu Tan, Yang Yang, Peng Hu, Qi Zhang, Fengwei Yu, Wei Wang, Shi Gu

HIGHLIGHT: In this work, we propose a novel PTQ framework, dubbed BRECQ, which pushes the limits of bitwidth in PTQ down to INT2 for the first time.

232, TITLE: No MCMC for me: Amortized sampling for fast and stable training of energy-based models
https://openreview.net/forum?id=ixpSxO9flk3
AUTHORS: Will Sussman Grathwohl, Jacob Jin Kelly, Milad Hashemi, Mohammad Norouzi, Kevin Swersky, David Duvenaud
HIGHLIGHT: We present a new generator-based approach for training EBMs and demonstrate that it trains models which obtain high likelihood and overcomes stability issues common in EBM training.

233, TITLE: GraphCodeBERT: Pre-training Code Representations with Data Flow
https://openreview.net/forum?id=jLoC4ez43PZ
AUTHORS: Daya Guo, Shuo Ren, Shuai Lu, Zhangyin Feng, Duyu Tang, Shujie LIU, Long Zhou, Nan Duan, Alexey Svyatkovskiy, Shengyu Fu, Michele Tufano, Shao Kun Deng, Colin Clement, Dawn Drain, Neel Sundaresan, Jian Yin, Daxin Jiang, Ming Zhou
HIGHLIGHT: We present GraphCodeBERT, a pre-trained model for programming language that considers the inherent structure of code.

234, TITLE: Conservative Safety Critics for Exploration
https://openreview.net/forum?id=iaO86DUuKi
AUTHORS: Homanga Bharadhwaj, Aviral Kumar, Nicholas Rhinehart, Sergey Levine, Florian Shkurti, Animesh Garg
HIGHLIGHT: Safe exploration in reinforcement learning can be achieved by constraining policy learning with conservative safety estimates of the environment.

235, TITLE: Improve Object Detection with Feature-based Knowledge Distillation: Towards Accurate and Efficient Detectors
https://openreview.net/forum?id=uKhGRvM8QNH
AUTHORS: Linfeng Zhang, Kaisheng Ma
HIGHLIGHT: We propose two knowledge distillation methods on object detection - attention-guided distillation and non-local distillation which lead to 4.1 AP improvements on Faster RCNN101 in MS COCO2017.

236, TITLE: A Temporal Kernel Approach for Deep Learning with Continuous-time Information
https://openreview.net/forum?id=whE31dn74cL
AUTHORS: Da Xu, Chuanwei Ruan, Evren Korpeoglu, Sushant Kumar, Kannan Achan
HIGHLIGHT: We propose a temporal kernel learning approach based on random features and reparameterization to characterize the continuous-time information in deep learning models.

237, TITLE: For self-supervised learning, Rationality implies generalization, provably
https://openreview.net/forum?id=Srmggo3b3X6
AUTHORS: Yamini Bansal, Gal Kaplun, Boaz Barak
HIGHLIGHT: We prove and empirically demonstrate generalization bounds for algorithms that fit a simple classifier on a representation learned via self-supervision; we obtain non-vacuous bounds for such top-performing algorithms on both CIFAR-10 and ImageNet.

238, TITLE: How to Find Your Friendly Neighborhood: Graph Attention Design with Self-Supervision
https://openreview.net/forum?id=Wi5KUNlqWty
AUTHORS: Dongkwan Kim, Alice Oh
HIGHLIGHT: We propose a method that self-supervise graph attention through edges and it should be designed according to the average degree and homophily of graphs.

239, TITLE: Interpretable Models for Granger Causality Using Self-explaining Neural Networks
https://openreview.net/forum?id=DEa4JdMWRHp
AUTHORS: Ricards Marcinkevics, Julia E Vogt
HIGHLIGHT: We propose an interpretable framework for inferring Granger causality based on self-explaining neural networks.

240, TITLE: Meta-learning Symmetries by Reparameterization
https://openreview.net/forum?id=-QxT4mJdijq
AUTHORS: Allan Zhou, Tom Knowles, Chelsea Finn
HIGHLIGHT: A method for automatically meta-learning and encoding equivariances into neural networks.

241, TITLE: Removing Undesirable Feature Contributions Using Out-of-Distribution Data
https://openreview.net/forum?id=eIHYL6fpbkA
AUTHORS: Saehyung Lee, Changhwa Park, Hyungyu Lee, Jihun Yi, Jonghyun Lee, Sungroh Yoon
HIGHLIGHT: We propose a simple method, Out-of-distribution data Augmented Training (OAT), to leverage OOD data for adversarial and standard learning.

242, TITLE: Mind the Gap when Conditioning Amortised Inference in Sequential Latent-Variable Models
https://openreview.net/forum?id=a2gqxKDvYys
AUTHORS: Justin Bayer, Maximilian Soelch, Atanas Mirchev, Baris Kayalibay, Patrick van der Smagt
HIGHLIGHT: We show how a common model assumption in amortised variational inference with sequential LVMS leads to a suboptimality and how to prevent it.

243, TITLE: On the Universality of the Double Descent Peak in Ridgeless Regression
https://openreview.net/forum?id=0IO5VdnSAaH
AUTHORS: David Holzm?ller
HIGHLIGHT: We prove a distribution-independent lower bound for the generalization error of ridgeless (random) features regression under weak assumptions, showing universal sensitivity to label noise around the interpolation threshold.

244, TITLE: Fair Mixup: Fairness via Interpolation
https://openreview.net/forum?id=DNl5s5BXeBn
AUTHORS: Ching-Yao Chuang, Youssef Mroueh
HIGHLIGHT: To improve the generalizability of fair classifiers, we propose fair mixup, a new data augmentation strategy for imposing the fairness constraint.

245, TITLE: Self-supervised Learning from a Multi-view Perspective
https://openreview.net/forum?id=-bdp_8Itjwp
AUTHORS: Yao-Hung Hubert Tsai, Yue Wu, Ruslan Salakhutdinov, Louis-Philippe Morency
HIGHLIGHT: From a multi-view learning perspective, this paper provides both theoretical and empirical analysis on self-supervised learning.

246, TITLE: Integrating Categorical Semantics into Unsupervised Domain Translation
https://openreview.net/forum?id=IMPA6MndSXU
AUTHORS: Samuel Lavoie-Marchildon, Faruk Ahmed, Aaron Courville
HIGHLIGHT: We present a method for learning domain invariant categorial semantics which enable UDT on two setups.

247, TITLE: The Unreasonable Effectiveness of Patches in Deep Convolutional Kernels Methods
https://openreview.net/forum?id=aYuZO9DIdnn
AUTHORS: Louis THIRY, Michael Arbel, Eugene Belilovsky, Edouard Oyallon
HIGHLIGHT: Patch-based representation is a key ingredient for competitive convolutional kernel methods.

248, TITLE: Open Question Answering over Tables and Text
https://openreview.net/forum?id=MmCRswl1UYl
AUTHORS: Wenhu Chen, Ming-Wei Chang, Eva Schlinger, William Yang Wang, William W. Cohen
HIGHLIGHT: We propose the new task of answering open-domain questions answering over web tables and text and design new techniques: 1) fused retrieval 2) cross-block reader to resolve the challenges posed in the new task.

249, TITLE: Evaluation of Similarity-based Explanations
https://openreview.net/forum?id=9uvhpyQwzM_
AUTHORS: Kazuaki Hanawa, Sho Yokoi, Satoshi Hara, Kentaro Inui
HIGHLIGHT: We investigated empirically which of the relevance metrics (e.g. similarity of hidden layer, influence function, etc.) are appropriate for similarity-based explanation.

250, TITLE: A Diffusion Theory For Deep Learning Dynamics: Stochastic Gradient Descent Exponentially Favors Flat Minima
https://openreview.net/forum?id=wXgk_iCiYGo
AUTHORS: Zeke Xie, Issei Sato, Masashi Sugiyama
HIGHLIGHT: We prove that, benefited from the Hessian-dependent covariance of stochastic gradient noise, SGD favors flat minima exponentially more than sharp minima.

251, TITLE: How Much Over-parameterization Is Sufficient to Learn Deep ReLU Networks?
https://openreview.net/forum?id=fgd7we_uZa6
AUTHORS: Zixiang Chen, Yuan Cao, Difan Zou, Quanquan Gu
HIGHLIGHT: We establish learning guarantees for deep ReLU networks with width polylogarithmic in sample size and the inverse of the target error.

252, TITLE: Auction Learning as a Two-Player Game
https://openreview.net/forum?id=YHdeAO61l6T
AUTHORS: Jad Rahme, Samy Jelassi, S. Matthew Weinberg
HIGHLIGHT: We formulate the auction learning problem as a two player game with a stationary utility functions and explore the advantages of such an approach.

253, TITLE: Robust Reinforcement Learning on State Observations with Learned Optimal Adversary
https://openreview.net/forum?id=sCZbhBvqQaU
AUTHORS: Huan Zhang, Hongge Chen, Duane S Boning, Cho-Jui Hsieh
HIGHLIGHT: We study the robustness of RL agents under perturbations on states and find that using an "optimal" adversary learned online in an alternating training manner can improve the robustness of agent policy.

254, TITLE: Optimizing Memory Placement using Evolutionary Graph Reinforcement Learning
https://openreview.net/forum?id=-6vS_4Kfz0
AUTHORS: Shauharda Khadka, Estelle Aflalo, Mattias Mardar, Avrech Ben-David, Santiago Miret, Shie Mannor, Tamir Hazan, Hanlin Tang, Somdeb Majumdar
HIGHLIGHT: We combine evolutionary and gradient-based reinforcement learning to tackle the large search spaces needed to map tensors to memory, yielding up to 78% speedup on BERT and ResNet on a deep learning inference chip.

255, TITLE: Hierarchical Autoregressive Modeling for Neural Video Compression
https://openreview.net/forum?id=TK_6nNb_C7q
AUTHORS: Ruihan Yang, Yibo Yang, Joseph Marino, Stephan Mandt
HIGHLIGHT: Specifically, we view recent neural video compression methods (Lu et al., 2019; Yang et al., 2020b; Agustssonet al., 2020) as instances of a generalized stochastic temporal autoregressive trans-form, and propose avenues for enhancement based on this insight.

256, TITLE: Individually Fair Rankings
https://openreview.net/forum?id=71zCSP_HuBN
AUTHORS: Amanda Bower, Hamid Eftekhari, Mikhail Yurochkin, Yuekai Sun
HIGHLIGHT: We present an algorithm for training individually fair learning-to-rank systems using optimal transport tools.

257, TITLE: Learning Neural Generative Dynamics for Molecular Conformation Generation
https://openreview.net/forum?id=pAbm1qfheGk
AUTHORS: Minkai Xu, Shitong Luo, Yoshua Bengio, Jian Peng, Jian Tang
HIGHLIGHT: A novel probabilistic framework to generate valid and diverse molecular conformations. Reaching state-of-the-art results on conformation generation and inter-atomic distance modeling.

258, TITLE: Efficient Certified Defenses Against Patch Attacks on Image Classifiers
https://openreview.net/forum?id=hr-3PMvDpil
AUTHORS: Jan Hendrik Metzen, Maksym Yatsura
HIGHLIGHT: We propose a method for certifying robustness against adversarial patches that combines high certified accuracy with efficient inference while maintaining strong performance on clean data.

259, TITLE: Convex Regularization behind Neural Reconstruction
https://openreview.net/forum?id=VErQxgyrbfn
AUTHORS: Arda Sahiner, Morteza Mardani, Batu Ozturkler, Mert Pilanci, John M. Pauly
HIGHLIGHT: This work proposes a finite-dimensional convex dual of a two-layer fully-convolutional ReLU network for denoising problems, and uses it for interpretation of neural network training and predictions.

260, TITLE: Targeted Attack against Deep Neural Networks via Flipping Limited Weight Bits
https://openreview.net/forum?id=iKQAk8a2kM0
AUTHORS: Jiawang Bai, Baoyuan Wu, Yong Zhang, Yiming Li, Zhifeng Li, Shu-Tao Xia
HIGHLIGHT: We propose a targeted attack method against the deployed DNN via flipping a few binary weight bits.

261, TITLE: Generalized Multimodal ELBO
https://openreview.net/forum?id=5Y21V0RDBV
AUTHORS: Thomas Marco Sutter, Imant Daunhawer, Julia E Vogt
HIGHLIGHT: We propose a generalized ELBO for modeling multiple data types in a scalable and self-supervised way.

262, TITLE: Large-width functional asymptotics for deep Gaussian neural networks
https://openreview.net/forum?id=0aW6lYOYB7d
AUTHORS: Daniele Bracale, Stefano Favaro, Sandra Fortini, Stefano Peluchetti
HIGHLIGHT: We establish the convergence of infinitely wide feed-forward deep neural networks in function space.

263, TITLE: Distributed Momentum for Byzantine-resilient Stochastic Gradient Descent
https://openreview.net/forum?id=H8UHdhWG6A3
AUTHORS: El Mahdi El Mhamdi, Rachid Guerraoui, S?bastien Rouault
HIGHLIGHT: An inexpensive method to substantially improve the effectiveness of existing Byzantine-resilient SGD defenses, assessed against state-of-the-art attacks and supported by theoretical insights.

264, TITLE: Fully Unsupervised Diversity Denoising with Convolutional Variational Autoencoders
https://openreview.net/forum?id=agHLCOBM5jP
AUTHORS: Mangal Prakash, Alexander Krull, Florian Jug
HIGHLIGHT: DivNoising performs fully unsupervised diversity denoising using fully convolutional variational autoencoders and achieves SOTA results for a number of well known datasets while also enabling VAE-like sampling

265, TITLE: Auxiliary Learning by Implicit Differentiation
https://openreview.net/forum?id=n7wIfYPdVet
AUTHORS: Aviv Navon, Idan Achituve, Haggai Maron, Gal Chechik, Ethan Fetaya
HIGHLIGHT: Learn to combine auxiliary tasks in a nonlinear fashion and to design them automatically.

266, TITLE: Balancing Constraints and Rewards with Meta-Gradient D4PG
https://openreview.net/forum?id=TQt98Ya7UMP
AUTHORS: Dan A. Calian, Daniel J Mankowitz, Tom Zahavy, Zhongwen Xu, Junhyuk Oh, Nir Levine, Timothy Mann
HIGHLIGHT: This paper uses meta-gradients to perform soft-constrained Reinforcement Learning (RL) optimization

267, TITLE: Adversarially Guided Actor-Critic
https://openreview.net/forum?id=_mQp5cr_iNy
AUTHORS: Yannis Flet-Berliac, Johan Ferret, Olivier Pietquin, Philippe Preux, Matthieu Geist
HIGHLIGHT: We introduce a third protagonist, the adversary. While this adversary mimics the actor by minimizing the KL-divergence between their respective action distributions, the actor maximizes the log-probability difference between its action and that of the adversary in combination with maximizing expected rewards.

268, TITLE: DARTS-: Robustly Stepping out of Performance Collapse Without Indicators
https://openreview.net/forum?id=KLH36ELmwIB
AUTHORS: Xiangxiang Chu, Xiaoxing Wang, Bo Zhang, Shun Lu, Xiaolin Wei, Junchi Yan
HIGHLIGHT: Indicator-free approach to stabilize DARTS

269, TITLE: Are wider nets better given the same number of parameters?
https://openreview.net/forum?id=_zx8Oka09eF
AUTHORS: Anna Golubeva, Guy Gur-Ari, Behnam Neyshabur
HIGHLIGHT: We show that increasing network width leads to better performance even when the number of weights remains fixed.

270, TITLE: Optimal Conversion of Conventional Artificial Neural Networks to Spiking Neural Networks
https://openreview.net/forum?id=FZ1oTwcXchK
AUTHORS: Shikuang Deng, Shi Gu
HIGHLIGHT: We propose and validate an optimal pipeline that efficiently converts conventional artificial neural networks to spiking neural networks with almost no accuracy loss in a fairly short simulation length.

271, TITLE: Deep Equals Shallow for ReLU Networks in Kernel Regimes
https://openreview.net/forum?id=aDjoksTpXOP
AUTHORS: Alberto Bietti, Francis Bach

HIGHLIGHT: We show that for ReLU activations, the kernels derived from deep fully-connected networks have essentially the same approximation properties as their shallow two-layer counterpart, namely the same eigenvalue decay for the corresponding integral operator.

272, TITLE: Graph Coarsening with Neural Networks
https://openreview.net/forum?id=uxpzitPEooJ
AUTHORS: Chen Cai, Dingkang Wang, Yusu Wang
HIGHLIGHT: We significantly improve the quality of existing graph coarsening algorithms with graph neural network.

273, TITLE: Early Stopping in Deep Networks: Double Descent and How to Eliminate it
https://openreview.net/forum?id=tlV90jvZbw
AUTHORS: Reinhard Heckel, Fatih Furkan Yilmaz
HIGHLIGHT: Epoch wise double descent can be explained as a superposition of two or more bias-variance tradeoffs that arise because different parts of the network are learned at different epochs.

274, TITLE: Efficient Inference of Flexible Interaction in Spiking-neuron Networks
https://openreview.net/forum?id=aGfU_xziEX8
AUTHORS: Feng Zhou, Yixuan Zhang, Jun Zhu
HIGHLIGHT: An efficient conjugate EM algorithm for nonlinear multivariate Hawkes processes based on auxiliary latent variables augmentation.

275, TITLE: DICE: Diversity in Deep Ensembles via Conditional Redundancy Adversarial Estimation
https://openreview.net/forum?id=R2ZlTVPx0Gk
AUTHORS: Alexandre Rame, Matthieu Cord
HIGHLIGHT: Driven by arguments from information theory, we introduce a new learning strategy for deep ensembles that increases diversity among members: we adversarially prevent features from being conditionally redundant, i.e., predictable from each other.

276, TITLE: Inductive Representation Learning in Temporal Networks via Causal Anonymous Walks
https://openreview.net/forum?id=KYPz4YsCPj
AUTHORS: Yanbang Wang, Yen-Yu Chang, Yunyu Liu, Jure Leskovec, Pan Li
HIGHLIGHT: The paper proposes Causal Anonymous Walks (CAW) as an effective way to encode the dynamic laws that govern the evolution of temporal networks, which significantly improves inductive representation learning on those networks.

277, TITLE: FairBatch: Batch Selection for Model Fairness
https://openreview.net/forum?id=YNnpaAKeCfx
AUTHORS: Yuji Roh, Kangwook Lee, Steven Euijong Whang, Changho Suh
HIGHLIGHT: We address model fairness via the lens of bilevel optimization and propose a batch selection algorithm called FairBatch, which is easy to adopt, has state-of-the-art performance, and is compatible with existing batch selection techniques.

278, TITLE: Representation Balancing Offline Model-based Reinforcement Learning
https://openreview.net/forum?id=QpNz8r_Ri2Y
AUTHORS: Byung-Jun Lee, Jongmin Lee, Kee-Eung Kim
HIGHLIGHT: We present RepB-SDE, a framework for balancing the model representation with stationary distribution estimation, aiming at obtaining a model robust to the distribution shift that arises in off-policy and offline RL.

279, TITLE: Accelerating Convergence of Replica Exchange Stochastic Gradient MCMC via Variance Reduction
https://openreview.net/forum?id=iOnhIy-a-0n
AUTHORS: Wei Deng, Qi Feng, Georgios P. Karagiannis, Guang Lin, Faming Liang
HIGHLIGHT: We propose a variance-reduced replica-exchange stochastic gradient Langevin dynamics to reduce the variance of the energy estimators to accelerate the convergence.

280, TITLE: The Importance of Pessimism in Fixed-Dataset Policy Optimization
https://openreview.net/forum?id=E3Ys6a1NTGT
AUTHORS: Jacob Buckman, Carles Gelada, Marc G Bellemare
HIGHLIGHT: A unified conceptual and mathematical framework for fixed-dataset policy optimization algorithms, revealing the importance of uncertainty and pessimism.

281, TITLE: Interpreting Knowledge Graph Relation Representation from Word Embeddings
https://openreview.net/forum?id=gLWj29369lW

AUTHORS: Carl Allen, Ivana Balazevic, Timothy Hospedales
HIGHLIGHT: Interpreting the structure of knowledge graph relation representation using insight from word embeddings.

282, TITLE: Hopfield Networks is All You Need
https://openreview.net/forum?id=tL89RnzIiCd
AUTHORS: Hubert Ramsauer, Bernhard Sch?fl, Johannes Lehner, Philipp Seidl, Michael Widrich, Lukas Gruber, Markus Holzleitner, Thomas Adler, David Kreil, Michael K Kopp, G?nter Klambauer, Johannes Brandstetter, Sepp Hochreiter
HIGHLIGHT: A novel continuous Hopfield network is proposed whose update rule is the attention mechanism of the transformer model and which can be integrated into deep learning architectures.

283, TITLE: Uncertainty Estimation and Calibration with Finite-State Probabilistic RNNs
https://openreview.net/forum?id=9EKHN1jOlA
AUTHORS: Cheng Wang, Carolin Lawrence, Mathias Niepert
HIGHLIGHT: A method to estimate and calibrate uncertainty in recurrent state transitions.

284, TITLE: Understanding the failure modes of out-of-distribution generalization
https://openreview.net/forum?id=fSTD6NFIW_b
AUTHORS: Vaishnavh Nagarajan, Anders Andreassen, Behnam Neyshabur
HIGHLIGHT: In this theoretical study, we explain why machine learning models rely on spuriously correlated features in the dataset and fail at out-of-distribution generalization.

285, TITLE: Generative Language-Grounded Policy in Vision-and-Language Navigation with Bayes' Rule
https://openreview.net/forum?id=45uOPa46Kh
AUTHORS: Shuhei Kurita, Kyunghyun Cho
HIGHLIGHT: We propose the novel generative language-grounded policy for vision-and-language navigation(VLN).

286, TITLE: Emergent Road Rules In Multi-Agent Driving Environments
https://openreview.net/forum?id=d8Q1mt2Ghw
AUTHORS: Avik Pal, Jonah Philion, Yuan-Hong Liao, Sanja Fidler
HIGHLIGHT: In multi-agent driving environments with noisy perception, driving conventions emerge

287, TITLE: Wasserstein-2 Generative Networks
https://openreview.net/forum?id=bEoxzW_EXsa
AUTHORS: Alexander Korotin, Vage Egiazarian, Arip Asadulaev, Alexander Safin, Evgeny Burnaev
HIGHLIGHT: We present a new end-to-end algorithm to compute optimal transport maps between continuous distributions without introducing bias or resorting to minimax optimization.

288, TITLE: Vulnerability-Aware Poisoning Mechanism for Online RL with Unknown Dynamics
https://openreview.net/forum?id=9r30XCjf5Dt
AUTHORS: Yanchao Sun, Da Huo, Furong Huang
HIGHLIGHT: We propose the first poisoning algorithm against deep policy-based RL methods, without any prior knowledge of the environment, covering heterogeneous poisoning models.

289, TITLE: Tomographic Auto-Encoder: Unsupervised Bayesian Recovery of Corrupted Data
https://openreview.net/forum?id=YtMG5ex0ou
AUTHORS: Francesco Tonolini, Andreas Damianou, Pablo Garcia Moreno, Roderick Murray-Smith
HIGHLIGHT: Recovering accurate posterior distributions for unsupervised data recovery

290, TITLE: Monotonic Kronecker-Factored Lattice
https://openreview.net/forum?id=0pxiMpCyBtr
AUTHORS: William Taylor Bakst, Nobuyuki Morioka, Erez Louidor
HIGHLIGHT: We show how to effectively and efficiently learn flexible and interpretable monotonic functions using Kronecker-Factored Lattice, an efficient reparameterization of flexible monotonic lattice regression via Kronecker product.

291, TITLE: LEAF: A Learnable Frontend for Audio Classification
https://openreview.net/forum?id=jM76BCb6F9m
AUTHORS: Neil Zeghidour, Olivier Teboul, F?lix de Chaumont Quitry, Marco Tagliasacchi
HIGHLIGHT: We propose a lightweight learnable frontend of audio that can replace fixed features over a wide range of audio classification tasks.

292, TITLE: Federated Learning via Posterior Averaging: A New Perspective and Practical Algorithms
https://openreview.net/forum?id=GFsU8a0sGB
AUTHORS: Maruan Al-Shedivat, Jennifer Gillenwater, Eric Xing, Afshin Rostamizadeh
HIGHLIGHT: A new approach to federated learning that generalizes federated optimization, combines local MCMC-based sampling with global optimization-based posterior inference, and achieves competitive results on challenging benchmarks.

293, TITLE: Rank the Episodes: A Simple Approach for Exploration in Procedurally-Generated Environments
https://openreview.net/forum?id=MtEE0CktZht
AUTHORS: Daochen Zha, Wenye Ma, Lei Yuan, Xia Hu, Ji Liu
HIGHLIGHT: Encouraging exploration via ranking the past episodes and reproducing past good exploration behaviors with imitation learning.

294, TITLE: Partitioned Learned Bloom Filters
https://openreview.net/forum?id=6BRLOfrMhW
AUTHORS: Kapil Vaidya, Eric Knorr, Michael Mitzenmacher, Tim Kraska
HIGHLIGHT: Here we show how to frame the problem of optimal model utilization as an optimization problem, and using our framework derive algorithms that can achieve near-optimal performance in many cases.

295, TITLE: Approximate Nearest Neighbor Negative Contrastive Learning for Dense Text Retrieval
https://openreview.net/forum?id=zeFrfgyZln
AUTHORS: Lee Xiong, Chenyan Xiong, Ye Li, Kwok-Fung Tang, Jialin Liu, Paul N. Bennett, Junaid Ahmed, Arnold Overwikj
HIGHLIGHT: This paper improves the learning of dense text retrieval using ANCE, which selects global negatives with bigger gradient norms using an asynchronously updated ANN index.

296, TITLE: AUXILIARY TASK UPDATE DECOMPOSITION: THE GOOD, THE BAD AND THE NEUTRAL
https://openreview.net/forum?id=1GTma8HwlYp
AUTHORS: Lucio M. Dery, Yann Dauphin, David Grangier
HIGHLIGHT: We improve how we use auxiliary task data for model pre-training by decomposing gradient updates into components guided by the primary task

297, TITLE: SSD: A Unified Framework for Self-Supervised Outlier Detection
https://openreview.net/forum?id=v5gjXpmR8J
AUTHORS: Vikash Sehwag, Mung Chiang, Prateek Mittal
HIGHLIGHT: We propose SSD, an outlier detector based on only unlabeled training data.

298, TITLE: Ask Your Humans: Using Human Instructions to Improve Generalization in Reinforcement Learning
https://openreview.net/forum?id=Y87Ri-GNHYu
AUTHORS: Valerie Chen, Abhinav Gupta, Kenneth Marino
HIGHLIGHT: To facilitate the automatic decomposition of hierarchical tasks, we propose the use of step-by-step human demonstrations in the form of natural language instructions and action trajectories.
We introduce a dataset of such demonstrations in a crafting-based grid world.

299, TITLE: Revisiting Few-sample BERT Fine-tuning
https://openreview.net/forum?id=cO1IH43yUF
AUTHORS: Tianyi Zhang, Felix Wu, Arzoo Katiyar, Kilian Q Weinberger, Yoav Artzi
HIGHLIGHT: This paper is a study of fine-tuning of BERT contextual representations, with focus on commonly observed instabilities in few-sample scenarios.

300, TITLE: Tilted Empirical Risk Minimization
https://openreview.net/forum?id=K5YasWXZT3O
AUTHORS: Tian Li, Ahmad Beirami, Maziar Sanjabi, Virginia Smith
HIGHLIGHT: We show that tilted empirical risk minimization (TERM) can be used for enforcing fairness between subgroups, mitigating the effect of outliers, and handling class imbalance, all in a unified framework.

301, TITLE: Deep Neural Tangent Kernel and Laplace Kernel Have the Same RKHS
https://openreview.net/forum?id=vK9WrZ0QYQ
AUTHORS: Lin Chen, Sheng Xu
HIGHLIGHT: We prove that the reproducing kernel Hilbert spaces of a deep neural tangent kernel and the Laplace kernel include the same set of functions.

302, TITLE: On the Transfer of Disentangled Representations in Realistic Settings
https://openreview.net/forum?id=8VXvj1QNRl1
AUTHORS: Andrea Dittadi, Frederik Tr?uble, Francesco Locatello, Manuel Wuthrich, Vaibhav Agrawal, Ole Winther, Stefan Bauer, Bernhard Sch?lkopf
HIGHLIGHT: We scale disentangled representation learning to a new realistic dataset and conduct a large-scale empirical study on OOD generalization.
We introduce a new high-resolution dataset with over 1M simulated images and 1k annotated real-world images of the same setup.

303, TITLE: Calibration tests beyond classification
https://openreview.net/forum?id=-bxf89v3Nx
AUTHORS: David Widmann, Fredrik Lindsten, Dave Zachariah
HIGHLIGHT: Unifying framework for calibration evaluations and tests of probabilistic predictive models

304, TITLE: Overparameterisation and worst-case generalisation: friend or foe?
https://openreview.net/forum?id=jphnJNOwe36
AUTHORS: Aditya Krishna Menon, Ankit Singh Rawat, Sanjiv Kumar
HIGHLIGHT: Overparameterised models' worst-subgroup performance can be improved via post-hoc processing.

305, TITLE: You Only Need Adversarial Supervision for Semantic Image Synthesis
https://openreview.net/forum?id=yvQKLaqNE6M
AUTHORS: Edgar Sch?nfeld, Vadim Sushko, Dan Zhang, Juergen Gall, Bernt Schiele, Anna Khoreva
HIGHLIGHT: We propose OASIS, a novel model for multi-modal semantic image synthesis, improving over previous methods in terms of image quality and diversity while only using adversarial supervision.

306, TITLE: Learning to Recombine and Resample Data For Compositional Generalization
https://openreview.net/forum?id=PS3IMnScugk
AUTHORS: Ekin Aky?rek, Afra Feyza Aky?rek, Jacob Andreas
HIGHLIGHT: This paper investigates a data augmentation procedure based on two weaker principles: recombination and resampling, and finds that it is sufficient to induce many of the compositional generalizations studied in previous work.

307, TITLE: A Critique of Self-Expressive Deep Subspace Clustering
https://openreview.net/forum?id=FOyuZ26emy
AUTHORS: Benjamin David Haeffele, Chong You, Rene Vidal
HIGHLIGHT: Here we show theoretically and experimentally that there are a number of flaws with many existing self-expressive deep subspace clustering models.

308, TITLE: INT: An Inequality Benchmark for Evaluating Generalization in Theorem Proving
https://openreview.net/forum?id=O6LPudowNQm
AUTHORS: Yuhuai Wu, Albert Jiang, Jimmy Ba, Roger Baker Grosse
HIGHLIGHT: We introduce INT, a synthetic INequality Theorem proving benchmark, to tackle the data sparsity and out-of-distribution problems for theorem proving and benchmarked transformer-based and GNN-based agents' generalization performance.

309, TITLE: Improved Estimation of Concentration Under lp-Norm Distance Metrics Using Half Spaces
https://openreview.net/forum?id=BUlyHkzjgmA
AUTHORS: Jack Prescott, Xiao Zhang, David Evans
HIGHLIGHT: We show that concentration of measure does not prohibit the existence of adversarially robust classifiers using a novel method of empirical concentration estimation.

310, TITLE: Adaptive Federated Optimization
https://openreview.net/forum?id=LkFG3lB13U5
AUTHORS: Sashank J. Reddi, Zachary Charles, Manzil Zaheer, Zachary Garrett, Keith Rush, Jakub Konecn?, Sanjiv Kumar, Hugh Brendan McMahan
HIGHLIGHT: We propose adaptive federated optimization techniques, and highlight their improved performance over popular methods such as FedAvg.

311, TITLE: On the Dynamics of Training Attention Models
https://openreview.net/forum?id=1OCTOShAmqB
AUTHORS: Haoye Lu, Yongyi Mao, Amiya Nayak

HIGHLIGHT: In this paper, we set up a simple text classification task and study the dynamics of training a simple attention-based classification model using gradient descent.

312, TITLE: Linear Convergent Decentralized Optimization with Compression
https://openreview.net/forum?id=84gjULz1t5
AUTHORS: Xiaorui Liu, Yao Li, Rongrong Wang, Jiliang Tang, Ming Yan
HIGHLIGHT: A Linear Convergent Decentralized Optimization with Communication Compression

313, TITLE: Efficient Transformers in Reinforcement Learning using Actor-Learner Distillation
https://openreview.net/forum?id=uR9LaO_QxF
AUTHORS: Emilio Parisotto, Russ Salakhutdinov
HIGHLIGHT: Actor-Learner Distillation uses an continual form of distillation to attain the high sample-efficiency of transformers while maintaining the reduced total training time of LSTMs in RL applications.

314, TITLE: Large Associative Memory Problem in Neurobiology and Machine Learning
https://openreview.net/forum?id=X4y_10OX-hX
AUTHORS: Dmitry Krotov, John J. Hopfield
HIGHLIGHT: Our paper proposes a microscopic biologically-plausible theory of modern Hopfield networks.

315, TITLE: Protecting DNNs from Theft using an Ensemble of Diverse Models
https://openreview.net/forum?id=LucJxySuJcE
AUTHORS: Sanjay Kariyappa, Atul Prakash, Moinuddin K Qureshi
HIGHLIGHT: Discontinuous predictions produced by an ensemble of diverse models can be used to create an effective defense against model stealing attacks.

316, TITLE: Proximal Gradient Descent-Ascent: Variable Convergence under KL Geometry
https://openreview.net/forum?id=LVotkZmYyDi
AUTHORS: Ziyi Chen, Yi Zhou, Tengyu Xu, Yingbin Liang
HIGHLIGHT: Specifically, we show that proximal-GDA admits a novel Lyapunov function, which monotonically decreases in the minimax optimization process and drives the variable sequences to a critical point.

317, TITLE: Contextual Dropout: An Efficient Sample-Dependent Dropout Module
https://openreview.net/forum?id=ct8_a9h1M
AUTHORS: XINJIE FAN, Shujian Zhang, Korawat Tanwisuth, Xiaoning Qian, Mingyuan Zhou
HIGHLIGHT: We propose contextual dropout as a scalable sample-dependent dropout method, which makes the dropout probabilities depend on the input covariates of each data sample.

318, TITLE: MIROSTAT: A NEURAL TEXT DECODING ALGORITHM THAT DIRECTLY CONTROLS PERPLEXITY
https://openreview.net/forum?id=W1G1JZEIy5_
AUTHORS: Sourya Basu, Govardana Sachitanandam Ramachandran, Nitish Shirish Keskar, Lav R. Varshney
HIGHLIGHT: We provide a new text decoding algorithm that directly controls perplexity and hence several important attributes of generated text.

319, TITLE: DialoGraph: Incorporating Interpretable Strategy-Graph Networks into Negotiation Dialogues
https://openreview.net/forum?id=kDnal_bbb-E
AUTHORS: Rishabh Joshi, Vidhisha Balachandran, Shikhar Vashishth, Alan Black, Yulia Tsvetkov
HIGHLIGHT: We propose DialoGraph, a negotiation dialogue system that leverages Graph Attention Networks to model complex negotiation strategies while providing interpretability for the model via intermediate structures.

320, TITLE: Multi-Time Attention Networks for Irregularly Sampled Time Series
https://openreview.net/forum?id=4c0J6lwQ4_
AUTHORS: Satya Narayan Shukla, Benjamin Marlin
HIGHLIGHT: This paper presents a new deep learning architecture for learning with sparse and irregularly sampled multivariate time series.

321, TITLE: Learning Energy-Based Generative Models via Coarse-to-Fine Expanding and Sampling
https://openreview.net/forum?id=aD1_5zowqV
AUTHORS: Yang Zhao, Jianwen Xie, Ping Li
HIGHLIGHT: We propose a coarse-to-fine energy-based model learning scheme for generative modeling.

322, TITLE: Unsupervised Audiovisual Synthesis via Exemplar Autoencoders
https://openreview.net/forum?id=43VKWxg_Sqr
AUTHORS: Kangle Deng, Aayush Bansal, Deva Ramanan
HIGHLIGHT: We present an unsupervised approach that converts the input speech of any individual into audiovisual streams of potentially-infinitely many output speakers.

323, TITLE: A Learning Theoretic Perspective on Local Explainability
https://openreview.net/forum?id=7aL-OtQrBWD
AUTHORS: Jeffrey Li, Vaishnavh Nagarajan, Gregory Plumb, Ameet Talwalkar
HIGHLIGHT: In this paper, we explore connections between interpretable machine learning and learning theory through the lens of local approximation explanations.

324, TITLE: SEED: Self-supervised Distillation For Visual Representation
https://openreview.net/forum?id=AHm3dbp7D1D
AUTHORS: Zhiyuan Fang, Jianfeng Wang, Lijuan Wang, Lei Zhang, Yezhou Yang, Zicheng Liu
HIGHLIGHT: We propose SEED, a self-supervised distillation technique for visual representation learning.

325, TITLE: Isometric Propagation Network for Generalized Zero-shot Learning
https://openreview.net/forum?id=-mWcQVLPSPy
AUTHORS: Lu Liu, Tianyi Zhou, Guodong Long, Jing Jiang, Xuanyi Dong, Chengqi Zhang
HIGHLIGHT: We improve the current zero-shot learning performance by a dynamic alignment between the semantic space and visual space that encourages the isometry of the class-prototype propagation procedures in the two spaces.
To evaluate the generalization capability of IPN, we further build two larger benchmarks with more diverse unseen classes, and demonstrate the advantages of IPN on them.

326, TITLE: Effective and Efficient Vote Attack on Capsule Networks
https://openreview.net/forum?id=33rtZ4Sjwjn
AUTHORS: Jindong Gu, Baoyuan Wu, Volker Tresp
HIGHLIGHT: We propose an effective and efficient vote attack to create adversarial examples and bypass adversarial example detection on Capsule Networks.

327, TITLE: Heteroskedastic and Imbalanced Deep Learning with Adaptive Regularization
https://openreview.net/forum?id=mEdwVCRJuX4
AUTHORS: Kaidi Cao, Yining Chen, Junwei Lu, Nikos Arechiga, Adrien Gaidon, Tengyu Ma
HIGHLIGHT: We propose a data-dependent regularization technique for learning heteroskedastic and imbalanced datasets.

328, TITLE: Continuous Wasserstein-2 Barycenter Estimation without Minimax Optimization
https://openreview.net/forum?id=3tFAs5E-Pe
AUTHORS: Alexander Korotin, Lingxiao Li, Justin Solomon, Evgeny Burnaev
HIGHLIGHT: We present a new algorithm to compute Wasserstein-2 barycenters of continuous distributions powered by a straightforward optimization procedure without introducing bias or a generative model.

329, TITLE: Neural Thompson Sampling
https://openreview.net/forum?id=tkAtoZkcUnm
AUTHORS: Weitong ZHANG, Dongruo Zhou, Lihong Li, Quanquan Gu
HIGHLIGHT: In this paper, we propose a new algorithm, called Neural Thompson Sampling, which adapts deep neural networks for both exploration and exploitation.

330, TITLE: Symmetry, Conservation Laws, and Learning Dynamics in Neural Networks
https://openreview.net/forum?id=q8qLAbQBupm
AUTHORS: Daniel Kunin, Javier Sagastuy-Brena, Surya Ganguli, Daniel LK Yamins, Hidenori Tanaka
HIGHLIGHT: By exploiting architectural symmetry, our work demonstrates that we can analytically describe the learning dynamics of various parameter combinations at finite learning rates and batch sizes for state of the art architectures trained on any dataset.

331, TITLE: Neural gradients are near-lognormal: improved quantized and sparse training
https://openreview.net/forum?id=EoFNy62JGd
AUTHORS: Brian Chmiel, Liad Ben-Uri, Moran Shkolnik, Elad Hoffer, Ron Banner, Daniel Soudry

HIGHLIGHT: Considering this, we suggest two closed-form analytical methods to reduce the computational and memory burdens of neural gradients.

332, TITLE: RODE: Learning Roles to Decompose Multi-Agent Tasks
https://openreview.net/forum?id=TTUVg6vkNjK
AUTHORS: Tonghan Wang, Tarun Gupta, Anuj Mahajan, Bei Peng, Shimon Whiteson, Chongjie Zhang
HIGHLIGHT: We propose a scalable role-based multi-agent learning method which effectively discovers roles based on joint action space decomposition according to action effects, establishing a new state of the art on the StarCraft multi-agent benchmark.

333, TITLE: Revisiting Hierarchical Approach for Persistent Long-Term Video Prediction
https://openreview.net/forum?id=3RLN4EPMdYd
AUTHORS: Wonkwang Lee, Whie Jung, Han Zhang, Ting Chen, Jing Yu Koh, Thomas Huang, Hyungsuk Yoon, Honglak Lee, Seunghoon Hong
HIGHLIGHT: We propose a simple yet effective hierarchical video prediction model that can synthesize future frames orders of magnitude longer than existing methods (thousands frames)

334, TITLE: Physics-aware, probabilistic model order reduction with guaranteed stability
https://openreview.net/forum?id=vyY0jnWG-tK
AUTHORS: Sebastian Kaltenbach, Phaedon Stelios Koutsourelakis
HIGHLIGHT: We propose a novel physics-aware, generative, probabilistic state-space model for learning an effective, lower-dimensional description that can produce long-term predictions.

335, TITLE: Modelling Hierarchical Structure between Dialogue Policy and Natural Language Generator with Option Framework for Task-oriented Dialogue System
https://openreview.net/forum?id=kLbhLJ8OT12
AUTHORS: Jianhong Wang, Yuan Zhang, Tae-Kyun Kim, Yunjie Gu
HIGHLIGHT: We propose a novel algorithm called HDNO for policy optimization for task-oriented dialogue system so that the performance on the comprehensibility of generated responses is improved compared with other RL-based algorithms.

336, TITLE: Learning explanations that are hard to vary
https://openreview.net/forum?id=hb1sDDSLbV
AUTHORS: Giambattista Parascandolo, Alexander Neitz, ANTONIO ORVIETO, Luigi Gresele, Bernhard Sch?lkopf
HIGHLIGHT: In this paper, we investigate the principle that good explanations are hard to vary in the context of deep learning.

337, TITLE: Efficient Generalized Spherical CNNs
https://openreview.net/forum?id=rWZz3sJfCkm
AUTHORS: Oliver Cobb, Christopher G. R. Wallis, Augustine N. Mavor-Parker, Augustin Marignier, Matthew A. Price, Mayeul d'Avezac, Jason McEwen
HIGHLIGHT: We present a generalized spherical CNN framework that encompasses various existing approaches and allows them to be leveraged alongside each other.

338, TITLE: Collective Robustness Certificates
https://openreview.net/forum?id=ULQdiUTHe3y
AUTHORS: Jan Schuchardt, Aleksandar Bojchevski, Johannes Klicpera, Stephan G?nnemann
HIGHLIGHT: We fuse multiple single-prediction certificates into a drastically stronger collective certificate leveraging the locality property of Graph Neural Networks.

339, TITLE: Entropic gradient descent algorithms and wide flat minima
https://openreview.net/forum?id=xjXg0bnoDmS
AUTHORS: Fabrizio Pittorino, Carlo Lucibello, Christoph Feinauer, Gabriele Perugini, Carlo Baldassi, Elizaveta Demyanenko, Riccardo Zecchina
HIGHLIGHT: Relation between local entropy, flat minima, entropic algorithms and good generalization.

340, TITLE: Achieving Linear Speedup with Partial Worker Participation in Non-IID Federated Learning
https://openreview.net/forum?id=jDdzh5ul-d
AUTHORS: Haibo Yang, Minghong Fang, Jia Liu
HIGHLIGHT: So far, it remains an open question whether or not the linear speedup for convergence is achievable under non-i.i.d. datasets with partial worker participation in FL. In this paper, we show that the answer is affirmative.

341, TITLE: Categorical Normalizing Flows via Continuous Transformations
https://openreview.net/forum?id=-GLNZeVDuik
AUTHORS: Phillip Lippe, Efstratios Gavves
HIGHLIGHT: In this paper, we investigate Categorical Normalizing Flows, that is normalizing flows for categorical data.

342, TITLE: Learning to Represent Action Values as a Hypergraph on the Action Vertices
https://openreview.net/forum?id=Xv_s64FiXTv
AUTHORS: Arash Tavakoli, Mehdi Fatemi, Petar Kormushev
HIGHLIGHT: This paper introduces a class of models for learning action representations in RL by leveraging the combinatorial structure of multidimensional discrete action spaces.

343, TITLE: Debiasing Concept-based Explanations with Causal Analysis
https://openreview.net/forum?id=6puUoArESGp
AUTHORS: Mohammad Taha Bahadori, David Heckerman
HIGHLIGHT: We use a technique from instrumental variables literature and remove the impact of noise and latent confounding from concept-based explanations.

344, TITLE: Lifelong Learning of Compositional Structures
https://openreview.net/forum?id=ADWd4TJO13G
AUTHORS: Jorge A Mendez, ERIC EATON
HIGHLIGHT: We create a general-purpose framework for lifelong learning of compositional structures that splits the learning process into two stages: assimilation of new tasks with existing components, and accommodation of new knowledge into the components.

345, TITLE: Rethinking Embedding Coupling in Pre-trained Language Models
https://openreview.net/forum?id=xpFFI_NtgpW
AUTHORS: Hyung Won Chung, Thibault Fevry, Henry Tsai, Melvin Johnson, Sebastian Ruder
HIGHLIGHT: Decoupling output embedding shapes leads to more transferable Transformer layers and prevents over-specialization of a Transformer's upper layers to the pre-training task.

346, TITLE: Creative Sketch Generation
https://openreview.net/forum?id=gwnoVHIES05
AUTHORS: Songwei Ge, Vedanuj Goswami, Larry Zitnick, Devi Parikh
HIGHLIGHT: We introduce two creative sketch datasets and DoodlerGAN -- a part-based GAN model that generates creative sketches.

347, TITLE: Concept Learners for Few-Shot Learning
https://openreview.net/forum?id=eJIJF3-LoZO
AUTHORS: Kaidi Cao, Maria Brbic, Jure Leskovec
HIGHLIGHT: COMET learns generalizable representations along human-understandable concept dimensions.

348, TITLE: Domain Generalization with MixStyle
https://openreview.net/forum?id=6xHJ37MVxxp
AUTHORS: Kaiyang Zhou, Yongxin Yang, Yu Qiao, Tao Xiang
HIGHLIGHT: MixStyle makes CNNs more domain-generalizable by mixing instance-level feature statistics of training samples across domains.

349, TITLE: DeLighT: Deep and Light-weight Transformer
https://openreview.net/forum?id=ujmgfuxSLrO
AUTHORS: Sachin Mehta, Marjan Ghazvininejad, Srinivasan Iyer, Luke Zettlemoyer, Hannaneh Hajishirzi
HIGHLIGHT: Deep and light-weight transformer that matches or improves the performance of baseline Transformers with 2 to 3 times fewer parameters on standard machine translation and language modeling tasks

350, TITLE: Single-Timescale Actor-Critic Provably Finds Globally Optimal Policy
https://openreview.net/forum?id=pqZV_srUVmK
AUTHORS: Zuyue Fu, Zhuoran Yang, Zhaoran Wang
HIGHLIGHT: We study the global convergence and global optimality of actor-critic, one of the most popular families of reinforcement learning algorithms.

351, TITLE: Mastering Atari with Discrete World Models

https://openreview.net/forum?id=0oabwyZbOu
AUTHORS: Danijar Hafner, Timothy P Lillicrap, Mohammad Norouzi, Jimmy Ba
HIGHLIGHT: We introduce DreamerV2, a reinforcement learning agent that learns behaviors purely from predictions in the compact latent space of a powerful world model.

352, TITLE: Learning Neural Event Functions for Ordinary Differential Equations
https://openreview.net/forum?id=kW_zpEmMLdP
AUTHORS: Ricky T. Q. Chen, Brandon Amos, Maximilian Nickel
HIGHLIGHT: We discuss how event handling in ODE solvers can be differentiated through, allowing us to extend Neural ODEs to cases of implicitly defined termination times and enabling learning of discrete events and discontinuous dynamics.

353, TITLE: Contemplating Real-World Object Classification
https://openreview.net/forum?id=Q4EUywJIkqr
AUTHORS: ali borji
HIGHLIGHT: We address whether current deep learning models are able to solve object recognition in real world and how robust they are to synthetic and natural distribution shifts.

354, TITLE: Neural Spatio-Temporal Point Processes
https://openreview.net/forum?id=XQQA6-So14
AUTHORS: Ricky T. Q. Chen, Brandon Amos, Maximilian Nickel
HIGHLIGHT: We motivate the use of Continuous-time Normalizing Flows for building spatio-temporal point processes, and discuss modeling conditional dependencies with recurrent- or attention-based Neural ODEs.

355, TITLE: Generative Time-series Modeling with Fourier Flows
https://openreview.net/forum?id=PpshD0AXfA
AUTHORS: Ahmed Alaa, Alex James Chan, Mihaela van der Schaar
HIGHLIGHT: In this paper, we propose an explicit likelihood model based on a novel class of normalizing flows that view time-series data in the frequency-domain rather than the time-domain.

356, TITLE: DOP: Off-Policy Multi-Agent Decomposed Policy Gradients
https://openreview.net/forum?id=6FqKiVAdI3Y
AUTHORS: Yihan Wang, Beining Han, Tonghan Wang, Heng Dong, Chongjie Zhang
HIGHLIGHT: We propose an off-policy multi-agent decomposed policy gradient method, addressing the drawbacks that prevent existing multi-agent policy gradient methods from achieving state-of-the-art performance.

357, TITLE: The Risks of Invariant Risk Minimization
https://openreview.net/forum?id=BbNIbVPJ-42
AUTHORS: Elan Rosenfeld, Pradeep Kumar Ravikumar, Andrej Risteski
HIGHLIGHT: We formally demonstrate that Invariant Risk Minimization and related alternative objectives often perform no better than standard ERM.

358, TITLE: DynaTune: Dynamic Tensor Program Optimization in Deep Neural Network Compilation
https://openreview.net/forum?id=GTGb3M_KcUl
AUTHORS: Minjia Zhang, Menghao Li, Chi Wang, Mingqin Li
HIGHLIGHT: We accelerate tensor program optimization by considering it as a multi-armed bandits problem and using Bayesian inference to achieve fast convergence.

359, TITLE: Bag of Tricks for Adversarial Training
https://openreview.net/forum?id=Xb8xvrtB8Ce
AUTHORS: Tianyu Pang, Xiao Yang, Yinpeng Dong, Hang Su, Jun Zhu
HIGHLIGHT: Empirical evaluation of basic training tricks used in adversarial training

360, TITLE: Learning with Instance-Dependent Label Noise: A Sample Sieve Approach
https://openreview.net/forum?id=2VXyy9mIyU3
AUTHORS: Hao Cheng, Zhaowei Zhu, Xingyu Li, Yifei Gong, Xing Sun, Yang Liu
HIGHLIGHT: This paper proposes a dynamic sample sieve method with strong theoretical guarantees to avoid overfitting to instance-based label noise.

361, TITLE: Efficient Reinforcement Learning in Factored MDPs with Application to Constrained RL
https://openreview.net/forum?id=fmtSg8591Q

AUTHORS: Xiaoyu Chen, Jiachen Hu, Lihong Li, Liwei Wang
HIGHLIGHT: We propose an efficient algorithm with near-optimal regret guarantee for factored MDP, and apply the algorithm to a new formulation of constrained RL.

362, TITLE: Unbiased Teacher for Semi-Supervised Object Detection
https://openreview.net/forum?id=MJIve1zgR_
AUTHORS: Yen-Cheng Liu, Chih-Yao Ma, Zijian He, Chia-Wen Kuo, Kan Chen, Peizhao Zhang, Bichen Wu, Zsolt Kira, Peter Vajda
HIGHLIGHT: We propose Unbiased Teacher to jointly address the pseudo-labeling bias issue and the overfitting issue in semi-supervised object detection, and our model performs favorably against existing works on COCO-standard, COCO-additional, and VOC.

363, TITLE: Neural Attention Distillation: Erasing Backdoor Triggers from Deep Neural Networks
https://openreview.net/forum?id=9l0K4OM-oXE
AUTHORS: Yige Li, Xingjun Ma, Nodens Koren, Lingjuan Lyu, Xixiang Lyu, Bo Li
HIGHLIGHT: A simple but effective nerual attention distillation method for backdoor defense.

364, TITLE: Contrastive Learning with Adversarial Perturbations for Conditional Text Generation
https://openreview.net/forum?id=Wga_hrCa3P3
AUTHORS: Seanie Lee, Dong Bok Lee, Sung Ju Hwang
HIGHLIGHT: We propose a contrastive learning with adversarial perturbation to tackle the exposure bias problem.

365, TITLE: When Optimizing f-Divergence is Robust with Label Noise
https://openreview.net/forum?id=WesiCoRVQ15
AUTHORS: Jiaheng Wei, Yang Liu
HIGHLIGHT: We show when maximizing a properly defined f-divergence measure with respect to a classifier's predictions and the supervised labels is robust with label noise.

366, TITLE: Conditional Generative Modeling via Learning the Latent Space
https://openreview.net/forum?id=VJnrYcnRc6
AUTHORS: Sameera Ramasinghe, Kanchana Nisal Ranasinghe, Salman Khan, Nick Barnes, Stephen Gould
HIGHLIGHT: Conditional generation in continuous multimodal spaces by learning the behavior of latent variables.

367, TITLE: Text Generation by Learning from Off-Policy Demonstrations
https://openreview.net/forum?id=RovX-uQ1Hua
AUTHORS: Richard Yuanzhe Pang, He He
HIGHLIGHT: We propose GOLD (generation by off-policy learning from demonstrations): an easy-to-optimize algorithm that learns from the off-policy demonstrations by importance weighting.

368, TITLE: Learning Long-term Visual Dynamics with Region Proposal Interaction Networks
https://openreview.net/forum?id=_X_4Akcd8Re
AUTHORS: Haozhi Qi, Xiaolong Wang, Deepak Pathak, Yi Ma, Jitendra Malik
HIGHLIGHT: We propose Region Proposal Interaction Networks for physical interaction prediction, which is applied across both simulation and real world environments for long-range prediction and planning.

369, TITLE: ChipNet: Budget-Aware Pruning with Heaviside Continuous Approximations
https://openreview.net/forum?id=xCxXwTzx4L1
AUTHORS: Rishabh Tiwari, Udbhav Bamba, Arnav Chavan, Deepak Gupta
HIGHLIGHT: A budget-aware deterministic strategy for structured pruning based on continuous Heaviside approximations and crispness loss.

370, TITLE: Learning to Deceive Knowledge Graph Augmented Models via Targeted Perturbation
https://openreview.net/forum?id=b7g3_ZMHnT0
AUTHORS: Mrigank Raman, Hansen Wang, PeiFeng Wang, Siddhant Agarwal, Sungchul Kim, Ryan Rossi, Handong Zhao, Nedim Lipka, Xiang Ren
HIGHLIGHT: Our findings raise doubts about KG-augmented models' ability to leverage KG information and provide plausible explanations.

371, TITLE: IEPT: Instance-Level and Episode-Level Pretext Tasks for Few-Shot Learning
https://openreview.net/forum?id=xzqLpqRzxLq

AUTHORS: Manli Zhang, Jianhong Zhang, Zhiwu Lu, Tao Xiang, Mingyu Ding, Songfang Huang
HIGHLIGHT: This paper proposes a novel Instance-level and Episode-level Pretext Task (IEPT) framework that seamlessly integrates SSL into FSL.

372, TITLE: The Role of Momentum Parameters in the Optimal Convergence of Adaptive Polyak's Heavy-ball Methods
https://openreview.net/forum?id=L7WD8ZdscQ5
AUTHORS: Wei Tao, Sheng Long, Gaowei Wu, Qing Tao
HIGHLIGHT: A theory-practice gap in convex optimization and deep learning is bridged by giving a novel convergence analysis of the last Iterate of adaptive Heavy-ball methods.

373, TITLE: Training with Quantization Noise for Extreme Model Compression
https://openreview.net/forum?id=dV19Yyi1fS3
AUTHORS: Angela Fan, Pierre Stock, Benjamin Graham, Edouard Grave, R?mi Gribonval, Herve Jegou, Armand Joulin
HIGHLIGHT: In this paper, we extend this approach to work with extreme compression methods where the approximations introduced by STE are severe.

374, TITLE: Adaptive Extra-Gradient Methods for Min-Max Optimization and Games
https://openreview.net/forum?id=R0a0kFI3dJx
AUTHORS: Kimon Antonakopoulos, Veronica Belmega, Panayotis Mertikopoulos
HIGHLIGHT: We develop an adaptive mirror-prox method for min-max problems and games that achieves order-optimal rates in both smooth and non-smooth problems.

375, TITLE: Distilling Knowledge from Reader to Retriever for Question Answering
https://openreview.net/forum?id=NTEz-6wysdb
AUTHORS: Gautier Izacard, Edouard Grave
HIGHLIGHT: In this paper, we propose a technique to learn retriever models for downstream tasks, inspired by knowledge distillation, and which does not require annotated pairs of query and documents.

376, TITLE: Discovering Diverse Multi-Agent Strategic Behavior via Reward Randomization
https://openreview.net/forum?id=lvRTC669EY_
AUTHORS: Zhenggang Tang, Chao Yu, Boyuan Chen, Huazhe Xu, Xiaolong Wang, Fei Fang, Simon Shaolei Du, Yu Wang, Yi Wu
HIGHLIGHT: We propose an MARL algorithm, RPG, which discovers diverse non-trivial strategic behavior in several challenging multi-agent games.

377, TITLE: not-MIWAE: Deep Generative Modelling with Missing not at Random Data
https://openreview.net/forum?id=tu29GQT0JFy
AUTHORS: Niels Bruun Ipsen, Pierre-Alexandre Mattei, Jes Frellsen
HIGHLIGHT: We present an approach for building and fitting deep latent variable models (DLVMs) in cases where the missing process is dependent on the missing data.

378, TITLE: IDF++: Analyzing and Improving Integer Discrete Flows for Lossless Compression
https://openreview.net/forum?id=MBOyiNnYthd
AUTHORS: Rianne van den Berg, Alexey A. Gritsenko, Mostafa Dehghani, Casper Kaae S?nderby, Tim Salimans
HIGHLIGHT: We analyze and improve integer discrete normalizing flows for lossless source compression.

379, TITLE: Explaining by Imitating: Understanding Decisions by Interpretable Policy Learning
https://openreview.net/forum?id=unI5ucw_Jk
AUTHORS: Alihan H?y?k, Daniel Jarrett, Cem Tekin, Mihaela van der Schaar
HIGHLIGHT: We present a method for learning interpretable representations of behavior to enable auditing, quantifying, and understanding human decision-making processes.

380, TITLE: Learning with AMIGo: Adversarially Motivated Intrinsic Goals
https://openreview.net/forum?id=ETBc_MIMgoX
AUTHORS: Andres Campero, Roberta Raileanu, Heinrich Kuttler, Joshua B. Tenenbaum, Tim Rockt?schel, Edward Grefenstette
HIGHLIGHT: A "constructively adversarial" teacher-student setup can augment on-policy algorithms to better solve difficult exploration tasks in RL.

381, TITLE: Incorporating Symmetry into Deep Dynamics Models for Improved Generalization

https://openreview.net/forum?id=wta_8Hx2KD
AUTHORS: Rui Wang, Robin Walters, Rose Yu
HIGHLIGHT: Integrate various symmetris into deep sequence models for forecasting turbulence and ocean currents with improved accuracy and physical consistency.

382, TITLE: CaPC Learning: Confidential and Private Collaborative Learning
https://openreview.net/forum?id=h2EbJ4_wMVq
AUTHORS: Christopher A. Choquette-Choo, Natalie Dullerud, Adam Dziedzic, Yunxiang Zhang, Somesh Jha, Nicolas Papernot, Xiao Wang
HIGHLIGHT: A method that enables parties to improve their own local heterogeneous machine learning models in a collaborative setting where both confidentiality and privacy need to be preserved to prevent both explicit and implicit sharing of private data.

383, TITLE: Heating up decision boundaries: isocapacitory saturation, adversarial scenarios and generalization bounds
https://openreview.net/forum?id=UwGY2qjqoLD
AUTHORS: Bogdan Georgiev, Lukas Franken, Mayukh Mukherjee
HIGHLIGHT: We propose a heat-theoretic/Brownian motion approach to evaluate decision boundary geometry (curvature and density) with further applications in adversarial defenses, compression and generalization estimates.

384, TITLE: A PAC-Bayesian Approach to Generalization Bounds for Graph Neural Networks
https://openreview.net/forum?id=TR-Nj6nFx42
AUTHORS: Renjie Liao, Raquel Urtasun, Richard Zemel
HIGHLIGHT: In this paper, we derived generalization bounds for two primary classes of graph neural networks (GNNs), namely graph convolutional networks (GCNs) and message passing GNNs, via a PAC-Bayesian approach.

385, TITLE: Clairvoyance: A Pipeline Toolkit for Medical Time Series
https://openreview.net/forum?id=xnC8YwKUE3k
AUTHORS: Daniel Jarrett, Jinsung Yoon, Ioana Bica, Zhaozhi Qian, Ari Ercole, Mihaela van der Schaar
HIGHLIGHT: We develop and present Clairvoyance: a pipeline toolkit for medical time series.

386, TITLE: Self-supervised Representation Learning with Relative Predictive Coding
https://openreview.net/forum?id=068E_JSq9O
AUTHORS: Yao-Hung Hubert Tsai, Martin Q. Ma, Muqiao Yang, Han Zhao, Louis-Philippe Morency, Ruslan Salakhutdinov
HIGHLIGHT: We present RPC, the Relative Predictive Coding, that achieves a good balance among the three challenges when modeling a contrastive learning objective: training stability, sensitivity to minibatch size, and downstream task performance.

387, TITLE: Offline Model-Based Optimization via Normalized Maximum Likelihood Estimation
https://openreview.net/forum?id=FmMKSO4e8JK
AUTHORS: Justin Fu, Sergey Levine
HIGHLIGHT: Offline, data-driven optimization using normalized maximum likelihood to produce robust function estimates.

388, TITLE: On the Impossibility of Global Convergence in Multi-Loss Optimization
https://openreview.net/forum?id=NQbnPjPYaG6
AUTHORS: Alistair Letcher
HIGHLIGHT: We prove that desirable convergence properties cannot simultaneously hold for any multi-loss optimization algorithm.

389, TITLE: A Block Minifloat Representation for Training Deep Neural Networks
https://openreview.net/forum?id=6zaTwpNSsQ2
AUTHORS: Sean Fox, Seyedramin Rasoulinezhad, Julian Faraone, david boland, Philip Leong
HIGHLIGHT: A new number representation, comparable to recently proposed 8-bit formats, for efficiently training a subset of DNN models.

390, TITLE: Selectivity considered harmful: evaluating the causal impact of class selectivity in DNNs
https://openreview.net/forum?id=8nl0k08uMi
AUTHORS: Matthew L Leavitt, Ari S. Morcos
HIGHLIGHT: Class selectivity in CNNs is neither sufficient nor strictly necessary for optimal test accuracy

391, TITLE: Discrete Graph Structure Learning for Forecasting Multiple Time Series

https://openreview.net/forum?id=WEHSlH5mOk
AUTHORS: Chao Shang, Jie Chen, Jinbo Bi
HIGHLIGHT: We propose a graph neural network approach that learns a graph structure to enhance the forecasting of multiple multivariate time series.

392, TITLE: Contrastive Learning with Hard Negative Samples
https://openreview.net/forum?id=CR1XOQ0UTh-
AUTHORS: Joshua David Robinson, Ching-Yao Chuang, Suvrit Sra, Stefanie Jegelka
HIGHLIGHT: We introduce an unsupervised method for sampling hard negatives for contrastive learning: the resulting embeddings have desirable theoretical properties, and have improved downstream performance on multiple different data modalities.

393, TITLE: Intraclass clustering: an implicit learning ability that regularizes DNNs
https://openreview.net/forum?id=tqOvYpjPax2
AUTHORS: Simon Carbonnelle, Christophe De Vleeschouwer
HIGHLIGHT: This paper provides empirical evidence that deep neural networks are implicitly regularized through their ability to extract meaningful clusters among the samples of a class.

394, TITLE: Sliced Kernelized Stein Discrepancy
https://openreview.net/forum?id=t0TaKv0Gx6Z
AUTHORS: Wenbo Gong, Yingzhen Li, Jos? Miguel Hern?ndez-Lobato
HIGHLIGHT: We proposed a method to tackle the curse-of-dimensionality issue of kernelized Stein discrepancy with RBF kernel, along with a novel particle inference algorithm resolving the vanishing repulsive issue of Stein variational gradient descent.

395, TITLE: Denoising Diffusion Implicit Models
https://openreview.net/forum?id=St1giarCHLP
AUTHORS: Jiaming Song, Chenlin Meng, Stefano Ermon
HIGHLIGHT: We show and justify a GAN-like iterative generative model with relatively fast sampling, high sample quality and without any adversarial training.

396, TITLE: Hierarchical Reinforcement Learning by Discovering Intrinsic Options
https://openreview.net/forum?id=r-gPPHEjpmw
AUTHORS: Jesse Zhang, Haonan Yu, Wei Xu
HIGHLIGHT: Hierarchical RL that discovers short-horizon task-agnostic options to perform well on sparse reward manipulation and navigation tasks.

397, TITLE: Answering Complex Open-Domain Questions with Multi-Hop Dense Retrieval
https://openreview.net/forum?id=EMHoBG0avc1
AUTHORS: Wenhan Xiong, Xiang Li, Srini Iyer, Jingfei Du, Patrick Lewis, William Yang Wang, Yashar Mehdad, Scott Yih, Sebastian Riedel, Douwe Kiela, Barlas Oguz
HIGHLIGHT: We propose a simple yet effective multi-hop dense retrieval approach for answering complex open-domain questions.

398, TITLE: Rethinking Soft Labels for Knowledge Distillation: A Bias?Variance Tradeoff Perspective
https://openreview.net/forum?id=gIHd-5X324
AUTHORS: Helong Zhou, Liangchen Song, Jiajie Chen, Ye Zhou, Guoli Wang, Junsong Yuan, Qian Zhang
HIGHLIGHT: For knowledge distillation, we analyze the regularization effect introduced by soft labels from a bias-variance perspective and propose weighted soft labels to handle the tradeoff.

399, TITLE: A Design Space Study for LISTA and Beyond
https://openreview.net/forum?id=GMgHyUPrXa
AUTHORS: Tianjian Meng, Xiaohan Chen, Yifan Jiang, Zhangyang Wang
HIGHLIGHT: Among all possible variations, we focus on extensively varying the connectivity patterns and neuron types, leading to a gigantic design space arising from LISTA.

400, TITLE: What Should Not Be Contrastive in Contrastive Learning
https://openreview.net/forum?id=CZ8Y3NzuVzO
AUTHORS: Tete Xiao, Xiaolong Wang, Alexei A Efros, Trevor Darrell
HIGHLIGHT: We introduce a contrastive learning framework which does not require prior knowledge of specific, task-dependent invariances.

401, TITLE: Do Wide and Deep Networks Learn the Same Things? Uncovering How Neural Network Representations Vary with Width and Depth
https://openreview.net/forum?id=KJNcAkY8tY4
AUTHORS: Thao Nguyen, Maithra Raghu, Simon Kornblith
HIGHLIGHT: We show that depth/width variations result in distinctive characteristics in the model internal representations, with resulting consequences for representations and output predictions across different model initializations and architectures.

402, TITLE: Learning to Set Waypoints for Audio-Visual Navigation
https://openreview.net/forum?id=cR91FAodFMe
AUTHORS: Changan Chen, Sagnik Majumder, Ziad Al-Halah, Ruohan Gao, Santhosh Kumar Ramakrishnan, Kristen Grauman
HIGHLIGHT: We introduce a hierarchical reinforcement learning approach to audio-visual navigation that learns to dynamically set waypoints in an end-to-end fashion

403, TITLE: Semi-supervised Keypoint Localization
https://openreview.net/forum?id=yFJ67zTeI2
AUTHORS: Olga Moskvyak, Frederic Maire, Feras Dayoub, Mahsa Baktashmotlagh
HIGHLIGHT: A novel method for semi-supervised keypoint localization via learning semantic keypoint representations.

404, TITLE: Neural Architecture Search on ImageNet in Four GPU Hours: A Theoretically Inspired Perspective
https://openreview.net/forum?id=Cnon5ezMHtu
AUTHORS: Wuyang Chen, Xinyu Gong, Zhangyang Wang
HIGHLIGHT: Our TE-NAS framework analyzes the spectrum of the neural tangent kernel (NTK) and the number of linear regions in the input space, achieving high-quality architecture search while dramatically reducing the search cost to four hours on ImageNet.

405, TITLE: Off-Dynamics Reinforcement Learning: Training for Transfer with Domain Classifiers
https://openreview.net/forum?id=eqBwg3AcIAK
AUTHORS: Benjamin Eysenbach, Shreyas Chaudhari, Swapnil Asawa, Sergey Levine, Ruslan Salakhutdinov
HIGHLIGHT: We propose a method for addressing domain adaptation in RL by using a (learned) modified reward, and prove that our method recovers a near-optimal policy for the target domain.

406, TITLE: Federated Semi-Supervised Learning with Inter-Client Consistency & Disjoint Learning
https://openreview.net/forum?id=ce6CFXBh30h
AUTHORS: Wonyong Jeong, Jaehong Yoon, Eunho Yang, Sung Ju Hwang
HIGHLIGHT: We introduce a new practical problem of federated learning with a deficiency of supervision and study two realistic scenarios with a novel method to tackle the problems, including inter-client consistency and disjoint learning.

407, TITLE: Learning Safe Multi-agent Control with Decentralized Neural Barrier Certificates
https://openreview.net/forum?id=P6_q1BRxY8Q
AUTHORS: Zengyi Qin, Kaiqing Zhang, Yuxiao Chen, Jingkai Chen, Chuchu Fan
HIGHLIGHT: We propose a safe and remarkably scalable multi-agent control approach via jointly learning the policy and decentralized control barrier certificates.

408, TITLE: Direction Matters: On the Implicit Regularization Effect of Stochastic Gradient Descent with Moderate Learning Rate
https://openreview.net/forum?id=3X64RLgzY6O
AUTHORS: Jingfeng Wu, Difan Zou, Vladimir Braverman, Quanquan Gu
HIGHLIGHT: We show a directional regularization effect for SGD with moderate learning rate

409, TITLE: Meta Attention Networks: Meta-Learning Attention to Modulate Information Between Recurrent Independent Mechanisms
https://openreview.net/forum?id=Lc28QAB4ypz
AUTHORS: Kanika Madan, Nan Rosemary Ke, Anirudh Goyal, Bernhard Sch?lkopf, Yoshua Bengio
HIGHLIGHT: Different time scale learning of independent mechanisms can lead to a better generalization.

410, TITLE: Policy-Driven Attack: Learning to Query for Hard-label Black-box Adversarial Examples
https://openreview.net/forum?id=pzpytjk3Xb2
AUTHORS: Ziang Yan, Yiwen Guo, Jian Liang, Changshui Zhang
HIGHLIGHT: A novel hard-label black-box adversarial attack that introduces a reinforcement learning based formulation with a pre-trained policy network

411, TITLE: A Mathematical Exploration of Why Language Models Help Solve Downstream Tasks
https://openreview.net/forum?id=vVjIW3sEc1s
AUTHORS: Nikunj Saunshi, Sadhika Malladi, Sanjeev Arora
HIGHLIGHT: We develop a mathematical framework for understanding why language model features help with downstream linear classification tasks of interest

412, TITLE: Representation Learning for Sequence Data with Deep Autoencoding Predictive Components
https://openreview.net/forum?id=Naqw7EHIfrv
AUTHORS: Junwen Bai, Weiran Wang, Yingbo Zhou, Caiming Xiong
HIGHLIGHT: We propose Deep Autoencoding Predictive Components (DAPC) -- a self-supervised representation learning method for sequence data, based on the intuition that useful representations of sequence data should exhibit a simple structure in the latent space.

413, TITLE: A unifying view on implicit bias in training linear neural networks
https://openreview.net/forum?id=ZsZM-4iMQkH
AUTHORS: Chulhee Yun, Shankar Krishnan, Hossein Mobahi
HIGHLIGHT: We propose a unifying framework for analyzing implicit bias of linear networks and show theorems that extend existing results with less convergence assumptions.

414, TITLE: What Makes Instance Discrimination Good for Transfer Learning?
https://openreview.net/forum?id=tC6iW2UUbJf
AUTHORS: Nanxuan Zhao, Zhirong Wu, Rynson W. H. Lau, Stephen Lin
HIGHLIGHT: Understanding why self-supervised contrastive learning outperforms supervised counterparts for image pretraining

415, TITLE: Learning Accurate Entropy Model with Global Reference for Image Compression
https://openreview.net/forum?id=cTbIjyrUVwJ
AUTHORS: Yichen Qian, Zhiyu Tan, Xiuyu Sun, Ming Lin, Dongyang Li, Zhenhong Sun, Li Hao, Rong Jin
HIGHLIGHT: In this paper, we propose a novel Reference-based Model for image compression to effectively leverage both the local and global context information, which yields an enhanced compression performance.

416, TITLE: Loss Function Discovery for Object Detection via Convergence-Simulation Driven Search
https://openreview.net/forum?id=5jzlpHvvRk
AUTHORS: Peidong Liu, Gengwei Zhang, Bochao Wang, Hang Xu, Xiaodan Liang, Yong Jiang, Zhenguo Li
HIGHLIGHT: We propose an effective convergence-simulation driven evolutionary search algorithm, called CSE-Autoloss, for object detection loss function discovery, which achieves 20x speedup via progressive convergence-simulation modules.

417, TITLE: Effective Abstract Reasoning with Dual-Contrast Network
https://openreview.net/forum?id=ldxlzGYWDmW
AUTHORS: Tao Zhuo, Mohan Kankanhalli
HIGHLIGHT: We propose a simple yet effective Dual-Contrast Network (DCNet) to solve Raven's progressive matrices without using auxiliary annotations and assumptions.

418, TITLE: Do not Let Privacy Overbill Utility: Gradient Embedding Perturbation for Private Learning
https://openreview.net/forum?id=7aogOj_VYO0
AUTHORS: Da Yu, Huishuai Zhang, Wei Chen, Tie-Yan Liu
HIGHLIGHT: A new algorithm for differentially private learning that advances state-of-the-art performance on several benchmark datasets.

419, TITLE: Set Prediction without Imposing Structure as Conditional Density Estimation
https://openreview.net/forum?id=04ArenGOz3
AUTHORS: David W Zhang, Gertjan J. Burghouts, Cees G. M. Snoek
HIGHLIGHT: A set prediction training & prediction framework that addresses tasks with ambiguous predictions.

420, TITLE: Clustering-friendly Representation Learning via Instance Discrimination and Feature Decorrelation
https://openreview.net/forum?id=e12NDM7wkEY
AUTHORS: Yaling Tao, Kentaro Takagi, Kouta Nakata

HIGHLIGHT: We present a clustering-friendly representation learning method using instance discrimination and feature decorrelation, which achieves accuracy of 81.5% and 95.4% on CIFAR-10 and ImageNet-10, respectively, far above state-of-the-art values.

421, TITLE: Language-Agnostic Representation Learning of Source Code from Structure and Context
https://openreview.net/forum?id=Xh5eMZVONGF
AUTHORS: Daniel Z?gner, Tobias Kirschstein, Michele Catasta, Jure Leskovec, Stephan G?nnemann
HIGHLIGHT: Language-agnostic learning from Structure and Context of programs improves learning.

422, TITLE: Training GANs with Stronger Augmentations via Contrastive Discriminator
https://openreview.net/forum?id=eo6U4CAwVmg
AUTHORS: Jongheon Jeong, Jinwoo Shin
HIGHLIGHT: We propose a novel discriminator of GAN showing that contrastive representation learning, e.g., SimCLR, and GAN can benefit each other when they are jointly trained.

423, TITLE: Influence Functions in Deep Learning Are Fragile
https://openreview.net/forum?id=xHKVVHGDOEk
AUTHORS: Samyadeep Basu, Phil Pope, Soheil Feizi
HIGHLIGHT: End-to-end investigation of the behaviour of influence functions in deep learning

424, TITLE: Separation and Concentration in Deep Networks
https://openreview.net/forum?id=8HhkbjrWLdE
AUTHORS: John Zarka, Florentin Guth, St?phane Mallat
HIGHLIGHT: Numerical experiments demonstrate that deep neural network classifiers progressively separate class distributions around their mean, achieving linear separability on the training set, and increasing the Fisher discriminant ratio. We explain this mechanism with two types of operators.

425, TITLE: Colorization Transformer
https://openreview.net/forum?id=5NA1PinlGFu
AUTHORS: Manoj Kumar, Dirk Weissenborn, Nal Kalchbrenner
HIGHLIGHT: Self-attention for colorization

426, TITLE: Autoregressive Dynamics Models for Offline Policy Evaluation and Optimization
https://openreview.net/forum?id=kmqjgSNXby
AUTHORS: Michael R Zhang, Thomas Paine, Ofir Nachum, Cosmin Paduraru, George Tucker, ziyu wang, Mohammad Norouzi
HIGHLIGHT: We demonstrate autoregressive dynamics models outperform standard feedforward models and other baselines in offline policy evaluation and optimization.

427, TITLE: FedBN: Federated Learning on Non-IID Features via Local Batch Normalization
https://openreview.net/forum?id=6YEQUn0QICG
AUTHORS: Xiaoxiao Li, Meirui JIANG, Xiaofei Zhang, Michael Kamp, Qi Dou
HIGHLIGHT: We propose a novel and efficient federated learning aggregation method, denoted FedBN, that uses local batch normalization to effectively tackle the underexplored non-iid problem of heterogeneous feature distributions, or feature shift.

428, TITLE: Learning Robust State Abstractions for Hidden-Parameter Block MDPs
https://openreview.net/forum?id=fmOOI2a3tQP
AUTHORS: Amy Zhang, Shagun Sodhani, Khimya Khetarpal, Joelle Pineau
HIGHLIGHT: In this work, we leverage ideas of common structure from the HiP-MDP setting, and extend it to enable robust state abstractions inspired by Block MDPs.

429, TITLE: Meta-Learning with Neural Tangent Kernels
https://openreview.net/forum?id=Ti87Pv5Oc8
AUTHORS: Yufan Zhou, Zhenyi Wang, Jiayi Xian, Changyou Chen, Jinhui Xu
HIGHLIGHT: First work to define meta learning in RKHS induced by Neural Tangent Kernel

430, TITLE: Continual learning in recurrent neural networks
https://openreview.net/forum?id=8xeBUgD8u9
AUTHORS: Benjamin Ehret, Christian Henning, Maria Cervera, Alexander Meulemans, Johannes Von Oswald, Benjamin F Grewe

HIGHLIGHT: This paper studies the behavior of established approaches to the problem of continual learning in the context of recurrent neural networks.

431, TITLE: A Trainable Optimal Transport Embedding for Feature Aggregation
https://openreview.net/forum?id=ZK6vTvb84s
AUTHORS: Gr?goire Mialon, Dexiong Chen, Alexandre d'Aspremont, Julien Mairal
HIGHLIGHT: We propose a new, trainable embedding for large sets of features such as biological sequences, and demonstrate its effectiveness.

432, TITLE: Learning "What-if" Explanations for Sequential Decision-Making
https://openreview.net/forum?id=h0de3QWtGG
AUTHORS: Ioana Bica, Daniel Jarrett, Alihan H?y?k, Mihaela van der Schaar
HIGHLIGHT: We propose explaining sequential decision-making by integrating counterfactual reasoning into batch inverse reinforcement learning and recovering the preferences of experts over "what-if" outcomes.

433, TITLE: Improving Transformation Invariance in Contrastive Representation Learning
https://openreview.net/forum?id=NomEDgIEBwE
AUTHORS: Adam Foster, Rattana Pukdee, Tom Rainforth
HIGHLIGHT: We propose methods to strengthen the invariance properties of representations obtained by contrastive learning using novel gradient regularization during training and feature averaging at test time.
Finally, we introduce the novel Spirograph dataset to explore our ideas in the context of a differentiable generative process with multiple downstream tasks, showing that our techniques for learning invariance are highly beneficial.

434, TITLE: Shapley explainability on the data manifold
https://openreview.net/forum?id=OPyWRrcjVQw
AUTHORS: Christopher Frye, Damien de Mijolla, Tom Begley, Laurence Cowton, Megan Stanley, Ilya Feige
HIGHLIGHT: We present drawbacks of model explanations that do not respect the data manifold, and introduce two methods for on-manifold explainability.

435, TITLE: Noise or Signal: The Role of Image Backgrounds in Object Recognition
https://openreview.net/forum?id=gl3D-xY7wLq
AUTHORS: Kai Yuanqing Xiao, Logan Engstrom, Andrew Ilyas, Aleksander Madry
HIGHLIGHT: We develop and use a toolkit to investigate models? use of (and reliance on) image backgrounds.

436, TITLE: Enjoy Your Editing: Controllable GANs for Image Editing via Latent Space Navigation
https://openreview.net/forum?id=HOFxeCutxZR
AUTHORS: Peiye Zhuang, Oluwasanmi O Koyejo, Alex Schwing
HIGHLIGHT: We propose a state-of-the-art approach to semantically edit images by transferring latent vectors towards meaningful latent space directions.

437, TITLE: Perceptual Adversarial Robustness: Generalizable Defenses Against Unforeseen Threat Models
https://openreview.net/forum?id=dFwBosAcJkN
AUTHORS: Cassidy Laidlaw, Sahil Singla, Soheil Feizi
HIGHLIGHT: Adversarial training against a perceptually-aligned attack gives high robustness against many diverse adversarial threat models.

438, TITLE: Zero-Cost Proxies for Lightweight NAS
https://openreview.net/forum?id=0cmMMy8J5q
AUTHORS: Mohamed S Abdelfattah, Abhinav Mehrotra, Lukasz Dudziak, Nicholas Donald Lane
HIGHLIGHT: A single minibatch of data is used to score neural networks for NAS instead of performing full training.

439, TITLE: Usable Information and Evolution of Optimal Representations During Training
https://openreview.net/forum?id=p8agn6bmTbr
AUTHORS: Michael Kleinman, Alessandro Achille, Daksh Idnani, Jonathan Kao
HIGHLIGHT: We introduce a notion of usable information contained in the representation learned by a deep network, and use it to study how optimal representations for the task emerge during training, and how they adapt to different tasks.

440, TITLE: Exploring the Uncertainty Properties of Neural Networks? Implicit Priors in the Infinite-Width Limit
https://openreview.net/forum?id=MjvduJCsE4
AUTHORS: Ben Adlam, Jaehoon Lee, Lechao Xiao, Jeffrey Pennington, Jasper Snoek

HIGHLIGHT: We study the uncertainty properties of infinitely-wide neural networks

441, TITLE: On the geometry of generalization and memorization in deep neural networks
https://openreview.net/forum?id=V8jrrnwGbuc
AUTHORS: Cory Stephenson, suchismita padhy, Abhinav Ganesh, Yue Hui, Hanlin Tang, SueYeon Chung
HIGHLIGHT: We analyze the representational geometry of deep neural networks during generalization and memorization, and find the structure across layers of when and where memorization occurs, as well as drivers for this emerging structure.

442, TITLE: Deep Partition Aggregation: Provable Defenses against General Poisoning Attacks
https://openreview.net/forum?id=YUGG2tFuPM
AUTHORS: Alexander Levine, Soheil Feizi
HIGHLIGHT: We propose novel certified defenses against label-flipping and general adversarial poisoning attacks.

443, TITLE: DC3: A learning method for optimization with hard constraints
https://openreview.net/forum?id=V1ZHVxJ6dSS
AUTHORS: Priya L. Donti, David Rolnick, J Zico Kolter
HIGHLIGHT: We describe a method, DC3, for fast approximate solutions to optimization problems with hard constraints, which enforces feasibility via a differentiable procedure incorporated into a neural network.

444, TITLE: Is Label Smoothing Truly Incompatible with Knowledge Distillation: An Empirical Study
https://openreview.net/forum?id=PObuuGVrGaZ
AUTHORS: Zhiqiang Shen, Zechun Liu, Dejia Xu, Zitian Chen, Kwang-Ting Cheng, Marios Savvides
HIGHLIGHT: This work empirically clarifies a recently discovered perspective that label smoothing is incompatible with knowledge distillation.

445, TITLE: Shape-Texture Debiased Neural Network Training
https://openreview.net/forum?id=Db4yerZTYkz
AUTHORS: Yingwei Li, Qihang Yu, Mingxing Tan, Jieru Mei, Peng Tang, Wei Shen, Alan Yuille, cihang xie
HIGHLIGHT: Training CNNs to acquire a debiased shape-texture representation improves image recognition.

446, TITLE: Using latent space regression to analyze and leverage compositionality in GANs
https://openreview.net/forum?id=sjuuTm4vj0
AUTHORS: Lucy Chai, Jonas Wulff, Phillip Isola
HIGHLIGHT: We use a latent regressor network to investigate compositional properties of image synthesis with GANs.

447, TITLE: Blending MPC & Value Function Approximation for Efficient Reinforcement Learning
https://openreview.net/forum?id=RqCC_00Bg7V
AUTHORS: Mohak Bhardwaj, Sanjiban Choudhury, Byron Boots
HIGHLIGHT: A framework for blending model-predictive control and model-free value function learning to systematically trade-off bias due to approximate dynamics models and value functions learned from real data

448, TITLE: Model Patching: Closing the Subgroup Performance Gap with Data Augmentation
https://openreview.net/forum?id=9YlaeLfuhJF
AUTHORS: Karan Goel, Albert Gu, Yixuan Li, Christopher Re
HIGHLIGHT: We describe how to fix classifiers that fail on subgroups of a class using a combination of learned data augmentation & consistency training to achieve subgroup invariance.

449, TITLE: Non-asymptotic Confidence Intervals of Off-policy Evaluation: Primal and Dual Bounds
https://openreview.net/forum?id=dKg5D1Z1Lm
AUTHORS: Yihao Feng, Ziyang Tang, na zhang, qiang liu
HIGHLIGHT: We propose an approach to constructing non-asymptotic confidence intervals of off-policy estimation.

450, TITLE: Linear Mode Connectivity in Multitask and Continual Learning
https://openreview.net/forum?id=Fmg_fQYUejf
AUTHORS: Seyed Iman Mirzadeh, Mehrdad Farajtabar, Dilan Gorur, Razvan Pascanu, Hassan Ghasemzadeh
HIGHLIGHT: We show that continual and multitask minima are connected by linear low-error paths and design an effective continual learning algorithm that exploits this property.

451, TITLE: Robust and Generalizable Visual Representation Learning via Random Convolutions

https://openreview.net/forum?id=BVSM0x3EDK6
AUTHORS: Zhenlin Xu, Deyi Liu, Junlin Yang, Colin Raffel, Marc Niethammer
HIGHLIGHT: We use random convolutions as data augmentation to train robust visual representation that generalize to new domains.

452, TITLE: Intrinsic-Extrinsic Convolution and Pooling for Learning on 3D Protein Structures
https://openreview.net/forum?id=l0mSUROpwY
AUTHORS: Pedro Hermosilla, Marco Sch?fer, Matej Lang, Gloria Fackelmann, Pere-Pau V?zquez, Barbora Kozlikova, Michael Krone, Tobias Ritschel, Timo Ropinski
HIGHLIGHT: We present a new neural network architecture to process 3D protein structures.
Second, we introduce a set of hierarchical pooling operators that enable multi-scale protein analysis.

453, TITLE: Variational State-Space Models for Localisation and Dense 3D Mapping in 6 DoF
https://openreview.net/forum?id=XAS3uKeFWj
AUTHORS: Atanas Mirchev, Baris Kayalibay, Patrick van der Smagt, Justin Bayer
HIGHLIGHT: We propose a variational state-space model with a latent map for 6-DoF localisation, 3D dense mapping and generative modelling for planning.

454, TITLE: AdamP: Slowing Down the Slowdown for Momentum Optimizers on Scale-invariant Weights
https://openreview.net/forum?id=Iz3zU3M316D
AUTHORS: Byeongho Heo, Sanghyuk Chun, Seong Joon Oh, Dongyoon Han, Sangdoo Yun, Gyuwan Kim, Youngjung Uh, Jung-Woo Ha
HIGHLIGHT: In this paper, we verify that the widely-adopted combination of the two ingredients lead to the premature decay of effective step sizes and sub-optimal model performances.

455, TITLE: MiCE: Mixture of Contrastive Experts for Unsupervised Image Clustering
https://openreview.net/forum?id=gV3wdEOGy_V
AUTHORS: Tsung Wei Tsai, Chongxuan Li, Jun Zhu
HIGHLIGHT: A principled probabilistic clustering method that exploits the discriminative representations learned by contrastive learning and the semantic structures captured by a latent mixture model in a unified framework.

456, TITLE: HalentNet: Multimodal Trajectory Forecasting with Hallucinative Intents
https://openreview.net/forum?id=9GBZBPn0Jx
AUTHORS: Deyao Zhu, Mohamed Zahran, Li Erran Li, Mohamed Elhoseiny
HIGHLIGHT: Towards this goal, we introduce HalentNet to better model the future motion distribution in addition to a traditional trajectory regression learning objective by incorporating generative augmentation losses.

457, TITLE: Model-based micro-data reinforcement learning: what are the crucial model properties and which model to choose?
https://openreview.net/forum?id=p5uylG94S68
AUTHORS: Bal?zs K?gl, Gabriel Hurtado, Albert Thomas
HIGHLIGHT: Crucial model properties for model-based reinforcement learning: multi-modal posterior predictives and heteroscedasticity.

458, TITLE: Private Image Reconstruction from System Side Channels Using Generative Models
https://openreview.net/forum?id=y06VOYLcQXa
AUTHORS: Yuanyuan Yuan, Shuai Wang, Junping Zhang
HIGHLIGHT: We present the first generative model-based side channel analysis (SCA) to reconstruct private user images.

459, TITLE: Contextual Transformation Networks for Online Continual Learning
https://openreview.net/forum?id=zx_uX-BO7CH
AUTHORS: Quang Pham, Chenghao Liu, Doyen Sahoo, Steven HOI
HIGHLIGHT: This paper develops a novel method that can model task-specific features with minimal complexity overhead.

460, TITLE: Towards A Unified Understanding and Improving of Adversarial Transferability
https://openreview.net/forum?id=X76iqnUbBjz
AUTHORS: Xin Wang, Jie Ren, Shuyun Lin, Xiangming Zhu, Yisen Wang, Quanshi Zhang
HIGHLIGHT: We prove the close relationship between the interaction and adversarial transferability, provide a unified explanation for previous transferability-boosting methods, and develop a loss to improve adversarial transferability.

461, TITLE: The inductive bias of ReLU networks on orthogonally separable data
https://openreview.net/forum?id=krz7T0xU9Z_
AUTHORS: Mary Phuong, Christoph H Lampert
HIGHLIGHT: We characterise the function learnt by two-layer ReLU nets trained on orthogonally separable data.

462, TITLE: A statistical theory of cold posteriors in deep neural networks
https://openreview.net/forum?id=Rd138pWXMvG
AUTHORS: Laurence Aitchison
HIGHLIGHT: We develop a generative model of dataset curation that explains the cold-posterior effect

463, TITLE: IOT: Instance-wise Layer Reordering for Transformer Structures
https://openreview.net/forum?id=ipUPfYxWZvM
AUTHORS: Jinhua Zhu, Lijun Wu, Yingce Xia, Shufang Xie, Tao Qin, Wengang Zhou, Houqiang Li, Tie-Yan Liu
HIGHLIGHT: Based on this observation, in this work, we break the assumption of the fixed layer order in Transformer and introduce instance-wise layer reordering into model structure.

464, TITLE: Counterfactual Generative Networks
https://openreview.net/forum?id=BXewfAYMmJw
AUTHORS: Axel Sauer, Andreas Geiger
HIGHLIGHT: A generative model structured into independent causal mechanisms produces images for training invariant classifiers.

465, TITLE: Conditionally Adaptive Multi-Task Learning: Improving Transfer Learning in NLP Using Fewer Parameters & Less Data
https://openreview.net/forum?id=de11dbHzAMF
AUTHORS: Jonathan Pilault, Amine El hattami, Christopher Pal
HIGHLIGHT: Can multi-task outperform single task fine-tuning? CA-MTL is a new method that shows that it is possible with task conditioned model adaption and uncertainty sampling.

466, TITLE: Towards Impartial Multi-task Learning
https://openreview.net/forum?id=IMPnRXEWpvr
AUTHORS: Liyang Liu, Yi Li, Zhanghui Kuang, Jing-Hao Xue, Yimin Chen, Wenming Yang, Qingmin Liao, Wayne Zhang
HIGHLIGHT: We propose an impartial multi-task learning method that treats all tasks equally without bias towards any task.

467, TITLE: Theoretical bounds on estimation error for meta-learning
https://openreview.net/forum?id=SZ3wtsXfzQR
AUTHORS: James Lucas, Mengye Ren, Irene Raissa KAMENI KAMENI, Toniann Pitassi, Richard Zemel
HIGHLIGHT: We prove novel minimax risk lower bounds and upper bounds for meta learners

468, TITLE: Domain-Robust Visual Imitation Learning with Mutual Information Constraints
https://openreview.net/forum?id=QubpWYfdNry
AUTHORS: Edoardo Cetin, Oya Celiktutan
HIGHLIGHT: Imitation of visual expert demonstrations robust to appearance and embodiment mismatch, working for high dimensional control problems.

469, TITLE: Unsupervised Representation Learning for Time Series with Temporal Neighborhood Coding
https://openreview.net/forum?id=8qDwejCuCN
AUTHORS: Sana Tonekaboni, Danny Eytan, Anna Goldenberg
HIGHLIGHT: An unsupervised representation learning framework for high-dimensional non-stationary time series

470, TITLE: Enforcing robust control guarantees within neural network policies
https://openreview.net/forum?id=5lhWG3Hj2By
AUTHORS: Priya L. Donti, Melrose Roderick, Mahyar Fazlyab, J Zico Kolter
HIGHLIGHT: We develop a generic nonlinear control policy class, parameterized by neural networks, that nonetheless enforces the same provable robustness criteria as robust control.

471, TITLE: Active Contrastive Learning of Audio-Visual Video Representations
https://openreview.net/forum?id=OMizHuea_HB
AUTHORS: Shuang Ma, Zhaoyang Zeng, Daniel McDuff, Yale Song

HIGHLIGHT: We propose an active learning approach to improve negative sampling for contrastive learning and demonstrate it on learning audio-visual representations from videos.

472, TITLE: Parameter Efficient Multimodal Transformers for Video Representation Learning
https://openreview.net/forum?id=6UdQLhqJyFD
AUTHORS: Sangho Lee, Youngjae Yu, Gunhee Kim, Thomas Breuel, Jan Kautz, Yale Song
HIGHLIGHT: We propose a technique to reduce the number of parameters in multimodal BERT models up to 80% (from 155 million to 31 million parameters).

473, TITLE: Robust Pruning at Initialization
https://openreview.net/forum?id=vXj_ucZQ4hA
AUTHORS: Soufiane Hayou, Jean-Francois Ton, Arnaud Doucet, Yee Whye Teh
HIGHLIGHT: Making pruning at initialization robust to Gradient vanishing/exploding

474, TITLE: Efficient Wasserstein Natural Gradients for Reinforcement Learning
https://openreview.net/forum?id=OHgnfSrn2jv
AUTHORS: Ted Moskovitz, Michael Arbel, Ferenc Huszar, Arthur Gretton
HIGHLIGHT: We develop novel, efficient estimators for the Wasserstein natural gradient applied to reinforcement learning that improve the efficiency and performance of advanced baselines.

475, TITLE: Probing BERT in Hyperbolic Spaces
https://openreview.net/forum?id=17VnwXYZyhH
AUTHORS: Boli Chen, Yao Fu, Guangwei Xu, Pengjun Xie, Chuanqi Tan, Mosha Chen, Liping Jing
HIGHLIGHT: We propose a Poincar? probe for finding syntax and sentiments from BERT in hyperbolic spaces.

476, TITLE: On Fast Adversarial Robustness Adaptation in Model-Agnostic Meta-Learning
https://openreview.net/forum?id=o81ZyBCojoA
AUTHORS: Ren Wang, Kaidi Xu, Sijia Liu, Pin-Yu Chen, Tsui-Wei Weng, Chuang Gan, Meng Wang
HIGHLIGHT: We propose a general but easily-optimized robustness-regularized meta-learning framework, which allows the use of unlabeled data augmentation, fast adversarial attack generation, and computationally-light fine-tuning.

477, TITLE: Anatomy of Catastrophic Forgetting: Hidden Representations and Task Semantics
https://openreview.net/forum?id=LhY8QdUGSuw
AUTHORS: Vinay Venkatesh Ramasesh, Ethan Dyer, Maithra Raghu
HIGHLIGHT: We study the layerwise change in representations due to catastrophic forgetting, and use our understanding to study how task similarity influences forgetting.

478, TITLE: Trusted Multi-View Classification
https://openreview.net/forum?id=OOsR8BzCnl5
AUTHORS: Zongbo Han, Changqing Zhang, Huazhu Fu, Joey Tianyi Zhou
HIGHLIGHT: To this end, we propose a novel multi-view classification method, termed trusted multi-view classification, which provides a new paradigm for multi-view learning by dynamically integrating different views at an evidence level.

479, TITLE: i-Mix: A Strategy for Regularizing Contrastive Representation Learning
https://openreview.net/forum?id=T6AxtOaWydQ
AUTHORS: Kibok Lee, Yian Zhu, Kihyuk Sohn, Chun-Liang Li, Jinwoo Shin, Honglak Lee
HIGHLIGHT: We propose i-Mix, a simple yet effective strategy for regularizing contrastive representation learning in both vision and non-vision domains.

480, TITLE: Initialization and Regularization of Factorized Neural Layers
https://openreview.net/forum?id=KTlJT1nof6d
AUTHORS: Mikhail Khodak, Neil A. Tenenholtz, Lester Mackey, Nicolo Fusi
HIGHLIGHT: Principled initialization and regularization of factorized neural layers leads to strong performance in compression, knowledge distillation, and language modeling tasks.

481, TITLE: Learning to Generate 3D Shapes with Generative Cellular Automata
https://openreview.net/forum?id=rABUmU3ulQh
AUTHORS: Dongsu Zhang, Changwoon Choi, Jeonghwan Kim, Young Min Kim
HIGHLIGHT: We present a Markov chain based 3D generative model named Generative Cellular Automata (GCA), that progressively evolves to a shape of the learned distribution using the local update rules of cellular automata.

482, TITLE: Self-Supervised Learning of Compressed Video Representations
https://openreview.net/forum?id=jMPcEkJpdD
AUTHORS: Youngjae Yu, Sangho Lee, Gunhee Kim, Yale Song
HIGHLIGHT: We propose a self-supervised approach to learning compressed video representations.

483, TITLE: Factorizing Declarative and Procedural Knowledge in Structured, Dynamical Environments
https://openreview.net/forum?id=VVdmjgu7pKM
AUTHORS: Anirudh Goyal, Alex Lamb, Phanideep Gampa, Philippe Beaudoin, Charles Blundell, Sergey Levine, Yoshua Bengio, Michael Curtis Mozer
HIGHLIGHT: We explore separate factorization of procedural and declarative knowledge in modular recurrent neural networks.

484, TITLE: Cut out the annotator, keep the cutout: better segmentation with weak supervision
https://openreview.net/forum?id=bjkX6Kzb5H
AUTHORS: Sarah Hooper, Michael Wornow, Ying Hang Seah, Peter Kellman, Hui Xue, Frederic Sala, Curtis Langlotz, Christopher Re
HIGHLIGHT: In this work, we present a weak supervision approach for segmentation tasks, allowing users to train high-performing segmentation CNNs with very few hand-labeled training points.

485, TITLE: FastSpeech 2: Fast and High-Quality End-to-End Text to Speech
https://openreview.net/forum?id=piLPYqxtWuA
AUTHORS: Yi Ren, Chenxu Hu, Xu Tan, Tao Qin, Sheng Zhao, Zhou Zhao, Tie-Yan Liu
HIGHLIGHT: We propose a non-autoregressive TTS model named FastSpeech 2 to better solve the one-to-many mapping problem in TTS and surpass autoregressive models in voice quality.

486, TITLE: On Learning Universal Representations Across Languages
https://openreview.net/forum?id=Uu1Nw-eeTxJ
AUTHORS: Xiangpeng Wei, Yue Hu, Rongxiang Weng, Luxi Xing, Heng Yu, Weihua Luo
HIGHLIGHT: In this work, we extend pre-trained language models to learn universal representations among multiple languages, and show the effectiveness on cross-lingual understanding and generation.

487, TITLE: Effective Distributed Learning with Random Features: Improved Bounds and Algorithms
https://openreview.net/forum?id=jxdXSW9Doc
AUTHORS: Yong Liu, Jiankun Liu, Shuqiang Wang
HIGHLIGHT: This papaer focuses on the studies of the statistical properties of distributed KRR together with random features

488, TITLE: Multi-Class Uncertainty Calibration via Mutual Information Maximization-based Binning
https://openreview.net/forum?id=AICNpd8ke-m
AUTHORS: Kanil Patel, William H. Beluch, Bin Yang, Michael Pfeiffer, Dan Zhang
HIGHLIGHT: We propose I-Max binning, a novel method for multi-class calibration, improving over previous methods in terms of various ECE measures

489, TITLE: Neural ODE Processes
https://openreview.net/forum?id=27acGyyI1BY
AUTHORS: Alexander Norcliffe, Cristian Bodnar, Ben Day, Jacob Moss, Pietro Li?
HIGHLIGHT: We introduce a stochastic process based on Neural ODEs.

490, TITLE: Conformation-Guided Molecular Representation with Hamiltonian Neural Networks
https://openreview.net/forum?id=q-cnWaaoUTH
AUTHORS: Ziyao Li, Shuwen Yang, Guojie Song, Lingsheng Cai
HIGHLIGHT: We propose a molecular representation algorithm, which preserves molecular conformations with a neural physics engine and generates fingerprints with an MPNN.

491, TITLE: An Unsupervised Deep Learning Approach for Real-World Image Denoising
https://openreview.net/forum?id=tIjRAiFmU3y
AUTHORS: Dihan Zheng, Sia Huat Tan, Xiaowen Zhang, Zuoqiang Shi, Kaisheng Ma, Chenglong Bao
HIGHLIGHT: We propose an unsupervised real-world image denoising approach that combines DNNs with classical MAP approaches.

492, TITLE: Uncertainty in Gradient Boosting via Ensembles
https://openreview.net/forum?id=1Jv6b0Zq3qi
AUTHORS: Andrey Malinin, Liudmila Prokhorenkova, Aleksei Ustimenko
HIGHLIGHT: Propose and analyze an ensemble-based framework for deriving uncertainty estimates in GBDT models.

493, TITLE: Lossless Compression of Structured Convolutional Models via Lifting
https://openreview.net/forum?id=oxnp2q-PGL4
AUTHORS: Gustav Sourek, Filip Zelezny, Ondrej Kuzelka
HIGHLIGHT: Speeding up weight-sharing dynamic neural computation graphs, such as GNNs, with lifted inference.

494, TITLE: Neural networks with late-phase weights
https://openreview.net/forum?id=C0qJUx5dxFb
AUTHORS: Johannes Von Oswald, Seijin Kobayashi, Joao Sacramento, Alexander Meulemans, Christian Henning, Benjamin F Grewe
HIGHLIGHT: Here, we show that the solutions found by SGD can be further improved by ensembling a subset of the weights in late stages of learning.

495, TITLE: Disambiguating Symbolic Expressions in Informal Documents
https://openreview.net/forum?id=K5j7D81ABvt
AUTHORS: Dennis M?ller, Cezary Kaliszyk
HIGHLIGHT: We propose the task of \emph{disambiguating} symbolic expressions in informal STEM documents in the form of \LaTeX files -- that is, determining their precise semantics and abstract syntax tree -- as a neural machine translation task.

496, TITLE: Learning Parametrised Graph Shift Operators
https://openreview.net/forum?id=0OlrLvrsHwQ
AUTHORS: George Dasoulas, Johannes F. Lutzeyer, Michalis Vazirgiannis
HIGHLIGHT: We propose a parametrised graph shift operator (PGSO) to encode graph structure, providing a unified view of the most common GSOs, and improve GNN performance by incorporating the PGSO into the model training in an end-to-end manner.

497, TITLE: Efficient Conformal Prediction via Cascaded Inference with Expanded Admission
https://openreview.net/forum?id=tnSo6VRLmT
AUTHORS: Adam Fisch, Tal Schuster, Tommi S. Jaakkola, Regina Barzilay
HIGHLIGHT: This work proposes two complementary techniques for improving the efficiency of conformal prediction in large-scale domains---while still retaining performance guarantees.

498, TITLE: GANs Can Play Lottery Tickets Too
https://openreview.net/forum?id=1AoMhc_9jER
AUTHORS: Xuxi Chen, Zhenyu Zhang, Yongduo Sui, Tianlong Chen
HIGHLIGHT: Winning tickets exist in the deep generative adversarial networks, which substantially outperform previous state-of-the-art compressed GAN

499, TITLE: ResNet After All: Neural ODEs and Their Numerical Solution
https://openreview.net/forum?id=HxzSxSxLOJZ
AUTHORS: Katharina Ott, Prateek Katiyar, Philipp Hennig, Michael Tiemann
HIGHLIGHT: We explain why some Neural ODE models do not permit a continuous-depth interpretation after training and how to fix it.

500, TITLE: Semantic Re-tuning with Contrastive Tension
https://openreview.net/forum?id=Ov_sMNau-PF
AUTHORS: Fredrik Carlsson, Magnus Sahlgren, Evangelia Gogoulou, Amaru Cuba Gyllensten, Erik Ylip?? Hellqvist
HIGHLIGHT: A self-supervised method for learning STS related sentence embedding using pre-trained language models, setting a new SOTA for STS related embedding tasks.

501, TITLE: Property Controllable Variational Autoencoder via Invertible Mutual Dependence
https://openreview.net/forum?id=tYxG_OMs9WE
AUTHORS: Xiaojie Guo, Yuanqi Du, Liang Zhao
HIGHLIGHT: A novel generative model for learning interpretable latent representation for generating data with desired properties.

502, TITLE: Latent Convergent Cross Mapping
https://openreview.net/forum?id=4TSiOTkKe5P
AUTHORS: Edward De Brouwer, Adam Arany, Jaak Simm, Yves Moreau
HIGHLIGHT: Latent CCM uses reconstruction between latent processes of dynamical systems to infer causality between short and sporadic time series.

503, TITLE: Adaptive Universal Generalized PageRank Graph Neural Network
https://openreview.net/forum?id=n6jl7fLxrP
AUTHORS: Eli Chien, Jianhao Peng, Pan Li, Olgica Milenkovic
HIGHLIGHT: We combine generalized PageRank with GNNs to adapt universal node label patterns and the over-smoothing problem.

504, TITLE: Neural Learning of One-of-Many Solutions for Combinatorial Problems in Structured Output Spaces
https://openreview.net/forum?id=ATp1nW2FuZL
AUTHORS: Yatin Nandwani, Deepanshu Jindal, Mausam ., Parag Singla
HIGHLIGHT: This work identifies and proposes a solution for handling solution multiplicity while learning neural methods for combinatorial problems in structured output spaces.

505, TITLE: My Body is a Cage: the Role of Morphology in Graph-Based Incompatible Control
https://openreview.net/forum?id=N3zUDGN5lO
AUTHORS: Vitaly Kurin, Maximilian Igl, Tim Rockt?schel, Wendelin Boehmer, Shimon Whiteson
HIGHLIGHT: Transformer-based approach to multitask incompatible continuous control inspired by a hypothesis that any benefits GNNs extract from the graph structure are outweighed by difficulties they create for message passing.

506, TITLE: FedBE: Making Bayesian Model Ensemble Applicable to Federated Learning
https://openreview.net/forum?id=dgtpE6gKjHn
AUTHORS: Hong-You Chen, Wei-Lun Chao
HIGHLIGHT: In this paper, we propose a novel aggregation algorithm named FedBE, which takes a Bayesian inference perspective by sampling higher-quality global models and combining them via Bayesian model Ensemble, leading to much robust aggregation.

507, TITLE: MALI: A memory efficient and reverse accurate integrator for Neural ODEs
https://openreview.net/forum?id=blfSjHeFM_e
AUTHORS: Juntang Zhuang, Nicha C Dvornek, sekhar tatikonda, James s Duncan
HIGHLIGHT: A solver for ODE that guarantees accuracy in reverse-time trajectory at a constant memory cost

508, TITLE: Reducing the Computational Cost of Deep Generative Models with Binary Neural Networks
https://openreview.net/forum?id=sTeoJiB4uR
AUTHORS: Thomas Bird, Friso Kingma, David Barber
HIGHLIGHT: We demonstrate that deep generative models can be effectively trained using binary weights and/or activations

509, TITLE: In-N-Out: Pre-Training and Self-Training using Auxiliary Information for Out-of-Distribution Robustness
https://openreview.net/forum?id=jznizqvr15J
AUTHORS: Sang Michael Xie, Ananya Kumar, Robbie Jones, Fereshte Khani, Tengyu Ma, Percy Liang
HIGHLIGHT: Using auxiliary information as inputs hurts OOD, but using auxiliary information by pretraining and self-training improves in-distribution and OOD accuracies on real-world datasets, with theoretical guarantees in a linear multi-task setting.

510, TITLE: Incremental few-shot learning via vector quantization in deep embedded space
https://openreview.net/forum?id=3SV-ZePhnZM
AUTHORS: Kuilin Chen, Chi-Guhn Lee
HIGHLIGHT: In this study, we propose a nonparametric method in deep embedded space to tackle incremental few-shot learning problems.

511, TITLE: Contrastive Syn-to-Real Generalization
https://openreview.net/forum?id=F8whUO8HNbP
AUTHORS: Wuyang Chen, Zhiding Yu, Shalini De Mello, Sifei Liu, Jose M. Alvarez, Zhangyang Wang, Anima Anandkumar
HIGHLIGHT: We propose a new contrastive synthetic-to-real generalization framework that achieves state-of-the-art performance on synthetic-to-real generalization problems.

512, TITLE: Remembering for the Right Reasons: Explanations Reduce Catastrophic Forgetting
https://openreview.net/forum?id=tHgJoMfy6nI
AUTHORS: Sayna Ebrahimi, Suzanne Petryk, Akash Gokul, William Gan, Joseph E. Gonzalez, Marcus Rohrbach, trevor darrell
HIGHLIGHT: Introducing a connection between continual learning and model explainability by regularizing saliency maps to avoid forgetting and showing its effect on memory and regularization-based continual learning approaches.

513, TITLE: Leaky Tiling Activations: A Simple Approach to Learning Sparse Representations Online
https://openreview.net/forum?id=zElset1Klrp
AUTHORS: Yangchen Pan, Kirby Banman, Martha White
HIGHLIGHT: A simple and efficient way to learn sparse feature in deep learning setting.

514, TITLE: High-Capacity Expert Binary Networks
https://openreview.net/forum?id=MxaY4FzOTa
AUTHORS: Adrian Bulat, Brais Martinez, Georgios Tzimiropoulos
HIGHLIGHT: (a) To increase model capacity, we propose Expert Binary Convolution, which, for the first time, tailors conditional computing to binary networks by learning to select one data-specific expert binary filter at a time conditioned on input features.

515, TITLE: Learning What To Do by Simulating the Past
https://openreview.net/forum?id=kBVJ2NtiY-
AUTHORS: David Lindner, Rohin Shah, Pieter Abbeel, Anca Dragan
HIGHLIGHT: Imitating policies given just a single state sampled from a rollout from an expert

516, TITLE: Progressive Skeletonization: Trimming more fat from a network at initialization
https://openreview.net/forum?id=9GsFOUyUPi
AUTHORS: Pau de Jorge, Amartya Sanyal, Harkirat Behl, Philip Torr, Gr?gory Rogez, Puneet K. Dokania
HIGHLIGHT: We find performance of current methods for pruning at initialization plummets at high sparsity levels, we study the possible reasons and present a more robust method overall.

517, TITLE: Filtered Inner Product Projection for Multilingual Embedding Alignment
https://openreview.net/forum?id=A2gNouoXE7
AUTHORS: Vin Sachidananda, Ziyi Yang, Chenguang Zhu
HIGHLIGHT: In this paper, we propose a method, Filtered Inner Product Projection (FIPP), for mapping embeddings to a common representation space.

518, TITLE: Learning Manifold Patch-Based Representations of Man-Made Shapes
https://openreview.net/forum?id=Gu5WqN9J3Fn
AUTHORS: Dmitriy Smirnov, Mikhail Bessmeltsev, Justin Solomon
HIGHLIGHT: We propose a parametrically defined patch-based 3D shape representation that is compatible both with traditional CAD modeling tools and modern deep learning pipelines.

519, TITLE: Aligning AI With Shared Human Values
https://openreview.net/forum?id=dNy_RKzJacY
AUTHORS: Dan Hendrycks, Collin Burns, Steven Basart, Andrew Critch, Jerry Li, Dawn Song, Jacob Steinhardt
HIGHLIGHT: We approach a longstanding problem in machine ethics provide evidence that it is soluble.

520, TITLE: Kanerva++: Extending the Kanerva Machine With Differentiable, Locally Block Allocated Latent Memory
https://openreview.net/forum?id=QoWatN-b8T
AUTHORS: Jason Ramapuram, Yan Wu, Alexandros Kalousis
HIGHLIGHT: Differentiable block allocated latent memory model for generative modeling.

521, TITLE: Measuring Massive Multitask Language Understanding
https://openreview.net/forum?id=d7KBjmI3GmQ
AUTHORS: Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, Jacob Steinhardt
HIGHLIGHT: We test language models on 57 different multiple-choice tasks.

522, TITLE: Towards Robust Neural Networks via Close-loop Control
https://openreview.net/forum?id=2AL06y9cDE-
AUTHORS: Zhuotong Chen, Qianxiao Li, Zheng Zhang

HIGHLIGHT: We propose a close-loop control framework to improve the robustness of neural networks under various data perturbations.

523, TITLE: Statistical inference for individual fairness
https://openreview.net/forum?id=z9k8BWL-_2u
AUTHORS: Subha Maity, Songkai Xue, Mikhail Yurochkin, Yuekai Sun
HIGHLIGHT: In this paper, we focus on the problem of detecting violations of individual fairness in ML models.

524, TITLE: HyperGrid Transformers: Towards A Single Model for Multiple Tasks
https://openreview.net/forum?id=hiq1rHO8pNT
AUTHORS: Yi Tay, Zhe Zhao, Dara Bahri, Donald Metzler, Da-Cheng Juan
HIGHLIGHT: State-of-the-art multi-task NLU performance with only a single model

525, TITLE: Greedy-GQ with Variance Reduction: Finite-time Analysis and Improved Complexity
https://openreview.net/forum?id=6t_dLShIUyZ
AUTHORS: Shaocong Ma, Ziyi Chen, Yi Zhou, Shaofeng Zou
HIGHLIGHT: In this paper, we propose a variance-reduced Greedy-GQ (VR-Greedy-GQ) algorithm for off-policy optimal control.

526, TITLE: On InstaHide, Phase Retrieval, and Sparse Matrix Factorization
https://openreview.net/forum?id=AhElGnhU2BV
AUTHORS: Sitan Chen, Xiaoxiao Li, Zhao Song, Danyang Zhuo
HIGHLIGHT: We examine the security of InstaHide, a recently proposed framework for private distributed learning, through the lens of phase retrieval and give an attack when the underlying datasets are Gaussian.

527, TITLE: VA-RED2: Video Adaptive Redundancy Reduction
https://openreview.net/forum?id=g21u6nlbPzn
AUTHORS: Bowen Pan, Rameswar Panda, Camilo Luciano Fosco, Chung-Ching Lin, Alex J Andonian, Yue Meng, Kate Saenko, Aude Oliva, Rogerio Feris
HIGHLIGHT: Here we present a redundancy reduction framework, termed VA-RED2, which is {\em input-dependent}.

528, TITLE: SEDONA: Search for Decoupled Neural Networks toward Greedy Block-wise Learning
https://openreview.net/forum?id=XLfdzwNKzch
AUTHORS: Myeongjang Pyeon, Jihwan Moon, Taeyoung Hahn, Gunhee Kim
HIGHLIGHT: Our approach is the first attempt to automate decoupling neural networks for greedy block-wise learning and outperforms both end-to-end backprop and state-of-the-art greedy-learning methods on CIFAR-10, Tiny-ImageNet and ImageNet classification.

529, TITLE: ALFWorld: Aligning Text and Embodied Environments for Interactive Learning
https://openreview.net/forum?id=0IOX0YcCdTn
AUTHORS: Mohit Shridhar, Xingdi Yuan, Marc-Alexandre Cote, Yonatan Bisk, Adam Trischler, Matthew Hausknecht
HIGHLIGHT: Pre-training in text-based environments allows agents to learn priors and policies that help solve embodied tasks.

530, TITLE: Learning Task Decomposition with Order-Memory Policy Network
https://openreview.net/forum?id=vcopnwZ7bC
AUTHORS: Yuchen Lu, Yikang Shen, Siyuan Zhou, Aaron Courville, Joshua B. Tenenbaum, Chuang Gan
HIGHLIGHT: We introduce an Ordered Memory Policy Network (OMPN) to discover task decomposition by imitation learning from demonstration.

531, TITLE: Adversarially-Trained Deep Nets Transfer Better
https://openreview.net/forum?id=ijJZbomCJIm
AUTHORS: Francisco Utrera, Evan Kravitz, N. Benjamin Erichson, Rajiv Khanna, Michael W. Mahoney
HIGHLIGHT: We demonstrate that adversarially-trained models transfer better to new domains than naturally-trained models, especially when only limited training data is available in the target domain.

532, TITLE: UMEC: Unified model and embedding compression for efficient recommendation systems
https://openreview.net/forum?id=BM---bH_RSh
AUTHORS: Jiayi Shen, Haotao Wang, Shupeng Gui, Jianchao Tan, Zhangyang Wang, Ji Liu

HIGHLIGHT: We propose a unified model and embedding compression (UMEC) framework to hammer an efficient neural network-based recommendation system.

533, TITLE: Exploring Balanced Feature Spaces for Representation Learning
https://openreview.net/forum?id=OqtLIabPTit
AUTHORS: Bingyi Kang, Yu Li, Sa Xie, Zehuan Yuan, Jiashi Feng
HIGHLIGHT: Inspired by these insights, we develop a novel representation learning method, called k-positive contrastive learning.

534, TITLE: Calibration of Neural Networks using Splines
https://openreview.net/forum?id=eQe8DEWNN2W
AUTHORS: Kartik Gupta, Amir Rahimi, Thalaiyasingam Ajanthan, Thomas Mensink, Cristian Sminchisescu, Richard Hartley
HIGHLIGHT: We introduce a binning-free calibration measure inspired by the classical Kolmogorov-Smirnov statistical test and obtain a recalibration function by approximating the empirical cumulative distribution using a differentiable function via splines.

535, TITLE: Improving relational regularized autoencoders with spherical sliced fused Gromov Wasserstein
https://openreview.net/forum?id=DiQD7FWL233
AUTHORS: Khai Nguyen, Son Nguyen, Nhat Ho, Tung Pham, Hung Bui
HIGHLIGHT: Improving relational regularized autoencoder by introducing new sliced optimal transport discrepancies between the prior and aggregated posterior distributions.

536, TITLE: Rethinking Positional Encoding in Language Pre-training
https://openreview.net/forum?id=09-528y2Fgf
AUTHORS: Guolin Ke, Di He, Tie-Yan Liu
HIGHLIGHT: A novel and better positional encoding method for Transformer-based language pre-training models.

537, TITLE: Discovering Autoregressive Orderings with Variational Inference
https://openreview.net/forum?id=jP1vTH3inC
AUTHORS: Xuanlin Li, Brandon Trabucco, Dong Huk Park, Yang Gao, Michael Luo, Sheng Shen, Trevor Darrell
HIGHLIGHT: The paper proposes an unsupervised learner that discovers high-quality autoregressive orders for language generation without domain-specific prior.

538, TITLE: Differentiable Trust Region Layers for Deep Reinforcement Learning
https://openreview.net/forum?id=qYZD-AO1Vn
AUTHORS: Fabian Otto, Philipp Becker, Vien Anh Ngo, Hanna Carolin Maria Ziesche, Gerhard Neumann
HIGHLIGHT: In this work, we propose differentiable neural network layers to enforce trust regions for deep Gaussian policies via closed-form projections.

539, TITLE: SaliencyMix: A Saliency Guided Data Augmentation Strategy for Better Regularization
https://openreview.net/forum?id=-M0QkvBGTTq
AUTHORS: A F M Shahab Uddin, Mst. Sirazam Monira, Wheemyung Shin, TaeChoong Chung, Sung-Ho Bae
HIGHLIGHT: The proposed method carefully selects a representative image patch with the help of a saliency map and mixes this indicative patch with the target image that leads the model to learn more appropriate feature representation

540, TITLE: Task-Agnostic Morphology Evolution
https://openreview.net/forum?id=CGQ6ENUMX6
AUTHORS: Donald Joseph Hejna III, Pieter Abbeel, Lerrel Pinto
HIGHLIGHT: We introduce TAME, a novel method for optimizing agent morphology using only randomly sampled action primitives and no task driven reward signals.

541, TITLE: Learning Associative Inference Using Fast Weight Memory
https://openreview.net/forum?id=TuK6agbdt27
AUTHORS: Imanol Schlag, Tsendsuren Munkhdalai, J?rgen Schmidhuber
HIGHLIGHT: We present a Recurrent Neural Network model which is augmented with an associative memory to generalise in a more systematically

542, TITLE: Boost then Convolve: Gradient Boosting Meets Graph Neural Networks
https://openreview.net/forum?id=ebS5NUfoMKL
AUTHORS: Sergei Ivanov, Liudmila Prokhorenkova

HIGHLIGHT: A novel strong architecture that combines advantages of GBDT and GNN for node-level prediction problems on graphs with tabular data.

543, TITLE: Degree-Quant: Quantization-Aware Training for Graph Neural Networks
https://openreview.net/forum?id=NSBrFgJAHg
AUTHORS: Shyam Anil Tailor, Javier Fernandez-Marques, Nicholas Donald Lane
HIGHLIGHT: We provide a training technique that enables graph neural networks to use low precision integer arithmetic at inference time, yielding up to 4.7x latency improvements on CPU

544, TITLE: Network Pruning That Matters: A Case Study on Retraining Variants
https://openreview.net/forum?id=Cb54AMqHQFP
AUTHORS: Duong Hoang Le, Binh-Son Hua
HIGHLIGHT: We study the effective of different retraining mechanisms while doing pruning

545, TITLE: Auto Seg-Loss: Searching Metric Surrogates for Semantic Segmentation
https://openreview.net/forum?id=MJAqnaC2vO1
AUTHORS: Hao Li, Chenxin Tao, Xizhou Zhu, Xiaogang Wang, Gao Huang, Jifeng Dai
HIGHLIGHT: Auto Seg-Loss is the first general framework for searching surrogate losses for mainstream semantic segmentation metrics.

546, TITLE: Differentiable Segmentation of Sequences
https://openreview.net/forum?id=4T489T4yav
AUTHORS: Erik Scharw?chter, Jonathan Lennartz, Emmanuel M?ller
HIGHLIGHT: We propose an architecture for effective gradient-based learning of segmented models for sequential data.

547, TITLE: Grounding Physical Object and Event Concepts Through Dynamic Visual Reasoning
https://openreview.net/forum?id=bhCDO_cEGCz
AUTHORS: Zhenfang Chen, Jiayuan Mao, Jiajun Wu, Kwan-Yee Kenneth Wong, Joshua B. Tenenbaum, Chuang Gan
HIGHLIGHT: We propose a neural-symbolic framework to learn physical object and event concepts for video causal reasoning.

548, TITLE: Learning Deep Features in Instrumental Variable Regression
https://openreview.net/forum?id=sy4Kg_ZQmS7
AUTHORS: Liyuan Xu, Yutian Chen, Siddarth Srinivasan, Nando de Freitas, Arnaud Doucet, Arthur Gretton
HIGHLIGHT: Propose a novel deep learning based method for instrumental variable regression

549, TITLE: Online Adversarial Purification based on Self-supervised Learning
https://openreview.net/forum?id=_i3ASPp12WS
AUTHORS: Changhao Shi, Chester Holtz, Gal Mishne
HIGHLIGHT: In this paper, we combine canonical supervised learning with self-supervised representation learning, and present Self-supervised Online Adversarial Purification (SOAP), a novel defense strategy that uses a self-supervised loss to purify adversarial examples at test-time.

550, TITLE: Graph Information Bottleneck for Subgraph Recognition
https://openreview.net/forum?id=bM4Iqfg8M2k
AUTHORS: Junchi Yu, Tingyang Xu, Yu Rong, Yatao Bian, Junzhou Huang, Ran He
HIGHLIGHT: In this paper, we propose a framework of Graph Information Bottleneck (GIB) for the subgraph recognition problem in deep graph learning.

551, TITLE: In Search of Lost Domain Generalization
https://openreview.net/forum?id=lQdXeXDoWtI
AUTHORS: Ishaan Gulrajani, David Lopez-Paz
HIGHLIGHT: Our ERM baseline achieves state-of-the-art performance across many domain generalization benchmarks

552, TITLE: Robust Curriculum Learning: from clean label detection to noisy label self-correction
https://openreview.net/forum?id=lmTWnm3coJJ
AUTHORS: Tianyi Zhou, Shengjie Wang, Jeff Bilmes
HIGHLIGHT: RoCL improves noisy label learning by periodical transitions from supervised learning of clean labeled data to self-supervision of wrongly-labeled data, where the data are selected according to training dynamics.

553, TITLE: Cross-Attentional Audio-Visual Fusion for Weakly-Supervised Action Localization
https://openreview.net/forum?id=hWr3e3r-oH5
AUTHORS: Jun-Tae Lee, Mihir Jain, Hyoungwoo Park, Sungrack Yun
HIGHLIGHT: An audio-visual fusion technique, called "multi-stage cross-attention", is developed to exploit the multi-modal representation in weakly-supervised action or event localization in untrimmed videos.

554, TITLE: CoCon: A Self-Supervised Approach for Controlled Text Generation
https://openreview.net/forum?id=VD_ozqvBy4W
AUTHORS: Alvin Chan, Yew-Soon Ong, Bill Pung, Aston Zhang, Jie Fu
HIGHLIGHT: We propose CoCon to control the content of text generation from LMs by conditioning on content inputs at an interleave layer.

555, TITLE: Group Equivariant Generative Adversarial Networks
https://openreview.net/forum?id=rgFNuJHHXv
AUTHORS: Neel Dey, Antong Chen, Soheil Ghafurian
HIGHLIGHT: Equivariance to symmetry groups improves the generative adversarial synthesis of globally-symmetric images.

556, TITLE: What they do when in doubt: a study of inductive biases in seq2seq learners
https://openreview.net/forum?id=YmA86Zo-P_t
AUTHORS: Eugene Kharitonov, Rahma Chaabouni
HIGHLIGHT: Standard seq2seq models can infer perfectly different rules when presented with as little as one training example, showing strikingly different inductive biases than can be studied via description length

557, TITLE: A teacher-student framework to distill future trajectories
https://openreview.net/forum?id=ECuvULjFQia
AUTHORS: Alexander Neitz, Giambattista Parascandolo, Bernhard Sch?lkopf
HIGHLIGHT: We explore meta-learning a teacher network to efficiently incorporate privileged information such as trajectories.

558, TITLE: Learning a Latent Search Space for Routing Problems using Variational Autoencoders
https://openreview.net/forum?id=90JprVrJBO
AUTHORS: Andr? Hottung, Bhanu Bhandari, Kevin Tierney
HIGHLIGHT: We use conditional variational autoencoders to learn continuous search spaces for routing problems that can be searched with any unconstrained continuous optimizer.

559, TITLE: Universal approximation power of deep residual neural networks via nonlinear control theory
https://openreview.net/forum?id=-IXhmY16R3M
AUTHORS: Paulo Tabuada, Bahman Gharesifard
HIGHLIGHT: Using nonlinear control theory, it is shown that deep residual neural networks have the power of universal approximation with respect to the supremum norm.

560, TITLE: On the Universality of Rotation Equivariant Point Cloud Networks
https://openreview.net/forum?id=6NFBvWlRXaG
AUTHORS: Nadav Dym, Haggai Maron
HIGHLIGHT: We provide sufficient conditions for universality of rotation equivariant point cloud networks and use these conditions to show that current models are universal as well as for devising new universal architectures.

561, TITLE: CT-Net: Channel Tensorization Network for Video Classification
https://openreview.net/forum?id=UoaQUQREMOs
AUTHORS: Kunchang Li, Xianhang Li, Yali Wang, Jun Wang, Yu Qiao
HIGHLIGHT: To achieve convolution efficiency and feature-interaction sufficiency, we propose a Channel Tensorization Network (CT-Net), by treating the channel dimension of input feature as a multiplication of K sub-dimensions.

562, TITLE: Learning to live with Dale's principle: ANNs with separate excitatory and inhibitory units
https://openreview.net/forum?id=eU776ZYxEpz
AUTHORS: Jonathan Cornford, Damjan Kalajdzievski, Marco Leite, Am?lie Lamarquette, Dimitri Michael Kullmann, Blake Aaron Richards
HIGHLIGHT: We present two insights that enable DANNs to learn well: (1) DANNs are related to normalization schemes, and can be initialized such that the inhibition centres and standardizes the excitatory activity, (2) updates to inhibitory neuron parameters should be scaled using corrections based on the Fisher Information matrix.

563, TITLE: Uncertainty Estimation in Autoregressive Structured Prediction
https://openreview.net/forum?id=jN5y-zb5Q7m
AUTHORS: Andrey Malinin, Mark Gales
HIGHLIGHT: A Deep Investigation of Ensemble-based Uncertainty Estimation for Autoregressive ASR and NMT models.

564, TITLE: Transformer protein language models are unsupervised structure learners
https://openreview.net/forum?id=fylclEqgvgd
AUTHORS: Roshan Rao, Joshua Meier, Tom Sercu, Sergey Ovchinnikov, Alexander Rives
HIGHLIGHT: Transformer attention maps directly represent protein contacts with state-of-the-art unsupervised precision.

565, TITLE: ANOCE: Analysis of Causal Effects with Multiple Mediators via Constrained Structural Learning
https://openreview.net/forum?id=7I12hXRi8F
AUTHORS: Hengrui Cai, Rui Song, Wenbin Lu
HIGHLIGHT: Analysis of causal effects on the level of individual mediators via constrained structural learning, with application to the COVID-19 Spread in China.

566, TITLE: Plan-Based Relaxed Reward Shaping for Goal-Directed Tasks
https://openreview.net/forum?id=w2Z2OwVNeK
AUTHORS: Ingmar Schubert, Ozgur S Oguz, Marc Toussaint
HIGHLIGHT: We introduce Final-Volume-Preserving Reward Shaping, and show in a plan-based setting that it significantly increases the sample efficiency of reinforcement learning.

567, TITLE: CcGAN: Continuous Conditional Generative Adversarial Networks for Image Generation
https://openreview.net/forum?id=PrzjugOsDeE
AUTHORS: Xin Ding, Yongwei Wang, Zuheng Xu, William J Welch, Z. Jane Wang
HIGHLIGHT: This work proposes the continuous conditional generative adversarial network (CcGAN), the first generative model for image generation conditional on continuous, scalar conditions (termed as regression labels).

568, TITLE: Single-Photon Image Classification
https://openreview.net/forum?id=CHLhSw9pSw8
AUTHORS: Thomas Fischbacher, Luciano Sbaiz
HIGHLIGHT: Mathematical proof that the classical accuracy limit for single-photon image classification can be exceeded very substantially by employing a problem-tailored quantum transformation on the photon state.

569, TITLE: Self-supervised Adversarial Robustness for the Low-label, High-data Regime
https://openreview.net/forum?id=bgQek2O63w
AUTHORS: Sven Gowal, Po-Sen Huang, Aaron van den Oord, Timothy Mann, Pushmeet Kohli
HIGHLIGHT: This paper builds of Bootstrap Your Own Latents and proposes a self-supervised learning technique that can learn robust representations that are competitive with fully-supervised techniques.

570, TITLE: Uncertainty-aware Active Learning for Optimal Bayesian Classifier
https://openreview.net/forum?id=Mu2ZxFctAI
AUTHORS: Guang Zhao, Edward Dougherty, Byung-Jun Yoon, Francis Alexander, Xiaoning Qian
HIGHLIGHT: We focus on pool-based Bayesian active learning, for which we have proposed a new weighted MOCU method and analyzed its theoretical properties to demonstrate the better convergence properties than existing methods in this category.

571, TITLE: Skill Transfer via Partially Amortized Hierarchical Planning
https://openreview.net/forum?id=jXe91kq3jAq
AUTHORS: Kevin Xie, Homanga Bharadhwaj, Danijar Hafner, Animesh Garg, Florian Shkurti
HIGHLIGHT: Partially amortized planning through hierarchy helps learn skills for complex control tasks

572, TITLE: Learning continuous-time PDEs from sparse data with graph neural networks
https://openreview.net/forum?id=aUX5Plaq7Oy
AUTHORS: Valerii Iakovlev, Markus Heinonen, Harri L?hdesm?ki
HIGHLIGHT: The paper introduces a method for learning partial differential equations on arbitrary spatial and temporal grids.

573, TITLE: Characterizing signal propagation to close the performance gap in unnormalized ResNets
https://openreview.net/forum?id=IX3Nnir2omJ

AUTHORS: Andrew Brock, Soham De, Samuel L Smith
HIGHLIGHT: We show how to train ResNets completely without normalization, and attain performance competitive with batch-normalized EfficientNets.

574, TITLE: Robust Overfitting may be mitigated by properly learned smoothening
https://openreview.net/forum?id=qZzy5urZw9
AUTHORS: Tianlong Chen, Zhenyu Zhang, Sijia Liu, Shiyu Chang, Zhangyang Wang
HIGHLIGHT: Mitigate robust overfitting by properly learned smoothening, establishing the new state-of-the-art bar in adversarial training

575, TITLE: Long Live the Lottery: The Existence of Winning Tickets in Lifelong Learning
https://openreview.net/forum?id=LXMSvPmsm0g
AUTHORS: Tianlong Chen, Zhenyu Zhang, Sijia Liu, Shiyu Chang, Zhangyang Wang
HIGHLIGHT: Proposed novel bottom-up lifelong pruning effectively identify the winning tickets, which significantly improve the performance of learning over continual tasks

576, TITLE: Symmetry-Aware Actor-Critic for 3D Molecular Design
https://openreview.net/forum?id=jEYKjPE1xYN
AUTHORS: Gregor N. C. Simm, Robert Pinsler, G?bor Cs?nyi, Jos? Miguel Hern?ndez-Lobato
HIGHLIGHT: Covariant actor-critic based on spherical harmonics that exploits symmetries to design molecules in 3D

577, TITLE: PseudoSeg: Designing Pseudo Labels for Semantic Segmentation
https://openreview.net/forum?id=-TwO99rbVRu
AUTHORS: Yuliang Zou, Zizhao Zhang, Han Zhang, Chun-Liang Li, Xiao Bian, Jia-Bin Huang, Tomas Pfister
HIGHLIGHT: This paper presents a new method that first demonstrates how well-calibrated soft pseudo labels obtained through wise fusion of predictions from diverse sources greatly improve consistency training for semantic segmentation.

578, TITLE: NAS-Bench-ASR: Reproducible Neural Architecture Search for Speech Recognition
https://openreview.net/forum?id=CU0APx9LMaL
AUTHORS: Abhinav Mehrotra, Alberto Gil C. P. Ramos, Sourav Bhattacharya, Lukasz Dudziak, Ravichander Vipperla, Thomas Chau, Mohamed S Abdelfattah, Samin Ishtiaq, Nicholas Donald Lane
HIGHLIGHT: In this work we apply NAS for finding cell architecture for ASR models and release a comprehensive NAS-Bench dataset for reproducible NAS research.

579, TITLE: Scaling the Convex Barrier with Active Sets
https://openreview.net/forum?id=uQfOy7LrlTR
AUTHORS: Alessandro De Palma, Harkirat Behl, Rudy R Bunel, Philip Torr, M. Pawan Kumar
HIGHLIGHT: We present a specialised dual solver for a tight ReLU convex relaxation and show that it speeds up formal network verification.

580, TITLE: Local Convergence Analysis of Gradient Descent Ascent with Finite Timescale Separation
https://openreview.net/forum?id=AWOSz_mMAPx
AUTHORS: Tanner Fiez, Lillian J Ratliff
HIGHLIGHT: We show that there exists a range of finite learning ratios such that gradient descent-ascent converges to a critical point if and only if it is a strict local minmax equilibrium

581, TITLE: Activation-level uncertainty in deep neural networks
https://openreview.net/forum?id=UvBPbpvHRj-
AUTHORS: Pablo Morales-Alvarez, Daniel Hern?ndez-Lobato, Rafael Molina, Jos? Miguel Hern?ndez-Lobato
HIGHLIGHT: We use 1D Gaussian Processes to introduce activation-level uncertainty in neural networks, which overcomes known limitations of (functional) Bayesian neural nets and obtains better results than the related deep Gaussian Processes.

582, TITLE: Efficient Continual Learning with Modular Networks and Task-Driven Priors
https://openreview.net/forum?id=EKV158tSfwv
AUTHORS: Tom Veniat, Ludovic Denoyer, MarcAurelio Ranzato
HIGHLIGHT: We propose a new benchmark allowing a detailed analysis of the properties of continual learning alogithms and a new modular neural network leveraging task-based priors to efficiently learn in the CL setting.

583, TITLE: Amending Mistakes Post-hoc in Deep Networks by Leveraging Class Hierarchies
https://openreview.net/forum?id=193sEnKY1ij

AUTHORS: Shyamgopal Karthik, Ameya Prabhu, Puneet K. Dokania, Vineet Gandhi
HIGHLIGHT: Conditional risk framework exploiting the label hierarchy outperforms state of the art and makes a strong baseline for future explorations.

584, TITLE: Ringing ReLUs: Harmonic Distortion Analysis of Nonlinear Feedforward Networks
https://openreview.net/forum?id=TaYhv-q1Xit
AUTHORS: Christian H.X. Ali Mehmeti-G?pel, David Hartmann, Michael Wand
HIGHLIGHT: Nonlinearities create high-frequency distortions that affect network trainability.

585, TITLE: Distance-Based Regularisation of Deep Networks for Fine-Tuning
https://openreview.net/forum?id=IFqrg1p5Bc
AUTHORS: Henry Gouk, Timothy Hospedales, massimiliano pontil
HIGHLIGHT: We derive generalisation bounds applicable to fine-tuning, then demonstrate an algorithm that regularises these bounds improves fine-tuning performance.

586, TITLE: Genetic Soft Updates for Policy Evolution in Deep Reinforcement Learning
https://openreview.net/forum?id=TGFO0DbD_pk
AUTHORS: Enrico Marchesini, Davide Corsi, Alessandro Farinelli
HIGHLIGHT: We present a novel mixed framework that combines the benefits of Evolutionary Algorithms and any DRL algorithms (including value-based ones); we support our claims on the beneficial policy improvement using recent formal verification tools.

587, TITLE: Towards Faster and Stabilized GAN Training for High-fidelity Few-shot Image Synthesis
https://openreview.net/forum?id=1Fqg133qRaI
AUTHORS: Bingchen Liu, Yizhe Zhu, Kunpeng Song, Ahmed Elgammal
HIGHLIGHT: A computational-efficient GAN for few-shot hi-fi image dataset (converge on single gpu with few hours' training, on 1024 resolution sub-hundred images).

588, TITLE: IsarStep: a Benchmark for High-level Mathematical Reasoning
https://openreview.net/forum?id=Pzj6fzU6wkj
AUTHORS: Wenda Li, Lei Yu, Yuhuai Wu, Lawrence C. Paulson
HIGHLIGHT: We present a benchmark for high-level mathematical reasoning and study the reasoning capabilities of neural sequence-to-sequence models.
We build a non-synthetic dataset from the largest repository of proofs written by human experts in a theorem prover.

589, TITLE: Multi-Prize Lottery Ticket Hypothesis: Finding Accurate Binary Neural Networks by Pruning A Randomly Weighted Network
https://openreview.net/forum?id=U_mat0b9iv
AUTHORS: James Diffenderfer, Bhavya Kailkhura
HIGHLIGHT: A new paradigm for learning compact yet accurate binary neural networks by pruning and quantizing randomly weighted full precision DNNs

590, TITLE: Average-case Acceleration for Bilinear Games and Normal Matrices
https://openreview.net/forum?id=H0syOoy3Ash
AUTHORS: Carles Domingo-Enrich, Fabian Pedregosa, Damien Scieur
HIGHLIGHT: We extend the framework of average-case optimal first-order methods to problems with non-symmetric matrices, which naturally arise in equilibrium finding for games.

591, TITLE: ECONOMIC HYPERPARAMETER OPTIMIZATION WITH BLENDED SEARCH STRATEGY
https://openreview.net/forum?id=VbLH04pRA3
AUTHORS: Chi Wang, Qingyun Wu, Silu Huang, Amin Saied
HIGHLIGHT: A low-cost hyperparameter optimization solution by blending global and local search methods with cost-based prioritization

592, TITLE: BSQ: Exploring Bit-Level Sparsity for Mixed-Precision Neural Network Quantization
https://openreview.net/forum?id=TiXl51SCNw8
AUTHORS: Huanrui Yang, Lin Duan, Yiran Chen, Hai Li
HIGHLIGHT: We propose bit-level sparsity inducing regularizer to induce mixed-percision quantization scheme in DNN with gradient-based training.

593, TITLE: AutoLRS: Automatic Learning-Rate Schedule by Bayesian Optimization on the Fly
https://openreview.net/forum?id=SlrqM9_lyju
AUTHORS: Yuchen Jin, Tianyi Zhou, Liangyu Zhao, Yibo Zhu, Chuanxiong Guo, Marco Canini, Arvind Krishnamurthy
HIGHLIGHT: In this paper, we consider the question: Can we automatically tune the LR over the course of training without human involvement?

594, TITLE: BERTology Meets Biology: Interpreting Attention in Protein Language Models
https://openreview.net/forum?id=YWtLZvLmud7
AUTHORS: Jesse Vig, Ali Madani, Lav R. Varshney, Caiming Xiong, richard socher, Nazneen Rajani
HIGHLIGHT: We analyze the internal representations of protein language models, and show that attention targets structural and functional properties of protein sequences.

595, TITLE: Learning Task-General Representations with Generative Neuro-Symbolic Modeling
https://openreview.net/forum?id=qzBUIzq5XR2
AUTHORS: Reuben Feinman, Brenden M. Lake
HIGHLIGHT: few-shot concept learning with generative neuro-symbolic models

596, TITLE: Zero-shot Synthesis with Group-Supervised Learning
https://openreview.net/forum?id=8wqCDnBmnrT
AUTHORS: Yunhao Ge, Sami Abu-El-Haija, Gan Xin, Laurent Itti
HIGHLIGHT: To aid neural networks to envision objects with different attributes, we propose GSL which allows us to decompose inputs into a disentangled representation with swappable components, that can be recombined to synthesize new samples.

597, TITLE: Selective Classification Can Magnify Disparities Across Groups
https://openreview.net/forum?id=N0M_4BkQ05i
AUTHORS: Erik Jones, Shiori Sagawa, Pang Wei Koh, Ananya Kumar, Percy Liang
HIGHLIGHT: In this paper, we find that while selective classification can improve average accuracies, it can simultaneously magnify existing accuracy disparities between various groups within a population, especially in the presence of spurious correlations.

598, TITLE: Better Fine-Tuning by Reducing Representational Collapse
https://openreview.net/forum?id=OQ08SN70M1V
AUTHORS: Armen Aghajanyan, Akshat Shrivastava, Anchit Gupta, Naman Goyal, Luke Zettlemoyer, Sonal Gupta
HIGHLIGHT: We present a lightweight augmentation to standard fine-tuning which outperforms previous methods across the board (i.e. SOTA on 3 summarization tasks, XNLI, RoBERTa on GLUE) while being computationally cheaper than other fine-tuning approaches.

599, TITLE: Training independent subnetworks for robust prediction
https://openreview.net/forum?id=OGg9XnKxFAH
AUTHORS: Marton Havasi, Rodolphe Jenatton, Stanislav Fort, Jeremiah Zhe Liu, Jasper Snoek, Balaji Lakshminarayanan, Andrew Mingbo Dai, Dustin Tran
HIGHLIGHT: We show that a deep neural network can be trained to give multiple independent predictions simultaneously, which results in a computationally efficient ensemble model.

600, TITLE: Meta-Learning of Structured Task Distributions in Humans and Machines
https://openreview.net/forum?id=--gvHfE3Xf5
AUTHORS: Sreejan Kumar, Ishita Dasgupta, Jonathan Cohen, Nathaniel Daw, Thomas Griffiths
HIGHLIGHT: We developed a novel meta-learning task with a structured task distribution and statistically equivalent "null" task distribution to show humans are more adept at the former whereas current meta-learning agents are more adept at the latter.

601, TITLE: BiPointNet: Binary Neural Network for Point Clouds
https://openreview.net/forum?id=9QLRCVysdlO
AUTHORS: Haotong Qin, Zhongang Cai, Mingyuan Zhang, Yifu Ding, Haiyu Zhao, Shuai Yi, Xianglong Liu, Hao Su
HIGHLIGHT: We present BiPointNet, the first model binarization approach to efficient deep learning on point clouds, targeting at extreme compression and acceleration.

602, TITLE: Benchmarks for Deep Off-Policy Evaluation
https://openreview.net/forum?id=kWSeGEeHvF8
AUTHORS: Justin Fu, Mohammad Norouzi, Ofir Nachum, George Tucker, ziyu wang, Alexander Novikov, Mengjiao Yang, Michael R Zhang, Yutian Chen, Aviral Kumar, Cosmin Paduraru, Sergey Levine, Thomas Paine
HIGHLIGHT: A benchmark proposal for off-policy evaluation and policy selection.

603, TITLE: Planning from Pixels using Inverse Dynamics Models
https://openreview.net/forum?id=V6BjBgku7Ro
AUTHORS: Keiran Paster, Sheila A. McIlraith, Jimmy Ba
HIGHLIGHT: GLAMOR learns a latent world model by learning to predict action sequences conditioned on task completion.

604, TITLE: Understanding the effects of data parallelism and sparsity on neural network training
https://openreview.net/forum?id=rsogjAnYs4z
AUTHORS: Namhoon Lee, Thalaiyasingam Ajanthan, Philip Torr, Martin Jaggi
HIGHLIGHT: We accurately measure the effects of data parallelism and sparsity on neural network training and develop a theoretical analysis to precisely account for their effects.

605, TITLE: NOVAS: Non-convex Optimization via Adaptive Stochastic Search for End-to-end Learning and Control
https://openreview.net/forum?id=Iw4ZGwenbXf
AUTHORS: Ioannis Exarchos, Marcus Aloysius Pereira, Ziyi Wang, Evangelos Theodorou
HIGHLIGHT: In this work we propose the use of adaptive stochastic search as a building block for general, non-convex optimization operations within deep neural network architectures.

606, TITLE: MoVie: Revisiting Modulated Convolutions for Visual Counting and Beyond
https://openreview.net/forum?id=8e6BrwU6AjQ
AUTHORS: Duy Kien Nguyen, Vedanuj Goswami, Xinlei Chen
HIGHLIGHT: 2020 VQA challenge winner; state-of-the-art performance on three counting benchmarks; can work beyond counting towards general visual reasoning

607, TITLE: NeMo: Neural Mesh Models of Contrastive Features for Robust 3D Pose Estimation
https://openreview.net/forum?id=pmj131uIL9H
AUTHORS: Angtian Wang, Adam Kortylewski, Alan Yuille
HIGHLIGHT: We introduce NeMo, a rendering-based approach to 3D pose estimation that models objects in terms of neural feature activations, instead of image intensities.

608, TITLE: On Graph Neural Networks versus Graph-Augmented MLPs
https://openreview.net/forum?id=tiqI7w64JG2
AUTHORS: Zhengdao Chen, Lei Chen, Joan Bruna
HIGHLIGHT: We establish a separation in representation power between GNNs and Graph-Augmented MLPs.

609, TITLE: Dual-mode ASR: Unify and Improve Streaming ASR with Full-context Modeling
https://openreview.net/forum?id=Pz_dcqfcKW8
AUTHORS: Jiahui Yu, Wei Han, Anmol Gulati, Chung-Cheng Chiu, Bo Li, Tara N Sainath, Yonghui Wu, Ruoming Pang
HIGHLIGHT: Dual-mode ASR unifies and improves Streaming ASR with full-context modeling, simplifying the development and deployment workflow and improving both latency and accuracy.

610, TITLE: Deep Learning meets Projective Clustering
https://openreview.net/forum?id=EQfpYwF3-b
AUTHORS: Alaa Maalouf, Harry Lang, Daniela Rus, Dan Feldman
HIGHLIGHT: We suggest a novel technique for compressing a fully connected layer (or an embedding layer).

611, TITLE: Reinforcement Learning with Random Delays
https://openreview.net/forum?id=QFYnKlBJYR
AUTHORS: Yann Bouteiller, Simon Ramstedt, Giovanni Beltrame, Christopher Pal, Jonathan Binas
HIGHLIGHT: We propose a framework for Reinforcement Learning with random action and observation delays.

612, TITLE: Isotropy in the Contextual Embedding Space: Clusters and Manifolds
https://openreview.net/forum?id=xYGNO86OWDH
AUTHORS: Xingyu Cai, Jiaji Huang, Yuchen Bian, Kenneth Church
HIGHLIGHT: This paper reveals isotropy in the clustered contextual embedding space, and found low-dimensional manifolds in there.

613, TITLE: Spatio-Temporal Graph Scattering Transform
https://openreview.net/forum?id=CF-ZIuSMXRz
AUTHORS: Chao Pan, Siheng Chen, Antonio Ortega

HIGHLIGHT: We put forth a novel mathematically designed framework "ST-GST" to analyze spatio-temporal data.

614, TITLE: Deployment-Efficient Reinforcement Learning via Model-Based Offline Optimization
https://openreview.net/forum?id=3hGNqpI4WS
AUTHORS: Tatsuya Matsushima, Hiroki Furuta, Yutaka Matsuo, Ofir Nachum, Shixiang Gu
HIGHLIGHT: We propose a novel method that achieves both high sample-efficiency in offline RL and "deployment-efficiency" in online RL.

615, TITLE: gradSim: Differentiable simulation for system identification and visuomotor control
https://openreview.net/forum?id=c_E8kFWfhp0
AUTHORS: J. Krishna Murthy, Miles Macklin, Florian Golemo, Vikram Voleti, Linda Petrini, Martin Weiss, Breandan Considine, J?r?me Parent-L?vesque, Kevin Xie, Kenny Erleben, Liam Paull, Florian Shkurti, Derek Nowrouzezahrai, Sanja Fidler
HIGHLIGHT: Differentiable models of time-varying dynamics and image formation pipelines result in highly accurate physical parameter estimation from video

616, TITLE: Evaluations and Methods for Explanation through Robustness Analysis
https://openreview.net/forum?id=4dXmpCDGNp7
AUTHORS: Cheng-Yu Hsieh, Chih-Kuan Yeh, Xuanqing Liu, Pradeep Kumar Ravikumar, Seungyeon Kim, Sanjiv Kumar, Cho-Jui Hsieh
HIGHLIGHT: We propose a suite of objective measurements for evaluating feature based explanations by the notion of robustness analysis; we further derive new explanation that captures different characteristics of explanation comparing to existing methods.

617, TITLE: RNNLogic: Learning Logic Rules for Reasoning on Knowledge Graphs
https://openreview.net/forum?id=tGZu6DlbreV
AUTHORS: Meng Qu, Junkun Chen, Louis-Pascal Xhonneux, Yoshua Bengio, Jian Tang
HIGHLIGHT: Learn Logic Rules for Reasoning on Knowledge Graphs.

618, TITLE: Can a Fruit Fly Learn Word Embeddings?
https://openreview.net/forum?id=xfmSoxdxFCG
AUTHORS: Yuchen Liang, Chaitanya Ryali, Benjamin Hoover, Saket Navlakha, Leopold Grinberg, Mohammed J Zaki, Dmitry Krotov
HIGHLIGHT: We show that a network motif from fruit fly brain can learn word embeddings.

619, TITLE: Neural representation and generation for RNA secondary structures
https://openreview.net/forum?id=snOgiCYZgJ7
AUTHORS: Zichao Yan, William L. Hamilton, Mathieu Blanchette
HIGHLIGHT: We investigate a new direction in computational drug discovery for designing large scale and complex macromolecular structures known as the RNAs.

620, TITLE: WaNet - Imperceptible Warping-based Backdoor Attack
https://openreview.net/forum?id=eEn8KTtJOx
AUTHORS: Tuan Anh Nguyen, Anh Tuan Tran
HIGHLIGHT: We propose an imperceptible backdoor attack based on image-warping, which can surpass both human and machine inspections.

621, TITLE: LowKey: Leveraging Adversarial Attacks to Protect Social Media Users from Facial Recognition
https://openreview.net/forum?id=hJmtwocEqzc
AUTHORS: Valeriia Cherepanova, Micah Goldblum, Harrison Foley, Shiyuan Duan, John P Dickerson, Gavin Taylor, Tom Goldstein
HIGHLIGHT: We leverage adversarial attacks in our tool, LowKey, which protects social media users from invasive mass surveillance systems.

622, TITLE: Learning from others' mistakes: Avoiding dataset biases without modeling them
https://openreview.net/forum?id=Hf3qXoiNkR
AUTHORS: Victor Sanh, Thomas Wolf, Yonatan Belinkov, Alexander M Rush
HIGHLIGHT: Reducing a model's reliance on dataset biases by encouraging a robust model to learn from a weak learner's mistakes.

623, TITLE: Prototypical Contrastive Learning of Unsupervised Representations

https://openreview.net/forum?id=KmykpuSrjcq
AUTHORS: Junnan Li, Pan Zhou, Caiming Xiong, Steven Hoi
HIGHLIGHT: We propose an unsupervised representation learning method that bridges contrastive learning with clustering in an EM framework.

624, TITLE: Extreme Memorization via Scale of Initialization
https://openreview.net/forum?id=Z4R1vxLbRLO
AUTHORS: Harsh Mehta, Ashok Cutkosky, Behnam Neyshabur
HIGHLIGHT: Through studying the effect of scale of initialization on generalization, we come up with an alignment measure that correlations with generalization of deep models.

625, TITLE: Interactive Weak Supervision: Learning Useful Heuristics for Data Labeling
https://openreview.net/forum?id=lDFQI9OY6K
AUTHORS: Benedikt Boecking, Willie Neiswanger, Eric Xing, Artur Dubrawski
HIGHLIGHT: We introduce a framework and method for training classifiers on datasets without ground truth annotation by interacting with domain experts to discover good weak supervision sources.

626, TITLE: Adaptive Procedural Task Generation for Hard-Exploration Problems
https://openreview.net/forum?id=8xLkv08d70T
AUTHORS: Kuan Fang, Yuke Zhu, Silvio Savarese, Fei-Fei Li
HIGHLIGHT: We propose a framework which creates tasks as curricula via procedural generation to expedite reinforcement learning in hard-exploration problems.

627, TITLE: Multi-timescale Representation Learning in LSTM Language Models
https://openreview.net/forum?id=9ITXiTrAoT
AUTHORS: Shivangi Mahto, Vy Ai Vo, Javier S. Turek, Alexander Huth
HIGHLIGHT: This work presents a theoretically-motivated analysis of memory and timescale in LSTM language models.

628, TITLE: Deep Encoder, Shallow Decoder: Reevaluating Non-autoregressive Machine Translation
https://openreview.net/forum?id=KpfasTaLUpq
AUTHORS: Jungo Kasai, Nikolaos Pappas, Hao Peng, James Cross, Noah Smith
HIGHLIGHT: We show that the speed disadvantage for autoregressive baselines to evaluate non-autoregressive machine translation is overestimated in three aspects: suboptimal layer allocation, insufficient speed measurement, and lack of knowledge distillation.

629, TITLE: Learning Better Structured Representations Using Low-rank Adaptive Label Smoothing
https://openreview.net/forum?id=5NsEIflpbSv
AUTHORS: Asish Ghoshal, Xilun Chen, Sonal Gupta, Luke Zettlemoyer, Yashar Mehdad
HIGHLIGHT: We propose an extension of label smoothing which improves generalization performance by adapting to the structure present in label space of structured prediction tasks.

630, TITLE: Information-theoretic Probing Explains Reliance on Spurious Features
https://openreview.net/forum?id=mNtmhaDkAr
AUTHORS: Charles Lovering, Rohan Jha, Tal Linzen, Ellie Pavlick
HIGHLIGHT: We find that feature extractability, measured by probing classifiers, can be viewed as an inductive bias: the more extractable a feature is after pre-training, the less statistical evidence needed during fine-tuning for the model to use the feature.

631, TITLE: Molecule Optimization by Explainable Evolution
https://openreview.net/forum?id=jHefDGsorp5
AUTHORS: Binghong Chen, Tianzhe Wang, Chengtao Li, Hanjun Dai, Le Song
HIGHLIGHT: We propose a novel EM-like evolution-by-explanation algorithm alternating between an explainable graph model and a conditional generative model for molecule optimization.

632, TITLE: Anchor & Transform: Learning Sparse Embeddings for Large Vocabularies
https://openreview.net/forum?id=Vd7lCMvtLqg
AUTHORS: Paul Pu Liang, Manzil Zaheer, Yuan Wang, Amr Ahmed
HIGHLIGHT: End-to-end learning of sparse embeddings for large vocabularies with a Bayesian nonparametric interpretation that results in up to 40x smaller embedding tables.

633, TITLE: PSTNet: Point Spatio-Temporal Convolution on Point Cloud Sequences

https://openreview.net/forum?id=O3bqkf_Puys
AUTHORS: Hehe Fan, Xin Yu, Yuhang Ding, Yi Yang, Mohan Kankanhalli
HIGHLIGHT: This paper proposes a point spatio-temporal (PST) convolution to learn representations of raw point cloud sequences.

634, TITLE: Group Equivariant Conditional Neural Processes
https://openreview.net/forum?id=e8W-hsu_q5
AUTHORS: Makoto Kawano, Wataru Kumagai, Akiyoshi Sannai, Yusuke Iwasawa, Yutaka Matsuo
HIGHLIGHT: A model for regression that learns conditional distributions of a stochastic process, by incorporating group equivariance into Conditional Neural Processes.

635, TITLE: When does preconditioning help or hurt generalization?
https://openreview.net/forum?id=S724o4_WB3
AUTHORS: Shun-ichi Amari, Jimmy Ba, Roger Baker Grosse, Xuechen Li, Atsushi Nitanda, Taiji Suzuki, Denny Wu, Ji Xu
HIGHLIGHT: Characterized the population risk of preconditioned least squares regression in the overparameterized regime and determined the optimal preconditioner for generalization.

636, TITLE: Learning Reasoning Paths over Semantic Graphs for Video-grounded Dialogues
https://openreview.net/forum?id=hPWj1qduVw8
AUTHORS: Hung Le, Nancy F. Chen, Steven Hoi
HIGHLIGHT: We introduce a novel approach to learn reasoning paths over semantic graphs which are built upon dialogue context at each turn, for video-grounded dialogues.

637, TITLE: Prototypical Representation Learning for Relation Extraction
https://openreview.net/forum?id=aCgLmfhIy_f
AUTHORS: Ning Ding, Xiaobin Wang, Yao Fu, Guangwei Xu, Rui Wang, Pengjun Xie, Ying Shen, Fei Huang, Hai-Tao Zheng, Rui Zhang
HIGHLIGHT: Instead of solely relying on the supervision from labels (which could be noisy), we propose to infer a latent prototype for each relation from contextual information to best explore the intrinsic semantics of relations.

638, TITLE: Layer-adaptive Sparsity for the Magnitude-based Pruning
https://openreview.net/forum?id=H6ATjJ0TKdf
AUTHORS: Jaeho Lee, Sejun Park, Sangwoo Mo, Sungsoo Ahn, Jinwoo Shin
HIGHLIGHT: We propose LAMP, a general-purpose layerwise sparsity selection scheme for magnitude pruning.

639, TITLE: Refining Deep Generative Models via Wasserstein Gradient Flows
https://openreview.net/forum?id=Zbc-ue9p_rE
AUTHORS: Abdul Fatir Ansari, Ming Liang Ang, Harold Soh
HIGHLIGHT: A method of refining samples from deep generative models using the discriminator gradient flow of f-divergences.

640, TITLE: Explaining the Efficacy of Counterfactually Augmented Data
https://openreview.net/forum?id=HHiiQKWsOcV
AUTHORS: Divyansh Kaushik, Amrith Setlur, Eduard H Hovy, Zachary Chase Lipton
HIGHLIGHT: We present a framework for thinking about counterfactually augmented data and make strides towards understanding its benefits in out-of-domain generalization.

641, TITLE: Lipschitz Recurrent Neural Networks
https://openreview.net/forum?id=-N7PBXqOUJZ
AUTHORS: N. Benjamin Erichson, Omri Azencot, Alejandro Queiruga, Liam Hodgkinson, Michael W. Mahoney
HIGHLIGHT: We develop a provably stable parameterization for continuous-time Lipschitz Recurrent Neural Networks that can employ high order integration schemes and outperform existing RNNs in performance, robustness, and conditioning.

642, TITLE: Learning Hyperbolic Representations of Topological Features
https://openreview.net/forum?id=yqPnIRhHtZv
AUTHORS: Panagiotis Kyriakis, Iordanis Fostiropoulos, Paul Bogdan
HIGHLIGHT: We develop a method to learn representations of topological features (i.e., persistence diagrams) on hyperbolic spaces.

643, TITLE: Bidirectional Variational Inference for Non-Autoregressive Text-to-Speech

https://openreview.net/forum?id=o3iritJHLfO
AUTHORS: Yoonhyung Lee, Joongbo Shin, Kyomin Jung
HIGHLIGHT: In this paper, a novel non-autoregressive text-to-speech model based on bidirectional-inference variational autoencoder called BVAE-TTS is proposed.

644, TITLE: Risk-Averse Offline Reinforcement Learning
https://openreview.net/forum?id=TBIzh9b5eaz
AUTHORS: N?ria Armengol Urp?, Sebastian Curi, Andreas Krause
HIGHLIGHT: We propose the first risk-averse reinforcement learning algorithm in the fully offline setting.

645, TITLE: Group Equivariant Stand-Alone Self-Attention For Vision
https://openreview.net/forum?id=JkfYjnOEo6M
AUTHORS: David W. Romero, Jean-Baptiste Cordonnier
HIGHLIGHT: We provide a general self-attention formulation to impose group equivariance to arbitrary symmetry groups.

646, TITLE: A Better Alternative to Error Feedback for Communication-Efficient Distributed Learning
https://openreview.net/forum?id=vYVI1CHPaQg
AUTHORS: Samuel Horv?th, Peter Richtarik
HIGHLIGHT: In this paper, we propose a new and theoretically and practically better alternative to EF for dealing with contractive compressors.

647, TITLE: Neural Delay Differential Equations
https://openreview.net/forum?id=Q1jmmQz72M2
AUTHORS: Qunxi Zhu, Yao Guo, Wei Lin
HIGHLIGHT: We propose a new class of continuous-depth neural networks with delay, named as Neural Delay Differential Equations and having better representation capability outperforming the Neural ODEs.

648, TITLE: Capturing Label Characteristics in VAEs
https://openreview.net/forum?id=wQRlSUZ5V7B
AUTHORS: Tom Joy, Sebastian Schmon, Philip Torr, Siddharth N, Tom Rainforth
HIGHLIGHT: We present a principled approach to incorporating labels in VAEs that captures the rich characteristic information associated with those labels.

649, TITLE: Graph Edit Networks
https://openreview.net/forum?id=dlEJsyHGeaL
AUTHORS: Benjamin Paassen, Daniele Grattarola, Daniele Zambon, Cesare Alippi, Barbara Eva Hammer
HIGHLIGHT: We show that graph neural networks can predict graph edits and are connected to the graph edit distance via graph mappings

650, TITLE: InfoBERT: Improving Robustness of Language Models from An Information Theoretic Perspective
https://openreview.net/forum?id=hpH98mK5Puk
AUTHORS: Boxin Wang, Shuohang Wang, Yu Cheng, Zhe Gan, Ruoxi Jia, Bo Li, Jingjing Liu
HIGHLIGHT: We propose a novel learning framework, InfoBERT, for robust fine-tuning of pre-trained language models from an information-theoretic perspective, and achieve state-of-the-art robust accuracy over several adversarial datasets on NLI and QA tasks.

651, TITLE: DrNAS: Dirichlet Neural Architecture Search
https://openreview.net/forum?id=9FWas6YbmB3
AUTHORS: Xiangning Chen, Ruochen Wang, Minhao Cheng, Xiaocheng Tang, Cho-Jui Hsieh
HIGHLIGHT: This paper proposes a novel differentiable architecture search method by formulating it into a distribution learning problem.

652, TITLE: Drop-Bottleneck: Learning Discrete Compressed Representation for Noise-Robust Exploration
https://openreview.net/forum?id=1rxHOBjeDUW
AUTHORS: Jaekyeom Kim, Minjung Kim, Dongyeon Woo, Gunhee Kim
HIGHLIGHT: Our novel IB method, Drop-Bottleneck, discretely drops task-irrelevant input features to build the compressed representation and shows state-of-the-art performance on noisy, sparse-reward navigation tasks in reinforcement learning.

653, TITLE: Monte-Carlo Planning and Learning with Language Action Value Estimates
https://openreview.net/forum?id=7_G8JySGecm

AUTHORS: Youngsoo Jang, Seokin Seo, Jongmin Lee, Kee-Eung Kim
HIGHLIGHT: We present Monte-Carlo planning with Language Action Value Estimates (MC-LAVE) that combines a Monte-Carlo tree search with language-driven exploration for Interactive Fiction games.

654, TITLE: Robust early-learning: Hindering the memorization of noisy labels
https://openreview.net/forum?id=Eql5b1_hTE4
AUTHORS: Xiaobo Xia, Tongliang Liu, Bo Han, Chen Gong, Nannan Wang, Zongyuan Ge, Yi Chang
HIGHLIGHT: In this paper, motivated by the \textit{lottery ticket hypothesis} which shows that only partial parameters are important for generalization, we find that only partial parameters are important for fitting clean labels and generalize well, which we term as \textit{critical parameters}; while the other parameters tend to fit noisy labels and cannot generalize well, which we term as \textit{non-critical parameters}.

655, TITLE: Identifying Physical Law of Hamiltonian Systems via Meta-Learning
https://openreview.net/forum?id=45NZvF1UHam
AUTHORS: Seungjun Lee, Haesang Yang, Woojae Seong
HIGHLIGHT: We introduce meta-learning algorithms to identify the shared representation of Hamiltonian systems.

656, TITLE: Reweighting Augmented Samples by Minimizing the Maximal Expected Loss
https://openreview.net/forum?id=9G5MIc-goqB
AUTHORS: Mingyang Yi, Lu Hou, Lifeng Shang, Xin Jiang, Qun Liu, Zhi-Ming Ma
HIGHLIGHT: a new reweighting strategy on augmented samples

657, TITLE: Understanding and Improving Encoder Layer Fusion in Sequence-to-Sequence Learning
https://openreview.net/forum?id=n1HD8M6WGn
AUTHORS: Xuebo Liu, Longyue Wang, Derek F. Wong, Liang Ding, Lidia S. Chao, Zhaopeng Tu
HIGHLIGHT: In this paper, our main contribution is to take a step further in understanding EncoderFusion.

658, TITLE: Spatial Dependency Networks: Neural Layers for Improved Generative Image Modeling
https://openreview.net/forum?id=I4c4K9vBNny
AUTHORS: ?orde Miladinovic, Aleksandar Stanic, Stefan Bauer, J?rgen Schmidhuber, Joachim M. Buhmann
HIGHLIGHT: A novel neural network layer for improved generative modeling of images, applied to variational autoencoders.

659, TITLE: Deep Repulsive Clustering of Ordered Data Based on Order-Identity Decomposition
https://openreview.net/forum?id=Yz-XtK5RBxB
AUTHORS: Seon-Ho Lee, Chang-Su Kim
HIGHLIGHT: A deep clustering algorithm for ordered data is proposed based on the order-identity decomposition.

660, TITLE: Revisiting Locally Supervised Learning: an Alternative to End-to-end Training
https://openreview.net/forum?id=fAbkE6ant2
AUTHORS: Yulin Wang, Zanlin Ni, Shiji Song, Le Yang, Gao Huang
HIGHLIGHT: We provide a deep understanding of locally supervised learning, and make it perform on par with end-to-end training, while with significantly reduced GPUs memory footprint.

661, TITLE: Nonseparable Symplectic Neural Networks
https://openreview.net/forum?id=B5VvQrI49Pa
AUTHORS: Shiying Xiong, Yunjin Tong, Xingzhe He, Cheng Yang, Shuqi Yang, Bo Zhu
HIGHLIGHT: To solve the problem, we propose a novel neural network architecture, Nonseparable Symplectic Neural Networks (NSSNNs), to uncover and embed the symplectic structure of a nonseparable Hamiltonian system from limited observation data.

662, TITLE: Gradient Origin Networks
https://openreview.net/forum?id=0O_cQfw6uEh
AUTHORS: Sam Bond-Taylor, Chris G. Willcocks
HIGHLIGHT: A new model that uses the negative gradient of the loss with respect to the origin as a latent vector is found to be superior to equivalent networks.

663, TITLE: Learning to Sample with Local and Global Contexts in Experience Replay Buffer
https://openreview.net/forum?id=gJYlaqL8i8
AUTHORS: Youngmin Oh, Kimin Lee, Jinwoo Shin, Eunho Yang, Sung Ju Hwang

HIGHLIGHT: We propose a learning-based neural replay which calculates the relative importance to sample experience for off-policy RL.

664, TITLE: Provable Rich Observation Reinforcement Learning with Combinatorial Latent States
https://openreview.net/forum?id=hx1IXFHAw7R
AUTHORS: Dipendra Misra, Qinghua Liu, Chi Jin, John Langford
HIGHLIGHT: We introduce a problem setup and a provable reinforcement learning algorithm for rich-observation problems with latent combinatorially large state space.

665, TITLE: Sharper Generalization Bounds for Learning with Gradient-dominated Objective Functions
https://openreview.net/forum?id=r28GdiQF7vM
AUTHORS: Yunwen Lei, Yiming Ying
HIGHLIGHT: We develop sharper generalization bounds for learning with gradient-dominated objective functions.

666, TITLE: Rapid Neural Architecture Search by Learning to Generate Graphs from Datasets
https://openreview.net/forum?id=rkQuFUmUOg3
AUTHORS: Hayeon Lee, Eunyoung Hyung, Sung Ju Hwang
HIGHLIGHT: We propose an efficient NAS framework that is trained once on a database consisting of datasets and pretrained networks and can rapidly generate a neural architecture for a novel dataset.

667, TITLE: Relating by Contrasting: A Data-efficient Framework for Multimodal Generative Models
https://openreview.net/forum?id=vhKe9UFbrJo
AUTHORS: Yuge Shi, Brooks Paige, Philip Torr, Siddharth N
HIGHLIGHT: To mitigate this, we develop a novel contrastive framework for generative model learning, allowing us to train the model not just by the commonality between modalities, but by the distinction between "related" and "unrelated" multimodal data.

668, TITLE: FedMix: Approximation of Mixup under Mean Augmented Federated Learning
https://openreview.net/forum?id=Ogga20D2HO-
AUTHORS: Tehrim Yoon, Sumin Shin, Sung Ju Hwang, Eunho Yang
HIGHLIGHT: We introduce a new federated framework, Mean Augmented Federated Learning (MAFL), and propose an efficient algorithm, Federated Mixup (FedMix), which shows good performance on difficult non-iid situations.

669, TITLE: Generalized Variational Continual Learning
https://openreview.net/forum?id=_IM-AfFhna9
AUTHORS: Noel Loo, Siddharth Swaroop, Richard E Turner
HIGHLIGHT: We generalize VCL and Online-EWC and combine with task-specific FiLM layers

670, TITLE: Understanding and Improving Lexical Choice in Non-Autoregressive Translation
https://openreview.net/forum?id=ZTFeSBIX9C
AUTHORS: Liang Ding, Longyue Wang, Xuebo Liu, Derek F. Wong, Dacheng Tao, Zhaopeng Tu
HIGHLIGHT: We reveal the side effect of knowledge distillation from lexical choice perspective for Non-autoregressvie machine translation, and then propose a simple yet effective approach to improve it.

671, TITLE: Bayesian Context Aggregation for Neural Processes
https://openreview.net/forum?id=ufZN2-aehFa
AUTHORS: Michael Volpp, Fabian Fl?renbrock, Lukas Grossberger, Christian Daniel, Gerhard Neumann
HIGHLIGHT: We propose a Bayesian Aggregation mechanism for Neural Process-based models which improves upon traditional mean aggregation.

672, TITLE: Variational Intrinsic Control Revisited
https://openreview.net/forum?id=P0p33rgyoE
AUTHORS: Taehwan Kwon
HIGHLIGHT: Revisitation of Variational Intrinsic Control (VIC) for the optimal behavior of implicit VIC under stochastic dynamics.

673, TITLE: Implicit Gradient Regularization
https://openreview.net/forum?id=3q5IqUrkcF
AUTHORS: David Barrett, Benoit Dherin
HIGHLIGHT: We have found a hidden form of regularization in gradient descent - Implicit Gradient Regularization - that biases overparameterized models towards flat, low test error solutions and helps us to understand why deep learning works so well.

674, TITLE: Return-Based Contrastive Representation Learning for Reinforcement Learning
https://openreview.net/forum?id=_TM6rT7tXke
AUTHORS: Guoqing Liu, Chuheng Zhang, Li Zhao, Tao Qin, Jinhua Zhu, Li Jian, Nenghai Yu, Tie-Yan Liu
HIGHLIGHT: We propose a novel contrastive learning based auxiliary task which forces the learnt representations to discriminate state-action pairs with different returns and achieve superior performance on complex tasks in terms of sample effiency.

675, TITLE: Scalable Bayesian Inverse Reinforcement Learning
https://openreview.net/forum?id=4qR3coiNaIv
AUTHORS: Alex James Chan, Mihaela van der Schaar
HIGHLIGHT: A variational inference approach to Bayesian inverse reinforcement learning.

676, TITLE: Exemplary Natural Images Explain CNN Activations Better than State-of-the-Art Feature Visualization
https://openreview.net/forum?id=QO9-y8also-
AUTHORS: Judy Borowski, Roland Simon Zimmermann, Judith Schepers, Robert Geirhos, Thomas S. A. Wallis, Matthias Bethge, Wieland Brendel
HIGHLIGHT: Using human psychophysical experiments, we show that natural images can be significantly more informative for interpreting neural network activations than synthetic feature visualizations.

677, TITLE: LiftPool: Bidirectional ConvNet Pooling
https://openreview.net/forum?id=kE3vd639uRW
AUTHORS: Jiaojiao Zhao, Cees G. M. Snoek
HIGHLIGHT: By adopting the philosophy of the classical Lifting Scheme from signal processing, we propose LiftPool for bidirectional pooling layers, including LiftDownPool and LiftUpPool.

678, TITLE: Adversarial score matching and improved sampling for image generation
https://openreview.net/forum?id=eLfqMl3z3lq
AUTHORS: Alexia Jolicoeur-Martineau, R?mi Pich?-Taillefer, Ioannis Mitliagkas, Remi Tachet des Combes
HIGHLIGHT: Combining GANs with score matching and using Consistent Sampling (as an alternative to Langevin dynamics) for improved generative modeling

679, TITLE: Transient Non-stationarity and Generalisation in Deep Reinforcement Learning
https://openreview.net/forum?id=Qun8fv4qSby
AUTHORS: Maximilian Igl, Gregory Farquhar, Jelena Luketina, Wendelin Boehmer, Shimon Whiteson
HIGHLIGHT: We find that transient non-stationarity can worsen generalization in reinforcement learning and propose a method to overcome this effeect.

680, TITLE: On the Origin of Implicit Regularization in Stochastic Gradient Descent
https://openreview.net/forum?id=rq_Qr0c1Hyo
AUTHORS: Samuel L Smith, Benoit Dherin, David Barrett, Soham De
HIGHLIGHT: For small learning rates, the iterates of SGD stay close to the path of gradient flow on a modified loss function containing an implicit regularizer.

681, TITLE: Analyzing the Expressive Power of Graph Neural Networks in a Spectral Perspective
https://openreview.net/forum?id=-qh0M9XWxnv
AUTHORS: Muhammet Balcilar, Guillaume Renton, Pierre H?roux, Benoit Ga?z?re, S?bastien Adam, Paul Honeine
HIGHLIGHT: This paper aims to analyse of the expressive power of Graph Neural Network in spectral domain.

682, TITLE: Pruning Neural Networks at Initialization: Why Are We Missing the Mark?
https://openreview.net/forum?id=Ig-VyQc-MLK
AUTHORS: Jonathan Frankle, Gintare Karolina Dziugaite, Daniel Roy, Michael Carbin
HIGHLIGHT: Methods for pruning neural nets at initialization perform the same or better when shuffling or reinitializing the weights they prune in each layer, a way in which they differ from SOTA weight-pruning methods after training.

683, TITLE: SkipW: Resource adaptable RNN with strict upper computational limit
https://openreview.net/forum?id=2CjEVW-RGOJ
AUTHORS: Tsiry Mayet, Anne Lambert, Pascal Leguyadec, Francoise Le Bolzer, Fran?ois Schnitzler
HIGHLIGHT: Skip-Window is a method to allow recurrent neural networks (RNNs) to trade off accuracy for computational cost during the analysis of a sequence while keeping a strict upper computational limit.

684, TITLE: Into the Wild with AudioScope: Unsupervised Audio-Visual Separation of On-Screen Sounds
https://openreview.net/forum?id=MDsQkFP1Aw
AUTHORS: Efthymios Tzinis, Scott Wisdom, Aren Jansen, Shawn Hershey, Tal Remez, Dan Ellis, John R. Hershey
HIGHLIGHT: We propose an open-domain unsupervised audio-visual on-screen separation system trained and tested on in-the-wild videos.

685, TITLE: Simple Augmentation Goes a Long Way: ADRL for DNN Quantization
https://openreview.net/forum?id=Qr0aRliE_Hb
AUTHORS: Lin Ning, Guoyang Chen, Weifeng Zhang, Xipeng Shen
HIGHLIGHT: Augments the neural networks in Deep Reinforcement Learning(DRL) with a complementary scheme to boost the performance of learning and solve the common low convergence problem in the early stage of DRL

686, TITLE: Few-Shot Bayesian Optimization with Deep Kernel Surrogates
https://openreview.net/forum?id=bJxgv5C3sYc
AUTHORS: Martin Wistuba, Josif Grabocka
HIGHLIGHT: Model-agnostic meta-learning meets Bayesian optimization to speed-up hyperparameter optimization by learning on metadata from different data sets.

687, TITLE: AdaSpeech: Adaptive Text to Speech for Custom Voice
https://openreview.net/forum?id=Drynvt7gg4L
AUTHORS: Mingjian Chen, Xu Tan, Bohan Li, Yanqing Liu, Tao Qin, sheng zhao, Tie-Yan Liu
HIGHLIGHT: We propose AdaSpeech, an adaptive TTS system for high-quality and efficient adaptation of new speaker in custom voice.

688, TITLE: HeteroFL: Computation and Communication Efficient Federated Learning for Heterogeneous Clients
https://openreview.net/forum?id=TNkPBBYFkXg
AUTHORS: Enmao Diao, Jie Ding, Vahid Tarokh
HIGHLIGHT: In this work, we propose a new federated learning framework named HeteroFL to train heterogeneous local models with varying computation complexities.

689, TITLE: DINO: A Conditional Energy-Based GAN for Domain Translation
https://openreview.net/forum?id=WAISmwsqDsb
AUTHORS: Konstantinos Vougioukas, Stavros Petridis, Maja Pantic
HIGHLIGHT: A framework for domain translation which uses a novel mechanism for conditioning Energy-based GANs.

690, TITLE: Universal Weakly Supervised Segmentation by Pixel-to-Segment Contrastive Learning
https://openreview.net/forum?id=N33d7wjgzde
AUTHORS: Tsung-Wei Ke, Jyh-Jing Hwang, Stella Yu
HIGHLIGHT: We propose a unified pixel-to-segment contrastive learning loss formulation for weakly supervised semantic segmentation with various types of annotations.

691, TITLE: PC2WF: 3D Wireframe Reconstruction from Raw Point Clouds
https://openreview.net/forum?id=8X2eaSZxTP
AUTHORS: Yujia Liu, Stefano D'Aronco, Konrad Schindler, Jan Dirk Wegner
HIGHLIGHT: An end-to-end trainable deep neural network for converting a 3D point cloud into a wireframe model.

692, TITLE: Multi-resolution modeling of a discrete stochastic process identifies causes of cancer
https://openreview.net/forum?id=KtH8W3S_RE
AUTHORS: Adam Uri Yaari, Maxwell Sherman, Oliver Clarke Priebe, Po-Ru Loh, Boris Katz, Andrei Barbu, Bonnie Berger
HIGHLIGHT: We integrate a deep learning framework with a probabilistic model to learn a discrete stochastic process at arbitrary length scales, the method accurately and efficiently model mutations load in a tumor and detect cancer driver mutations genome-wide

693, TITLE: C-Learning: Horizon-Aware Cumulative Accessibility Estimation
https://openreview.net/forum?id=W3Wf_wKmqm9
AUTHORS: Panteha Naderian, Gabriel Loaiza-Ganem, Harry J. Braviner, Anthony L. Caterini, Jesse C. Cresswell, Tong Li, Animesh Garg
HIGHLIGHT: We introduce C-learning, a Q-learning inspired method to learn horizon-dependent policies for goal reaching.

694, TITLE: Shapley Explanation Networks
https://openreview.net/forum?id=vsU0efpivw
AUTHORS: Rui Wang, Xiaoqian Wang, David I. Inouye
HIGHLIGHT: To enable new capabilities, we propose to use Shapley values as inter-layer representations in deep neural networks rather than as post-hoc explanations.

695, TITLE: The role of Disentanglement in Generalisation
https://openreview.net/forum?id=qbH974jKUVy
AUTHORS: Milton Llera Montero, Casimir JH Ludwig, Rui Ponte Costa, Gaurav Malhotra, Jeffrey Bowers
HIGHLIGHT: Disentangled models do not achieve compositional generalization when tested systematically.

696, TITLE: Learning N:M Fine-grained Structured Sparse Neural Networks From Scratch
https://openreview.net/forum?id=K9bw7vqp_s
AUTHORS: Aojun Zhou, Yukun Ma, Junnan Zhu, Jianbo Liu, Zhijie Zhang, Kun Yuan, Wenxiu Sun, Hongsheng Li
HIGHLIGHT: a simple yet universal recipe to learn N:M sparse neural networks from scratch

697, TITLE: Unsupervised Meta-Learning through Latent-Space Interpolation in Generative Models
https://openreview.net/forum?id=XOjv2HxIF6i
AUTHORS: Siavash Khodadadeh, Sharare Zehtabian, Saeed Vahidian, Weijia Wang, Bill Lin, Ladislau Boloni
HIGHLIGHT: We use interpolation in generative models latent space to generate tasks for unsupervised meta-learninig.

698, TITLE: On Data-Augmentation and Consistency-Based Semi-Supervised Learning
https://openreview.net/forum?id=7FNqrcPtieT
AUTHORS: Atin Ghosh, Alexandre H. Thiery
HIGHLIGHT: We propose a simple and natural framework leveraging the Hidden Manifold Model to study modern SSL methods.

699, TITLE: Learning from Demonstration with Weakly Supervised Disentanglement
https://openreview.net/forum?id=Ldau9eHU-qO
AUTHORS: Yordan Hristov, Subramanian Ramamoorthy
HIGHLIGHT: We propose a generative model-based approach to learning interpretable robot trajectory representations from demonstrations (image embeddings and end-effector trajectories) paired with coarse labels, which provide a form of weak supervision.

700, TITLE: Neurally Augmented ALISTA
https://openreview.net/forum?id=q_S44KLQ_Aa
AUTHORS: Freya Behrens, Jonathan Sauder, Peter Jung
HIGHLIGHT: We introduce Neurally Augmented ALISTA, extending ALISTA to compute adaptive parameters to achieve improved recovery of individual sparse target vectors.

701, TITLE: Shape or Texture: Understanding Discriminative Features in CNNs
https://openreview.net/forum?id=NcFEZOi-rLa
AUTHORS: Md Amirul Islam, Matthew Kowal, Patrick Esser, Sen Jia, Bj?rn Ommer, Konstantinos G. Derpanis, Neil Bruce
HIGHLIGHT: Exploring and quantifying shape information encoded in CNNs.

702, TITLE: Convex Potential Flows: Universal Probability Distributions with Optimal Transport and Convex Optimization
https://openreview.net/forum?id=te7PVH1sPxJ
AUTHORS: Chin-Wei Huang, Ricky T. Q. Chen, Christos Tsirigotis, Aaron Courville
HIGHLIGHT: We propose to use an input-convex neural network to parameterize an invertible model with universal density approximation guarantees.

703, TITLE: Wasserstein Embedding for Graph Learning
https://openreview.net/forum?id=AAes_3W-2z
AUTHORS: Soheil Kolouri, Navid Naderializadeh, Gustavo K. Rohde, Heiko Hoffmann
HIGHLIGHT: Wasserstein Embedding for Graph Learning (WEGL) is a novel and fast framework for embedding entire graphs into a vector space in which the Euclidean distance between representations approximates the 2-Wasserstein distance.

704, TITLE: Meta-learning with negative learning rates
https://openreview.net/forum?id=60j5LygnmD
AUTHORS: Alberto Bernacchia

HIGHLIGHT: We show theoretically that the optimal inner learning rate of MAML during training is always negative in a family of models

705, TITLE: Representing Partial Programs with Blended Abstract Semantics
https://openreview.net/forum?id=mCtadqIxOJ
AUTHORS: Maxwell Nye, Yewen Pu, Matthew Bowers, Jacob Andreas, Joshua B. Tenenbaum, Armando Solar-Lezama
HIGHLIGHT: We use a combination of concrete execution and learned neural semantics to represent partial programs, resulting in more accurate program synthesis.

706, TITLE: Fast convergence of stochastic subgradient method under interpolation
https://openreview.net/forum?id=w2mYg3d0eot
AUTHORS: Huang Fang, Zhenan Fan, Michael Friedlander
HIGHLIGHT: This paper studies the behaviour of the stochastic subgradient descent (SSGD) method applied to over-parameterized nonsmooth optimization problems that satisfy an interpolation condition.

707, TITLE: A Hypergradient Approach to Robust Regression without Correspondence
https://openreview.net/forum?id=l35SB-_raSQ
AUTHORS: Yujia Xie, Yixiu Mao, Simiao Zuo, Hongteng Xu, Xiaojing Ye, Tuo Zhao, Hongyuan Zha
HIGHLIGHT: We propose a differentiable programming framework for the regression without correspondence problem.

708, TITLE: On the role of planning in model-based deep reinforcement learning
https://openreview.net/forum?id=IrM64DGB21
AUTHORS: Jessica B Hamrick, Abram L. Friesen, Feryal Behbahani, Arthur Guez, Fabio Viola, Sims Witherspoon, Thomas Anthony, Lars Holger Buesing, Petar Velickovic, Theophane Weber
HIGHLIGHT: An empirical investigation into how planning drives performance in model-based RL algorithms.

709, TITLE: Trajectory Prediction using Equivariant Continuous Convolution
https://openreview.net/forum?id=J8_GttYLFgr
AUTHORS: Robin Walters, Jinxi Li, Rose Yu
HIGHLIGHT: Our model, ECCO, uses rotationally-equivariant continuous convolution to improve generalization in trajectory prediction.

710, TITLE: Grounding Language to Autonomously-Acquired Skills via Goal Generation
https://openreview.net/forum?id=chPj_I5KMHG
AUTHORS: Ahmed Akakzia, C?dric Colas, Pierre-Yves Oudeyer, Mohamed CHETOUANI, Olivier Sigaud
HIGHLIGHT: We propose a new RL architecture called Language-Goal-Behavior that proposes to decouple skill learning and language grounding via the introduction of an intermediate semantic goal representation.

711, TITLE: Chaos of Learning Beyond Zero-sum and Coordination via Game Decompositions
https://openreview.net/forum?id=a3wKPZpGtCF
AUTHORS: Yun Kuen Cheung, Yixin Tao
HIGHLIGHT: We characterize games in which popular learning algorithms exhibit Lyapunov chaos.

712, TITLE: Isometric Transformation Invariant and Equivariant Graph Convolutional Networks
https://openreview.net/forum?id=FX0vR39SJ5q
AUTHORS: Masanobu Horie, Naoki Morita, Toshiaki Hishinuma, Yu Ihara, Naoto Mitsume
HIGHLIGHT: We developed isometric transformation invariant and equivariant graph convolutional networks, which shows high prediction performance and computational efficiency.

713, TITLE: R-GAP: Recursive Gradient Attack on Privacy
https://openreview.net/forum?id=RSU17UoKfJF
AUTHORS: Junyi Zhu, Matthew B. Blaschko
HIGHLIGHT: However, there is a fundamental lack of theoretical understanding of how and when gradients can lead to unique recovery of original data. Our research fills this gap by providing a closed-form recursive procedure to recover data from gradients in deep neural networks.

714, TITLE: Multi-Level Local SGD: Distributed SGD for Heterogeneous Hierarchical Networks
https://openreview.net/forum?id=C70cp4Cn32
AUTHORS: Timothy Castiglia, Anirban Das, Stacy Patterson

HIGHLIGHT: We propose Multi-Level Local SGD, a distributed stochastic gradient method for learning a smooth, non-convex objective in a multi-level communication network with heterogeneous workers.

715, TITLE: GShard: Scaling Giant Models with Conditional Computation and Automatic Sharding
https://openreview.net/forum?id=qrwe7XHTmYb
AUTHORS: Dmitry Lepikhin, HyoukJoong Lee, Yuanzhong Xu, Dehao Chen, Orhan Firat, Yanping Huang, Maxim Krikun, Noam Shazeer, Zhifeng Chen
HIGHLIGHT: In this paper we demonstrate conditional computation as a remedy to the above mentioned impediments, and demonstrate its efficacy and utility.

716, TITLE: Representation learning for improved interpretability and classification accuracy of clinical factors from EEG
https://openreview.net/forum?id=TVjLza1t4hI
AUTHORS: Garrett Honke, Irina Higgins, Nina Thigpen, Vladimir Miskovic, Katie Link, Sunny Duan, Pramod Gupta, Julia Klawohn, Greg Hajcak
HIGHLIGHT: We use disentangled representations of EEG signals to improve performance on clinical classification tasks, provide interpretable recommendations for post-hoc analysis and allow for extraction of ERPs from novel single EEG trajectories.

717, TITLE: Multiplicative Filter Networks
https://openreview.net/forum?id=OmtmcPkkhT
AUTHORS: Rizal Fathony, Anit Kumar Sahu, Devin Willmott, J Zico Kolter
HIGHLIGHT: In this paper, we propose and empirically demonstrate that an arguably simpler class of function approximators can work just as well for such problems: multiplicative filter networks.

718, TITLE: Are Neural Nets Modular? Inspecting Functional Modularity Through Differentiable Weight Masks
https://openreview.net/forum?id=7uVcpu-gMD
AUTHORS: R?bert Csord?s, Sjoerd van Steenkiste, J?rgen Schmidhuber
HIGHLIGHT: We develop a method for analyzing emerging functional modularity in neural networks based on differentiable weight masks and use it to point out important issues in current-day neural networks.

719, TITLE: Modeling the Second Player in Distributionally Robust Optimization
https://openreview.net/forum?id=ZDnzZrTqU9N
AUTHORS: Paul Michel, Tatsunori Hashimoto, Graham Neubig
HIGHLIGHT: We use generative neural models to define the uncertainty set in distributionally robust optimization, and show that this helps train more robust classifiers

720, TITLE: Private Post-GAN Boosting
https://openreview.net/forum?id=6isfR3JCbi
AUTHORS: Marcel Neunhoeffer, Steven Wu, Cynthia Dwork
HIGHLIGHT: We propose Private post-GAN boosting (Private PGB), a differentially private method that combines samples produced by the sequence of generators obtained during GAN training to create a high-quality synthetic dataset.

721, TITLE: Flowtron: an Autoregressive Flow-based Generative Network for Text-to-Speech Synthesis
https://openreview.net/forum?id=Ig53hpHxS4
AUTHORS: Rafael Valle, Kevin J. Shih, Ryan Prenger, Bryan Catanzaro
HIGHLIGHT: In this paper we propose Flowtron: an autoregressive flow-based generative network for text-to-speech synthesis with style transfer and speech varation.

722, TITLE: Learning Structural Edits via Incremental Tree Transformations
https://openreview.net/forum?id=v9hAX77--cZ
AUTHORS: Ziyu Yao, Frank F. Xu, Pengcheng Yin, Huan Sun, Graham Neubig
HIGHLIGHT: A generic incremental editing model for tree-structured data

723, TITLE: Sample-Efficient Automated Deep Reinforcement Learning
https://openreview.net/forum?id=hSjxQ3B7GWq
AUTHORS: J?rg K.H. Franke, Gregor Koehler, Andr? Biedenkapp, Frank Hutter
HIGHLIGHT: SEARL trains a population of off-policy RL agents while simultaneously optimizing the hyperparameters and the neural architecture sample-efficiently.

724, TITLE: Unsupervised Discovery of 3D Physical Objects
https://openreview.net/forum?id=lf7st0bJIA5

AUTHORS: Yilun Du, Kevin A. Smith, Tomer Ullman, Joshua B. Tenenbaum, Jiajun Wu
HIGHLIGHT: We propose an unsupervised framework for discovery 3D physical objects and show that these 3D objects to be used for tasks mimicking early infant cognition.

725, TITLE: Global optimality of softmax policy gradient with single hidden layer neural networks in the mean-field regime
https://openreview.net/forum?id=bB2drc7DPuB
AUTHORS: Andrea Agazzi, Jianfeng Lu
HIGHLIGHT: We prove that softmax policy gradient algorithms with single hidden layer neural networks in the mean-field regime can be expressed as a gradient flow in Wasserstein space and prove that all the fixed points of such dynamics are global optimizers

726, TITLE: Extracting Strong Policies for Robotics Tasks from zero-order trajectory optimizers
https://openreview.net/forum?id=Nc3TJqbcl3
AUTHORS: Cristina Pinneri, Shambhuraj Sawant, Sebastian Blaes, Georg Martius
HIGHLIGHT: We propose an adaptively guided imitation learning method that is able to extract strong policies for hard robotic tasks from zero-order trajectory optimizers.

727, TITLE: Temporally-Extended e-Greedy Exploration
https://openreview.net/forum?id=ONBPHFZ7zG4
AUTHORS: Will Dabney, Georg Ostrovski, Andre Barreto
HIGHLIGHT: We discuss a new framework for option-based exploration, present a thorough empirical study of a simple, generally applicable set of options within this framework, and observe improved performance over state-of-the-art agents and exploration methods.

728, TITLE: Rapid Task-Solving in Novel Environments
https://openreview.net/forum?id=F-mvpFpn_0q
AUTHORS: Samuel Ritter, Ryan Faulkner, Laurent Sartran, Adam Santoro, Matthew Botvinick, David Raposo
HIGHLIGHT: Our agents meta-learn to explore, build models on-the-fly, and plan, enabling them to rapidly solve sequences of tasks in unfamiliar environments.
To enable progress toward RTS, we introduce two challenge domains: (1) a minimal RTS challenge called the Memory\&Planning Game and (2) One-Shot StreetLearn Navigation, which introduces scale and complexity from real-world data.

729, TITLE: Tradeoffs in Data Augmentation: An Empirical Study
https://openreview.net/forum?id=ZcKPWuhG6wy
AUTHORS: Raphael Gontijo-Lopes, Sylvia Smullin, Ekin Dogus Cubuk, Ethan Dyer
HIGHLIGHT: We quantify mechanisms of how data augmentation works with two metrics we introduce: Affinity and Diversity.

730, TITLE: Multiscale Score Matching for Out-of-Distribution Detection
https://openreview.net/forum?id=xoHdgbQJohv
AUTHORS: Ahsan Mahmood, Junier Oliva, Martin Andreas Styner
HIGHLIGHT: Using score estimates at multiple noise scales outperforms state-of-the-art in out-of-distribution detection.

731, TITLE: Understanding Over-parameterization in Generative Adversarial Networks
https://openreview.net/forum?id=C3qvk5IQIJY
AUTHORS: Yogesh Balaji, Mohammadmahdi Sajedi, Neha Mukund Kalibhat, Mucong Ding, Dominik St?ger, Mahdi Soltanolkotabi, Soheil Feizi
HIGHLIGHT: We present an analysis of over-paramterization in GANs both theoretically and empirically.

732, TITLE: Go with the flow: Adaptive control for Neural ODEs
https://openreview.net/forum?id=giit4HdDNa
AUTHORS: Mathieu Chalvidal, Matthew Ricci, Rufin VanRullen, Thomas Serre
HIGHLIGHT: This paper presents a new method to enhance Neural ODEs representation power by dynamically controlling their weight parametrization.

733, TITLE: Linear Last-iterate Convergence in Constrained Saddle-point Optimization
https://openreview.net/forum?id=dx11_7vm5_r
AUTHORS: Chen-Yu Wei, Chung-Wei Lee, Mengxiao Zhang, Haipeng Luo

HIGHLIGHT: We prove Optimistic Gradient Descent Ascent (OGDA) and Optimistic Multiplicative Weights Update (OMWU) converge exponentially fast to the Nash equilibrium in the sense of last-iterate in various game settings including matrix games.

734, TITLE: Learning advanced mathematical computations from examples
https://openreview.net/forum?id=-gfhS00XfKj
AUTHORS: Francois Charton, Amaury Hayat, Guillaume Lample
HIGHLIGHT: We train transformers to predict qualitative and numerical properties of differential equations

735, TITLE: WaveGrad: Estimating Gradients for Waveform Generation
https://openreview.net/forum?id=NsMLjcFaO8O
AUTHORS: Nanxin Chen, Yu Zhang, Heiga Zen, Ron J Weiss, Mohammad Norouzi, William Chan
HIGHLIGHT: This paper introduces WaveGrad, a conditional model for waveform generation through estimating gradients of the data density.

736, TITLE: SALD: Sign Agnostic Learning with Derivatives
https://openreview.net/forum?id=7EDgLu9reQD
AUTHORS: Matan Atzmon, Yaron Lipman
HIGHLIGHT: Sign agnostic learning with derivatives for learning high fidelity 3D implicit neural representations shape space from raw data.

737, TITLE: Generalized Energy Based Models
https://openreview.net/forum?id=0PtUPB9z6qK
AUTHORS: Michael Arbel, Liang Zhou, Arthur Gretton
HIGHLIGHT: We introduce the Generalized Energy Based Model (GEBM) for generative modelling.

738, TITLE: Long Range Arena : A Benchmark for Efficient Transformers
https://openreview.net/forum?id=qVyeW-grC2k
AUTHORS: Yi Tay, Mostafa Dehghani, Samira Abnar, Yikang Shen, Dara Bahri, Philip Pham, Jinfeng Rao, Liu Yang, Sebastian Ruder, Donald Metzler
HIGHLIGHT: Better benchmarking for Xformers

739, TITLE: Beyond Categorical Label Representations for Image Classification
https://openreview.net/forum?id=MyHwDabUHZm
AUTHORS: Boyuan Chen, Yu Li, Sunand Raghupathi, Hod Lipson
HIGHLIGHT: We study the role of label representations for standard image classification task and found high-dimensional hign-entropy labes generally lead to more robust and data-efficient networks.

740, TITLE: CoCo: Controllable Counterfactuals for Evaluating Dialogue State Trackers
https://openreview.net/forum?id=eom0IUrF__F
AUTHORS: SHIYANG LI, Semih Yavuz, Kazuma Hashimoto, Jia Li, Tong Niu, Nazneen Rajani, Xifeng Yan, Yingbo Zhou, Caiming Xiong
HIGHLIGHT: CoCo is a principled method to more flexibly evaluate the robustness of the DST component of TOD systems.

741, TITLE: Stochastic Security: Adversarial Defense Using Long-Run Dynamics of Energy-Based Models
https://openreview.net/forum?id=gwFTuzxJW0
AUTHORS: Mitch Hill, Jonathan Craig Mitchell, Song-Chun Zhu
HIGHLIGHT: Our defensive transformation using long-run MCMC sampling with a convergent EBM is the first method to successfully defend naturally-trained classifiers against adversarial attacks.

742, TITLE: X2T: Training an X-to-Text Typing Interface with Online Learning from User Feedback
https://openreview.net/forum?id=LiX3ECzDPHZ
AUTHORS: Jensen Gao, Siddharth Reddy, Glen Berseth, Anca Dragan, Sergey Levine
HIGHLIGHT: We use online learning from user feedback to train an adaptive interface for typing words using inputs like eye gaze.

743, TITLE: Mapping the Timescale Organization of Neural Language Models
https://openreview.net/forum?id=J3OUycKwz-
AUTHORS: Hsiang-Yun Sherry Chien, Jinhan Zhang, Christopher Honey

HIGHLIGHT: We demonstrated a model-free technique for mapping the timescale organization in neural network models, and we applied this method to reveal a hierarchical timescale organization within LSTM language models.

744, TITLE: PDE-Driven Spatiotemporal Disentanglement
https://openreview.net/forum?id=vLaHRtHvfFp
AUTHORS: J?r?mie Don?, Jean-Yves Franceschi, sylvain lamprier, patrick gallinari
HIGHLIGHT: We introduce a novel interpretation of spatiotemporal disentanglement, inducing a simple and performant disentangled prediction model.

745, TITLE: OPAL: Offline Primitive Discovery for Accelerating Offline Reinforcement Learning
https://openreview.net/forum?id=V69LGwJ0lIN
AUTHORS: Anurag Ajay, Aviral Kumar, Pulkit Agrawal, Sergey Levine, Ofir Nachum
HIGHLIGHT: An effective way to leverage multimodal offline behavioral data is to extract a continuous space of primitives, and use it for downstream task learning.

746, TITLE: Does enhanced shape bias improve neural network robustness to common corruptions?
https://openreview.net/forum?id=yUxUNaj2Sl
AUTHORS: Chaithanya Kumar Mummadi, Ranjitha Subramaniam, Robin Hutmacher, Julien Vitay, Volker Fischer, Jan Hendrik Metzen
HIGHLIGHT: We show that robustness on common corruptions donot correlate with strong shape bias but with the effective data augmentation strategies like stylization

747, TITLE: Directed Acyclic Graph Neural Networks
https://openreview.net/forum?id=JbuYF437WB6
AUTHORS: Veronika Thost, Jie Chen
HIGHLIGHT: We propose DAGNN, a graph neural network tailored to directed acyclic graphs that outperforms conventional GNNs by leveraging the partial order as strong inductive bias besides other suitable architectural features.

748, TITLE: QPLEX: Duplex Dueling Multi-Agent Q-Learning
https://openreview.net/forum?id=Rcmk0xxIQV
AUTHORS: Jianhao Wang, Zhizhou Ren, Terry Liu, Yang Yu, Chongjie Zhang
HIGHLIGHT: A novel multi-agent Q-learning algorithm with a complete IGM (Individual-Global-Max) function class.

749, TITLE: Learning Energy-Based Models by Diffusion Recovery Likelihood
https://openreview.net/forum?id=v_1Soh8QUNc
AUTHORS: Ruiqi Gao, Yang Song, Ben Poole, Ying Nian Wu, Diederik P Kingma
HIGHLIGHT: We present a diffusion recovery likelihood method to tractably learn and sample from a sequence of EBMs based on a diffusion process. We achieves high sample quality, stable long-run sampling chains and estimation of likelihood.

750, TITLE: Neural Network Extrapolations with G-invariances from a Single Environment
https://openreview.net/forum?id=7t1FcJUWhi3
AUTHORS: S Chandra Mouli, Bruno Ribeiro
HIGHLIGHT: This work introduces a novel learning framework for single-environment extrapolations, where invariance to transformation groups is mandatory even without evidence, unless the learner deems it inconsistent with the training data.

751, TITLE: Model-Based Offline Planning
https://openreview.net/forum?id=OMNB1G5xzd4
AUTHORS: Arthur Argenson, Gabriel Dulac-Arnold
HIGHLIGHT: This approach adapts model-based reinforcement learning to offline regimes with little data, and shows state of the art control in offline scenarios.

752, TITLE: On Dyadic Fairness: Exploring and Mitigating Bias in Graph Connections
https://openreview.net/forum?id=xgGS6PmzNq6
AUTHORS: Peizhao Li, Yifei Wang, Han Zhao, Pengyu Hong, Hongfu Liu
HIGHLIGHT: A new method on the fairness of predictive relationships in graph-structured data

753, TITLE: Coping with Label Shift via Distributionally Robust Optimisation
https://openreview.net/forum?id=BtZhsSGNRNi
AUTHORS: Jingzhao Zhang, Aditya Krishna Menon, Andreas Veit, Srinadh Bhojanapalli, Sanjiv Kumar, Suvrit Sra
HIGHLIGHT: We propose an objective to cope with label shift, and provide an adversarial algorithm to effectively optimize it.

754, TITLE: Faster Binary Embeddings for Preserving Euclidean Distances
https://openreview.net/forum?id=YCXrx6rRCXO
AUTHORS: Jinjie Zhang, Rayan Saab
HIGHLIGHT: We propose a fast binary embedding algorithm to preserve Euclidean distances among well-spread vectors and it achieves optimal bit complexity.

755, TITLE: Learning and Evaluating Representations for Deep One-Class Classification
https://openreview.net/forum?id=HCSgyPUfeDj
AUTHORS: Kihyuk Sohn, Chun-Liang Li, Jinsung Yoon, Minho Jin, Tomas Pfister
HIGHLIGHT: We present a two-stage framework for deep one-class classification, composed of state-of-the-art self-supervised representation learning followed by generative or discriminative one-class classifiers.

756, TITLE: Conditional Negative Sampling for Contrastive Learning of Visual Representations
https://openreview.net/forum?id=v8b3e5jN66j
AUTHORS: Mike Wu, Milan Mosse, Chengxu Zhuang, Daniel Yamins, Noah Goodman
HIGHLIGHT: Theoretical and experimental evidence that choosing difficult negative examples in contrastive learning can learn stronger representations as measured by several downstream tasks and image distributions.

757, TITLE: On Position Embeddings in BERT
https://openreview.net/forum?id=onxoVA9FxMw
AUTHORS: Benyou Wang, Lifeng Shang, Christina Lioma, Xin Jiang, Hao Yang, Qun Liu, Jakob Grue Simonsen
HIGHLIGHT: This paper amis to understand and evaluate position embeddings, especially in pretrain language models

758, TITLE: Repurposing Pretrained Models for Robust Out-of-domain Few-Shot Learning
https://openreview.net/forum?id=qkLMTphG5-h
AUTHORS: Namyeong Kwon, Hwidong Na, Gabriel Huang, Simon Lacoste-Julien
HIGHLIGHT: We propose an alternative meta-testing procedure and combine MAML gradient steps with adversarial training and uncertainty-based stepsize adaptation.

759, TITLE: Dataset Meta-Learning from Kernel-Ridge Regression
https://openreview.net/forum?id=l-PrrQrK0QR
AUTHORS: Timothy Nguyen, Zhourong Chen, Jaehoon Lee
HIGHLIGHT: We introduce a meta-learning approach to distilling datasets, achieving state of the art performance for kernel-ridge regression and neural networks.

760, TITLE: AdaFuse: Adaptive Temporal Fusion Network for Efficient Action Recognition
https://openreview.net/forum?id=bM3L3I_853
AUTHORS: Yue Meng, Rameswar Panda, Chung-Ching Lin, Prasanna Sattigeri, Leonid Karlinsky, Kate Saenko, Aude Oliva, Rogerio Feris
HIGHLIGHT: In this paper, we introduce an adaptive temporal fusion network, called AdaFuse, that dynamically fuses channels from current and past feature maps for strong temporal modelling.

761, TITLE: One Network Fits All? Modular versus Monolithic Task Formulations in Neural Networks
https://openreview.net/forum?id=uz5uw6gM0m
AUTHORS: Atish Agarwala, Abhimanyu Das, Brendan Juba, Rina Panigrahy, Vatsal Sharan, Xin Wang, Qiuyi Zhang
HIGHLIGHT: Theoretical bounds and experimental results showing that neural networks trained with SGD can provably solve multiple, very different tasks simultaneously.

762, TITLE: Vector-output ReLU Neural Network Problems are Copositive Programs: Convex Analysis of Two Layer Networks and Polynomial-time Algorithms
https://openreview.net/forum?id=fGF8qAqpXXG
AUTHORS: Arda Sahiner, Tolga Ergen, John M. Pauly, Mert Pilanci
HIGHLIGHT: We demonstrate that two-layer vector-output ReLU networks can be expressed as copositive programs, and introduce algorithms for provably finding their global optima, which are polynomial in the number of samples for a fixed data rank.

763, TITLE: In Defense of Pseudo-Labeling: An Uncertainty-Aware Pseudo-label Selection Framework for Semi-Supervised Learning
https://openreview.net/forum?id=-ODN6SbiUU
AUTHORS: Mamshad Nayeem Rizve, Kevin Duarte, Yogesh S Rawat, Mubarak Shah

HIGHLIGHT: We present an uncertainty-aware pseudo-label selection framework for semi-supervised learning which greatly reduces the noise introduced by the pseudo-labeling process.

764, TITLE: MELR: Meta-Learning via Modeling Episode-Level Relationships for Few-Shot Learning
https://openreview.net/forum?id=D3PcGLdMx0
AUTHORS: Nanyi Fei, Zhiwu Lu, Tao Xiang, Songfang Huang
HIGHLIGHT: This is the first work on explicitly modeling episode-level relationships for few-shot learning.

765, TITLE: Why resampling outperforms reweighting for correcting sampling bias with stochastic descents
https://openreview.net/forum?id=iQQK02mxVIT
AUTHORS: Jing An, Lexing Ying, Yuhua Zhu
HIGHLIGHT: We explain why resampling outperforms reweighting for correcting sampling bias when stochastic gradient algorithms are used.

766, TITLE: Prediction and generalisation over directed actions by grid cells
https://openreview.net/forum?id=Ptaz_zIFbX
AUTHORS: Changmin Yu, Timothy Behrens, Neil Burgess
HIGHLIGHT: Extending existing normative prediction models of grid cells to directed transitions, and provide a unifying framework for mechanistic and normative models of grid cells.

767, TITLE: Hopper: Multi-hop Transformer for Spatiotemporal Reasoning
https://openreview.net/forum?id=MaZFq7bJif7
AUTHORS: Honglu Zhou, Asim Kadav, Farley Lai, Alexandru Niculescu-Mizil, Martin Renqiang Min, Mubbasir Kapadia, Hans Peter Graf
HIGHLIGHT: We propose Hopper, Multi-Hop Transformer, and CATER-h dataset to approach object-centric spatiotemporal reasoning in videos.

768, TITLE: Sparse encoding for more-interpretable feature-selecting representations in probabilistic matrix factorization
https://openreview.net/forum?id=D_KeYoqCYC
AUTHORS: Joshua C Chang, Patrick Fletcher, Jungmin Han, Ted L Chang, Shashaank Vattikuti, Ayah Zirikly, Bart Desmet, Carson C Chow
HIGHLIGHT: We introduce a simple modification to existing sparse matrix factorization methods to rectify widespread erroneous interpretation of the factors.

769, TITLE: Dance Revolution: Long-Term Dance Generation with Music via Curriculum Learning
https://openreview.net/forum?id=xGZG2kS5bFk
AUTHORS: Ruozi Huang, Huang Hu, Wei Wu, Kei Sawada, Mi Zhang, Daxin Jiang
HIGHLIGHT: In this paper, we formalize the music-driven dance generation as a sequence-to-sequence learning problem and devise a novel seq2seq architecture to efficiently process long sequences of music features and capture the fine-grained correspondence between music and dance.

770, TITLE: PAC Confidence Predictions for Deep Neural Network Classifiers
https://openreview.net/forum?id=Qk-Wq5AIjpq
AUTHORS: Sangdon Park, Shuo Li, Osbert Bastani, Insup Lee
HIGHLIGHT: We propose a novel algorithm for constructing predicted classification confidences for DNNs that comes with provable correctness guarantees, and demonstrate how our predicted confidences can be used to enable downstream guarantees in two settings.

771, TITLE: BREEDS: Benchmarks for Subpopulation Shift
https://openreview.net/forum?id=mQPBmvyAuk
AUTHORS: Shibani Santurkar, Dimitris Tsipras, Aleksander Madry
HIGHLIGHT: We develop a methodology for constructing large-scale subpopulation shift benchmarks and use them to assess model robustness as well as the effectiveness existing robustness interventions.

772, TITLE: Bypassing the Ambient Dimension: Private SGD with Gradient Subspace Identification
https://openreview.net/forum?id=7dpmlkBuJFC
AUTHORS: Yingxue Zhou, Steven Wu, Arindam Banerjee
HIGHLIGHT: In this paper, we circumvent the dependence on the ambient dimension by leveraging a low-dimensional structure of gradient space in deep networks---that is, the stochastic gradients for deep nets usually stay in a low dimensional subspace in the training process.

773, TITLE: End-to-End Egospheric Spatial Memory
https://openreview.net/forum?id=rRFIni1CYmy
AUTHORS: Daniel James Lenton, Stephen James, Ronald Clark, Andrew Davison
HIGHLIGHT: End-to-End Egospheric Spatial Memory (ESM) forms a bridge between real-time mapping and differentiable memory

774, TITLE: Evaluating the Disentanglement of Deep Generative Models through Manifold Topology
https://openreview.net/forum?id=djwS0m4Ft_A
AUTHORS: Sharon Zhou, Eric Zelikman, Fred Lu, Andrew Y. Ng, Gunnar E. Carlsson, Stefano Ermon
HIGHLIGHT: Evaluate disentanglement of generative models by measuring manifold topology using persistent homology

775, TITLE: SCoRe: Pre-Training for Context Representation in Conversational Semantic Parsing
https://openreview.net/forum?id=oyZxhRI2RiE
AUTHORS: Tao Yu, Rui Zhang, Alex Polozov, Christopher Meek, Ahmed Hassan Awadallah
HIGHLIGHT: In this work, we present SCORE, a new pre-training approach for CSP tasks designed to induce representations that capture the alignment between the dialogue flow and the structural context.

776, TITLE: Decoupling Global and Local Representations via Invertible Generative Flows
https://openreview.net/forum?id=iWLByfvUhN
AUTHORS: Xuezhe Ma, Xiang Kong, Shanghang Zhang, Eduard H Hovy
HIGHLIGHT: Generative Flow for Decoupled Representation Learning

777, TITLE: Pre-training Text-to-Text Transformers for Concept-centric Common Sense
https://openreview.net/forum?id=3k20LAiHYL2
AUTHORS: Wangchunshu Zhou, Dong-Ho Lee, Ravi Kiran Selvam, Seyeon Lee, Xiang Ren
HIGHLIGHT: We propose self-supervised objectives and a joint training framework to augment pre-trained language models with common sense without relying on external knowledge bases.

778, TITLE: Local Search Algorithms for Rank-Constrained Convex Optimization
https://openreview.net/forum?id=tH6_VWZjoq
AUTHORS: Kyriakos Axiotis, Maxim Sviridenko
HIGHLIGHT: Efficient greedy and local search algorithms for optimizing a convex objective under a rank constraint.

779, TITLE: Combining Label Propagation and Simple Models out-performs Graph Neural Networks
https://openreview.net/forum?id=8E1-f3VhX1o
AUTHORS: Qian Huang, Horace He, Abhay Singh, Ser-Nam Lim, Austin Benson
HIGHLIGHT: Here, we show that for many standard transductive node classification benchmarks, we can ex-ceed or match the performance of state-of-the-art GNNs by combining shallow multilayer perceptrons models (that ignore the graph structure entirely) with two simple postprocessings for correlation in the label structure: (i) an ?error correlation? that spreads residual errors in training data to correct errors in test data and(ii) an ?prediction correlation? that smooths the predictions on the test data.

780, TITLE: Supervised Contrastive Learning for Pre-trained Language Model Fine-tuning
https://openreview.net/forum?id=cu7IUiOhujH
AUTHORS: Beliz Gunel, Jingfei Du, Alexis Conneau, Veselin Stoyanov
HIGHLIGHT: Driven by the intuition that good generalization requires capturing the similarity between examples in one class and contrasting them with examples in other classes, we propose a supervised contrastive learning (SCL) objective for the fine-tuning stage.

781, TITLE: SAFENet: A Secure, Accurate and Fast Neural Network Inference
https://openreview.net/forum?id=Cz3dbFm5u-
AUTHORS: Qian Lou, Yilin Shen, Hongxia Jin, Lei Jiang
HIGHLIGHT: We propose SAFENet that supports automatic channel-wise activation approximation to enable a Secure, Accurate and Fast nEural Network inference service.

782, TITLE: Provably robust classification of adversarial examples with detection
https://openreview.net/forum?id=sRA5rLNpmQc
AUTHORS: Fatemeh Sheikholeslami, Ali Lotfi, J Zico Kolter
HIGHLIGHT: We propose a joint classifier/detector training scheme with provable performance guarantees against adversarial perturbations.

783, TITLE: Saliency is a Possible Red Herring When Diagnosing Poor Generalization
https://openreview.net/forum?id=c9-WeM-ceB
AUTHORS: Joseph D Viviano, Becks Simpson, Francis Dutil, Yoshua Bengio, Joseph Paul Cohen
HIGHLIGHT: We controlled feature construction on images using masks to help models generalize to test distributions with covariate shift and noticed that it didn't affect the saliency maps in the way one would expect even though it improved generalization.

784, TITLE: Fourier Neural Operator for Parametric Partial Differential Equations
https://openreview.net/forum?id=c8P9NQVtmnO
AUTHORS: Zongyi Li, Nikola Borislavov Kovachki, Kamyar Azizzadenesheli, Burigede liu, Kaushik Bhattacharya, Andrew Stuart, Anima Anandkumar
HIGHLIGHT: A novel neural operator based on Fourier transformation for learning partial differential equations.

785, TITLE: Combining Ensembles and Data Augmentation Can Harm Your Calibration
https://openreview.net/forum?id=g11CZSghXyY
AUTHORS: Yeming Wen, Ghassen Jerfel, Rafael Muller, Michael W Dusenberry, Jasper Snoek, Balaji Lakshminarayanan, Dustin Tran
HIGHLIGHT: We found that combining ensembles and data augmentation worsens calibration than applying them individually, and we proposed a simple fix to it.

786, TITLE: SOLAR: Sparse Orthogonal Learned and Random Embeddings
https://openreview.net/forum?id=fw-BHZ1KjxJ
AUTHORS: Tharun Medini, Beidi Chen, Anshumali Shrivastava
HIGHLIGHT: We propose a distributed training scheme to learn high dimensional sparse embeddings that are much better than dense embeddings on both precision and speed.

787, TITLE: Efficient Empowerment Estimation for Unsupervised Stabilization
https://openreview.net/forum?id=u2YNJPcQlwq
AUTHORS: Ruihan Zhao, Kevin Lu, Pieter Abbeel, Stas Tiomkin
HIGHLIGHT: We propose an efficient estimation of empowerment which is demonstrated on unsupervised stabilization of dynamical systems, and compared to the existing relevant methods.

788, TITLE: More or Less: When and How to Build Convolutional Neural Network Ensembles
https://openreview.net/forum?id=z5Z023VBmDZ
AUTHORS: Abdul Wasay, Stratos Idreos
HIGHLIGHT: We show that when we perform a holistic assessment, we uncover a wide design space, where ensembles not only provide better accuracy but also train and deploy with fewer resources than comparable single convolutional network models.

789, TITLE: VTNet: Visual Transformer Network for Object Goal Navigation
https://openreview.net/forum?id=DlLxQP08O3B
AUTHORS: Heming Du, Xin Yu, Liang Zheng
HIGHLIGHT: In this paper, we introduce a Visual Transformer Network (VTNet) for learning informative visual representation in navigation.

790, TITLE: Class Normalization for Zero-Shot Learning
https://openreview.net/forum?id=7pgFL2Dkyyy
AUTHORS: Ivan Skorokhodov, Mohamed Elhoseiny
HIGHLIGHT: We develop theoretical understanding of signal normalization inside zero-shot learning models, propose a novel normalization scheme and use it to achieve SotA ZSL performance with a simple MLP

791, TITLE: Batch Reinforcement Learning Through Continuation Method
https://openreview.net/forum?id=po-DLlBuAuz
AUTHORS: Yijie Guo, Shengyu Feng, Nicolas Le Roux, Ed Chi, Honglak Lee, Minmin Chen
HIGHLIGHT: In this work, we propose a simple yet effective policy iteration approach to batch RL using global optimization techniques known as continuation.

792, TITLE: Towards Resolving the Implicit Bias of Gradient Descent for Matrix Factorization: Greedy Low-Rank Learning
https://openreview.net/forum?id=AHOs7Sm5H7R
AUTHORS: Zhiyuan Li, Yuping Luo, Kaifeng Lyu
HIGHLIGHT: We prove that for depth-2 matrix factorization, gradient flow with infinitesimal initialization is mathematically equivalent to a simple heuristic rank minimization algorithm, Greedy Low-Rank Learning, under some reasonable assumptions.

793, TITLE: Training BatchNorm and Only BatchNorm: On the Expressive Power of Random Features in CNNs
https://openreview.net/forum?id=vYeQQ29Tbvx
AUTHORS: Jonathan Frankle, David J. Schwab, Ari S. Morcos
HIGHLIGHT: We study the role and expressive power of learned affine parameters that transform features by freezing all weights at their random initializations and training only BatchNorm.

794, TITLE: CompOFA ? Compound Once-For-All Networks for Faster Multi-Platform Deployment
https://openreview.net/forum?id=IgIk8RRT-Z
AUTHORS: Manas Sahni, Shreya Varshini, Alind Khare, Alexey Tumanov
HIGHLIGHT: CNN design-space and system insights for faster latency-guided training and searching of models for diverse deployment targets.

795, TITLE: Adaptive and Generative Zero-Shot Learning
https://openreview.net/forum?id=ahAUv8TI2Mz
AUTHORS: Yu-Ying Chou, Hsuan-Tien Lin, Tyng-Luh Liu
HIGHLIGHT: We address the problem of generalized zero-shot learning (GZSL) where the task is to predict the class label of a target image whether its label belongs to the seen or unseen category.

796, TITLE: Learning to Make Decisions via Submodular Regularization
https://openreview.net/forum?id=ac288vnG_7U
AUTHORS: Ayya Alieva, Aiden Aceves, Jialin Song, Stephen Mayo, Yisong Yue, Yuxin Chen
HIGHLIGHT: A data-driven sequential decision making framework based on a novel submodular-regularized loss.

797, TITLE: Generating Furry Cars: Disentangling Object Shape and Appearance across Multiple Domains
https://openreview.net/forum?id=M88oFvqp_9
AUTHORS: Utkarsh Ojha, Krishna Kumar Singh, Yong Jae Lee
HIGHLIGHT: We present a framework for multi-domain disentanglement, facilitating transfer of appearance from one domain to another.

798, TITLE: Implicit Under-Parameterization Inhibits Data-Efficient Deep Reinforcement Learning
https://openreview.net/forum?id=O9bnihsFfXU
AUTHORS: Aviral Kumar, Rishabh Agarwal, Dibya Ghosh, Sergey Levine
HIGHLIGHT: Studies feature matrix rank collapse (i.e. implicit regularization) in deep Q-learning methods.

799, TITLE: Disentangling 3D Prototypical Networks for Few-Shot Concept Learning
https://openreview.net/forum?id=-Lr-u0b42he
AUTHORS: Mihir Prabhudesai, Shamit Lal, Darshan Patil, Hsiao-Yu Tung, Adam W Harley, Katerina Fragkiadaki
HIGHLIGHT: We present neural architectures that disentangle RGB-D images into objects? shapes and styles and a map of the background scene, and explore their applications for few-shot 3D object detection and few-shot concept classification.

800, TITLE: Anytime Sampling for Autoregressive Models via Ordered Autoencoding
https://openreview.net/forum?id=TSRTzJnuEBS
AUTHORS: Yilun Xu, Yang Song, Sahaj Garg, Linyuan Gong, Rui Shu, Aditya Grover, Stefano Ermon
HIGHLIGHT: We propose a new family of autoregressive model that enables anytime sampling

801, TITLE: HyperDynamics: Generating Expert Dynamics Models by Observation
https://openreview.net/forum?id=pHXfe1cOmA
AUTHORS: Zhou Xian, Shamit Lal, Hsiao-Yu Tung, Emmanouil Antonios Platanios, Katerina Fragkiadaki
HIGHLIGHT: We propose HyperDynamics, a framework that conditions on an agent?s interactions with the environment and optionally its visual observations, and generates the parameters of neural dynamics models based on inferred properties of the dynamical system.

802, TITLE: Improving Zero-Shot Voice Style Transfer via Disentangled Representation Learning
https://openreview.net/forum?id=TgSVWXw22FQ
AUTHORS: Siyang Yuan, Pengyu Cheng, Ruiyi Zhang, Weituo Hao, Zhe Gan, Lawrence Carin
HIGHLIGHT: An information-theoretic disentangled representation learning framework for zero-shot voice style transfer.

803, TITLE: GraPPa: Grammar-Augmented Pre-Training for Table Semantic Parsing

https://openreview.net/forum?id=kyaIeYj4zZ
AUTHORS: Tao Yu, Chien-Sheng Wu, Xi Victoria Lin, bailin wang, Yi Chern Tan, Xinyi Yang, Dragomir Radev, richard socher, Caiming Xiong
HIGHLIGHT: Language model pre-training for table semantic parsing.

804, TITLE: Estimating Lipschitz constants of monotone deep equilibrium models
https://openreview.net/forum?id=VcB4QkSfyO
AUTHORS: Chirag Pabbaraju, Ezra Winston, J Zico Kolter
HIGHLIGHT: Monotone deep equilibrium models have Lipschitz constants which are simple to bound and small relative to those of standard DNNs, which suffer with depth.

805, TITLE: Estimating informativeness of samples with Smooth Unique Information
https://openreview.net/forum?id=kEnBH98BGs5
AUTHORS: Hrayr Harutyunyan, Alessandro Achille, Giovanni Paolini, Orchid Majumder, Avinash Ravichandran, Rahul Bhotika, Stefano Soatto
HIGHLIGHT: We define, both in weight-space and function-space, a notion of unique information that an individual sample provides to the training of a deep network and show how to compute it efficiently large networks using a linearization of the model.

806, TITLE: NBDT: Neural-Backed Decision Tree
https://openreview.net/forum?id=mCLVeEpplNE
AUTHORS: Alvin Wan, Lisa Dunlap, Daniel Ho, Jihan Yin, Scott Lee, Suzanne Petryk, Sarah Adel Bargal, Joseph E. Gonzalez
HIGHLIGHT: Neural-Backed Decision Trees improve interpretability and accuracy: (1) out-generalize, improve, and match or outperform baseline neural networks; (2) show visual evidence of generalization, reveal ambiguous ImageNet labels, and improve human trust.

807, TITLE: Accurate Learning of Graph Representations with Graph Multiset Pooling
https://openreview.net/forum?id=JHcqXGaqiGn
AUTHORS: Jinheon Baek, Minki Kang, Sung Ju Hwang
HIGHLIGHT: A novel graph pooling method for graph representation learning, that considers multiset with attention-based operations.

808, TITLE: Byzantine-Resilient Non-Convex Stochastic Gradient Descent
https://openreview.net/forum?id=PbEHqvFtcS
AUTHORS: Dan Alistarh, Zeyuan Allen-Zhu, Faeze Ebrahimianghazani, Jerry Li
HIGHLIGHT: We give a new algorithm for Byzantine-resilient non-convex distributed optimization, with strong theoretical guarantees, which improves on the performance of prior methods for training deep neural networks against Byzantine attacks.

809, TITLE: MetaNorm: Learning to Normalize Few-Shot Batches Across Domains
https://openreview.net/forum?id=9z_dNsC4B5t
AUTHORS: Yingjun Du, Xiantong Zhen, Ling Shao, Cees G. M. Snoek
HIGHLIGHT: We propose MetaNorm, a simple yet effective meta-learning normalization approach that learns adaptive statistics for few-shot classification and domain generalization

810, TITLE: Large Batch Simulation for Deep Reinforcement Learning
https://openreview.net/forum?id=cP5IcoAkfKa
AUTHORS: Brennan Shacklett, Erik Wijmans, Aleksei Petrenko, Manolis Savva, Dhruv Batra, Vladlen Koltun, Kayvon Fatahalian
HIGHLIGHT: The key idea of our approach is to design a 3D renderer and environment simulator around the principle of "batch simulation": accepting and executing large batches of requests simultaneously.

811, TITLE: Personalized Federated Learning with First Order Model Optimization
https://openreview.net/forum?id=ehJqJQk9cw
AUTHORS: Michael Zhang, Karan Sapra, Sanja Fidler, Serena Yeung, Jose M. Alvarez
HIGHLIGHT: We propose a new federated learning framework that efficiently computes a personalized weighted combination of available models for each client, outperforming existing work for personalized federated learning.

812, TITLE: Combining Physics and Machine Learning for Network Flow Estimation
https://openreview.net/forum?id=l0V53bErniB
AUTHORS: Arlei Lopes da Silva, Furkan Kocayusufoglu, Saber Jafarpour, Francesco Bullo, Ananthram Swami, Ambuj Singh

HIGHLIGHT: Paper solves missing flow estimation problem using bilevel optimization and deep learning.

813, TITLE: Knowledge Distillation as Semiparametric Inference
https://openreview.net/forum?id=m4UCf24r0Y
AUTHORS: Tri Dao, Govinda M Kamath, Vasilis Syrgkanis, Lester Mackey
HIGHLIGHT: Viewing knowledge distillation as a semiparametric inference problem leads to improved generalization guarantees of the distillation process

814, TITLE: Learning Value Functions in Deep Policy Gradients using Residual Variance
https://openreview.net/forum?id=NX1He-aFO_F
AUTHORS: Yannis Flet-Berliac, reda ouhamma, odalric-ambrym maillard, Philippe Preux
HIGHLIGHT: We introduce a method to improve the learning of the critic in the actor-critic framework.

815, TITLE: Randomized Ensembled Double Q-Learning: Learning Fast Without a Model
https://openreview.net/forum?id=AY8zfZm0tDd
AUTHORS: Xinyue Chen, Che Wang, Zijian Zhou, Keith W. Ross
HIGHLIGHT: We propose and analyze a novel model-free algorithm that achieves strong performance with a high update-to-data ratio.

816, TITLE: DECENTRALIZED ATTRIBUTION OF GENERATIVE MODELS
https://openreview.net/forum?id=_kxlwvhOodK
AUTHORS: Changhoon Kim, Yi Ren, Yezhou Yang
HIGHLIGHT: This paper investigates the feasibility of decentralized attribution of generative models.

817, TITLE: Constellation Nets for Few-Shot Learning
https://openreview.net/forum?id=vujTf_I8Kmc
AUTHORS: Weijian Xu, yifan xu, Huaijin Wang, Zhuowen Tu
HIGHLIGHT: We tackle the few-shot learning problem by introducing an explicit cell feature clustering procedure with relation learning via self-attention.

818, TITLE: Adapting to Reward Progressivity via Spectral Reinforcement Learning
https://openreview.net/forum?id=dyjPVUc2KB
AUTHORS: Michael Dann, John Thangarajah
HIGHLIGHT: In this paper, we identify a problem with value-based deep reinforcement learning that has not previously been investigated -- namely, reward progressivity -- and propose an approach that addresses it via magnitudinal decomposition of the reward.

819, TITLE: TropEx: An Algorithm for Extracting Linear Terms in Deep Neural Networks
https://openreview.net/forum?id=IqtonxWI0V3
AUTHORS: Martin Trimmel, Henning Petzka, Cristian Sminchisescu
HIGHLIGHT: We propose an algorithm for extracting linear terms of piecewise linear deep neural network functions and apply it to study differences between convolutional and fully-connected networks.

820, TITLE: Reset-Free Lifelong Learning with Skill-Space Planning
https://openreview.net/forum?id=HIGSa_3kOx3
AUTHORS: Kevin Lu, Aditya Grover, Pieter Abbeel, Igor Mordatch
HIGHLIGHT: We propose \textit{Lifelong Skill Planning} (LiSP), an algorithmic framework for lifelong RL based on planning in an abstract space of higher-order skills.

821, TITLE: Robust Learning of Fixed-Structure Bayesian Networks in Nearly-Linear Time
https://openreview.net/forum?id=euDnVs0Ynts
AUTHORS: Yu Cheng, Honghao Lin
HIGHLIGHT: We give the first nearly-linear time algorithm for the robust learning of fixed-structure Bayesian networks.

822, TITLE: Teaching Temporal Logics to Neural Networks
https://openreview.net/forum?id=dOcQK-f4byz
AUTHORS: Christopher Hahn, Frederik Schmitt, Jens U. Kreber, Markus Norman Rabe, Bernd Finkbeiner
HIGHLIGHT: We study two fundamental questions in neuro-symbolic computing: can deep learning tackle challenging problems in logics end-to-end, and can neural networks learn the semantics of logics.

823, TITLE: Spatially Structured Recurrent Modules
https://openreview.net/forum?id=5l9zj5G7vDY
AUTHORS: Nasim Rahaman, Anirudh Goyal, Muhammad Waleed Gondal, Manuel Wuthrich, Stefan Bauer, Yash Sharma, Yoshua Bengio, Bernhard Sch?lkopf
HIGHLIGHT: We model a dynamical system as a collection of recurrent modules that interact according to a spatially informed but learned topology.

824, TITLE: Bayesian Few-Shot Classification with One-vs-Each P?lya-Gamma Augmented Gaussian Processes
https://openreview.net/forum?id=lgNx56yZh8a
AUTHORS: Jake Snell, Richard Zemel
HIGHLIGHT: We propose a Gaussian process approach to few-shot classification based on the one-vs-each softmax approximation and P?lya-gamma augmentation, and demonstrate competitive few-shot accuracy and strong uncertainty quantification.

825, TITLE: Parameter-based Value Functions
https://openreview.net/forum?id=tV6oBfuyLTQ
AUTHORS: Francesco Faccio, Louis Kirsch, J?rgen Schmidhuber
HIGHLIGHT: We propose value functions whose inputs include the policy parameters and which can generalize across different policies

826, TITLE: Hyperbolic Neural Networks++
https://openreview.net/forum?id=Ec85b0tUwbA
AUTHORS: Ryohei Shimizu, YUSUKE Mukuta, Tatsuya Harada
HIGHLIGHT: We present novel methods for constructing hyperbolic neural network architectures in the Poincar? ball model, including a parameter-reduced MLR, geodesic-aware FC layers, convolutional layers, and attention mechanisms.

827, TITLE: Neural Jump Ordinary Differential Equations
https://openreview.net/forum?id=JFKR3WqwyXR
AUTHORS: Calypso Herrera, Florian Krach, Josef Teichmann
HIGHLIGHT: On-line prediction of irregularly-observed time series using neural ODE with jumps with theoretical and empirical justification.

828, TITLE: Seq2Tens: An Efficient Representation of Sequences by Low-Rank Tensor Projections
https://openreview.net/forum?id=dx4b7lm8jMM
AUTHORS: Csaba Toth, Patric Bonnier, Harald Oberhauser
HIGHLIGHT: An Efficient Representation of Sequences by Low-Rank Tensor Projections

829, TITLE: Efficient Fully-Offline Meta-Reinforcement Learning via Distance Metric Learning and Behavior Regularization
https://openreview.net/forum?id=8cpHIfgY4Dj
AUTHORS: Lanqing Li, Rui Yang, Dijun Luo
HIGHLIGHT: A model-free, end-to-end fully-offline meta-RL algorithm designed to maximize practicality, performance and sample/computational efficiency.

830, TITLE: On the Curse of Memory in Recurrent Neural Networks: Approximation and Optimization Analysis
https://openreview.net/forum?id=8Sqhl-nF50
AUTHORS: Zhong Li, Jiequn Han, Weinan E, Qianxiao Li
HIGHLIGHT: We study the approximation properties and optimization dynamics of RNNs in the linear setting, where we uncover precisely the adverse effect of memory on learning.

831, TITLE: Generating Adversarial Computer Programs using Optimized Obfuscations
https://openreview.net/forum?id=PH5PH9ZO_4
AUTHORS: Shashank Srikant, Sijia Liu, Tamara Mitrovska, Shiyu Chang, Quanfu Fan, Gaoyuan Zhang, Una-May O'Reilly
HIGHLIGHT: A differentiable generator of adversarial computer programs which can deceive ML models trained on computer programs.

832, TITLE: BOIL: Towards Representation Change for Few-shot Learning
https://openreview.net/forum?id=umIdUL8rMH
AUTHORS: Jaehoon Oh, Hyungjun Yoo, ChangHwan Kim, Se-Young Yun
HIGHLIGHT: We propose a novel meta-learning algorithm, BOIL, based on representation change.

833, TITLE: Towards Understanding and Improving Dropout in Game Theory
https://openreview.net/forum?id=Jacdvfjicf7
AUTHORS: Hao Zhang, Sen Li, YinChao Ma, Mingjie Li, Yichen Xie, Quanshi Zhang
HIGHLIGHT: We prove and improve the utility of the dropout operation from a game-theoretic view.

834, TITLE: Representation Learning via Invariant Causal Mechanisms
https://openreview.net/forum?id=9p2ekP904Rs
AUTHORS: Jovana Mitrovic, Brian McWilliams, Jacob C Walker, Lars Holger Buesing, Charles Blundell
HIGHLIGHT: We propose a new self-supervised objective with an explicit invariance regularizer and provide an alternative explanation for the success of contrastive learning using causality; we outperform competing methods on ImageNet and Atari.

835, TITLE: Fooling a Complete Neural Network Verifier
https://openreview.net/forum?id=4IwieFS44l
AUTHORS: D?niel Zombori, Bal?zs B?nhelyi, Tibor Csendes, Istv?n Megyeri, M?rk Jelasity
HIGHLIGHT: We propose an attack (along with a defense) to fool complete verification based on exploiting numerical errors.

836, TITLE: CPR: Classifier-Projection Regularization for Continual Learning
https://openreview.net/forum?id=F2v4aqEL6ze
AUTHORS: Sungmin Cha, Hsiang Hsu, Taebaek Hwang, Flavio Calmon, Taesup Moon
HIGHLIGHT: We devise wide local minima promoting regularization term for continual learning.

837, TITLE: CO2: Consistent Contrast for Unsupervised Visual Representation Learning
https://openreview.net/forum?id=U4XLJhqwNF1
AUTHORS: Chen Wei, Huiyu Wang, Wei Shen, Alan Yuille
HIGHLIGHT: We introduce a consistency regularization term into the current contrastive learning framework.

838, TITLE: GAN2GAN: Generative Noise Learning for Blind Denoising with Single Noisy Images
https://openreview.net/forum?id=SHvF5xaueVn
AUTHORS: Sungmin Cha, Taeeon Park, Byeongjoon Kim, Jongduk Baek, Taesup Moon
HIGHLIGHT: We devise GAN2GAN method that trains a blind denoiser solely based on the single noisy images.

839, TITLE: Learning Subgoal Representations with Slow Dynamics
https://openreview.net/forum?id=wxRwhSdORKG
AUTHORS: Siyuan Li, Lulu Zheng, Jianhao Wang, Chongjie Zhang
HIGHLIGHT: We propose a slowness objective to learn subgoal representations in hierarchical reinforcement learning.

840, TITLE: Bowtie Networks: Generative Modeling for Joint Few-Shot Recognition and Novel-View Synthesis
https://openreview.net/forum?id=ESG-DMKQKsD
AUTHORS: Zhipeng Bao, Yu-Xiong Wang, Martial Hebert
HIGHLIGHT: We propose a novel feedback-based bowtie network to learn a shared generative model for joint few-shot recognition and novel-view synthesis, consistently and significantly improving performance for both tasks, especially in the low-data regime.

841, TITLE: Taming GANs with Lookahead-Minmax
https://openreview.net/forum?id=ZW0yXJyNmoG
AUTHORS: Tatjana Chavdarova, Matteo Pagliardini, Sebastian U Stich, Fran?ois Fleuret, Martin Jaggi
HIGHLIGHT: A novel optimizer for GANs and games in general.

842, TITLE: Certify or Predict: Boosting Certified Robustness with Compositional Architectures
https://openreview.net/forum?id=USCNapootw
AUTHORS: Mark Niklas Mueller, Mislav Balunovic, Martin Vechev
HIGHLIGHT: We propose a compositional network architecture boosting the certified robustness of an accurate state-of-the-art network, by combining it with a shallow, provable network using a certified, adaptive selection mechanism.

843, TITLE: New Bounds For Distributed Mean Estimation and Variance Reduction
https://openreview.net/forum?id=t86MwoUCCNe
AUTHORS: Peter Davies, Vijaykrishna Gurunanthan, Niusha Moshrefi, Saleh Ashkboos, Dan Alistarh
HIGHLIGHT: We provide optimal algorithms and lower bounds for distributed mean estimation and variance reduction via a new connection to lattice theory, and show that this technique can be used to improve upon current approaches in practice.

844, TITLE: Interpretable Neural Architecture Search via Bayesian Optimisation with Weisfeiler-Lehman Kernels
https://openreview.net/forum?id=j9Rv7qdXjd
AUTHORS: Xingchen Wan, Binxin Ru, Xiaowen Dong, Michael Osborne
HIGHLIGHT: We propose a NAS method that is sample-efficient, highly performant and interpretable.

845, TITLE: A Discriminative Gaussian Mixture Model with Sparsity
https://openreview.net/forum?id=-_Zp7r2-cGK
AUTHORS: Hideaki Hayashi, Seiichi Uchida
HIGHLIGHT: A sparse classifier based on a discriminative Gaussian mixture model, which can also be embedded into a neural network.

846, TITLE: Communication in Multi-Agent Reinforcement Learning: Intention Sharing
https://openreview.net/forum?id=qpsl2dR9twy
AUTHORS: Woojun Kim, Jongeui Park, Youngchul Sung
HIGHLIGHT: This paper propose a new communication scheme named intention sharing to enhance the coordination among agents.

847, TITLE: Is Attention Better Than Matrix Decomposition?
https://openreview.net/forum?id=1FvkSpWosOl
AUTHORS: Zhengyang Geng, Meng-Hao Guo, Hongxu Chen, Xia Li, Ke Wei, Zhouchen Lin
HIGHLIGHT: Our intriguing finding is that self-attention is not better than the matrix decomposition (MD) model developed 20 years ago regarding the performance and computational cost for encoding the long-distance dependencies.

848, TITLE: Fast and Complete: Enabling Complete Neural Network Verification with Rapid and Massively Parallel Incomplete Verifiers
https://openreview.net/forum?id=nVZtXBI6LNn
AUTHORS: Kaidi Xu, Huan Zhang, Shiqi Wang, Yihan Wang, Suman Jana, Xue Lin, Cho-Jui Hsieh
HIGHLIGHT: We use fast bound propagation methods on GPUs for complete neural network verification and achieve large speedup compared to SOTA.

849, TITLE: A Geometric Analysis of Deep Generative Image Models and Its Applications
https://openreview.net/forum?id=GH7QRzUDdXG
AUTHORS: Binxu Wang, Carlos R Ponce
HIGHLIGHT: We developed tools to compute the metric tensor of image manifold learnt by GANs, empirically analyzed their geometry, and found this knowledge useful to GAN inversion and finding interpretable axes.

850, TITLE: Solving Compositional Reinforcement Learning Problems via Task Reduction
https://openreview.net/forum?id=9SS69KwomAM
AUTHORS: Yunfei Li, Huazhe Xu, Yilin Wu, Xiaolong Wang, Yi Wu
HIGHLIGHT: We propose a deep RL algorithm for learning compositional strategies to solve sparse-reward continuous-control problems.

851, TITLE: ARMOURED: Adversarially Robust MOdels using Unlabeled data by REgularizing Diversity
https://openreview.net/forum?id=JoCR4h9O3Ew
AUTHORS: Kangkang Lu, Cuong Manh Nguyen, Xun Xu, Kiran Chari, Yu Jing Goh, Chuan-Sheng Foo
HIGHLIGHT: ARMOURED is a novel technique for adversarially robust learning that elegantly unifies semi-supervised learning and diversity regularization through a multi-view learning framework.

852, TITLE: Acting in Delayed Environments with Non-Stationary Markov Policies
https://openreview.net/forum?id=j1RMMKeP2gR
AUTHORS: Gal Dalal, Esther Derman, Shie Mannor
HIGHLIGHT: In this paper, we derive theoretical results on execution-delay MDPs, and devise a DQN-based algorithm to empirically tackle this setup.

853, TITLE: Overfitting for Fun and Profit: Instance-Adaptive Data Compression
https://openreview.net/forum?id=oFp8Mx_V5FL
AUTHORS: Ties van Rozendaal, Iris AM Huijben, Taco Cohen
HIGHLIGHT: We show that we can finetune an entire data compression model on a single instance, and improve the rate-distortion performance, taking into account the additional costs for sending the model updates.

854, TITLE: LEARNABLE EMBEDDING SIZES FOR RECOMMENDER SYSTEMS
https://openreview.net/forum?id=vQzcqQWIS0q
AUTHORS: Siyi Liu, Chen Gao, Yihong Chen, Depeng Jin, Yong Li
HIGHLIGHT: Learning flexible feature-aware embedding sizes effectively and efficiently for recommendation models.

855, TITLE: Generative Scene Graph Networks
https://openreview.net/forum?id=RmcPm9m3tnk
AUTHORS: Fei Deng, Zhuo Zhi, Donghun Lee, Sungjin Ahn
HIGHLIGHT: We propose the first object-centric generative model capable of learning part-object relationships from multi-object scenes without access to pre-segmented parts.

856, TITLE: Deconstructing the Regularization of BatchNorm
https://openreview.net/forum?id=d-XzF81Wg1
AUTHORS: Yann Dauphin, Ekin Dogus Cubuk
HIGHLIGHT: We deconstruct the regularization effect of batch normalization and show that preventing explosive growth at the final layer at initialization and during training can explain most of BatchNorm's generalization boost.

857, TITLE: PolarNet: Learning to Optimize Polar Keypoints for Keypoint Based Object Detection
https://openreview.net/forum?id=TYXs_y84xRj
AUTHORS: Wu Xiongwei, Steven HOI, Doyen Sahoo
HIGHLIGHT: In this paper, we propose a new anchor-free keypoint based detector ``"PolarNet", where keypoints are represented as a set of Polar coordinates instead of Cartesian coordinates.

858, TITLE: Simple Spectral Graph Convolution
https://openreview.net/forum?id=CYO5T-YjWZV
AUTHORS: Hao Zhu, Piotr Koniusz
HIGHLIGHT: It is a simple and efficient method for graph convolution based on the Markov Diffusion Kernel which works well in different tasks under unsupervised, semi-supervised and supervised setting.

859, TITLE: Explainable Subgraph Reasoning for Forecasting on Temporal Knowledge Graphs
https://openreview.net/forum?id=pGIHq1m7PU
AUTHORS: Zhen Han, Peng Chen, Yunpu Ma, Volker Tresp
HIGHLIGHT: Especially, we propose a temporal relational attention mechanism and a novel reverse representation update scheme to guide the extraction of an enclosing subgraph around the query.

860, TITLE: EVALUATION OF NEURAL ARCHITECTURES TRAINED WITH SQUARE LOSS VS CROSS-ENTROPY IN CLASSIFICATION TASKS
https://openreview.net/forum?id=hsFN92eQEla
AUTHORS: Like Hui, Mikhail Belkin
HIGHLIGHT: An experimental evaluation of neural architectures in classification tasks shows that training with square loss produces better results than the cross-entropy in the majority of NLP and ASR experiments.