

- 1, TITLE: PAC-Bayesian Bounds on Rate-Efficient Classifiers
<https://proceedings.mlr.press/v162/abbas22a.html>
AUTHORS: Alhabib Abbas, Yiannis Andreopoulos
HIGHLIGHT: We derive analytic bounds on the noise invariance of majority vote classifiers operating on compressed inputs.

- 2, TITLE: Sharp-MAML: Sharpness-Aware Model-Agnostic Meta Learning
<https://proceedings.mlr.press/v162/abbas22b.html>
AUTHORS: Momin Abbas, Quan Xiao, Lisha Chen, Pin-Yu Chen, Tianyi Chen
HIGHLIGHT: Specifically, the loss landscape of MAML is much complex with possibly many more saddle points and local minima than its empirical risk minimization counterpart. To address this challenge, we leverage the recently invented sharpness-aware minimization and develop a sharpness-aware MAML approach that we term Sharp-MAML.

- 3, TITLE: An Initial Alignment between Neural Network and Target is Needed for Gradient Descent to Learn
<https://proceedings.mlr.press/v162/abbe22a.html>
AUTHORS: Emmanuel Abbe, Elisabetta Cornacchia, Jan Hazla, Christopher Marquis
HIGHLIGHT: This paper introduces the notion of “Initial Alignment” (INAL) between a neural network at initialization and a target function.

- 4, TITLE: Active Sampling for Min-Max Fairness
<https://proceedings.mlr.press/v162/abernethy22a.html>
AUTHORS: Jacob D Abernethy, Pranjali Awasthi, Matth?us Kleindessner, Jamie Morgenstern, Chris Russell, Jie Zhang
HIGHLIGHT: We propose simple active sampling and reweighting strategies for optimizing min-max fairness that can be applied to any classification or regression model learned via loss minimization.

- 5, TITLE: Meaningfully debugging model mistakes using conceptual counterfactual explanations
<https://proceedings.mlr.press/v162/abid22a.html>
AUTHORS: Abubakar Abid, Mert Yuksekogunul, James Zou
HIGHLIGHT: In this paper, we propose a systematic approach, conceptual counterfactual explanations (CCE), that explains why a classifier makes a mistake on a particular test sample(s) in terms of human-understandable concepts (e.g. this zebra is misclassified as a dog because of faint stripes).

- 6, TITLE: Batched Dueling Bandits
<https://proceedings.mlr.press/v162/agarwal22a.html>
AUTHORS: Arpit Agarwal, Rohan Ghuge, Viswanath Nagarajan
HIGHLIGHT: We study the batched K-armed dueling bandit problem under two standard settings: (i) existence of a Condorcet winner, and (ii) strong stochastic transitivity and stochastic triangle inequality.

- 7, TITLE: Hierarchical Shrinkage: Improving the accuracy and interpretability of tree-based models.
<https://proceedings.mlr.press/v162/agarwal22b.html>
AUTHORS: Abhinav Agarwal, Yan Shuo Tan, Omer Ronen, Chandan Singh, Bin Yu
HIGHLIGHT: We introduce Hierarchical Shrinkage (HS), a post-hoc algorithm which regularizes the tree not by altering its structure, but by shrinking the prediction over each leaf toward the sample means over each of its ancestors, with weights depending on a single regularization parameter and the number of samples in each ancestor.

- 8, TITLE: Deep equilibrium networks are sensitive to initialization statistics
<https://proceedings.mlr.press/v162/agarwala22a.html>
AUTHORS: Atish Agarwala, Samuel S Schoenholz
HIGHLIGHT: We show that DEQs are sensitive to the higher order statistics of the matrix families from which they are initialized.

- 9, TITLE: Learning of Cluster-based Feature Importance for Electronic Health Record Time-series
<https://proceedings.mlr.press/v162/aguilar22a.html>
AUTHORS: Henrique Aguiar, Mauro Santos, Peter Watkinson, Tingting Zhu
HIGHLIGHT: We propose a supervised deep learning model to cluster EHR data based on the identification of clinically understandable phenotypes with regard to both outcome prediction and patient trajectory.

- 10, TITLE: On the Convergence of the Shapley Value in Parametric Bayesian Learning Games
<https://proceedings.mlr.press/v162/agussurja22a.html>
AUTHORS: Lucas Agussurja, Xinyi Xu, Bryan Kian Hsiang Low
HIGHLIGHT: In this paper, we establish the convergence property of the Shapley value in parametric Bayesian learning games where players perform a Bayesian inference using their combined data, and the posterior-prior KL divergence is used as the characteristic function.

- 11, TITLE: Individual Preference Stability for Clustering
<https://proceedings.mlr.press/v162/ahmadi22a.html>
AUTHORS: Saba Ahmadi, Pranjali Awasthi, Samir Khuller, Matth?us Kleindessner, Jamie Morgenstern, Pattara Sukprasert, Ali Vakilian
HIGHLIGHT: In this paper, we propose a natural notion of individual preference (IP) stability for clustering, which asks that every data point, on average, is closer to the points in its own cluster than to the points in any other cluster.

- 12, TITLE: Understanding the unstable convergence of gradient descent
<https://proceedings.mlr.press/v162/ahn22a.html>
AUTHORS: Kwangjun Ahn, Jingzhao Zhang, Suvrit Sra
HIGHLIGHT: However, many works have observed that in machine learning applications step sizes often do not fulfill this condition, yet (stochastic) gradient descent still converges, albeit in an unstable manner. We investigate this unstable convergence phenomenon from first principles, and discuss key causes behind it.
- 13, TITLE: Minimum Cost Intervention Design for Causal Effect Identification
<https://proceedings.mlr.press/v162/akbari22a.html>
AUTHORS: Sina Akbari, Jalal Etesami, Negar Kiyavash
HIGHLIGHT: In this work, we consider the problem of designing the collection of interventions with the minimum cost to identify the desired effect.
- 14, TITLE: How Faithful is your Synthetic Data? Sample-level Metrics for Evaluating and Auditing Generative Models
<https://proceedings.mlr.press/v162/alaa22a.html>
AUTHORS: Ahmed Alaa, Boris Van Breugel, Evgeny S. Saveliev, Mihaela van der Schaar
HIGHLIGHT: In this paper, we introduce a 3-dimensional evaluation metric, (α -Precision, β -Recall, Authenticity), that characterizes the fidelity, diversity and generalization performance of any generative model in a domain-agnostic fashion.
- 15, TITLE: A Natural Actor-Critic Framework for Zero-Sum Markov Games
<https://proceedings.mlr.press/v162/alacaoglu22a.html>
AUTHORS: Ahmet Alacaoglu, Luca Viano, Niao He, Volkan Cevher
HIGHLIGHT: We introduce algorithms based on natural actor-critic and analyze their sample complexity for solving two player zero-sum Markov games in the tabular case.
- 16, TITLE: Deploying Convolutional Networks on Untrusted Platforms Using 2D Holographic Reduced Representations
<https://proceedings.mlr.press/v162/alam22a.html>
AUTHORS: Mohammad Mahmudul Alam, Edward Raff, Tim Oates, James Holt
HIGHLIGHT: By leveraging Holographic Reduced Representations (HRRs), we create a neural network with a pseudo-encryption style defense that empirically shows robustness to attack, even under threat models that unrealistically favor the adversary.
- 17, TITLE: Optimistic Linear Support and Successor Features as a Basis for Optimal Policy Transfer
<https://proceedings.mlr.press/v162/alegre22a.html>
AUTHORS: Lucas Nunes Alegre, Ana Bazzan, Bruno C. Da Silva
HIGHLIGHT: However, the identified solutions are not guaranteed to be optimal. We introduce a novel algorithm that addresses this limitation.
- 18, TITLE: Structured Stochastic Gradient MCMC
<https://proceedings.mlr.press/v162/alexos22a.html>
AUTHORS: Antonios Alexos, Alex J Boyd, Stephan Mandt
HIGHLIGHT: Unfortunately, VI makes strong assumptions on both the factorization and functional form of the posterior. To relax these assumptions, this work proposes a new non-parametric variational inference scheme that combines ideas from both SGCMC and coordinate-ascent VI.
- 19, TITLE: XAI for Transformers: Better Explanations through Conservative Propagation
<https://proceedings.mlr.press/v162/ali22a.html>
AUTHORS: Ameen Ali, Thomas Schnake, Oliver Eberle, Gr?goire Montavon, Klaus-Robert M?ller, Lior Wolf
HIGHLIGHT: We show that the gradient in a Transformer reflects the function only locally, and thus fails to reliably identify the contribution of input features to the prediction. We identify Attention Heads and LayerNorm as main reasons for such unreliable explanations and propose a more stable way for propagation through these layers.
- 20, TITLE: RUMs from Head-to-Head Contests
<https://proceedings.mlr.press/v162/almanza22a.html>
AUTHORS: Matteo Almanza, Flavio Chierichetti, Ravi Kumar, Alessandro Panconesi, Andrew Tomkins
HIGHLIGHT: In this paper, we focus on slates of size two representing head-to-head contests.
- 21, TITLE: Neuro-Symbolic Language Modeling with Automaton-augmented Retrieval
<https://proceedings.mlr.press/v162/alon22a.html>
AUTHORS: Uri Alon, Frank Xu, Junxian He, Sudipta Sengupta, Dan Roth, Graham Neubig
HIGHLIGHT: In this paper, we present RetoMaton - retrieval automaton - which approximates the datastore search, based on (1) saving pointers between consecutive datastore entries, and (2) clustering of entries into "states".
- 22, TITLE: Minimax Classification under Concept Drift with Multidimensional Adaptation and Performance Guarantees
<https://proceedings.mlr.press/v162/alvarez22a.html>
AUTHORS: Ver?nica ?lvarez, Santiago Mazuelas, Jose A Lozano
HIGHLIGHT: This paper presents adaptive minimax risk classifiers (AMRCs) that account for multidimensional time changes by means of a multivariate and high-order tracking of the time-varying underlying distribution.
- 23, TITLE: Scalable First-Order Bayesian Optimization via Structured Automatic Differentiation

<https://proceedings.mlr.press/v162/ament22a.html>

AUTHORS: Sebastian E Ament, Carla P Gomes

HIGHLIGHT: Here, we observe that a wide range of kernels gives rise to structured matrices, enabling an exact $O(n^2d)$ matrix-vector multiply for gradient observations and $O(n^2d^2)$ for Hessian observations. Beyond canonical kernel classes, we derive a programmatic approach to leveraging this type of structure for transformations and combinations of the discussed kernel classes, which constitutes a structure-aware automatic differentiation algorithm.

24, TITLE: Public Data-Assisted Mirror Descent for Private Model Training

<https://proceedings.mlr.press/v162/amid22a.html>

AUTHORS: Ehsan Amid, Arun Ganesh, Rajiv Mathews, Swaroop Ramaswamy, Shuang Song, Thomas Steinke, Thomas Steinke, Vinith M Suriyakumar, Om Thakkar, Abhradeep Thakurta

HIGHLIGHT: In this paper, we revisit the problem of using in-distribution public data to improve the privacy/utility trade-offs for differentially private (DP) model training.

25, TITLE: On Last-Iterate Convergence Beyond Zero-Sum Games

<https://proceedings.mlr.press/v162/anagnostides22a.html>

AUTHORS: Ioannis Anagnostides, Ioannis Panageas, Gabriele Farina, Tuomas Sandholm

HIGHLIGHT: In this paper we provide new results and techniques that apply to broader families of games and learning dynamics.

26, TITLE: Online Algorithms with Multiple Predictions

<https://proceedings.mlr.press/v162/anand22a.html>

AUTHORS: Keerti Anand, Rong Ge, Amit Kumar, Debmalya Panigrahi

HIGHLIGHT: We give a generic algorithmic framework for online covering problems with multiple predictions that obtains an online solution that is competitive against the performance of the best solution obtained from the predictions.

27, TITLE: Learning to Hash Robustly, Guaranteed

<https://proceedings.mlr.press/v162/andoni22a.html>

AUTHORS: Alexandr Andoni, Daniel Beaglehole

HIGHLIGHT: In this paper, we design an NNS algorithm for the Hamming space that has worst-case guarantees essentially matching that of theoretical algorithms, while optimizing the hashing to the structure of the dataset (think instance-optimal algorithms) for performance on the minimum-performing query.

28, TITLE: Set Based Stochastic Subsampling

<https://proceedings.mlr.press/v162/andreis22a.html>

AUTHORS: Bruno Andreis, Seanie Lee, A. Tuan Nguyen, Juho Lee, Eunho Yang, Sung Ju Hwang

HIGHLIGHT: Deep models are designed to operate on huge volumes of high dimensional data such as images. In order to reduce the volume of data these models must process, we propose a set-based two-stage end-to-end neural subsampling model that is jointly optimized with an arbitrary downstream task network (e.g. classifier).

29, TITLE: Towards Understanding Sharpness-Aware Minimization

<https://proceedings.mlr.press/v162/andriushchenko22a.html>

AUTHORS: Maksym Andriushchenko, Nicolas Flammarion

HIGHLIGHT: We argue that the existing justifications for the success of SAM which are based on a PAC-Bayes generalization bound and the idea of convergence to flat minima are incomplete.

30, TITLE: Fair and Fast k-Center Clustering for Data Summarization

<https://proceedings.mlr.press/v162/angelidakis22a.html>

AUTHORS: Haris Angelidakis, Adam Kurpisz, Leon Sering, Rico Zenklusen

HIGHLIGHT: We consider two key issues faced by many clustering methods when used for data summarization, namely (a) an unfair representation of “demographic groups” and (b) distorted summarizations, where data points in the summary represent subsets of the original data of vastly different sizes.

31, TITLE: Interactive Correlation Clustering with Existential Cluster Constraints

<https://proceedings.mlr.press/v162/angell22a.html>

AUTHORS: Rico Angell, Nicholas Monath, Nishant Yadav, Andrew McCallum

HIGHLIGHT: In this paper, we introduce existential cluster constraints: a new form of feedback where users indicate the features of desired clusters.

32, TITLE: Image-to-Image Regression with Distribution-Free Uncertainty Quantification and Applications in Imaging

<https://proceedings.mlr.press/v162/angelopoulos22a.html>

AUTHORS: Anastasios N Angelopoulos, Amit Pal Kohli, Stephen Bates, Michael Jordan, Jitendra Malik, Thayer Alshaabi, Srigokul Upadhyayula, Yaniv Romano

HIGHLIGHT: Current algorithms, however, do not generally offer statistical guarantees that protect against a model’s mistakes and hallucinations. To address this, we develop uncertainty quantification techniques with rigorous statistical guarantees for image-to-image regression problems.

33, TITLE: AdaGrad Avoids Saddle Points

<https://proceedings.mlr.press/v162/antonakopoulos22a.html>

AUTHORS: Kimon Antonakopoulos, Panayotis Mertikopoulos, Georgios Piliouras, Xiao Wang

HIGHLIGHT: In this paper, we focus on the AdaGrad family of algorithms - from scalar to full-matrix preconditioning - and we examine the question of whether the method's trajectories avoid saddle points.

34, **TITLE:** UnderGrad: A Universal Black-Box Optimization Method with Almost Dimension-Free Convergence Rate Guarantees

<https://proceedings.mlr.press/v162/antonakopoulos22b.html>

AUTHORS: Kimon Antonakopoulos, Dong Quan Vu, Volkan Cevher, Kfir Levy, Panayotis Mertikopoulos

HIGHLIGHT: Our paper aims to bridge this gap by providing a scalable universal method - dubbed UnDERGrad - which enjoys an almost dimension-free oracle complexity in problems with a favorable geometry (like the simplex, ℓ_1 -ball or trace-constraints), while retaining the order-optimal dependence on T described above.

35, **TITLE:** Adapting the Linearised Laplace Model Evidence for Modern Deep Learning

<https://proceedings.mlr.press/v162/antoran22a.html>

AUTHORS: Javier Antoran, David Janz, James U Allingham, Erik Daxberger, Riccardo Rb Barbano, Eric Nalisnick, Jose Miguel Hernandez-Lobato

HIGHLIGHT: In this work, we examine the assumptions behind this method, particularly in conjunction with model selection.

36, **TITLE:** EAT-C: Environment-Adversarial sub-Task Curriculum for Efficient Reinforcement Learning

<https://proceedings.mlr.press/v162/ao22a.html>

AUTHORS: Shuang Ao, Tianyi Zhou, Jing Jiang, Guodong Long, Xuan Song, Chengqi Zhang

HIGHLIGHT: Reinforcement learning (RL) is inefficient on long-horizon tasks due to sparse rewards and its policy can be fragile to slightly perturbed environments. We address these challenges via a curriculum of tasks with coupled environments, generated by two policies trained jointly with RL: (1) a co-operative planning policy recursively decomposing a hard task into a coarse-to-fine sub-task tree; and (2) an adversarial policy modifying the environment in each sub-task.

37, **TITLE:** Online Balanced Experimental Design

<https://proceedings.mlr.press/v162/arbour22a.html>

AUTHORS: David Arbour, Drew Dimmery, Tung Mai, Anup Rao

HIGHLIGHT: In this work, we present algorithms that build on recent advances in online discrepancy minimization which accommodate both arbitrary treatment probabilities and multiple treatments.

38, **TITLE:** VariGrow: Variational Architecture Growing for Task-Agnostic Continual Learning based on Bayesian Novelty

<https://proceedings.mlr.press/v162/ardywibowo22a.html>

AUTHORS: Randy Ardywibowo, Zepeng Huo, Zhangyang Wang, Bobak J Mortazavi, Shuai Huang, Xiaoning Qian

HIGHLIGHT: This paper proposes a variational architecture growing framework dubbed VariGrow.

39, **TITLE:** Thresholded Lasso Bandit

<https://proceedings.mlr.press/v162/ariu22a.html>

AUTHORS: Kaito Ariu, Kenshi Abe, Alexandre Proutiere

HIGHLIGHT: In this paper, we revisit the regret minimization problem in sparse stochastic contextual linear bandits, where feature vectors may be of large dimension d , but where the reward function depends on a few, say d_0 , of these features only.

40, **TITLE:** Gradient Based Clustering

<https://proceedings.mlr.press/v162/armacki22a.html>

AUTHORS: Aleksandar Armacki, Dragana Bajovic, Dusan Jakovetic, Soumya Kar

HIGHLIGHT: We propose a general approach for distance based clustering, using the gradient of the cost function that measures clustering quality with respect to cluster assignments and cluster center positions.

41, **TITLE:** Understanding Gradient Descent on the Edge of Stability in Deep Learning

<https://proceedings.mlr.press/v162/arora22a.html>

AUTHORS: Sanjeev Arora, Zhiyuan Li, Abhishek Panigrahi

HIGHLIGHT: The current paper mathematically analyzes a new mechanism of implicit regularization in the EoS phase, whereby GD updates due to non-smooth loss landscape turn out to evolve along some deterministic flow on the manifold of minimum loss.

42, **TITLE:** Private optimization in the interpolation regime: faster rates and hardness results

<https://proceedings.mlr.press/v162/asi22a.html>

AUTHORS: Hilal Asi, Karan Chadha, Gary Cheng, John Duchi

HIGHLIGHT: In this paper, we investigate differentially private stochastic optimization in the interpolation regime.

43, **TITLE:** Optimal Algorithms for Mean Estimation under Local Differential Privacy

<https://proceedings.mlr.press/v162/asi22b.html>

AUTHORS: Hilal Asi, Vitaly Feldman, Kunal Talwar

HIGHLIGHT: In this work, we investigate the question of designing the randomizer with the smallest variance.

44, **TITLE:** Asymptotically-Optimal Gaussian Bandits with Side Observations

<https://proceedings.mlr.press/v162/atsidakou22a.html>

AUTHORS: Alexia Atsidakou, Orestis Papadigenopoulos, Constantine Caramanis, Sujay Sanghavi, Sanjay Shakkottai

HIGHLIGHT: The LP optimizes the cost (regret) required to reliably estimate the suboptimality gap of each arm. This LP lower bound motivates our main contribution: the first known asymptotically optimal algorithm for this general setting.

- 45, TITLE: Congested Bandits: Optimal Routing via Short-term Resets
<https://proceedings.mlr.press/v162/awasthi22a.html>
AUTHORS: Pranjali Awasthi, Kush Bhatia, Sreenivas Gollapudi, Kostas Kollias
HIGHLIGHT: For the multi-armed setup, we propose a UCB style algorithm and show that its policy regret scales as $\tilde{O}(\sqrt{K \Delta T})$. Motivated by this, we introduce the problem of Congested Bandits where each arm's reward is allowed to depend on the number of times it was played in the past Δ timesteps.
- 46, TITLE: Do More Negative Samples Necessarily Hurt In Contrastive Learning?
<https://proceedings.mlr.press/v162/awasthi22b.html>
AUTHORS: Pranjali Awasthi, Nishanth Dikkala, Pritish Kamath
HIGHLIGHT: We show in a simple theoretical setting, where positive pairs are generated by sampling from the underlying latent class (introduced by Saunshi et al. (ICML 2019)), that the downstream performance of the representation optimizing the (population) contrastive loss in fact does not degrade with the number of negative samples.
- 47, TITLE: H-Consistency Bounds for Surrogate Loss Minimizers
<https://proceedings.mlr.press/v162/awasthi22c.html>
AUTHORS: Pranjali Awasthi, Anqi Mao, Mehryar Mohri, Yutao Zhong
HIGHLIGHT: We present a detailed study of estimation errors in terms of surrogate loss estimation errors.
- 48, TITLE: Iterative Hard Thresholding with Adaptive Regularization: Sparser Solutions Without Sacrificing Runtime
<https://proceedings.mlr.press/v162/axiotis22a.html>
AUTHORS: Kyriakos Axiotis, Maxim Sviridenko
HIGHLIGHT: We propose a simple modification to the iterative hard thresholding (IHT) algorithm, which recovers asymptotically sparser solutions as a function of the condition number.
- 49, TITLE: Proving Theorems using Incremental Learning and Hindsight Experience Replay
<https://proceedings.mlr.press/v162/aygun22a.html>
AUTHORS: Eser Aygün, Ankit Anand, Laurent Orseau, Xavier Glorot, Stephen M McAleer, Vlad Firoiu, Lei M Zhang, Doina Precup, Shihab Mourad
HIGHLIGHT: In this paper, we adapt the idea of hindsight experience replay from reinforcement learning to the automated theorem proving domain, so as to use the intermediate data generated during unsuccessful proof attempts.
- 50, TITLE: Near-optimal rate of consistency for linear models with missing values
<https://proceedings.mlr.press/v162/ayme22a.html>
AUTHORS: Alexis Ayme, Claire Boyer, Aymeric Dieuleveut, Erwan Scornet
HIGHLIGHT: In this paper, we focus on the extensively-studied linear models, but in presence of missing values, which turns out to be quite a challenging task.
- 51, TITLE: How Tempering Fixes Data Augmentation in Bayesian Neural Networks
<https://proceedings.mlr.press/v162/bachmann22a.html>
AUTHORS: Gregor Bachmann, Lorenzo Noci, Thomas Hofmann
HIGHLIGHT: In this work we identify two interlaced factors concurrently influencing the strength of the cold posterior effect, namely the correlated nature of augmentations and the degree of invariance of the employed model to such transformations.
- 52, TITLE: ASAP.SGD: Instance-based Adaptiveness to Staleness in Asynchronous SGD
<https://proceedings.mlr.press/v162/backstrom22a.html>
AUTHORS: Karl Bockström, Marina Papatriantafidou, Philippas Tsigas
HIGHLIGHT: We introduce (i) ASAP.SGD , an analytical framework capturing necessary and desired properties of staleness-adaptive step size functions and (ii) $\text{tail-}\tau$, a method for utilizing key properties of the execution instance, generating a tailored strategy that not only dampens the impact of stale updates, but also leverages fresh ones.
- 53, TITLE: From Noisy Prediction to True Label: Noisy Prediction Calibration via Generative Model
<https://proceedings.mlr.press/v162/bae22a.html>
AUTHORS: Heesun Bae, Seungjae Shin, Byeonghu Na, Joonho Jang, Kyungwoo Song, Il-Chul Moon
HIGHLIGHT: We suggest a new branch of method, Noisy Prediction Calibration (NPC) in learning with noisy labels.
- 54, TITLE: data2vec: A General Framework for Self-supervised Learning in Speech, Vision and Language
<https://proceedings.mlr.press/v162/baevski22a.html>
AUTHORS: Alexei Baevski, Wei-Ning Hsu, Qiantong Xu, Arun Babu, Jiatao Gu, Michael Auli
HIGHLIGHT: To get us closer to general self-supervised learning, we present data2vec, a framework that uses the same learning method for either speech, NLP or computer vision.
- 55, TITLE: End-to-End Balancing for Causal Continuous Treatment-Effect Estimation
<https://proceedings.mlr.press/v162/bahadori22a.html>
AUTHORS: Taha Bahadori, Eric Tchetgen Tchetgen, David Heckerman
HIGHLIGHT: We propose a new theory for consistency of entropy balancing for continuous treatments.
- 56, TITLE: A Hierarchical Transitive-Aligned Graph Kernel for Un-attributed Graphs
<https://proceedings.mlr.press/v162/bai22a.html>

- AUTHORS: Lu Bai, Lixin Cui, Hancock Edwin
HIGHLIGHT: In this paper, we develop a new graph kernel, namely the Hierarchical Transitive-Aligned Kernel, by transitively aligning the vertices between graphs through a family of hierarchical prototype graphs.
- 57, TITLE: Near-Optimal Learning of Extensive-Form Games with Imperfect Information
<https://proceedings.mlr.press/v162/bai22b.html>
AUTHORS: Yu Bai, Chi Jin, Song Mei, Tiancheng Yu
HIGHLIGHT: We present the first line of algorithms that require only $\tilde{O}((XA+YB)^{\epsilon})$ episodes of play to find an ϵ -approximate Nash equilibrium in two-player zero-sum games, where X, Y are the number of information sets and A, B are the number of actions for the two players.
- 58, TITLE: Gaussian Mixture Variational Autoencoder with Contrastive Learning for Multi-Label Classification
<https://proceedings.mlr.press/v162/bai22c.html>
AUTHORS: Junwen Bai, Shufeng Kong, Carla P Gomes
HIGHLIGHT: We propose a novel contrastive learning boosted multi-label prediction model based on a Gaussian mixture variational autoencoder (C-GMVAE), which learns a multimodal prior space and employs a contrastive loss.
- 59, TITLE: AS³T: Alignment-Aware Acoustic and Text Pretraining for Speech Synthesis and Editing
<https://proceedings.mlr.press/v162/bai22d.html>
AUTHORS: He Bai, Renjie Zheng, Junkun Chen, Mingbo Ma, Xintong Li, Liang Huang
HIGHLIGHT: However, all the above tasks are in the direction of speech understanding, but for the inverse direction, speech synthesis, the potential of representation learning is yet to be realized, due to the challenging nature of generating high-quality speech. To address this problem, we propose our framework, Alignment-Aware Acoustic-Text Pretraining (AS³T), which reconstructs masked acoustic signals with text input and acoustic-text alignment during training.
- 60, TITLE: Stability Based Generalization Bounds for Exponential Family Langevin Dynamics
<https://proceedings.mlr.press/v162/banerjee22a.html>
AUTHORS: Arindam Banerjee, Tiancong Chen, Xinyan Li, Yingxue Zhou
HIGHLIGHT: In this paper, we unify and substantially generalize stability based generalization bounds and make three technical contributions.
- 61, TITLE: Certified Neural Network Watermarks with Randomized Smoothing
<https://proceedings.mlr.press/v162/bansal22a.html>
AUTHORS: Arpit Bansal, Ping-Yeh Chiang, Michael J Curry, Rajiv Jain, Curtis Wigington, Varun Manjunatha, John P Dickerson, Tom Goldstein
HIGHLIGHT: In this paper, we propose the first certifiable watermarking method.
- 62, TITLE: Data Scaling Laws in NMT: The Effect of Noise and Architecture
<https://proceedings.mlr.press/v162/bansal22b.html>
AUTHORS: Yamini Bansal, Behrooz Ghorbani, Ankush Garg, Biao Zhang, Colin Cherry, Behnam Neyshabur, Orhan Firat
HIGHLIGHT: In this work, we study the effect of varying the architecture and training data quality on the data scaling properties of Neural Machine Translation (NMT).
- 63, TITLE: Learning Stable Classifiers by Transferring Unstable Features
<https://proceedings.mlr.press/v162/bao22a.html>
AUTHORS: Yujia Bao, Shiyu Chang, Dr.Regina Barzilay
HIGHLIGHT: In this work, we explicitly inform the target classifier about unstable features in the source tasks.
- 64, TITLE: Fast Composite Optimization and Statistical Recovery in Federated Learning
<https://proceedings.mlr.press/v162/bao22b.html>
AUTHORS: Yajie Bao, Michael Crawshaw, Shan Luo, Mingrui Liu
HIGHLIGHT: From optimization upfront, we propose a new algorithm named Fast Federated Dual Averaging for strongly convex and smooth loss and establish state-of-the-art iteration and communication complexity in the composite setting.
- 65, TITLE: Generative Modeling for Multi-task Visual Learning
<https://proceedings.mlr.press/v162/bao22c.html>
AUTHORS: Zhipeng Bao, Martial Hebert, Yu-Xiong Wang
HIGHLIGHT: In this paper, motivated by multi-task learning of shareable feature representations, we consider a novel problem of learning a shared generative model that is useful across various visual perception tasks.
- 66, TITLE: Estimating the Optimal Covariance with Imperfect Mean in Diffusion Probabilistic Models
<https://proceedings.mlr.press/v162/bao22d.html>
AUTHORS: Fan Bao, Chongxuan Li, Jiacheng Sun, Jun Zhu, Bo Zhang
HIGHLIGHT: In this work, we consider diagonal and full covariances to improve the expressive power of DPMs.
- 67, TITLE: On the Surrogate Gap between Contrastive and Supervised Losses
<https://proceedings.mlr.press/v162/bao22e.html>
AUTHORS: Han Bao, Yoshihiro Nagano, Kento Nozawa
HIGHLIGHT: Following the simplified setting where positive pairs are drawn from the true distribution (not generated by data augmentation; as supposed in previous studies), this study establishes surrogate upper and lower bounds for the downstream

classification loss for all negative sample sizes that best explain the empirical observations on the negative sample size in the earlier studies.

- 68, TITLE: Representation Topology Divergence: A Method for Comparing Neural Network Representations.
<https://proceedings.mlr.press/v162/barannikov22a.html>
AUTHORS: Serguei Barannikov, Ilya Trofimov, Nikita Balabin, Evgeny Burnaev
HIGHLIGHT: We propose a method for comparing two data representations.
- 69, TITLE: Sparse Mixed Linear Regression with Guarantees: Taming an Intractable Problem with Invex Relaxation
<https://proceedings.mlr.press/v162/barik22a.html>
AUTHORS: Adarsh Barik, Jean Honorio
HIGHLIGHT: In this paper, we study the problem of sparse mixed linear regression on an unlabeled dataset that is generated from linear measurements from two different regression parameter vectors.
- 70, TITLE: Neural Fisher Discriminant Analysis: Optimal Neural Network Embeddings in Polynomial Time
<https://proceedings.mlr.press/v162/bartan22a.html>
AUTHORS: Burak Bartan, Mert Pilanci
HIGHLIGHT: We introduce a natural extension of FLDA that employs neural networks, called Neural Fisher Discriminant Analysis (NFDA).
- 71, TITLE: Fictitious Play and Best-Response Dynamics in Identical Interest and Zero-Sum Stochastic Games
<https://proceedings.mlr.press/v162/audin22a.html>
AUTHORS: Lucas Baudin, Rida Laraki
HIGHLIGHT: This paper proposes an extension of a popular decentralized discrete-time learning procedure when repeating a static game called fictitious play (FP) (Brown, 1951; Robinson, 1951) to a dynamic model called discounted stochastic game (Shapley, 1953).
- 72, TITLE: Information Discrepancy in Strategic Learning
<https://proceedings.mlr.press/v162/bechavod22a.html>
AUTHORS: Yahav Bechavod, Chara Podimata, Steven Wu, Juba Ziani
HIGHLIGHT: We initiate the study of the effects of non-transparency in decision rules on individuals' ability to improve in strategic learning settings.
- 73, TITLE: On the Hidden Biases of Policy Mirror Ascent in Continuous Action Spaces
<https://proceedings.mlr.press/v162/bedi22a.html>
AUTHORS: Amrit Singh Bedi, Souradip Chakraborty, Anjali Parayil, Brian M Sadler, Pratap Tokekar, Alec Koppel
HIGHLIGHT: To mitigate this hidden bias, heavy-tailed policy parameterizations may be used, which exhibit a bounded score function, but doing so can cause instability in algorithmic updates. To address these issues, in this work, we study the convergence of policy gradient algorithms under heavy-tailed parameterizations, which we propose to stabilize with a combination of mirror ascent-type updates and gradient tracking.
- 74, TITLE: Imitation Learning by Estimating Expertise of Demonstrators
<https://proceedings.mlr.press/v162/beliaev22a.html>
AUTHORS: Mark Beliaev, Andy Shih, Stefano Ermon, Dorsa Sadigh, Ramtin Pedarsani
HIGHLIGHT: In this work, we show that unsupervised learning over demonstrator expertise can lead to a consistent boost in the performance of imitation learning algorithms.
- 75, TITLE: Matching Normalizing Flows and Probability Paths on Manifolds
<https://proceedings.mlr.press/v162/ben-hamu22a.html>
AUTHORS: Heli Ben-Hamu, Samuel Cohen, Joey Bose, Brandon Amos, Maximilian Nickel, Aditya Grover, Ricky T. Q. Chen, Yaron Lipman
HIGHLIGHT: We propose to train CNFs on manifolds by minimizing probability path divergence (PPD), a novel family of divergences between the probability density path generated by the CNF and a target probability density path.
- 76, TITLE: Stochastic Contextual Dueling Bandits under Linear Stochastic Transitivity Models
<https://proceedings.mlr.press/v162/bengs22a.html>
AUTHORS: Viktor Bengs, Aadirupa Saha, Eyke Hüllermeier
HIGHLIGHT: We propose a computationally efficient algorithm, $\text{Algo}\{\text{CoLSTIM}\}$, which makes its choice based on imitating the feedback process using perturbed context-dependent utility estimates of the underlying CoLST model.
- 77, TITLE: Neural Inverse Kinematic
<https://proceedings.mlr.press/v162/bensadoun22a.html>
AUTHORS: Raphael Bensadoun, Shir Gur, Nitsan Blau, Lior Wolf
HIGHLIGHT: In this work, we propose a neural IK method that employs the hierarchical structure of the problem to sequentially sample valid joint angles conditioned on the desired position and on the preceding joints along the chain.
- 78, TITLE: Volatility Based Kernels and Moving Average Means for Accurate Forecasting with Gaussian Processes
<https://proceedings.mlr.press/v162/benton22a.html>
AUTHORS: Gregory Benton, Wesley Maddox, Andrew Gordon Wilson

HIGHLIGHT: To address this fundamental limitation, we show how to re-cast a class of stochastic volatility models as a hierarchical Gaussian process (GP) model with specialized covariance functions.

79, **TITLE:** Gradient Descent on Neurons and its Link to Approximate Second-order Optimization
<https://proceedings.mlr.press/v162/benzing22a.html>

AUTHORS: Frederik Benzing

HIGHLIGHT: This challenges widely held beliefs and immediately raises the question why KFAC performs so well. Towards answering this question we present evidence strongly suggesting that KFAC approximates a first-order algorithm, which performs gradient descent on neurons rather than weights.

80, **TITLE:** Safe Learning in Tree-Form Sequential Decision Making: Handling Hard and Soft Constraints
<https://proceedings.mlr.press/v162/bernasconi22a.html>

AUTHORS: Martino Bernasconi, Federico Cacciamani, Matteo Castiglioni, Alberto Marchesi, Nicola Gatti, Francesco Trov?

HIGHLIGHT: We study the hard-threshold problem of achieving sublinear regret while guaranteeing that the threshold constraint is satisfied at every iteration with high probability.

81, **TITLE:** Skin Deep Unlearning: Artefact and Instrument Debiasing in the Context of Melanoma Classification

<https://proceedings.mlr.press/v162/bevan22a.html>

AUTHORS: Peter Bevan, Amir Atapour-Abarghouei

HIGHLIGHT: In this work, we robustly remove bias and spurious variation from an automated melanoma classification pipeline using two leading bias unlearning techniques.

82, **TITLE:** Approximate Bayesian Computation with Domain Expert in the Loop

<https://proceedings.mlr.press/v162/bharti22a.html>

AUTHORS: Ayush Bharti, Louis Filstroff, Samuel Kaski

HIGHLIGHT: In this work, we introduce an active learning method for ABC statistics selection which reduces the domain expert's work considerably.

83, **TITLE:** Minimax M-estimation under Adversarial Contamination

<https://proceedings.mlr.press/v162/bhatt22a.html>

AUTHORS: Sujay Bhatt, Guanhua Fang, Ping Li, Gennady Samorodnitsky

HIGHLIGHT: To illustrate the usefulness of the derived robust M-estimator in an online setting, we present a bandit algorithm for the partially identifiable best arm identification problem that improves upon the sample complexity of the state of the art algorithms.

84, **TITLE:** Nearly Optimal Catoni's M-estimator for Infinite Variance

<https://proceedings.mlr.press/v162/bhatt22b.html>

AUTHORS: Sujay Bhatt, Guanhua Fang, Ping Li, Gennady Samorodnitsky

HIGHLIGHT: In this paper, we extend the remarkable M-estimator of Catoni & Cat12 to situations where the variance is infinite.

85, **TITLE:** Personalization Improves Privacy-Accuracy Tradeoffs in Federated Learning

<https://proceedings.mlr.press/v162/bietti22a.html>

AUTHORS: Alberto Bietti, Chen-Yu Wei, Miroslav Dudik, John Langford, Steven Wu

HIGHLIGHT: In this paper, we study stochastic optimization algorithms for a personalized federated learning setting involving local and global models subject to user-level (joint) differential privacy.

86, **TITLE:** Non-Vacuous Generalisation Bounds for Shallow Neural Networks

<https://proceedings.mlr.press/v162/biggs22a.html>

AUTHORS: Felix Biggs, Benjamin Guedj

HIGHLIGHT: We focus on a specific class of shallow neural networks with a single hidden layer, namely those with SL_2 -normalised data and either a sigmoid-shaped Gaussian error function ("erf") activation or a Gaussian Error Linear Unit (GELU) activation. For these networks, we derive new generalisation bounds through the PAC-Bayesian theory; unlike most existing such bounds they apply to neural networks with deterministic rather than randomised parameters.

87, **TITLE:** Structure-preserving GANs

<https://proceedings.mlr.press/v162/birrell22a.html>

AUTHORS: Jeremiah Birrell, Markos Katsoulakis, Luc Rey-Bellet, Wei Zhu

HIGHLIGHT: We introduce structure-preserving GANs as a data-efficient framework for learning distributions with additional structure such as group symmetry, by developing new variational representations for divergences.

88, **TITLE:** Scalable Spike-and-Slab

<https://proceedings.mlr.press/v162/biswas22a.html>

AUTHORS: Niloy Biswas, Lester Mackey, Xiao-Li Meng

HIGHLIGHT: In this article, we propose Scalable Spike-and-Slab (S^3), a scalable Gibbs sampling implementation for high-dimensional Bayesian regression with the continuous spike-and-slab prior of George & McCulloch (1993).

89, **TITLE:** Breaking Down Out-of-Distribution Detection: Many Methods Based on OOD Training Data Estimate a Combination of the Same Core Quantities

- <https://proceedings.mlr.press/v162/bitterwolf22a.html>
AUTHORS: Julian Bitterwolf, Alexander Meinke, Maximilian Augustin, Matthias Hein
HIGHLIGHT: The goal of this paper is to recognize common objectives as well as to identify the implicit scoring functions of different OOD detection methods.
- 90, TITLE: A query-optimal algorithm for finding counterfactuals
<https://proceedings.mlr.press/v162/blanc22a.html>
AUTHORS: Guy Blanc, Caleb Koch, Jane Lange, Li-Yang Tan
HIGHLIGHT: We design an algorithm for finding counterfactuals with strong theoretical guarantees on its performance.
- 91, TITLE: Popular decision tree algorithms are provably noise tolerant
<https://proceedings.mlr.press/v162/blanc22b.html>
AUTHORS: Guy Blanc, Jane Lange, Ali Malik, Li-Yang Tan
HIGHLIGHT: Using the framework of boosting, we prove that all impurity-based decision tree learning algorithms, including the classic ID3, C4.5, and CART, are highly noise tolerant.
- 92, TITLE: Optimizing Sequential Experimental Design with Deep Reinforcement Learning
<https://proceedings.mlr.press/v162/blau22a.html>
AUTHORS: Tom Blau, Edwin V. Bonilla, Iadine Chades, Amir Dezfouli
HIGHLIGHT: However, these methods may not sufficiently explore the design space, require access to a differentiable probabilistic model and can only optimize over continuous design spaces. Here, we address these limitations by showing that the problem of optimizing policies can be reduced to solving a Markov decision process (MDP).
- 93, TITLE: Lagrangian Method for Q-Function Learning (with Applications to Machine Translation)
<https://proceedings.mlr.press/v162/bojun22a.html>
AUTHORS: Huang Bojun
HIGHLIGHT: This paper discusses a new approach to the fundamental problem of learning optimal Q-functions.
- 94, TITLE: Generalized Results for the Existence and Consistency of the MLE in the Bradley-Terry-Luce Model
<https://proceedings.mlr.press/v162/bong22a.html>
AUTHORS: Heejong Bong, Alessandro Rinaldo
HIGHLIGHT: In this paper, we study the performance of the Bradley-Terry-Luce model for ranking from pairwise comparison data under more realistic settings than those considered in the literature so far.
- 95, TITLE: How to Train Your Wide Neural Network Without Backprop: An Input-Weight Alignment Perspective
<https://proceedings.mlr.press/v162/boopathy22a.html>
AUTHORS: Akhilan Boopathy, Ila Fiete
HIGHLIGHT: Leveraging NTK theory, we show theoretically that gradient descent drives layerwise weight updates that are aligned with their input activity correlations weighted by error, and demonstrate empirically that the result also holds in finite-width wide networks.
- 96, TITLE: Improving Language Models by Retrieving from Trillions of Tokens
<https://proceedings.mlr.press/v162/borgeaud22a.html>
AUTHORS: Sebastian Borgeaud, Arthur Mensch, Jordan Hoffmann, Trevor Cai, Eliza Rutherford, Katie Millican, George Bm Van Den Driessche, Jean-Baptiste Lespiau, Bogdan Damoc, Aidan Clark, Diego De Las Casas, Aurelia Guy, Jacob Menick, Roman Ring, Tom Hennigan, Saffron Huang, Loren Maggiore, Chris Jones, Albin Cassirer, Andy Brock, Michela Paganini, Geoffrey Irving, Oriol Vinyals, Simon Osindero, Karen Simonyan, Jack Rae, Erich Elsen, Laurent Sifre
HIGHLIGHT: We enhance auto-regressive language models by conditioning on document chunks retrieved from a large corpus, based on local similarity with preceding tokens.
- 97, TITLE: Lie Point Symmetry Data Augmentation for Neural PDE Solvers
<https://proceedings.mlr.press/v162/brandstetter22a.html>
AUTHORS: Johannes Brandstetter, Max Welling, Daniel E Worrall
HIGHLIGHT: Thus, we are presented with a proverbial chicken-and-egg problem. In this paper, we present a method, which can partially alleviate this problem, by improving neural PDE solver sample complexity—Lie point symmetry data augmentation (LPSDA).
- 98, TITLE: An iterative clustering algorithm for the Contextual Stochastic Block Model with optimality guarantees
<https://proceedings.mlr.press/v162/braun22a.html>
AUTHORS: Guillaume Braun, Hemant Tyagi, Christophe Biernacki
HIGHLIGHT: We propose a new iterative algorithm to cluster networks with side information for nodes (in the form of covariates) and show that our algorithm is optimal under the Contextual Symmetric Stochastic Block Model.
- 99, TITLE: Tractable Dendritic RNNs for Reconstructing Nonlinear Dynamical Systems
<https://proceedings.mlr.press/v162/brenner22a.html>
AUTHORS: Manuel Brenner, Florian Hess, Jonas M Mikhaeil, Leonard F Bereska, Zahra Monfared, Po-Chen Kuo, Daniel Durstewitz
HIGHLIGHT: Motivated by the emerging principles of dendritic computation, we augment a dynamically interpretable and mathematically tractable piecewise-linear (PL) recurrent neural network (RNN) by a linear spline basis expansion.

- 100, TITLE: Learning to Predict Graphs with Fused Gromov-Wasserstein Barycenters
<https://proceedings.mlr.press/v162/brogat-motte22a.html>
AUTHORS: Luc Brogat-Motte, R?mi Flamary, Celine Brouard, Juho Rousu, Florence D?Alch?-Buc
HIGHLIGHT: This paper introduces a novel and generic framework to solve the flagship task of supervised labeled graph prediction by leveraging Optimal Transport tools.
- 101, TITLE: Efficient Learning of CNNs using Patch Based Features
<https://proceedings.mlr.press/v162/brutzkus22a.html>
AUTHORS: Alon Brutzkus, Amir Globerson, Eran Malach, Alon Regev Netser, Shai Shalev-Schwartz
HIGHLIGHT: Recent work has demonstrated the effectiveness of using patch based representations when learning from image data. Here we provide theoretical support for this observation, by showing that a simple semi-supervised algorithm that uses patch statistics can efficiently learn labels produced by a one-hidden-layer Convolutional Neural Network (CNN).
- 102, TITLE: Causal structure-based root cause analysis of outliers
<https://proceedings.mlr.press/v162/budhathoki22a.html>
AUTHORS: Kailash Budhathoki, Lenon Minorics, Patrick Bloebaum, Dominik Janzing
HIGHLIGHT: We present a formal method to identify "root causes" of outliers, amongst variables.
- 103, TITLE: IGLUE: A Benchmark for Transfer Learning across Modalities, Tasks, and Languages
<https://proceedings.mlr.press/v162/bugliarello22a.html>
AUTHORS: Emanuele Bugliarello, Fangyu Liu, Jonas Pfeiffer, Siva Reddy, Desmond Elliott, Edoardo Maria Ponti, Ivan Vulic
HIGHLIGHT: Due to the lack of a multilingual benchmark, however, vision-and-language research has mostly focused on English language tasks. To fill this gap, we introduce the Image-Grounded Language Understanding Evaluation benchmark.
- 104, TITLE: Interactive Inverse Reinforcement Learning for Cooperative Games
<https://proceedings.mlr.press/v162/buning22a.html>
AUTHORS: Thomas Kleine B?ning, Anne-Marie George, Christos Dimitrakakis
HIGHLIGHT: We study the problem of designing autonomous agents that can learn to cooperate effectively with a potentially suboptimal partner while having no access to the joint reward function.
- 105, TITLE: Convolutional and Residual Networks Provably Contain Lottery Tickets
<https://proceedings.mlr.press/v162/burkholz22a.html>
AUTHORS: Rebekka Burkholz
HIGHLIGHT: We prove that also modern architectures consisting of convolutional and residual layers that can be equipped with almost arbitrary activation functions can contain lottery tickets with high probability.
- 106, TITLE: Near-Optimal Algorithms for Autonomous Exploration and Multi-Goal Stochastic Shortest Path
<https://proceedings.mlr.press/v162/cai22a.html>
AUTHORS: Haoyuan Cai, Tengyu Ma, Simon Du
HIGHLIGHT: We introduce a new algorithm with stronger sample complexity bounds than existing ones.
- 107, TITLE: Convergence of Invariant Graph Networks
<https://proceedings.mlr.press/v162/cai22b.html>
AUTHORS: Chen Cai, Yusu Wang
HIGHLIGHT: In this paper, we investigate the convergence of one powerful GNN, Invariant Graph Network (IGN) over graphs sampled from graphons.
- 108, TITLE: Reinforcement Learning from Partial Observation: Linear Function Approximation with Provable Sample Efficiency
<https://proceedings.mlr.press/v162/cai22c.html>
AUTHORS: Qi Cai, Zhuoran Yang, Zhaoran Wang
HIGHLIGHT: In detail, we propose a reinforcement learning algorithm (Optimistic Exploration via Adversarial Integral Equation or OP-TENET) that attains an ϵ -optimal policy within $O(1/\epsilon^2)$ episodes.
- 109, TITLE: Scaling Gaussian Process Optimization by Evaluating a Few Unique Candidates Multiple Times
<https://proceedings.mlr.press/v162/calandriello22a.html>
AUTHORS: Daniele Calandriello, Luigi Carratino, Alessandro Lazaric, Michal Valko, Lorenzo Rosasco
HIGHLIGHT: We show that sequential black-box optimization based on GPs (GP-Opt) can be made efficient by sticking to a candidate solution for multiple evaluation steps and switch only when necessary.
- 110, TITLE: Adaptive Gaussian Process Change Point Detection
<https://proceedings.mlr.press/v162/caldarelli22a.html>
AUTHORS: Edoardo Caldarelli, Philippe Wenk, Stefan Bauer, Andreas Krause
HIGHLIGHT: Detecting change points in time series, i.e., points in time at which some observed process suddenly changes, is a fundamental task that arises in many real-world applications, with consequences for safety and reliability. In this work, we propose ADAGA, a novel Gaussian process-based solution to this problem, that leverages a powerful heuristics we developed based on statistical hypothesis testing.
- 111, TITLE: Measuring dissimilarity with diffeomorphism invariance

- <https://proceedings.mlr.press/v162/cantelobre22a.html>
AUTHORS: Th?ophile Cantelobre, Carlo Ciliberto, Benjamin Guedj, Alessandro Rudi
HIGHLIGHT: We introduce DID, a pairwise dissimilarity measure applicable to a wide range of data spaces, which leverages the data's internal structure to be invariant to diffeomorphisms.
- 112, TITLE: A Model-Agnostic Randomized Learning Framework based on Random Hypothesis Subspace Sampling
<https://proceedings.mlr.press/v162/cao22a.html>
AUTHORS: Yiting Cao, Chao Lan
HIGHLIGHT: We propose a model-agnostic randomized learning framework based on Random Hypothesis Subspace Sampling (RHSS).
- 113, TITLE: Gaussian Process Uniform Error Bounds with Unknown Hyperparameters for Safety-Critical Applications
<https://proceedings.mlr.press/v162/capone22a.html>
AUTHORS: Alexandre Capone, Armin Lederer, Sandra Hirche
HIGHLIGHT: However, state-of-the-art techniques for safety-critical settings hinge on the assumption that the kernel hyperparameters are known, which does not apply in general. To mitigate this, we introduce robust Gaussian process uniform error bounds in settings with unknown hyperparameters.
- 114, TITLE: Burst-Dependent Plasticity and Dendritic Amplification Support Target-Based Learning and Hierarchical Imitation Learning
<https://proceedings.mlr.press/v162/capone22b.html>
AUTHORS: Cristiano Capone, Cosimo Lupo, Paolo Muratore, Pier Stanislaw Paolucci
HIGHLIGHT: We propose a multi-compartment model of pyramidal neuron, in which bursts and dendritic input segregation give the possibility to plausibly support a biological target-based learning.
- 115, TITLE: A Marriage between Adversarial Team Games and 2-player Games: Enabling Abstractions, No-regret Learning, and Subgame Solving
<https://proceedings.mlr.press/v162/carminati22a.html>
AUTHORS: Luca Carminati, Federico Cacciamani, Marco Ciccone, Nicola Gatti
HIGHLIGHT: In particular, we propose a new, suitable game representation that we call team-public-information, in which a team is represented as a single coordinator who only knows information common to the whole team and prescribes to each member an action for any possible private state.
- 116, TITLE: RECAPP: Crafting a More Efficient Catalyst for Convex Optimization
<https://proceedings.mlr.press/v162/carmon22a.html>
AUTHORS: Yair Carmon, Arun Jambulapati, Yujia Jin, Aaron Sidford
HIGHLIGHT: In this work, we propose a novel Relaxed Error Criterion for Accelerated Proximal Point (RECAPP) that eliminates the need for high accuracy subproblem solutions.
- 117, TITLE: Estimating and Penalizing Induced Preference Shifts in Recommender Systems
<https://proceedings.mlr.press/v162/carroll22a.html>
AUTHORS: Micah D Carroll, Anca Dragan, Stuart Russell, Dylan Hadfield-Menell
HIGHLIGHT: We focus on induced preference shifts in users.
- 118, TITLE: YourTTS: Towards Zero-Shot Multi-Speaker TTS and Zero-Shot Voice Conversion for Everyone
<https://proceedings.mlr.press/v162/casanova22a.html>
AUTHORS: Edresson Casanova, Julian Weber, Christopher D Shulby, Arnaldo Candido Junior, Eren G?lge, Moacir A Ponti
HIGHLIGHT: Our method builds upon the VITS model and adds several novel modifications for zero-shot multi-speaker and multilingual training.
- 119, TITLE: The Infinite Contextual Graph Markov Model
<https://proceedings.mlr.press/v162/castellana22a.html>
AUTHORS: Daniele Castellana, Federico Errica, Davide Bacciu, Alessio Micheli
HIGHLIGHT: As with most Deep Graph Networks, an inherent limitation is the need to perform an extensive model selection to choose the proper size of each layer's latent representation. In this paper, we address this problem by introducing the Infinite Contextual Graph Markov Model (iCGMM), the first deep Bayesian nonparametric model for graph learning.
- 120, TITLE: Compressed-VFL: Communication-Efficient Learning with Vertically Partitioned Data
<https://proceedings.mlr.press/v162/castiglia22a.html>
AUTHORS: Timothy J Castiglia, Anirban Das, Shiqiang Wang, Stacy Patterson
HIGHLIGHT: We propose Compressed Vertical Federated Learning (C-VFL) for communication-efficient training on vertically partitioned data.
- 121, TITLE: Online Learning with Knapsacks: the Best of Both Worlds
<https://proceedings.mlr.press/v162/castiglioni22a.html>
AUTHORS: Matteo Castiglioni, Andrea Celli, Christian Kroer
HIGHLIGHT: We study online learning problems in which a decision maker wants to maximize their expected reward without violating a finite set of resource constraints.
- 122, TITLE: Stabilizing Off-Policy Deep Reinforcement Learning from Pixels

- <https://proceedings.mlr.press/v162/cetin22a.html>
AUTHORS: Edoardo Cetin, Philip J Ball, Stephen Roberts, Oya Celiktutan
HIGHLIGHT: As a result, many successful algorithms must combine different domain-specific practices and auxiliary losses to learn meaningful behaviors in complex environments. In this work, we provide novel analysis demonstrating that these instabilities arise from performing temporal-difference learning with a convolutional encoder and low-magnitude rewards.
- 123, TITLE: Accelerated, Optimal and Parallel: Some results on model-based stochastic optimization
<https://proceedings.mlr.press/v162/chadha22a.html>
AUTHORS: Karan Chadha, Gary Cheng, John Duchi
HIGHLIGHT: In this paper, we propose an acceleration scheme for the APROX family and provide non-asymptotic convergence guarantees, which are order-optimal in all problem-dependent constants and provide even larger minibatching speedups.
- 124, TITLE: Robust Imitation Learning against Variations in Environment Dynamics
<https://proceedings.mlr.press/v162/cha22a.html>
AUTHORS: Jongseong Chae, Seungyul Han, Whiyoung Jung, Myungsik Cho, Sungho Choi, Youngchul Sung
HIGHLIGHT: In this paper, we propose a robust imitation learning (IL) framework that improves the robustness of IL when environment dynamics are perturbed.
- 125, TITLE: Fairness with Adaptive Weights
<https://proceedings.mlr.press/v162/chai22a.html>
AUTHORS: Junyi Chai, Xiaoqian Wang
HIGHLIGHT: In this paper, we propose a novel adaptive reweighing method to address representation bias.
- 126, TITLE: UNIREX: A Unified Learning Framework for Language Model Rationale Extraction
<https://proceedings.mlr.press/v162/chan22a.html>
AUTHORS: Aaron Chan, Maziar Sanjabi, Lambert Mathias, Liang Tan, Shaoliang Nie, Xiaochang Peng, Xiang Ren, Hamed Firooz
HIGHLIGHT: Although attribution algorithms and select-predict pipelines are commonly used in rationale extraction, they both rely on certain heuristics that hinder them from satisfying all three desiderata. In light of this, we propose UNIREX, a flexible learning framework which generalizes rationale extractor optimization as follows: (1) specify architecture for a learned rationale extractor; (2) select explainability objectives (ie faithfulness and plausibility criteria); and (3) jointly train the task model and rationale extractor on the task using selected objectives.
- 127, TITLE: Revisiting Label Smoothing and Knowledge Distillation Compatibility: What was Missing?
<https://proceedings.mlr.press/v162/chandrasegaran22a.html>
AUTHORS: Keshigeyan Chandrasegaran, Ngoc-Trung Tran, Yunqing Zhao, Ngai-Man Cheung
HIGHLIGHT: The main contributions of our work are the discovery, analysis and validation of systematic diffusion as the missing concept which is instrumental in understanding and resolving these contradictory findings.
- 128, TITLE: Style Equalization: Unsupervised Learning of Controllable Generative Sequence Models
<https://proceedings.mlr.press/v162/chang22a.html>
AUTHORS: Jen-Hao Rick Chang, Ashish Shrivastava, Hema Koppula, Xiaoshuai Zhang, Oncel Tuzel
HIGHLIGHT: In this paper, we tackle the training-inference mismatch encountered during unsupervised learning of controllable generative sequence models.
- 129, TITLE: Learning Bellman Complete Representations for Offline Policy Evaluation
<https://proceedings.mlr.press/v162/chang22b.html>
AUTHORS: Jonathan Chang, Kaiwen Wang, Nathan Kallus, Wen Sun
HIGHLIGHT: In this work, we propose BCRL, which directly learns from data an approximately linear Bellman complete representation with good coverage.
- 130, TITLE: Sample Efficient Learning of Predictors that Complement Humans
<https://proceedings.mlr.press/v162/charusaie22a.html>
AUTHORS: Mohammad-Amin Charusaie, Hussein Mozannar, David Sontag, Samira Samadi
HIGHLIGHT: In this work, we provide the first theoretical analysis of the benefit of learning complementary predictors in expert deferral.
- 131, TITLE: Nystrom Kernel Mean Embeddings
<https://proceedings.mlr.press/v162/chatalic22a.html>
AUTHORS: Antoine Chatalic, Nicolas Schreuder, Lorenzo Rosasco, Alessandro Rudi
HIGHLIGHT: We propose an efficient approximation procedure based on the Nyström method, which exploits a small random subset of the dataset.
- 132, TITLE: Coarsening the Granularity: Towards Structurally Sparse Lottery Tickets
<https://proceedings.mlr.press/v162/chen22a.html>
AUTHORS: Tianlong Chen, Xuxi Chen, Xiaolong Ma, Yanzhi Wang, Zhangyang Wang
HIGHLIGHT: In this paper, we demonstrate the first positive result that a structurally sparse winning ticket can be effectively found in general.
- 133, TITLE: Learning Domain Adaptive Object Detection with Probabilistic Teacher

- <https://proceedings.mlr.press/v162/chen22b.html>
AUTHORS: Meilin Chen, Weijie Chen, Shicai Yang, Jie Song, Xinchao Wang, Lei Zhang, Yunfeng Yan, Donglian Qi, Yueting Zhuang, Di Xie, Shiliang Pu
HIGHLIGHT: In this paper, we present a simple yet effective framework, termed as Probabilistic Teacher (PT), which aims to capture the uncertainty of unlabeled target data from a gradually evolving teacher and guides the learning of a student in a mutually beneficial manner.
- 134, TITLE: The Fundamental Price of Secure Aggregation in Differentially Private Federated Learning
<https://proceedings.mlr.press/v162/chen22c.html>
AUTHORS: Wei-Ning Chen, Christopher A Choquette Choo, Peter Kairouz, Ananda Theertha Suresh
HIGHLIGHT: We consider the problem of training a d dimensional model with distributed differential privacy (DP) where secure aggregation (SecAgg) is used to ensure that the server only sees the noisy sum of n model updates in every training round.
- 135, TITLE: Perfectly Balanced: Improving Transfer and Robustness of Supervised Contrastive Learning
<https://proceedings.mlr.press/v162/chen22d.html>
AUTHORS: Mayee Chen, Daniel Y Fu, Avanika Narayan, Michael Zhang, Zhao Song, Kayvon Fatahalian, Christopher Re
HIGHLIGHT: We argue that creating spread alone is insufficient for better representations, since spread is invariant to permutations within classes.
- 136, TITLE: Strategies for Safe Multi-Armed Bandits with Logarithmic Regret and Risk
<https://proceedings.mlr.press/v162/chen22e.html>
AUTHORS: Tianrui Chen, Aditya Gangrade, Venkatesh Saligrama
HIGHLIGHT: We investigate a natural but surprisingly unstudied approach to the multi-armed bandit problem under safety risk constraints.
- 137, TITLE: On the Sample Complexity of Learning Infinite-horizon Discounted Linear Kernel MDPs
<https://proceedings.mlr.press/v162/chen22f.html>
AUTHORS: Yuanzhou Chen, Jiafan He, Quanquan Gu
HIGHLIGHT: In this paper, we extend the uniform-PAC sample complexity from episodic setting to the infinite-horizon discounted setting, and propose a novel algorithm dubbed UPAC-UCLK that achieves an $\tilde{O}\left(d^2 \left((1-\gamma)^4 \epsilon^2 + 1 / \left((1-\gamma)^6 \epsilon^2 \right) \right)\right)$ uniform-PAC sample complexity, where d is the dimension of the feature mapping, $\gamma \in (0,1)$ is the discount factor of the MDP and ϵ is the accuracy parameter.
- 138, TITLE: Streaming Algorithms for Support-Aware Histograms
<https://proceedings.mlr.press/v162/chen22g.html>
AUTHORS: Justin Chen, Piotr Indyk, Tal Wagner
HIGHLIGHT: As a result, even relatively simple distributions cannot be approximated by succinct histograms without incurring large error. In this paper, we address this issue by adapting the definition of approximation so that only the errors of the items that belong to the support of the distribution are considered.
- 139, TITLE: Improved No-Regret Algorithms for Stochastic Shortest Path with Linear MDP
<https://proceedings.mlr.press/v162/chen22h.html>
AUTHORS: Liyu Chen, Rahul Jain, Haipeng Luo
HIGHLIGHT: We introduce two new no-regret algorithms for the stochastic shortest path (SSP) problem with a linear MDP that significantly improve over the only existing results of (Vial et al., 2021).
- 140, TITLE: Learning Infinite-horizon Average-reward Markov Decision Process with Constraints
<https://proceedings.mlr.press/v162/chen22i.html>
AUTHORS: Liyu Chen, Rahul Jain, Haipeng Luo
HIGHLIGHT: We study regret minimization for infinite-horizon average-reward Markov Decision Processes (MDPs) under cost constraints.
- 141, TITLE: Active Multi-Task Representation Learning
<https://proceedings.mlr.press/v162/chen22j.html>
AUTHORS: Yifang Chen, Kevin Jamieson, Simon Du
HIGHLIGHT: We propose an algorithm that iteratively estimates the relevance of each source task to the target task and samples from each source task based on the estimated relevance.
- 142, TITLE: On Collective Robustness of Bagging Against Data Poisoning
<https://proceedings.mlr.press/v162/chen22k.html>
AUTHORS: Ruoxin Chen, Zenan Li, Jie Li, Junchi Yan, Chentao Wu
HIGHLIGHT: Based on this analysis, we propose hash bagging to improve the robustness of vanilla bagging almost for free.
- 143, TITLE: Online Active Regression
<https://proceedings.mlr.press/v162/chen22l.html>
AUTHORS: Cheng Chen, Yi Li, Yiming Sun
HIGHLIGHT: The goal is to efficiently maintain the regression of received data points with a small budget of label queries. We propose novel algorithms for this problem under ℓ_p loss where $p \in [1,2]$.
- 144, TITLE: Selling Data To a Machine Learner: Pricing via Costly Signaling

<https://proceedings.mlr.press/v162/chen22m.html>
AUTHORS: Junjie Chen, Minming Li, Haifeng Xu
HIGHLIGHT: We consider a new problem of selling data to a machine learner who looks to purchase data to train his machine learning model.

145, TITLE: ME-GAN: Learning Panoptic Electrocardio Representations for Multi-view ECG Synthesis Conditioned on Heart Diseases
<https://proceedings.mlr.press/v162/chen22n.html>
AUTHORS: Jintai Chen, Kuanlun Liao, Kun Wei, Haochao Ying, Danny Z Chen, Jian Wu
HIGHLIGHT: In this paper, we propose a novel disease-aware generative adversarial network for multi-view ECG synthesis called ME-GAN, which attains panoptic electrocardio representations conditioned on heart diseases and projects the representations onto multiple standard views to yield ECG signals.

146, TITLE: Weisfeiler-Lehman Meets Gromov-Wasserstein
<https://proceedings.mlr.press/v162/chen22o.html>
AUTHORS: Samantha Chen, Sunhyuk Lim, Facundo Memoli, Zhengchao Wan, Yusu Wang
HIGHLIGHT: In this paper, we propose the Weisfeiler-Lehman (WL) distance, a notion of distance between labeled measure Markov chains (LMMCs), of which labeled graphs are special cases.

147, TITLE: On Non-local Convergence Analysis of Deep Linear Networks
<https://proceedings.mlr.press/v162/chen22p.html>
AUTHORS: Kun Chen, Dachao Lin, Zhihua Zhang
HIGHLIGHT: In this paper, we study the non-local convergence properties of deep linear networks.

148, TITLE: Flow-based Recurrent Belief State Learning for POMDPs
<https://proceedings.mlr.press/v162/chen22q.html>
AUTHORS: Xiaoyu Chen, Yao Mark Mu, Ping Luo, Shengbo Li, Jianyu Chen
HIGHLIGHT: In this paper, we introduce the F -based R -recurrent BE S -tate model (FORBES), which incorporates normalizing flows into the variational inference to learn general continuous belief states for POMDPs.

149, TITLE: Structure-Aware Transformer for Graph Representation Learning
<https://proceedings.mlr.press/v162/chen22r.html>
AUTHORS: Dexiong Chen, Leslie O'Bray, Karsten Borgwardt
HIGHLIGHT: We propose several methods for automatically generating the subgraph representation and show theoretically that the resulting representations are at least as expressive as the subgraph representations.

150, TITLE: The Poisson Binomial Mechanism for Unbiased Federated Learning with Secure Aggregation
<https://proceedings.mlr.press/v162/chen22s.html>
AUTHORS: Wei-Ning Chen, Ayfer Ozgur, Peter Kairouz
HIGHLIGHT: We introduce the Poisson Binomial mechanism (PBM), a discrete differential privacy mechanism for distributed mean estimation (DME) with applications to federated learning and analytics.

151, TITLE: Learning Mixtures of Linear Dynamical Systems
<https://proceedings.mlr.press/v162/chen22t.html>
AUTHORS: Yanxi Chen, H. Vincent Poor
HIGHLIGHT: We study the problem of learning a mixture of multiple linear dynamical systems (LDSs) from unlabeled short sample trajectories, each generated by one of the LDS models.

152, TITLE: On Well-posedness and Minimax Optimal Rates of Nonparametric Q-function Estimation in Off-policy Evaluation
<https://proceedings.mlr.press/v162/chen22u.html>
AUTHORS: Xiaohong Chen, Zhengling Qi
HIGHLIGHT: We study the off-policy evaluation (OPE) problem in an infinite-horizon Markov decision process with continuous states and actions.

153, TITLE: Faster Fundamental Graph Algorithms via Learned Predictions
<https://proceedings.mlr.press/v162/chen22v.html>
AUTHORS: Justin Chen, Sandeep Silwal, Ali Vakilian, Fred Zhang
HIGHLIGHT: We consider the question of speeding up classic graph algorithms with machine-learned predictions.

154, TITLE: Improve Single-Point Zeroth-Order Optimization Using High-Pass and Low-Pass Filters
<https://proceedings.mlr.press/v162/chen22w.html>
AUTHORS: Xin Chen, Yujie Tang, Na Li
HIGHLIGHT: In this work, we borrow the idea of high-pass and low-pass filters from extremum seeking control (continuous-time version of SZO) and develop a novel SZO method called HLF-SZO by integrating these filters.

155, TITLE: Deep Variational Graph Convolutional Recurrent Network for Multivariate Time Series Anomaly Detection
<https://proceedings.mlr.press/v162/chen22x.html>
AUTHORS: Wenchao Chen, Long Tian, Bo Chen, Liang Dai, Zhibin Duan, Mingyuan Zhou

HIGHLIGHT: In this paper, we model sensor dependency and stochasticity within MTS by developing an embedding-guided probabilistic generative network.

156, **TITLE:** Auxiliary Learning with Joint Task and Data Scheduling

<https://proceedings.mlr.press/v162/chen22y.html>

AUTHORS: Hong Chen, Xin Wang, Chaoyu Guan, Yue Liu, Wenwu Zhu

HIGHLIGHT: In this paper, we propose to learn a joint task and data schedule for auxiliary learning, which captures the importance of different data samples in each auxiliary task to the target task.

157, **TITLE:** Optimization-Induced Graph Implicit Nonlinear Diffusion

<https://proceedings.mlr.press/v162/chen22z.html>

AUTHORS: Qi Chen, Yifei Wang, Yisen Wang, Jiansheng Yang, Zhouchen Lin

HIGHLIGHT: Due to the over-smoothing issue, most existing graph neural networks can only capture limited dependencies with their inherently finite aggregation layers. To overcome this limitation, we propose a new kind of graph convolution, called Graph Implicit Nonlinear Diffusion (GIND), which implicitly has access to infinite hops of neighbors while adaptively aggregating features with nonlinear diffusion to prevent over-smoothing.

158, **TITLE:** Robust Meta-learning with Sampling Noise and Label Noise via Eigen-Reptile

<https://proceedings.mlr.press/v162/chen22aa.html>

AUTHORS: Dong Chen, Lingfei Wu, Siliang Tang, Xiao Yun, Bo Long, Yueting Zhuang

HIGHLIGHT: Besides, when handling the data with noisy labels, the meta-learner could be extremely sensitive to label noise on a corrupted dataset. To address these two challenges, we present Eigen-Reptile (ER) that updates the meta-parameters with the main direction of historical task-specific parameters.

159, **TITLE:** Adaptive Model Design for Markov Decision Process

<https://proceedings.mlr.press/v162/chen22ab.html>

AUTHORS: Siyu Chen, Donglin Yang, Jiayang Li, Senmiao Wang, Zhuoran Yang, Zhaoran Wang

HIGHLIGHT: Hence, appropriate regulations are often required, if we hope to take the external costs/benefits of its actions into consideration. In this paper, we study how to regulate such an agent by redesigning model parameters that can affect the rewards and/or the transition kernels.

160, **TITLE:** State Transition of Dendritic Spines Improves Learning of Sparse Spiking Neural Networks

<https://proceedings.mlr.press/v162/chen22ac.html>

AUTHORS: Yanqi Chen, Zhaoifei Yu, Wei Fang, Zhengyu Ma, Tiejun Huang, Yonghong Tian

HIGHLIGHT: Inspired by the state transition of dendritic spines in the filopodial model of spinogenesis, we model different states of SNN weights, facilitating weight optimization for pruning.

161, **TITLE:** Efficient Online ML API Selection for Multi-Label Classification Tasks

<https://proceedings.mlr.press/v162/chen22ad.html>

AUTHORS: Lingjiao Chen, Matei Zaharia, James Zou

HIGHLIGHT: In this work, we propose FrugalMCT, a principled framework that adaptively selects the APIs to use for different data in an online fashion while respecting the user's budget.

162, **TITLE:** Data-Efficient Double-Win Lottery Tickets from Robust Pre-training

<https://proceedings.mlr.press/v162/chen22ae.html>

AUTHORS: Tianlong Chen, Zhenyu Zhang, Sijia Liu, Yang Zhang, Shiyu Chang, Zhangyang Wang

HIGHLIGHT: In this paper, we formulate a more rigorous concept, Double-Win Lottery Tickets, in which a located subnetwork from a pre-trained model can be independently transferred on diverse downstream tasks, to reach BOTH the same standard and robust generalization, under BOTH standard and adversarial training regimes, as the full pre-trained model can do.

163, **TITLE:** Linearity Grafting: Relaxed Neuron Pruning Helps Certifiable Robustness

<https://proceedings.mlr.press/v162/chen22af.html>

AUTHORS: Tianlong Chen, Huan Zhang, Zhenyu Zhang, Shiyu Chang, Sijia Liu, Pin-Yu Chen, Zhangyang Wang

HIGHLIGHT: To trade off the DNN expressiveness (which calls for more non-linearity) and robustness certification scalability (which prefers more linearity), we propose a novel solution to strategically manipulate neurons, by "grafting" appropriate levels of linearity.

164, **TITLE:** Human-in-the-loop: Provably Efficient Preference-based Reinforcement Learning with General Function Approximation

<https://proceedings.mlr.press/v162/chen22ag.html>

AUTHORS: Xiaoyu Chen, Han Zhong, Zhuoran Yang, Zhaoran Wang, Liwei Wang

HIGHLIGHT: In this paper, we propose the first optimistic model-based algorithm for PbRL with general function approximation, which estimates the model using value-targeted regression and calculates the exploratory policies by solving an optimistic planning problem.

165, **TITLE:** Sample and Communication-Efficient Decentralized Actor-Critic Algorithms with Finite-Time Analysis

<https://proceedings.mlr.press/v162/chen22ah.html>

AUTHORS: Ziyi Chen, Yi Zhou, Rong-Rong Chen, Shaofeng Zou

HIGHLIGHT: In this work, we develop decentralized AC and natural AC (NAC) algorithms that avoid sharing agents' local information and are sample and communication-efficient.

- 166, TITLE: Task-aware Privacy Preservation for Multi-dimensional Data
<https://proceedings.mlr.press/v162/cheng22a.html>
AUTHORS: Jiangnan Cheng, Ao Tang, Sandeep Chinchali
HIGHLIGHT: In this paper, we address how to significantly improve the ultimate task performance with multi-dimensional user data by considering a task-aware privacy preservation problem.
- 167, TITLE: Adversarially Trained Actor Critic for Offline Reinforcement Learning
<https://proceedings.mlr.press/v162/cheng22b.html>
AUTHORS: Ching-An Cheng, Tengyang Xie, Nan Jiang, Alekh Agarwal
HIGHLIGHT: We propose Adversarially Trained Actor Critic (ATAC), a new model-free algorithm for offline reinforcement learning (RL) under insufficient data coverage, based on the concept of relative pessimism.
- 168, TITLE: Quantum-Inspired Algorithms from Randomized Numerical Linear Algebra
<https://proceedings.mlr.press/v162/chepurko22a.html>
AUTHORS: Nadiia Chepurko, Kenneth Clarkson, Lior Horesh, Honghao Lin, David Woodruff
HIGHLIGHT: We create classical (non-quantum) dynamic data structures supporting queries for recommender systems and least-squares regression that are comparable to their quantum analogues.
- 169, TITLE: RieszNet and ForestRiesz: Automatic Debaised Machine Learning with Neural Nets and Random Forests
<https://proceedings.mlr.press/v162/chernozhukov22a.html>
AUTHORS: Victor Chernozhukov, Whitney Newey, Victor M Quintas-Martinez, Vasilis Syrgkanis
HIGHLIGHT: We propose a multitasking Neural Net debiasing method with stochastic gradient descent minimization of a combined Riesz representer and regression loss, while sharing representation layers for the two functions.
- 170, TITLE: Self-supervised learning with random-projection quantizer for speech recognition
<https://proceedings.mlr.press/v162/chiu22a.html>
AUTHORS: Chung-Cheng Chiu, James Qin, Yu Zhang, Jiahui Yu, Yonghui Wu
HIGHLIGHT: We present a simple and effective self-supervised learning approach for speech recognition.
- 171, TITLE: Discrete Probabilistic Inverse Optimal Transport
<https://proceedings.mlr.press/v162/chiu22b.html>
AUTHORS: Wei-Ting Chiu, Pei Wang, Patrick Shafto
HIGHLIGHT: We formalize and systematically analyze the properties of IOT using tools from the study of entropy-regularized OT.
- 172, TITLE: Selective Network Linearization for Efficient Private Inference
<https://proceedings.mlr.press/v162/cho22a.html>
AUTHORS: Minsu Cho, Ameya Joshi, Brandon Reagen, Siddharth Garg, Chinmay Hegde
HIGHLIGHT: To reduce PI latency we propose a gradient-based algorithm that selectively linearizes ReLUs while maintaining prediction accuracy.
- 173, TITLE: From block-Toeplitz matrices to differential equations on graphs: towards a general theory for scalable masked Transformers
<https://proceedings.mlr.press/v162/choromanski22a.html>
AUTHORS: Krzysztof Choromanski, Han Lin, Haoxian Chen, Tianyi Zhang, Arijitohanobish, Valerii Likhoshesterov, Jack Parker-Holder, Tamas Sarlos, Adrian Weller, Thomas Weingarten
HIGHLIGHT: In this paper we provide, to the best of our knowledge, the first comprehensive approach for incorporating various masking mechanisms into Transformers architectures in a scalable way.
- 174, TITLE: Shuffle Private Linear Contextual Bandits
<https://proceedings.mlr.press/v162/chowdhury22a.html>
AUTHORS: Sayak Ray Chowdhury, Xingyu Zhou
HIGHLIGHT: We propose a general algorithmic framework for linear contextual bandits under the shuffle trust model, where there exists a trusted shuffler – in between users and the central server– that randomly permutes a batch of users data before sending those to the server.
- 175, TITLE: DNA: Domain Generalization with Diversified Neural Averaging
<https://proceedings.mlr.press/v162/chu22a.html>
AUTHORS: Xu Chu, Yujie Jin, Wenwu Zhu, Yasha Wang, Xin Wang, Shanghang Zhang, Hong Mei
HIGHLIGHT: Methodologically, we propose a diversified neural averaging (DNA) method for DG, which optimizes the proposed PAC-Bayes bound approximately.
- 176, TITLE: TPC: Transformation-Specific Smoothing for Point Cloud Models
<https://proceedings.mlr.press/v162/chu22b.html>
AUTHORS: Wenda Chu, Linyi Li, Bo Li
HIGHLIGHT: In this paper, we propose a transformation-specific smoothing framework TPC, which provides tight and scalable robustness guarantees for point cloud models against semantic transformation attacks.
- 177, TITLE: Unified Scaling Laws for Routed Language Models

<https://proceedings.mlr.press/v162/clark22a.html>

AUTHORS: Aidan Clark, Diego De Las Casas, Aurelia Guy, Arthur Mensch, Michela Paganini, Jordan Hoffmann, Bogdan Damoc, Blake Hechtman, Trevor Cai, Sebastian Borgeaud, George Bm Van Den Driessche, Eliza Rutherford, Tom Hennigan, Matthew J Johnson, Albin Cassirer, Chris Jones, Elena Buchatskaya, David Budden, Laurent Sifre, Simon Osindero, Oriol Vinyals, Marc?Aurelio Ranzato, Jack Rae, Erich Elsen, Koray Kavukcuoglu, Karen Simonyan
HIGHLIGHT: For these models, parameter count and computational requirement form two independent axes along which an increase leads to better performance. In this work we derive and justify scaling laws defined on these two variables which generalize those known for standard language models and describe the performance of a wide range of routing architectures trained via three different techniques.

178, TITLE: Context-Aware Drift Detection

<https://proceedings.mlr.press/v162/cobb22a.html>

AUTHORS: Oliver Cobb, Arnaud Van Looveren

HIGHLIGHT: Instead we may wish to test for differences in the distributions conditional on context that is permitted to change. To facilitate this we borrow machinery from the causal inference domain to develop a more general drift detection framework built upon a foundation of two-sample tests for conditional distributional treatment effects.

179, TITLE: On the Robustness of CountSketch to Adaptive Inputs

<https://proceedings.mlr.press/v162/cohen22a.html>

AUTHORS: Edith Cohen, Xin Lyu, Jelani Nelson, Tamas Sarlos, Moshe Shechner, Uri Stemmer

HIGHLIGHT: We propose a robust estimator (for a slightly modified sketch) that allows for quadratic number of queries in the sketch size, which is an improvement factor of \sqrt{k} (for k heavy hitters) over prior “blackbox” approaches.

180, TITLE: Diffusion bridges vector quantized variational autoencoders

<https://proceedings.mlr.press/v162/cohen22b.html>

AUTHORS: Max Cohen, Guillaume Quispe, Sylvain Le Corff, Charles Ollion, Eric Moulines

HIGHLIGHT: In this work, we propose a new model to train the prior and the encoder/decoder networks simultaneously.

181, TITLE: Online and Consistent Correlation Clustering

<https://proceedings.mlr.press/v162/cohen-addad22a.html>

AUTHORS: Vincent Cohen-Addad, Silvio Lattanzi, Andreas Maggiori, Nikos Parotsidis

HIGHLIGHT: In this paper we study the problem in the classic online setting with recourse; The vertices of the graphs arrive in an online manner and the goal is to maintain an approximate clustering while minimizing the number of times each vertex changes cluster.

182, TITLE: Massively Parallel k -Means Clustering for Perturbation Resilient Instances

<https://proceedings.mlr.press/v162/cohen-addad22b.html>

AUTHORS: Vincent Cohen-Addad, Vahab Mirrokni, Peilin Zhong

HIGHLIGHT: We consider k -means clustering of n data points in Euclidean space in the Massively Parallel Computation (MPC) model, a computational model which is an abstraction of modern massively parallel computing system such as MapReduce.

183, TITLE: One-Pass Diversified Sampling with Application to Terabyte-Scale Genomic Sequence Streams

<https://proceedings.mlr.press/v162/coleman22a.html>

AUTHORS: Benjamin Coleman, Benito Geordie, Li Chou, R. A. Leo Elworth, Todd Treangen, Anshumali Shrivastava

HIGHLIGHT: We propose an efficient sampling routine that uses an online representation of the data distribution as a prefilter to retain elements from rare groups.

184, TITLE: Transfer and Marginalize: Explaining Away Label Noise with Privileged Information

<https://proceedings.mlr.press/v162/collier22a.html>

AUTHORS: Mark Collier, Rodolphe Jenatton, Effrosyni Kokiopoulou, Jesse Berent

HIGHLIGHT: We develop a simple and efficient method for supervised learning with neural networks: it transfers via weight sharing the knowledge learned with privileged information and approximately marginalizes over privileged information at test time.

185, TITLE: MAML and ANIL Provably Learn Representations

<https://proceedings.mlr.press/v162/collins22a.html>

AUTHORS: Liam Collins, Aryan Mokhtari, Sewoong Oh, Sanjay Shakkottai

HIGHLIGHT: In this paper, we prove that two well-known GBML methods, MAML and ANIL, as well as their first-order approximations, are capable of learning common representation among a set of given tasks.

186, TITLE: Entropic Causal Inference: Graph Identifiability

<https://proceedings.mlr.press/v162/compton22a.html>

AUTHORS: Spencer Compton, Kristjan Greenewald, Dmitriy A Katz, Murat Kocaoglu

HIGHLIGHT: In our work, we first extend the causal graph identifiability result in the two-variable setting under relaxed assumptions. We then show the first identifiability result using the entropic approach for learning causal graphs with more than two nodes.

187, TITLE: Mitigating Gender Bias in Face Recognition using the von Mises-Fisher Mixture Model

<https://proceedings.mlr.press/v162/conti22a.html>

AUTHORS: Jean-Remy Conti, Nathan Noiry, Stephan Clemencon, Vincent Despiegel, Stéphane Gentric

HIGHLIGHT: In this work, we investigate the gender bias of deep Face Recognition networks.

- 188, TITLE: Counterfactual Transportability: A Formal Approach
<https://proceedings.mlr.press/v162/correa22a.html>
AUTHORS: Juan D Correa, Sanghack Lee, Elias Bareinboim
HIGHLIGHT: In this paper, we investigate the transportability of counterfactuals from an arbitrary combination of observational and experimental distributions coming from disparate domains.
- 189, TITLE: Label-Free Explainability for Unsupervised Models
<https://proceedings.mlr.press/v162/crabbe22a.html>
AUTHORS: Jonathan Crabb?, Mihaela van der Schaar
HIGHLIGHT: Hence, choosing which component(s) to interpret in a label-free unsupervised/self-supervised setting is an important, yet unsolved problem. To bridge this gap in the literature, we introduce two crucial extensions of post-hoc explanation techniques: (1) label-free feature importance and (2) label-free example importance that respectively highlight influential features and training examples for a black-box to construct representations at inference time.
- 190, TITLE: Evaluating the Adversarial Robustness of Adaptive Test-time Defenses
<https://proceedings.mlr.press/v162/croce22a.html>
AUTHORS: Francesco Croce, Sven Gowal, Thomas Brunner, Evan Shelhamer, Matthias Hein, Taylan Cemgil
HIGHLIGHT: While these results are disappointing, we still believe that adaptive test-time defenses are a promising avenue of research and, as such, we provide recommendations for their thorough evaluation.
- 191, TITLE: Adversarial Robustness against Multiple and Single ℓ_p -Threat Models via Quick Fine-Tuning of Robust Classifiers
<https://proceedings.mlr.press/v162/croce22b.html>
AUTHORS: Francesco Croce, Matthias Hein
HIGHLIGHT: In this paper we propose Extreme norm Adversarial Training (E-AT) for multiple-norm robustness which is based on geometric properties of ℓ_p -balls.
- 192, TITLE: Self-conditioning Pre-Trained Language Models
<https://proceedings.mlr.press/v162/cuadros22a.html>
AUTHORS: Xavier Suau Cuadros, Luca Zappella, Nicholas Apostoloff
HIGHLIGHT: In this paper we aim to investigate the mechanisms that guide text generation with pre-trained Transformer-based Language Models (TLMs).
- 193, TITLE: Only tails matter: Average-Case Universality and Robustness in the Convex Regime
<https://proceedings.mlr.press/v162/cunha22a.html>
AUTHORS: Leonardo Cunha, Gauthier Gidel, Fabian Pedregosa, Damien Scieur, Courtney Paquette
HIGHLIGHT: This work shows that the concentration of eigenvalues near the edges of the ESD determines a problem's asymptotic average complexity.
- 194, TITLE: Principal Component Flows
<https://proceedings.mlr.press/v162/cunningham22a.html>
AUTHORS: Edmond Cunningham, Adam D Cobb, Susmit Jha
HIGHLIGHT: In this paper we characterize the geometric structure of flows using principal manifolds and understand the relationship between latent variables and samples using contours.
- 195, TITLE: Deep symbolic regression for recurrence prediction
<https://proceedings.mlr.press/v162/d-ascoli22a.html>
AUTHORS: Stéphane D'Ascoli, Pierre-Alexandre Kamienny, Guillaume Lample, François Charton
HIGHLIGHT: In this paper, we train Transformers to infer the function or recurrence relation underlying sequences of integers or floats, a typical task in human IQ tests which has hardly been tackled in the machine learning literature.
- 196, TITLE: Continuous Control with Action Quantization from Demonstrations
<https://proceedings.mlr.press/v162/dadashi22a.html>
AUTHORS: Robert Dadashi, L'onard Hussenot, Damien Vincent, Sertan Girgin, Anton Raichuk, Matthieu Geist, Olivier Pietquin
HIGHLIGHT: In this paper, we propose a novel Reinforcement Learning (RL) framework for problems with continuous action spaces: Action Quantization from Demonstrations (AQuaDem).
- 197, TITLE: Dialog Inpainting: Turning Documents into Dialogs
<https://proceedings.mlr.press/v162/dai22a.html>
AUTHORS: Zhuyun Dai, Arun Tejasvi Chaganty, Vincent Y Zhao, Aida Amini, Qazi Mamunur Rashid, Mike Green, Kelvin Guu
HIGHLIGHT: However, conversational question answering (ConvQA) systems have long been stymied by scarce training data that is expensive to collect. To address this problem, we propose a new technique for synthetically generating diverse and high-quality dialog data: dialog inpainting.
- 198, TITLE: DisPFL: Towards Communication-Efficient Personalized Federated Learning via Decentralized Sparse Training
<https://proceedings.mlr.press/v162/dai22b.html>
AUTHORS: Rong Dai, Li Shen, Fengxiang He, Xinmei Tian, Dacheng Tao

- HIGHLIGHT:** In this work, we propose a novel personalized federated learning framework in a decentralized (peer-to-peer) communication protocol named DisPFL, which employs personalized sparse masks to customize sparse local models on the edge.
- 199, **TITLE:** Marginal Distribution Adaptation for Discrete Sets via Module-Oriented Divergence Minimization
<https://proceedings.mlr.press/v162/dai22c.html>
AUTHORS: Hanjun Dai, Mengjiao Yang, Yuan Xue, Dale Schuurmans, Bo Dai
HIGHLIGHT: We develop a general framework to adapt a generative model subject to a (possibly counterfactual) target data distribution with both sampling and computation efficiency.
- 200, **TITLE:** Balancing Sample Efficiency and Suboptimality in Inverse Reinforcement Learning
<https://proceedings.mlr.press/v162/damiani22a.html>
AUTHORS: Angelo Damiani, Giorgio Manganini, Alberto Maria Metelli, Marcello Restelli
HIGHLIGHT: We propose a novel formulation for the Inverse Reinforcement Learning (IRL) problem, which jointly accounts for the compatibility with the expert behavior of the identified reward and its effectiveness for the subsequent forward learning phase.
- 201, **TITLE:** Understanding Robust Generalization in Learning Regular Languages
<https://proceedings.mlr.press/v162/dan22a.html>
AUTHORS: Soham Dan, Osbert Bastani, Dan Roth
HIGHLIGHT: We hypothesize that standard end-to-end modeling strategies cannot generalize well to systematic distribution shifts and propose a compositional strategy to address this.
- 202, **TITLE:** Unsupervised Image Representation Learning with Deep Latent Particles
<https://proceedings.mlr.press/v162/daniel22a.html>
AUTHORS: Tal Daniel, Aviv Tamar
HIGHLIGHT: We propose a new representation of visual data that disentangles object position from appearance.
- 203, **TITLE:** Guarantees for Epsilon-Greedy Reinforcement Learning with Function Approximation
<https://proceedings.mlr.press/v162/dann22a.html>
AUTHORS: Chris Dann, Yishay Mansour, Mehryar Mohri, Ayush Sekhari, Karthik Sridharan
HIGHLIGHT: These crucial questions have been scarcely investigated, despite the prominent practical importance of these policies. This paper presents a theoretical analysis of such policies and provides the first regret and sample-complexity bounds for reinforcement learning with myopic exploration.
- 204, **TITLE:** Monarch: Expressive Structured Matrices for Efficient and Accurate Training
<https://proceedings.mlr.press/v162/dao22a.html>
AUTHORS: Tri Dao, Beidi Chen, Nimit S Sohoni, Arjun Desai, Michael Poli, Jessica Grogan, Alexander Liu, Aniruddh Rao, Atri Rudra, Christopher Re
HIGHLIGHT: These methods have not seen widespread adoption (1) in end-to-end training due to unfavorable efficiency-quality tradeoffs, and (2) in dense-to-sparse fine-tuning due to lack of tractable algorithms to approximate a given dense weight matrix. To address these issues, we propose a class of matrices (Monarch) that is hardware-efficient (they are parameterized as products of two block-diagonal matrices for better hardware utilization) and expressive (they can represent many commonly used transforms).
- 205, **TITLE:** Score-Guided Intermediate Level Optimization: Fast Langevin Mixing for Inverse Problems
<https://proceedings.mlr.press/v162/daras22a.html>
AUTHORS: Giannis Daras, Yuval Dagan, Alex Dimakis, Constantinos Daskalakis
HIGHLIGHT: In practice, to allow for increased expressivity, we propose to do posterior sampling in the latent space of a pre-trained generative model.
- 206, **TITLE:** Test-Time Training Can Close the Natural Distribution Shift Performance Gap in Deep Learning Based Compressed Sensing
<https://proceedings.mlr.press/v162/darestani22a.html>
AUTHORS: Mohammad Zalbagi Darestani, Jiayu Liu, Reinhard Heckel
HIGHLIGHT: In this work, we propose a domain adaptation method for deep learning based compressive sensing that relies on self-supervision during training paired with test-time training at inference.
- 207, **TITLE:** Knowledge Base Question Answering by Case-based Reasoning over Subgraphs
<https://proceedings.mlr.press/v162/das22a.html>
AUTHORS: Rajarshi Das, Ameya Godbole, Ankita Naik, Elliot Tower, Manzil Zaheer, Hannaneh Hajishirzi, Robin Jia, Andrew McCallum
HIGHLIGHT: Leveraging this structural similarity between local neighborhoods of different subgraphs, we introduce a semiparametric model (CBR-SUBG) with (i) a nonparametric component that for each query, dynamically retrieves other similar k -nearest neighbor (KNN) training queries along with query-specific subgraphs and (ii) a parametric component that is trained to identify the (latent) reasoning patterns from the subgraphs of KNN queries and then apply them to the subgraph of the target query.
- 208, **TITLE:** Framework for Evaluating Faithfulness of Local Explanations
<https://proceedings.mlr.press/v162/dasgupta22a.html>
AUTHORS: Sanjoy Dasgupta, Nave Frost, Michal Moshkovitz
HIGHLIGHT: We study the faithfulness of an explanation system to the underlying prediction model.

- 209, TITLE: Distinguishing rule and exemplar-based generalization in learning systems
<https://proceedings.mlr.press/v162/dasgupta22b.html>
AUTHORS: Ishita Dasgupta, Erin Grant, Tom Griffiths
HIGHLIGHT: The trade-off between exemplar- and rule-based generalization has been studied extensively in cognitive psychology; in this work, we present a protocol inspired by these experimental approaches to probe the inductive biases that control this trade-off in category-learning systems such as artificial neural networks.
- 210, TITLE: Robust Multi-Objective Bayesian Optimization Under Input Noise
<https://proceedings.mlr.press/v162/daulton22a.html>
AUTHORS: Samuel Daulton, Sait Cakmak, Maximilian Balandat, Michael A. Osborne, Enlu Zhou, Eytan Bakshy
HIGHLIGHT: Since directly optimizing MVaR is computationally infeasible in many settings, we propose a scalable, theoretically-grounded approach for optimizing MVaR using random scalarizations.
- 211, TITLE: Attentional Meta-learners for Few-shot Polythetic Classification
<https://proceedings.mlr.press/v162/day22a.html>
AUTHORS: Ben J Day, Ramon Vi?as Torn?, Nikola Simidjievski, Pietro Li?
HIGHLIGHT: However, we find that in the presence of task-irrelevant features, inherent to meta-learning problems, attentional models are susceptible to misclassification. To address this challenge, we propose a self-attention feature-selection mechanism that adaptively dilutes non-discriminative features.
- 212, TITLE: Adversarial Vulnerability of Randomized Ensembles
<https://proceedings.mlr.press/v162/dbouk22a.html>
AUTHORS: Hassan Dbouk, Naresh Shanbhag
HIGHLIGHT: However, this impressive performance raises the question: Are these robustness gains provided by randomized ensembles real? In this work we address this question both theoretically and empirically.
- 213, TITLE: Born-Infeld (BI) for AI: Energy-Conserving Descent (ECD) for Optimization
<https://proceedings.mlr.press/v162/de-luca22a.html>
AUTHORS: Giuseppe Bruno De Luca, Eva Silverstein
HIGHLIGHT: We introduce a novel framework for optimization based on energy-conserving Hamiltonian dynamics in a strongly mixing (chaotic) regime and establish its key properties analytically and numerically.
- 214, TITLE: Error-driven Input Modulation: Solving the Credit Assignment Problem without a Backward Pass
<https://proceedings.mlr.press/v162/dellaferreira22a.html>
AUTHORS: Giorgia Dellaferreira, Gabriel Kreiman
HIGHLIGHT: Here, we propose to replace the backward pass with a second forward pass in which the input signal is modulated based on the error of the network.
- 215, TITLE: DreamerPro: Reconstruction-Free Model-Based Reinforcement Learning with Prototypical Representations
<https://proceedings.mlr.press/v162/deng22a.html>
AUTHORS: Fei Deng, Ingook Jang, Sungjin Ahn
HIGHLIGHT: In this paper, we propose a reconstruction-free MBRL agent, called DreamerPro, that can enhance robustness to distractions.
- 216, TITLE: NeuralEF: Deconstructing Kernels by Deep Neural Networks
<https://proceedings.mlr.press/v162/deng22b.html>
AUTHORS: Zhijie Deng, Jiaxin Shi, Jun Zhu
HIGHLIGHT: However, the existing method relies on an expensive orthogonalization step and is difficult to implement. We show that these problems can be fixed by using a new series of objective functions that generalizes the EigenGame to function space.
- 217, TITLE: Deep Causal Metric Learning
<https://proceedings.mlr.press/v162/deng22c.html>
AUTHORS: Xiang Deng, Zhongfei Zhang
HIGHLIGHT: However, this can lead the model to recklessly learn all the correlated distances found in training data including the spurious distance (e.g., background differences) that is not the distance of interest and can harm the generalization of the learned metric. To address this issue, we study metric learning from a causality perspective and accordingly propose deep causal metric learning (DCML) that pursues the true causality of the distance between samples.
- 218, TITLE: On the Convergence of Inexact Predictor-Corrector Methods for Linear Programming
<https://proceedings.mlr.press/v162/dexter22a.html>
AUTHORS: Gregory Dexter, Agniva Chowdhury, Haim Avron, Petros Drineas
HIGHLIGHT: To remedy this, we theoretically and empirically analyze (slightly modified) predictor-corrector IPMs when using approximate linear solvers: our approach guarantees that, when certain conditions are satisfied, the number of IPM iterations does not increase and that the final solution remains feasible.
- 219, TITLE: Analysis of Stochastic Processes through Replay Buffers
<https://proceedings.mlr.press/v162/di-castro22a.html>
AUTHORS: Shirli Di-Castro, Shie Mannor, Dotan Di Castro
HIGHLIGHT: In this paper we analyze a system where a stochastic process X is pushed into a replay buffer and then randomly sampled to generate a stochastic process Y from the replay buffer.

- 220, TITLE: Streaming Algorithms for High-Dimensional Robust Statistics
<https://proceedings.mlr.press/v162/diakonikolas22a.html>
AUTHORS: Ilias Diakonikolas, Daniel M. Kane, Ankit Pensia, Thanasis Pittas
HIGHLIGHT: In this work, we develop the first efficient streaming algorithms for high-dimensional robust statistics with near-optimal memory requirements (up to logarithmic factors).
- 221, TITLE: Learning General Halfspaces with Adversarial Label Noise via Online Gradient Descent
<https://proceedings.mlr.press/v162/diakonikolas22b.html>
AUTHORS: Ilias Diakonikolas, Vasilis Kontonis, Christos Tzamos, Nikos Zarifis
HIGHLIGHT: In this work, we show that the problem can be solved directly via online gradient descent applied to a sequence of natural non-convex surrogates.
- 222, TITLE: Variational Feature Pyramid Networks
<https://proceedings.mlr.press/v162/dimitrakopoulos22a.html>
AUTHORS: Panagiotis Dimitrakopoulos, Giorgos Sfikas, Christophoros Nikou
HIGHLIGHT: In this work, we opt to learn a dataset-specific architecture for Feature Pyramid Networks.
- 223, TITLE: Understanding Doubly Stochastic Clustering
<https://proceedings.mlr.press/v162/ding22a.html>
AUTHORS: Tianjiao Ding, Derek Lim, Rene Vidal, Benjamin D Haeffele
HIGHLIGHT: However, the analysis of why this projection improves clustering has been limited. In this paper we present theoretical conditions on the given affinity matrix under which its doubly stochastic projection is an ideal affinity matrix (i.e., it has no false connections between clusters, and is well-connected within each cluster).
- 224, TITLE: Independent Policy Gradient for Large-Scale Markov Potential Games: Sharper Rates, Function Approximation, and Game-Agnostic Convergence
<https://proceedings.mlr.press/v162/ding22b.html>
AUTHORS: Dongsheng Ding, Chen-Yu Wei, Kaiqing Zhang, Mihailo Jovanovic
HIGHLIGHT: To learn a Nash equilibrium of an MPG in which the size of state space and/or the number of players can be very large, we propose new independent policy gradient algorithms that are run by all players in tandem.
- 225, TITLE: Generalization and Robustness Implications in Object-Centric Learning
<https://proceedings.mlr.press/v162/dittadi22a.html>
AUTHORS: Andrea Dittadi, Samuele S Papa, Michele De Vita, Bernhard Schölkopf, Ole Winther, Francesco Locatello
HIGHLIGHT: In this paper, we train state-of-the-art unsupervised models on five common multi-object datasets and evaluate segmentation metrics and downstream object property prediction.
- 226, TITLE: Fair Generalized Linear Models with a Convex Penalty
<https://proceedings.mlr.press/v162/do22a.html>
AUTHORS: Hyungrok Do, Preston Putzel, Axel S Martin, Padhraic Smyth, Judy Zhong
HIGHLIGHT: In this paper we introduce two fairness criteria for GLMs based on equalizing expected outcomes or log-likelihoods.
- 227, TITLE: Bayesian Learning with Information Gain Provably Bounds Risk for a Robust Adversarial Defense
<https://proceedings.mlr.press/v162/doan22a.html>
AUTHORS: Bao Gia Doan, Ehsan M Abbasnejad, Javen Qinfeng Shi, Damith Ranasinghe
HIGHLIGHT: We present a new algorithm to learn a deep neural network model robust against adversarial attacks.
- 228, TITLE: On the Adversarial Robustness of Causal Algorithmic Recourse
<https://proceedings.mlr.press/v162/dominguez-olmedo22a.html>
AUTHORS: Ricardo Dominguez-Olmedo, Amir H Karimi, Bernhard Schölkopf
HIGHLIGHT: In this work, we formulate the adversarially robust recourse problem and show that recourse methods that offer minimally costly recourse fail to be robust.
- 229, TITLE: Finding the Task-Optimal Low-Bit Sub-Distribution in Deep Neural Networks
<https://proceedings.mlr.press/v162/dong22a.html>
AUTHORS: Runpei Dong, Zhanhong Tan, Mengdi Wu, Linfeng Zhang, Kaisheng Ma
HIGHLIGHT: In this paper, we present an adaptive-mapping quantization method to learn an optimal latent sub-distribution that is inherent within models and smoothly approximated with a concrete Gaussian Mixture (GM).
- 230, TITLE: PACE: A Parallelizable Computation Encoder for Directed Acyclic Graphs
<https://proceedings.mlr.press/v162/dong22b.html>
AUTHORS: Zehao Dong, Muhan Zhang, Fuhai Li, Yixin Chen
HIGHLIGHT: In this work, we propose a Parallelizable Attention-based Computation structure Encoder (PACE) that processes nodes simultaneously and encodes DAGs in parallel.
- 231, TITLE: Privacy for Free: How does Dataset Condensation Help Privacy?
<https://proceedings.mlr.press/v162/dong22c.html>
AUTHORS: Tian Dong, Bo Zhao, Lingjuan Lyu

- HIGHLIGHT:** In this work, we for the first time identify that dataset condensation (DC) which is originally designed for improving training efficiency is also a better solution to replace the traditional data generators for private data generation, thus providing privacy for free.
- 232, **TITLE:** Fast rates for noisy interpolation require rethinking the effect of inductive bias
<https://proceedings.mlr.press/v162/donhauser22a.html>
AUTHORS: Konstantin Donhauser, Nicolò Ruggeri, Stefan Stojanovic, Fanny Yang
HIGHLIGHT: Even though this intuition is valid for regularized models, in this paper we caution against a strong inductive bias for interpolation in the presence of noise: While a stronger inductive bias encourages a simpler structure that is more aligned with the ground truth, it also increases the detrimental effect of noise.
- 233, **TITLE:** Adapting to Mixing Time in Stochastic Optimization with Markovian Data
<https://proceedings.mlr.press/v162/dorfman22a.html>
AUTHORS: Ron Dorfman, Kfir Yehuda Levy
HIGHLIGHT: We propose the first optimization method that does not require the knowledge of the mixing time, yet obtains the optimal asymptotic convergence rate when applied to convex problems.
- 234, **TITLE:** TACTiS: Transformer-Attentional Copulas for Time Series
<https://proceedings.mlr.press/v162/drouin22a.html>
AUTHORS: Alexandre Drouin, ?tienne Marcotte, Nicolas Chapados
HIGHLIGHT: In this work, we address the problem of estimating the joint predictive distribution of high-dimensional multivariate time series.
- 235, **TITLE:** Branching Reinforcement Learning
<https://proceedings.mlr.press/v162/du22a.html>
AUTHORS: Yihan Du, Wei Chen
HIGHLIGHT: In this paper, we propose a novel Branching Reinforcement Learning (Branching RL) model, and investigate both Regret Minimization (RM) and Reward-Free Exploration (RFE) metrics for this model.
- 236, **TITLE:** Bayesian Imitation Learning for End-to-End Mobile Manipulation
<https://proceedings.mlr.press/v162/du22b.html>
AUTHORS: Yuqing Du, Daniel Ho, Alex Alemi, Eric Jang, Mohi Khansari
HIGHLIGHT: In this work we investigate and demonstrate benefits of a Bayesian approach to imitation learning from multiple sensor inputs, as applied to the task of opening office doors with a mobile manipulator.
- 237, **TITLE:** GLaM: Efficient Scaling of Language Models with Mixture-of-Experts
<https://proceedings.mlr.press/v162/du22c.html>
AUTHORS: Nan Du, Yanping Huang, Andrew M Dai, Simon Tong, Dmitry Lepikhin, Yuanzhong Xu, Maxim Krikun, Yanqi Zhou, Adams Wei Yu, Orhan Firat, Barret Zoph, Liam Fedus, Maarten P Bosma, Zongwei Zhou, Tao Wang, Emma Wang, Kellie Webster, Marie Pellat, Kevin Robinson, Kathleen Meier-Hellstern, Toju Duke, Lucas Dixon, Kun Zhang, Quoc Le, Yonghui Wu, Zhifeng Chen, Claire Cui
HIGHLIGHT: In this paper, we propose and develop a family of language models named glam (G eneralist L anguage M odel), which uses a sparsely activated mixture-of-experts architecture to scale the model capacity while also incurring substantially less training cost compared to dense variants.
- 238, **TITLE:** Learning Iterative Reasoning through Energy Minimization
<https://proceedings.mlr.press/v162/du22d.html>
AUTHORS: Yilun Du, Shuang Li, Joshua Tenenbaum, Igor Mordatch
HIGHLIGHT: In this work, we present a new framework for iterative reasoning with neural networks.
- 239, **TITLE:** SE(3) Equivariant Graph Neural Networks with Complete Local Frames
<https://proceedings.mlr.press/v162/du22e.html>
AUTHORS: Weitao Du, He Zhang, Yuanqi Du, Qi Meng, Wei Chen, Nanning Zheng, Bin Shao, Tie-Yan Liu
HIGHLIGHT: In this paper, we propose a framework to construct SE(3) equivariant graph neural networks that can approximate the geometric quantities efficiently.
- 240, **TITLE:** A Context-Integrated Transformer-Based Neural Network for Auction Design
<https://proceedings.mlr.press/v162/duan22a.html>
AUTHORS: Zhijian Duan, Jingwu Tang, Yutong Yin, Zhe Feng, Xiang Yan, Manzil Zaheer, Xiaotie Deng
HIGHLIGHT: However, these works either focus on a fixed set of bidders and items, or restrict the auction to be symmetric. In this work, we overcome such limitations by factoring public contextual information of bidders and items into the auction learning framework.
- 241, **TITLE:** Augment with Care: Contrastive Learning for Combinatorial Problems
<https://proceedings.mlr.press/v162/duan22b.html>
AUTHORS: Haonan Duan, Pashootan Vaezipoor, Max B Paulus, Yangjun Ruan, Chris Maddison
HIGHLIGHT: We find that label-preserving augmentations are critical for the success of contrastive pre-training.
- 242, **TITLE:** Parametric Visual Program Induction with Function Modularization
<https://proceedings.mlr.press/v162/duan22c.html>

- AUTHORS: Xuguang Duan, Xin Wang, Ziwei Zhang, Wenwu Zhu
HIGHLIGHT: In this paper, we propose the concept of parametric visual program induction.
- 243, TITLE: Bayesian Deep Embedding Topic Meta-Learner
<https://proceedings.mlr.press/v162/duan22d.html>
AUTHORS: Zhibin Duan, Yishi Xu, Jianqiao Sun, Bo Chen, Wenchao Chen, Chaojie Wang, Mingyuan Zhou
HIGHLIGHT: In this paper, we propose a novel framework that efficiently solves the problem of topic modeling under the small data regime.
- 244, TITLE: Deletion Robust Submodular Maximization over Matroids
<https://proceedings.mlr.press/v162/duetting22a.html>
AUTHORS: Paul Duetting, Federico Fusco, Silvio Lattanzi, Ashkan Norouzi-Fard, Morteza Zadimoghaddam
HIGHLIGHT: In this paper we study the deletion robust version of the problem under the classic matroids constraint.
- 245, TITLE: From data to functa: Your data point is a function and you can treat it like one
<https://proceedings.mlr.press/v162/dupont22a.html>
AUTHORS: Emilien Dupont, Hyunjik Kim, S. M. Ali Eslami, Danilo Jimenez Rezende, Dan Rosenbaum
HIGHLIGHT: A powerful continuous alternative is then to represent these measurements using an implicit neural representation, a neural function trained to output the appropriate measurement value for any input spatial location. In this paper, we take this idea to its next level: what would it take to perform deep learning on these functions instead, treating them as data?
- 246, TITLE: Efficient Low Rank Convex Bounds for Pairwise Discrete Graphical Models
<https://proceedings.mlr.press/v162/durante22a.html>
AUTHORS: Valentin Durante, George Katsirelos, Thomas Schiex
HIGHLIGHT: In this paper, we extend a Burer-Monteiro style method to compute low rank Semi-Definite Programming (SDP) bounds for the MAP problem on discrete graphical models with an arbitrary number of states and arbitrary pairwise potentials.
- 247, TITLE: Robust Counterfactual Explanations for Tree-Based Ensembles
<https://proceedings.mlr.press/v162/dutta22a.html>
AUTHORS: Sanghamitra Dutta, Jason Long, Saumitra Mishra, Cecilia Tilli, Daniele Magazzeni
HIGHLIGHT: In this work, we propose a novel strategy - that we call RobX - to generate robust counterfactuals for tree-based ensembles, e.g., XGBoost.
- 248, TITLE: On the Difficulty of Defending Self-Supervised Learning against Model Extraction
<https://proceedings.mlr.press/v162/dziedzic22a.html>
AUTHORS: Adam Dziedzic, Nikita Dhawan, Muhammad Ahmad Kaleem, Jonas Guan, Nicolas Papernot
HIGHLIGHT: We thus explore model stealing attacks against SSL.
- 249, TITLE: LIMO: Latent Inceptionism for Targeted Molecule Generation
<https://proceedings.mlr.press/v162/eckmann22a.html>
AUTHORS: Peter Eckmann, Kunyang Sun, Bo Zhao, Mudong Feng, Michael Gilson, Rose Yu
HIGHLIGHT: We present Latent Inceptionism on Molecules (LIMO), which significantly accelerates molecule generation with an inceptionism-like technique.
- 250, TITLE: Inductive Biases and Variable Creation in Self-Attention Mechanisms
<https://proceedings.mlr.press/v162/edelman22a.html>
AUTHORS: Benjamin L Edelman, Surbhi Goel, Sham Kakade, Cyril Zhang
HIGHLIGHT: To support our analysis, we present synthetic experiments to probe the sample complexity of learning sparse Boolean functions with Transformers.
- 251, TITLE: Provable Reinforcement Learning with a Short-Term Memory
<https://proceedings.mlr.press/v162/efroni22a.html>
AUTHORS: Yonathan Efroni, Chi Jin, Akshay Krishnamurthy, Sobhan Miryoosefi
HIGHLIGHT: Motivated by the problem structure in several physical applications, as well as a commonly used technique known as “frame stacking”, this paper proposes to study a new subclass of POMDPs, whose latent states can be decoded by the most recent history of a short length m .
- 252, TITLE: Sparsity in Partially Controllable Linear Systems
<https://proceedings.mlr.press/v162/efroni22b.html>
AUTHORS: Yonathan Efroni, Sham Kakade, Akshay Krishnamurthy, Cyril Zhang
HIGHLIGHT: In particular, our structural results characterize those state variables which are irrelevant for optimal control, an analysis which departs from classical control techniques.
- 253, TITLE: FedNew: A Communication-Efficient and Privacy-Preserving Newton-Type Method for Federated Learning
<https://proceedings.mlr.press/v162/elgabri22a.html>
AUTHORS: Anis Elgabri, Chaouki Ben Issaid, Amrit Singh Bedi, Ketan Rajawat, Mehdi Bennis, Vaneet Aggarwal
HIGHLIGHT: In this work, we introduced a novel framework called FedNew in which there is no need to transmit Hessian information from clients to PS, hence resolving the bottleneck to improve communication efficiency.
- 254, TITLE: pathGCN: Learning General Graph Spatial Operators from Paths

- <https://proceedings.mlr.press/v162/eliasof22a.html>
AUTHORS: Moshe Eliasof, Eldad Haber, Eran Treister
HIGHLIGHT: In this paper we propose pathGCN, a novel approach to learn the spatial operator from random paths on the graph.
- 255, TITLE: Discrete Tree Flows via Tree-Structured Permutations
<https://proceedings.mlr.press/v162/elkady22a.html>
AUTHORS: Mai Elkady, Hyung Zin Lim, David I Inouye
HIGHLIGHT: Our approach seeks to reduce computational burden and remove the need for pseudo-gradients by developing a discrete flow based on decision trees—building upon the success of efficient tree-based methods for classification and regression for discrete data.
- 256, TITLE: For Learning in Symmetric Teams, Local Optima are Global Nash Equilibria
<https://proceedings.mlr.press/v162/emmons22a.html>
AUTHORS: Scott Emmons, Caspar Oesterheld, Andrew Critch, Vincent Conitzer, Stuart Russell
HIGHLIGHT: In this work, we show that any locally optimal symmetric strategy profile is also a (global) Nash equilibrium.
- 257, TITLE: Streaming Algorithm for Monotone k -Submodular Maximization with Cardinality Constraints
<https://proceedings.mlr.press/v162/ene22a.html>
AUTHORS: Alina Ene, Huy Nguyen
HIGHLIGHT: In this work, we develop a new streaming algorithm for maximizing a monotone k -submodular function subject to a per-coordinate cardinality constraint attaining an approximation guarantee close to the state of the art guarantee in the offline setting.
- 258, TITLE: Towards Scaling Difference Target Propagation by Learning Backprop Targets
<https://proceedings.mlr.press/v162/ernout22a.html>
AUTHORS: Maxence M Ernout, Fabrice Normandin, Abhinav Moudgil, Sean Spinney, Eugene Belilovsky, Irina Rish, Blake Richards, Yoshua Bengio
HIGHLIGHT: In this paper, we propose a novel feedback weight training scheme that ensures both that DTP approximates BP and that layer-wise feedback weight training can be restored without sacrificing any theoretical guarantees.
- 259, TITLE: Understanding Dataset Difficulty with V -Usable Information
<https://proceedings.mlr.press/v162/ethayarajh22a.html>
AUTHORS: Kavin Ethayarajh, Yejin Choi, Swabha Swayamdipta
HIGHLIGHT: To address these questions, we frame dataset difficulty-w.r.t. a model V -as the lack of V -usable information (Xu et al., 2019), where a lower value indicates a more difficult dataset for V .
- 260, TITLE: Head2Toe: Utilizing Intermediate Representations for Better Transfer Learning
<https://proceedings.mlr.press/v162/evci22a.html>
AUTHORS: Utku Evci, Vincent Dumoulin, Hugo Larochelle, Michael C Mozer
HIGHLIGHT: We propose a method, Head-to-Toe probing (Head2Toe), that selects features from all layers of the source model to train a classification head for the target-domain.
- 261, TITLE: Variational Sparse Coding with Learned Thresholding
<https://proceedings.mlr.press/v162/fallah22a.html>
AUTHORS: Kion Fallah, Christopher J Rozell
HIGHLIGHT: In this work, we propose a new approach to variational sparse coding that allows us to learn sparse distributions by thresholding samples, avoiding the use of problematic relaxations.
- 262, TITLE: Training Discrete Deep Generative Models via Gapped Straight-Through Estimator
<https://proceedings.mlr.press/v162/fan22a.html>
AUTHORS: Ting-Han Fan, Ta-Chung Chi, Alexander I. Rudnicky, Peter J Ramadge
HIGHLIGHT: We propose a Gapped Straight-Through (GST) estimator to reduce the variance without incurring resampling overhead.
- 263, TITLE: DRIBO: Robust Deep Reinforcement Learning via Multi-View Information Bottleneck
<https://proceedings.mlr.press/v162/fan22b.html>
AUTHORS: Jiameng Fan, Wenchao Li
HIGHLIGHT: Specifically, we introduce a novel contrastive version of the Multi-View Information Bottleneck (MIB) objective for temporal data.
- 264, TITLE: Generalized Data Distribution Iteration
<https://proceedings.mlr.press/v162/fan22c.html>
AUTHORS: Jiajun Fan, Changnan Xiao
HIGHLIGHT: To obtain higher sample efficiency and superior final performance simultaneously has been one of the major challenges for deep reinforcement learning (DRL). Previous work could handle one of these challenges but typically failed to address them concurrently. In this paper, we try to tackle these two challenges simultaneously.
- 265, TITLE: Variational Wasserstein gradient flow
<https://proceedings.mlr.press/v162/fan22d.html>

- AUTHORS: Jiaojiao Fan, Qinsheng Zhang, Amirhossein Taghvaei, Yongxin Chen
HIGHLIGHT: This paper builds on the recent works with a slight but crucial difference: we propose to utilize a variational formulation of the objective function formulated as maximization over a parametric class of functions.
- 266, TITLE: Data Determines Distributional Robustness in Contrastive Language Image Pre-training (CLIP)
<https://proceedings.mlr.press/v162/fang22a.html>
AUTHORS: Alex Fang, Gabriel Ilharco, Mitchell Wortsman, Yuhao Wan, Vaishaal Shankar, Achal Dave, Ludwig Schmidt
HIGHLIGHT: Since these language-image models differ from previous training approaches in several ways, an important question is what causes the large robustness gains. We answer this question via a systematic experimental investigation.
- 267, TITLE: Bayesian Continuous-Time Tucker Decomposition
<https://proceedings.mlr.press/v162/fang22b.html>
AUTHORS: Shikai Fang, Akil Narayan, Robert Kirby, Shandian Zhe
HIGHLIGHT: They either drop the timestamps or bin them into crude steps and hence ignore the temporal dynamics within each step or use simple parametric time coefficients. To overcome these limitations, we propose Bayesian Continuous-Time Tucker Decomposition.
- 268, TITLE: Byzantine Machine Learning Made Easy By Resilient Averaging of Momentums
<https://proceedings.mlr.press/v162/farhadkhani22a.html>
AUTHORS: Sadegh Farhadkhani, Rachid Guerraoui, Nirupam Gupta, Rafael Pinot, John Stephan
HIGHLIGHT: We present RESAM (RESilient Averaging of Momentums), a unified framework that makes it simple to establish optimal Byzantine resilience, relying only on standard machine learning assumptions.
- 269, TITLE: An Equivalence Between Data Poisoning and Byzantine Gradient Attacks
<https://proceedings.mlr.press/v162/farhadkhani22b.html>
AUTHORS: Sadegh Farhadkhani, Rachid Guerraoui, L? Nguy?n Hoang, Oscar Villedmaud
HIGHLIGHT: In this paper, we show a surprising equivalence between this model and data poisoning, a threat considered much more realistic.
- 270, TITLE: Investigating Generalization by Controlling Normalized Margin
<https://proceedings.mlr.press/v162/farhang22a.html>
AUTHORS: Alexander R Farhang, Jeremy D Bernstein, Kushal Tirumala, Yang Liu, Yisong Yue
HIGHLIGHT: The paper finds that yes {—} in a standard training setup, test performance closely tracks normalized margin. The paper suggests a Gaussian process model as a promising explanation for this behavior.
- 271, TITLE: Kernelized Multiplicative Weights for 0/1-Polyhedral Games: Bridging the Gap Between Learning in Extensive-Form and Normal-Form Games
<https://proceedings.mlr.press/v162/farina22a.html>
AUTHORS: Gabriele Farina, Chung-Wei Lee, Haipeng Luo, Christian Kroer
HIGHLIGHT: In this paper we show that the Optimistic Multiplicative Weights Update (OMWU) algorithm—the premier learning algorithm for NFGs—can be simulated on the normal-form equivalent of an EFG in linear time per iteration in the game tree size using a kernel trick.
- 272, TITLE: Local Linear Convergence of Douglas-Rachford for Linear Programming: a Probabilistic Analysis
<https://proceedings.mlr.press/v162/faust22a.html>
AUTHORS: Oisin Faust, Hamza Fawzi
HIGHLIGHT: In this paper we analyze the local linear convergence rate ρ of the DRS method for random linear programs, and give explicit and tight bounds on ρ .
- 273, TITLE: Matching Structure for Dual Learning
<https://proceedings.mlr.press/v162/fei22a.html>
AUTHORS: Hao Fei, Shengqiong Wu, Yafeng Ren, Meishan Zhang
HIGHLIGHT: In this work, we propose to further enhance dual learning with structure matching that explicitly builds structural connections in between.
- 274, TITLE: Cascaded Gaps: Towards Logarithmic Regret for Risk-Sensitive Reinforcement Learning
<https://proceedings.mlr.press/v162/fei22b.html>
AUTHORS: Yingjie Fei, Ruitu Xu
HIGHLIGHT: In this paper, we study gap-dependent regret guarantees for risk-sensitive reinforcement learning based on the entropic risk measure.
- 275, TITLE: Private frequency estimation via projective geometry
<https://proceedings.mlr.press/v162/feldman22a.html>
AUTHORS: Vitaly Feldman, Jelani Nelson, Huy Nguyen, Kunal Talwar
HIGHLIGHT: In this work, we propose a new algorithm ProjectiveGeometryResponse (PGR) for locally differentially private (LDP) frequency estimation.
- 276, TITLE: An Intriguing Property of Geophysics Inversion
<https://proceedings.mlr.press/v162/feng22a.html>
AUTHORS: Yinan Feng, Yinpeng Chen, Shihang Feng, Peng Jin, Zicheng Liu, Youzuo Lin

HIGHLIGHT: To alleviate those issues, recent studies leverage deep neural networks to learn the inversion mappings from measurements to the property directly. In this paper, we show that such a mapping can be well modeled by a very shallow (but not wide) network with only five layers.

277, **TITLE:** Principled Knowledge Extrapolation with GANs
<https://proceedings.mlr.press/v162/feng22b.html>

AUTHORS: Ruili Feng, Jie Xiao, Kecheng Zheng, Deli Zhao, Jingren Zhou, Qibin Sun, Zheng-Jun Zha
HIGHLIGHT: In this paper, we propose to study counterfactual synthesis from a new perspective of knowledge extrapolation, where a given knowledge dimension of the data distribution is extrapolated, but the remaining knowledge is kept indistinguishable from the original distribution.

278, **TITLE:** A Resilient Distributed Boosting Algorithm
<https://proceedings.mlr.press/v162/filmus22a.html>

AUTHORS: Yuval Filmus, Idan Mehalel, Shay Moran
HIGHLIGHT: We present a distributed boosting algorithm which is resilient to a limited amount of noise.

279, **TITLE:** Model-Value Inconsistency as a Signal for Epistemic Uncertainty
<https://proceedings.mlr.press/v162/filos22a.html>

AUTHORS: Angelos Filos, Eszter V?rtes, Zita Marinho, Gregory Farquhar, Diana Borsa, Abram Friesen, Feryal Behbahani, Tom Schaul, Andre Barreto, Simon Osindero
HIGHLIGHT: Using a model of the environment and a value function, an agent can construct many estimates of a state's value, by unrolling the model for different lengths and bootstrapping with its value function. Our key insight is that one can treat this set of value estimates as a type of ensemble, which we call an implicit value ensemble (IVE).

280, **TITLE:** Coordinated Double Machine Learning
<https://proceedings.mlr.press/v162/fingerhut22a.html>

AUTHORS: Nitai Fingerhut, Matteo Sesia, Yaniv Romano
HIGHLIGHT: While this methodology is flexible and can accommodate arbitrary predictive models, typically trained independently of one another, this paper argues that a carefully coordinated learning algorithm for deep neural networks may reduce the estimation bias.

281, **TITLE:** Conformal Prediction Sets with Limited False Positives
<https://proceedings.mlr.press/v162/fisch22a.html>

AUTHORS: Adam Fisch, Tal Schuster, Tommi Jaakkola, Dr.Regina Barzilay
HIGHLIGHT: We develop a new approach to multi-label conformal prediction in which we aim to output a precise set of promising prediction candidates with a bounded number of incorrect answers.

282, **TITLE:** Fast Population-Based Reinforcement Learning on a Single Machine
<https://proceedings.mlr.press/v162/flajolet22a.html>

AUTHORS: Arthur Flajolet, Claire Bizon Monroc, Karim Beguir, Thomas Pierrot
HIGHLIGHT: In this work, we compare implementations and revisit previous studies to show that the judicious use of compilation and vectorization allows population-based training to be performed on a single machine with one accelerator with minimal overhead compared to training a single agent.

283, **TITLE:** Fast Relative Entropy Coding with A* coding
<https://proceedings.mlr.press/v162/flamich22a.html>

AUTHORS: Gergely Flamich, Stratis Markou, Jose Miguel Hernandez-Lobato
HIGHLIGHT: We introduce AS* and AD* coding, two REC algorithms based on A* sampling.

284, **TITLE:** Contrastive Mixture of Posteriors for Counterfactual Inference, Data Integration and Fairness
<https://proceedings.mlr.press/v162/foster22a.html>

AUTHORS: Adam Foster, Arpi Vezer, Craig A. Glastonbury, Paidi Creed, Samer Abujudeh, Aaron Sim
HIGHLIGHT: Adopting a Conditional VAE framework, we show that marginal independence between the representation and a condition variable plays a key role in both of these challenges. We propose the Contrastive Mixture of Posteriors (CoMP) method that uses a novel misalignment penalty defined in terms of mixtures of the variational posteriors to enforce this independence in latent space.

285, **TITLE:** Label Ranking through Nonparametric Regression
<https://proceedings.mlr.press/v162/fotakis22a.html>

AUTHORS: Dimitris Fotakis, Alkis Kalavasis, Eleni Psaroudaki
HIGHLIGHT: We introduce a generative model for Label Ranking, in noiseless and noisy nonparametric regression settings, and provide sample complexity bounds for learning algorithms in both cases.

286, **TITLE:** A Neural Tangent Kernel Perspective of GANs
<https://proceedings.mlr.press/v162/franceschi22a.html>

AUTHORS: Jean-Yves Franceschi, Emmanuel De B?zenac, Ibrahim Ayed, Mickael Chen, Sylvain Lamprier, Patrick Gallinari
HIGHLIGHT: We propose a novel theoretical framework of analysis for Generative Adversarial Networks (GANs).

287, **TITLE:** Extracting Latent State Representations with Linear Dynamics from Rich Observations

- <https://proceedings.mlr.press/v162/frandsen22a.html>
AUTHORS: Abraham Frandsen, Rong Ge, Holden Lee
HIGHLIGHT: We consider a setting where there is a hidden linear subspace of the high-dimensional feature space in which the dynamics are linear.
- 288, TITLE: SPDY: Accurate Pruning with Speedup Guarantees
<https://proceedings.mlr.press/v162/frantar22a.html>
AUTHORS: Elias Frantar, Dan Alistarh
HIGHLIGHT: Yet, most existing pruning methods minimize just the number of remaining weights, i.e. the size of the model, rather than optimizing for inference time. We address this gap by introducing SPDY, a new compression method which automatically determines layer-wise sparsity targets achieving a desired inference speedup on a given system, while minimizing accuracy loss.
- 289, TITLE: Revisiting the Effects of Stochasticity for Hamiltonian Samplers
<https://proceedings.mlr.press/v162/franzese22a.html>
AUTHORS: Giulio Franzese, Dimitrios Miliotis, Maurizio Filippone, Pietro Michiardi
HIGHLIGHT: We revisit the theoretical properties of Hamiltonian stochastic differential equations (SDES) for Bayesian posterior sampling, and we study the two types of errors that arise from numerical SDE simulation: the discretization error and the error due to noisy gradient estimates in the context of data subsampling.
- 290, TITLE: Bregman Neural Networks
<https://proceedings.mlr.press/v162/frecon22a.html>
AUTHORS: Jordan Frecon, Gilles Gasso, Massimiliano Pontil, Saverio Salzo
HIGHLIGHT: We present a framework based on bilevel optimization for learning multilayer, deep data representations.
- 291, TITLE: (Non-)Convergence Results for Predictive Coding Networks
<https://proceedings.mlr.press/v162/frieder22a.html>
AUTHORS: Simon Frieder, Thomas Lukasiewicz
HIGHLIGHT: One major open problem around PCNs is their convergence behavior. In this paper, we use dynamical systems theory to formally investigate the convergence of PCNs as they are used in machine learning.
- 292, TITLE: Scaling Structured Inference with Randomization
<https://proceedings.mlr.press/v162/fu22a.html>
AUTHORS: Yao Fu, John Cunningham, Mirella Lapata
HIGHLIGHT: Here, we propose a family of randomized dynamic programming (RDP) algorithms for scaling structured models to tens of thousands of latent states.
- 293, TITLE: Greedy when Sure and Conservative when Uncertain about the Opponents
<https://proceedings.mlr.press/v162/fu22b.html>
AUTHORS: Haobo Fu, Ye Tian, Hongxiang Yu, Weiming Liu, Shuang Wu, Jiechao Xiong, Ying Wen, Kai Li, Junliang Xing, Qiang Fu, Wei Yang
HIGHLIGHT: We develop a new approach, named Greedy when Sure and Conservative when Uncertain (GSCU), to competing online against unknown and nonstationary opponents.
- 294, TITLE: DepthShrinker: A New Compression Paradigm Towards Boosting Real-Hardware Efficiency of Compact Neural Networks
<https://proceedings.mlr.press/v162/fu22c.html>
AUTHORS: Yonggan Fu, Haichuan Yang, Jiayi Yuan, Meng Li, Cheng Wan, Raghuraman Krishnamoorthi, Vikas Chandra, Yingyan Lin
HIGHLIGHT: In this work, we open up a new compression paradigm for developing real-hardware efficient DNNs, leading to boosted hardware efficiency while maintaining model accuracy.
- 295, TITLE: Revisiting Some Common Practices in Cooperative Multi-Agent Reinforcement Learning
<https://proceedings.mlr.press/v162/fu22d.html>
AUTHORS: Wei Fu, Chao Yu, Zelai Xu, Jiaqi Yang, Yi Wu
HIGHLIGHT: Inspired by our theoretical analysis, we present practical suggestions on implementing multi-agent PG algorithms for either high rewards or diverse emergent behaviors and empirically validate our findings on a variety of domains, ranging from the simplified matrix and grid-world games to complex benchmarks such as StarCraft Multi-Agent Challenge and Google Research Football.
- 296, TITLE: $\mathcal{S}p$ -Laplacian Based Graph Neural Networks
<https://proceedings.mlr.press/v162/fu22e.html>
AUTHORS: Guoji Fu, Peilin Zhao, Yatao Bian
HIGHLIGHT: Moreover, when the topology is non-informative for label prediction, ordinary GNNs may work significantly worse than simply applying multi-layer perceptrons (MLPs) on each node. To tackle the above problem, we propose a new $\mathcal{S}p$ -Laplacian based GNN model, termed as \mathcal{S}^p GNN, whose message passing mechanism is derived from a discrete regularization framework and could be theoretically explained as an approximation of a polynomial graph filter defined on the spectral domain of $\mathcal{S}p$ -Laplacians.
- 297, TITLE: Why Should I Trust You, Bellman? The Bellman Error is a Poor Replacement for Value Error
<https://proceedings.mlr.press/v162/fujimoto22a.html>

- AUTHORS: Scott Fujimoto, David Meger, Doina Precup, Ofir Nachum, Shixiang Shane Gu
HIGHLIGHT: In this work, we study the use of the Bellman equation as a surrogate objective for value prediction accuracy.
- 298, TITLE: Robin Hood and Matthew Effects: Differential Privacy Has Disparate Impact on Synthetic Data
<https://proceedings.mlr.press/v162/ganev22a.html>
AUTHORS: Georgi Ganev, Bristena Oprisanu, Emiliano De Cristofaro
HIGHLIGHT: We analyze the impact of DP on these models vis-a-vis underrepresented classes/subgroups of data, specifically, studying: 1) the size of classes/subgroups in the synthetic data, and 2) the accuracy of classification tasks run on them.
- 299, TITLE: The Complexity of k-Means Clustering when Little is Known
<https://proceedings.mlr.press/v162/ganian22a.html>
AUTHORS: Robert Galian, Thekla Hamm, Viktoriia Korchemna, Karolina Okrasa, Kirill Simonov
HIGHLIGHT: Here, we study the complexity of k-means clustering in settings where most of the data is not known or simply irrelevant.
- 300, TITLE: IDYNO: Learning Nonparametric DAGs from Interventional Dynamic Data
<https://proceedings.mlr.press/v162/gao22a.html>
AUTHORS: Tian Gao, Debarun Bhattacharjya, Elliot Nelson, Miao Liu, Yue Yu
HIGHLIGHT: We propose a new algorithm, IDYNO, to learn the DAG structure from potentially nonlinear times series data by using a continuous optimization framework that includes a recent formulation for continuous acyclicity constraint.
- 301, TITLE: Loss Function Learning for Domain Generalization by Implicit Gradient
<https://proceedings.mlr.press/v162/gao22b.html>
AUTHORS: Boyan Gao, Henry Gouk, Yongxin Yang, Timothy Hospedales
HIGHLIGHT: In particular, we introduce a novel meta-learning approach to loss function search based on implicit gradient.
- 302, TITLE: On the Convergence of Local Stochastic Compositional Gradient Descent with Momentum
<https://proceedings.mlr.press/v162/gao22c.html>
AUTHORS: Hongchang Gao, Junyi Li, Heng Huang
HIGHLIGHT: In this paper, we developed a novel local stochastic compositional gradient descent with momentum method, which facilitates Federated Learning for the stochastic compositional problem.
- 303, TITLE: Deep Reference Priors: What is the best way to pretrain a model?
<https://proceedings.mlr.press/v162/gao22d.html>
AUTHORS: Yansong Gao, Rahul Ramesh, Pratik Chaudhari
HIGHLIGHT: This paper presents the first demonstration of reference priors for medium-scale deep networks and image-based data.
- 304, TITLE: On the Equivalence Between Temporal and Static Equivariant Graph Representations
<https://proceedings.mlr.press/v162/gao22e.html>
AUTHORS: Jianfei Gao, Bruno Ribeiro
HIGHLIGHT: This work formalizes the associational task of predicting node attribute evolution in temporal graphs from the perspective of learning equivariant representations.
- 305, TITLE: Generalizing Gaussian Smoothing for Random Search
<https://proceedings.mlr.press/v162/gao22f.html>
AUTHORS: Katelyn Gao, Ozan Sener
HIGHLIGHT: Based on an analysis of DFO for non-convex functions, we propose to choose a distribution for perturbations that minimizes the mean squared error (MSE) of the gradient estimate.
- 306, TITLE: Rethinking Image-Scaling Attacks: The Interplay Between Vulnerabilities in Machine Learning Systems
<https://proceedings.mlr.press/v162/gao22g.html>
AUTHORS: Yue Gao, Iliia Shumailov, Kassem Fawaz
HIGHLIGHT: In this paper, we investigate the interplay between vulnerabilities of the image scaling procedure and machine learning models in the decision-based black-box setting.
- 307, TITLE: Lazy Estimation of Variable Importance for Large Neural Networks
<https://proceedings.mlr.press/v162/gao22h.html>
AUTHORS: Yue Gao, Abby Stevens, Garvesh Raskutti, Rebecca Willett
HIGHLIGHT: In this work, we propose a fast and flexible method for approximating the reduced model with important inferential guarantees.
- 308, TITLE: Fast and Reliable Evaluation of Adversarial Robustness with Minimum-Margin Attack
<https://proceedings.mlr.press/v162/gao22i.html>
AUTHORS: Ruize Gao, Jiong Xiao Wang, Kaiwen Zhou, Feng Liu, Binghui Xie, Gang Niu, Bo Han, James Cheng
HIGHLIGHT: In this paper, we propose a novel method, minimum-margin (MM) attack, to fast and reliably evaluate adversarial robustness.
- 309, TITLE: Value Function based Difference-of-Convex Algorithm for Bilevel Hyperparameter Selection Problems
<https://proceedings.mlr.press/v162/gao22j.html>

AUTHORS: Lucy L Gao, Jane Ye, Haian Yin, Shangzhi Zeng, Jin Zhang
HIGHLIGHT: In this work, we develop a sequentially convergent Value Function based Difference-of-Convex Algorithm with inexactness (VF-iDCA).

310, TITLE: Learning to Incorporate Texture Saliency Adaptive Attention to Image Cartoonization
<https://proceedings.mlr.press/v162/gao22k.html>

AUTHORS: Xiang Gao, Yuqi Zhang, Yingjie Tian
HIGHLIGHT: To this end, a novel cartoon-texture-saliency-sampler (CTSS) module is proposed to adaptively sample cartoon-texture-salient patches from training data.

311, TITLE: Stochastic smoothing of the top-K calibrated hinge loss for deep imbalanced classification
<https://proceedings.mlr.press/v162/garcin22a.html>

AUTHORS: Camille Garcin, Maximilien Servajean, Alexis Joly, Joseph Salmon
HIGHLIGHT: In this paper we introduce a stochastic top-K hinge loss inspired by recent developments on top-K calibrated losses.

312, TITLE: PAGE-PG: A Simple and Loopless Variance-Reduced Policy Gradient Method with Probabilistic Gradient Estimation
<https://proceedings.mlr.press/v162/gargiani22a.html>

AUTHORS: Matilde Gargiani, Andrea Zanelli, Andrea Martinelli, Tyler Summers, John Lygeros
HIGHLIGHT: After a compact survey on some of the main variance-reduced REINFORCE-type methods, we propose Probabilistic Gradient Estimation for Policy Gradient (PAGE-PG), a novel loopless variance-reduced policy gradient method based on a probabilistic switch between two types of update.

313, TITLE: The power of first-order smooth optimization for black-box non-smooth problems
<https://proceedings.mlr.press/v162/gasnikov22a.html>

AUTHORS: Alexander Gasnikov, Anton Novitskii, Vasilii Novitskii, Farshed Abdukhakimov, Dmitry Kamzolov, Aleksandr Beznosikov, Martin Takac, Pavel Dvurechensky, Bin Gu
HIGHLIGHT: In this paper, besides the oracle complexity, we focus also on iteration complexity, and propose a generic approach that, based on optimal first-order methods, allows to obtain in a black-box fashion new zeroth-order algorithms for non-smooth convex optimization problems.

314, TITLE: A Functional Information Perspective on Model Interpretation
<https://proceedings.mlr.press/v162/gat22a.html>

AUTHORS: Itai Gat, Nitay Calderon, Roi Reichart, Tamir Hazan
HIGHLIGHT: This work suggests a theoretical framework for model interpretability by measuring the contribution of relevant features to the functional entropy of the network with respect to the input.

315, TITLE: UniRank: Unimodal Bandit Algorithms for Online Ranking
<https://proceedings.mlr.press/v162/gauthier22a.html>

AUTHORS: Camille-Sovanneary Gauthier, Romaric Gaudel, Elisa Fromont
HIGHLIGHT: We propose a generic algorithm, UniRank, that tackles state-of-the-art click models.

316, TITLE: Variational Inference with Locally Enhanced Bounds for Hierarchical Models
<https://proceedings.mlr.press/v162/geffner22a.html>

AUTHORS: Tomas Geffner, Justin Domke
HIGHLIGHT: We propose a new family of variational bounds for hierarchical models, based on the application of tightening methods (e.g. importance weighting) separately for each group of local random variables.

317, TITLE: Inducing Causal Structure for Interpretable Neural Networks
<https://proceedings.mlr.press/v162/geiger22a.html>

AUTHORS: Atticus Geiger, Zhengxuan Wu, Hanson Lu, Josh Rozner, Elisa Kreiss, Thomas Icard, Noah Goodman, Christopher Potts
HIGHLIGHT: In many areas, we have well-founded insights about causal structure that would be useful to bring into our trained models while still allowing them to learn in a data-driven fashion. To achieve this, we present the new method of interchange intervention training (IIT).

318, TITLE: Achieving Minimax Rates in Pool-Based Batch Active Learning
<https://proceedings.mlr.press/v162/gentile22a.html>

AUTHORS: Claudio Gentile, Zhilei Wang, Tong Zhang
HIGHLIGHT: In this paper we propose a solution which requires a careful trade off between the informativeness of the queried points and their diversity.

319, TITLE: Near-Exact Recovery for Tomographic Inverse Problems via Deep Learning
<https://proceedings.mlr.press/v162/genzel22a.html>

AUTHORS: Martin Genzel, Ingo G?hring, Jan Macdonald, Maximilian M?rz
HIGHLIGHT: This work is concerned with the following fundamental question in scientific machine learning: Can deep-learning-based methods solve noise-free inverse problems to near-perfect accuracy?

320, TITLE: Online Learning for Min Sum Set Cover and Pandora's Box

- <https://proceedings.mlr.press/v162/gergatsouli22a.html>
AUTHORS: Evangelia Gergatsouli, Christos Tzamos
HIGHLIGHT: We present a computationally efficient algorithm that is constant-competitive against the cost of the optimal search order.
- 321, TITLE: Equivariance versus Augmentation for Spherical Images
<https://proceedings.mlr.press/v162/gerken22a.html>
AUTHORS: Jan Gerken, Oscar Carlsson, Hampus Linander, Fredrik Ohlsson, Christoffer Petersson, Daniel Persson
HIGHLIGHT: We analyze the role of rotational equivariance in convolutional neural networks (CNNs) applied to spherical images.
- 322, TITLE: A Regret Minimization Approach to Multi-Agent Control
<https://proceedings.mlr.press/v162/ghai22a.html>
AUTHORS: Udaya Ghai, Udari Madhushani, Naomi Leonard, Elad Hazan
HIGHLIGHT: We study the problem of multi-agent control of a dynamical system with known dynamics and adversarial disturbances.
- 323, TITLE: Blocks Assemble! Learning to Assemble with Large-Scale Structured Reinforcement Learning
<https://proceedings.mlr.press/v162/ghasemipour22a.html>
AUTHORS: Seyed Kamyar Seyed Ghasemipour, Satoshi Kataoka, Byron David, Daniel Freeman, Shixiang Shane Gu, Igor Mordatch
HIGHLIGHT: We introduce a naturalistic physics-based environment with a set of connectable magnet blocks inspired by children's toy kits.
- 324, TITLE: Faster Privacy Accounting via Evolving Discretization
<https://proceedings.mlr.press/v162/ghazi22a.html>
AUTHORS: Badih Ghazi, Pritish Kamath, Ravi Kumar, Pasin Manurangsi
HIGHLIGHT: We introduce a new algorithm for numerical composition of privacy random variables, useful for computing the accurate differential privacy parameters for compositions of mechanisms.
- 325, TITLE: Plug-In Inversion: Model-Agnostic Inversion for Vision with Data Augmentations
<https://proceedings.mlr.press/v162/ghiasi22a.html>
AUTHORS: Amin Ghiasi, Hamid Kazemi, Steven Reich, Chen Zhu, Micah Goldblum, Tom Goldstein
HIGHLIGHT: In this work, we introduce Plug-In Inversion, which relies on a simple set of augmentations and does not require excessive hyper-parameter tuning.
- 326, TITLE: Offline RL Policies Should Be Trained to be Adaptive
<https://proceedings.mlr.press/v162/ghosh22a.html>
AUTHORS: Dibya Ghosh, Anurag Ajay, Pulkit Agrawal, Sergey Levine
HIGHLIGHT: In this work, we propose that offline RL methods should instead be adaptive in the presence of uncertainty.
- 327, TITLE: Breaking the \sqrt{T} Barrier: Instance-Independent Logarithmic Regret in Stochastic Contextual Linear Bandits
<https://proceedings.mlr.press/v162/ghosh22b.html>
AUTHORS: Avishkek Ghosh, Abishek Sankararaman
HIGHLIGHT: In this paper, we show that stochastic contexts indeed help to reduce the regret from \sqrt{T} to $\text{polylog}(T)$.
- 328, TITLE: SCHA-VAE: Hierarchical Context Aggregation for Few-Shot Generation
<https://proceedings.mlr.press/v162/giannone22a.html>
AUTHORS: Giorgio Giannone, Ole Winther
HIGHLIGHT: We extend current latent variable models for sets to a fully hierarchical approach with an attention-based point to set-level aggregation and call our method SCHA-VAE for Set-Context-Hierarchical-Aggregation Variational Autoencoder.
- 329, TITLE: A Joint Exponential Mechanism For Differentially Private Top- k
<https://proceedings.mlr.press/v162/gillenwater22a.html>
AUTHORS: Jennifer Gillenwater, Matthew Joseph, Andres Munoz, Monica Ribero Diaz
HIGHLIGHT: We present a differentially private algorithm for releasing the sequence of k elements with the highest counts from a data domain of d elements.
- 330, TITLE: Neuro-Symbolic Hierarchical Rule Induction
<https://proceedings.mlr.press/v162/glanois22a.html>
AUTHORS: Claire Glanois, Zhaohui Jiang, Xuening Feng, Paul Weng, Matthieu Zimmer, Dong Li, Wulong Liu, Jianye Hao
HIGHLIGHT: We propose Neuro-Symbolic Hierarchical Rule Induction, an efficient interpretable neuro-symbolic model, to solve Inductive Logic Programming (ILP) problems.
- 331, TITLE: It's Raw! Audio Generation with State-Space Models
<https://proceedings.mlr.press/v162/goel22a.html>
AUTHORS: Karan Goel, Albert Gu, Chris Donahue, Christopher Re

- HIGHLIGHT:** We propose SaShiMi, a new multi-scale architecture for waveform modeling built around the recently introduced S4 model for long sequence modeling.
- 332, **TITLE:** RankSim: Ranking Similarity Regularization for Deep Imbalanced Regression
<https://proceedings.mlr.press/v162/gong22a.html>
AUTHORS: Yu Gong, Greg Mori, Fred Tung
HIGHLIGHT: This paper presents the RankSim (ranking similarity) regularizer for deep imbalanced regression, which encodes an inductive bias that samples that are closer in label space should also be closer in feature space.
- 333, **TITLE:** How to Fill the Optimum Set? Population Gradient Descent with Harmless Diversity
<https://proceedings.mlr.press/v162/gong22b.html>
AUTHORS: Chengyue Gong, Lemeng Wu, Qiang Liu
HIGHLIGHT: Therefore, it is useful to consider the problem of finding a set of diverse points in the optimum set of an objective function. In this work, we frame this problem as a bi-level optimization problem of maximizing a diversity score inside the optimum set of the main loss function, and solve it with a simple population gradient descent framework that iteratively updates the points to maximize the diversity score in a fashion that does not hurt the optimization of the main loss.
- 334, **TITLE:** Partial Label Learning via Label Influence Function
<https://proceedings.mlr.press/v162/gong22c.html>
AUTHORS: Xiuwen Gong, Dong Yuan, Wei Bao
HIGHLIGHT: In this paper, inspired by influence function, we develop a novel PLL framework called Partial Label Learning via Label Influence Function (PLL-IF).
- 335, **TITLE:** Secure Distributed Training at Scale
<https://proceedings.mlr.press/v162/gorbunov22a.html>
AUTHORS: Eduard Gorbunov, Alexander Borzunov, Michael Diskin, Max Ryabinin
HIGHLIGHT: In this work, we propose a novel protocol for secure (Byzantine-tolerant) decentralized training that emphasizes communication efficiency.
- 336, **TITLE:** Retrieval-Augmented Reinforcement Learning
<https://proceedings.mlr.press/v162/goyal22a.html>
AUTHORS: Anirudh Goyal, Abram Friesen, Andrea Banino, Theophane Weber, Nan Rosemary Ke, Adrià Puigdomènech Badia, Arthur Guez, Mehdi Mirza, Peter C Humphreys, Ksenia Konyushova, Michal Valko, Simon Osindero, Timothy Lillicrap, Nicolas Heess, Charles Blundell
HIGHLIGHT: In this paper we explore an alternative paradigm in which we train a network to map a dataset of past experiences to optimal behavior.
- 337, **TITLE:** The State of Sparse Training in Deep Reinforcement Learning
<https://proceedings.mlr.press/v162/graesser22a.html>
AUTHORS: Laura Graesser, Utku Evci, Erich Elsen, Pablo Samuel Castro
HIGHLIGHT: In this work we perform a systematic investigation into applying a number of existing sparse training techniques on a variety of DRL agents and environments.
- 338, **TITLE:** Causal Inference Through the Structural Causal Marginal Problem
<https://proceedings.mlr.press/v162/gresele22a.html>
AUTHORS: Luigi Gresele, Julius Von Kumbler, Jonas Kumbler, Elke Kirschbaum, Bernhard Schölkopf, Dominik Janzing
HIGHLIGHT: We introduce an approach to counterfactual inference based on merging information from multiple datasets.
- 339, **TITLE:** Mirror Learning: A Unifying Framework of Policy Optimisation
<https://proceedings.mlr.press/v162/grudzien22a.html>
AUTHORS: Jakub Grudzien, Christian A Schroeder De Witt, Jakob Foerster
HIGHLIGHT: In contrast, in this paper, we introduce a novel theoretical framework, named Mirror Learning, which provides theoretical guarantees to a large class of algorithms, including TRPO and PPO.
- 340, **TITLE:** Adapting k-means Algorithms for Outliers
<https://proceedings.mlr.press/v162/grunau22a.html>
AUTHORS: Christoph Grunau, Vlad Rozhn
HIGHLIGHT: In this paper, we build on their ideas and show how to adapt several sequential and distributed k-means algorithms to the setting with outliers, but with substantially stronger theoretical guarantees: our algorithms output $(1 + \epsilon)z$ outliers while achieving an $O(1/\epsilon)$ -approximation to the objective function.
- 341, **TITLE:** Variational Mixtures of ODEs for Inferring Cellular Gene Expression Dynamics
<https://proceedings.mlr.press/v162/gu22a.html>
AUTHORS: Yichen Gu, David T Blauw, Joshua Welch
HIGHLIGHT: Additionally, a single progenitor cell type often bifurcates into multiple child cell types, further complicating the problem of modeling the dynamics. To address this problem, we developed an approach called variational mixtures of ordinary differential equations.
- 342, **TITLE:** Learning Pseudometric-based Action Representations for Offline Reinforcement Learning
<https://proceedings.mlr.press/v162/gu22b.html>

AUTHORS: Pengjie Gu, Mengchen Zhao, Chen Chen, Dong Li, Jianye Hao, Bo An
HIGHLIGHT: This paper proposes an action representation learning framework for offline RL based on a pseudometric, which measures both the behavioral relation and the data-distributional relation between actions.

343, **TITLE:** NeuroFluid: Fluid Dynamics Grounding with Particle-Driven Neural Radiance Fields
<https://proceedings.mlr.press/v162/guan22a.html>

AUTHORS: Shanyan Guan, Huayu Deng, Yunbo Wang, Xiaokang Yang
HIGHLIGHT: In this paper, we consider a partially observable scenario known as fluid dynamics grounding, that is, inferring the state transitions and interactions within the fluid particle systems from sequential visual observations of the fluid surface.

344, **TITLE:** Fast-Rate PAC-Bayesian Generalization Bounds for Meta-Learning
<https://proceedings.mlr.press/v162/guan22b.html>

AUTHORS: Jiechao Guan, Zhiwu Lu
HIGHLIGHT: In this work, we propose a general PAC-Bayesian framework to cope with single-task learning and meta-learning uniformly.

345, **TITLE:** Leveraging Approximate Symbolic Models for Reinforcement Learning via Skill Diversity
<https://proceedings.mlr.press/v162/guan22c.html>

AUTHORS: Lin Guan, Sarath Sreedharan, Subbarao Kambhampati
HIGHLIGHT: Symbolic models of real world tasks are however often incomplete. To this end, we introduce Approximate Symbolic-Model Guided Reinforcement Learning, wherein we will formalize the relationship between the symbolic model and the underlying MDP that will allow us to characterize the incompleteness of the symbolic model.

346, **TITLE:** Large-Scale Graph Neural Architecture Search
<https://proceedings.mlr.press/v162/guan22d.html>

AUTHORS: Chaoyu Guan, Xin Wang, Hong Chen, Ziwei Zhang, Wenwu Zhu
HIGHLIGHT: However, existing approaches fail to handle large-scale graphs because current performance estimation strategies in GNAS are computationally expensive for large-scale graphs and suffer from consistency collapse issues. To tackle these problems, we propose the Graph Architecture Search at Scale (GAUSS) method that can handle large-scale graphs by designing an efficient light-weight supernet and the joint architecture-graph sampling.

347, **TITLE:** Identifiability Conditions for Domain Adaptation
<https://proceedings.mlr.press/v162/gulrajani22a.html>

AUTHORS: Ishaan Gulrajani, Tatsunori Hashimoto
HIGHLIGHT: Unfortunately, it is unclear under what conditions this identifiability assumption holds, even when restricting ourselves to the case where a correct bijective map between domains exists. We study this bijective domain mapping problem and provide several new sufficient conditions for the identifiability of linear domain maps.

348, **TITLE:** A Parametric Class of Approximate Gradient Updates for Policy Optimization
<https://proceedings.mlr.press/v162/gummadi22a.html>

AUTHORS: Ramki Gummadi, Saurabh Kumar, Junfeng Wen, Dale Schuurmans
HIGHLIGHT: To better capture the commonalities and identify key differences between policy optimization methods, we develop a unified perspective that re-expresses the underlying updates in terms of a limited choice of gradient form and scaling function.

349, **TITLE:** Provably Efficient Offline Reinforcement Learning for Partially Observable Markov Decision Processes
<https://proceedings.mlr.press/v162/guo22a.html>

AUTHORS: Hongyi Guo, Qi Cai, Yufeng Zhang, Zhuoran Yang, Zhaoran Wang
HIGHLIGHT: In the offline setting, estimating these operators directly is challenging due to (i) the large observation space and (ii) insufficient coverage of the offline dataset. To tackle these challenges, we propose a novel algorithm that constructs confidence regions for these Bellman operators via offline estimation of their RKHS embeddings, and returns the final policy via pessimistic planning within the confidence regions.

350, **TITLE:** No-Regret Learning in Partially-Informed Auctions
<https://proceedings.mlr.press/v162/guo22b.html>

AUTHORS: Wenshuo Guo, Michael Jordan, Ellen Vitercik
HIGHLIGHT: Auctions with partially-revealed information about items are broadly employed in real-world applications, but the underlying mechanisms have limited theoretical support. In this work, we study a machine learning formulation of these types of mechanisms, presenting algorithms that are no-regret from the buyer's perspective.

351, **TITLE:** Bounding Training Data Reconstruction in Private (Deep) Learning
<https://proceedings.mlr.press/v162/guo22c.html>

AUTHORS: Chuan Guo, Brian Karrer, Kamalika Chaudhuri, Laurens van der Maaten
HIGHLIGHT: In this paper, we derive the first semantic guarantees for DP mechanisms against training data reconstruction attacks under a formal threat model.

352, **TITLE:** Adversarially trained neural representations are already as robust as biological neural representations
<https://proceedings.mlr.press/v162/guo22d.html>

AUTHORS: Chong Guo, Michael Lee, Guillaume Leclerc, Joel Dapello, Yug Rao, Aleksander Madry, James Dicarlo
HIGHLIGHT: In this work, we develop a method for performing adversarial visual attacks directly on primate brain activity.

- 353, TITLE: Class-Imbalanced Semi-Supervised Learning with Adaptive Thresholding
<https://proceedings.mlr.press/v162/guo22e.html>
AUTHORS: Lan-Zhe Guo, Yu-Feng Li
HIGHLIGHT: In this paper, we develop a simple yet effective framework, which only involves adaptive thresholding for different classes in SSL algorithms, and achieves remarkable performance improvement on more than twenty imbalance ratios.
- 354, TITLE: Deep Squared Euclidean Approximation to the Levenshtein Distance for DNA Storage
<https://proceedings.mlr.press/v162/guo22f.html>
AUTHORS: Alan J.X. Guo, Cong Liang, Qing-Hu Hou
HIGHLIGHT: In this work, we propose a novel deep squared Euclidean embedding for DNA sequences using Siamese neural network, squared Euclidean embedding, and chi-squared regression.
- 355, TITLE: Online Continual Learning through Mutual Information Maximization
<https://proceedings.mlr.press/v162/guo22g.html>
AUTHORS: Yiduo Guo, Bing Liu, Dongyan Zhao
HIGHLIGHT: This paper proposed a new online continual learning approach called OCMM based on mutual information (MI) maximization.
- 356, TITLE: Fast Provably Robust Decision Trees and Boosting
<https://proceedings.mlr.press/v162/guo22h.html>
AUTHORS: Jun-Qi Guo, Ming-Zhuo Teng, Wei Gao, Zhi-Hua Zhou
HIGHLIGHT: This work proposes the Fast Provably Robust Decision Tree (FPRDT) with the smallest computational complexity $O(n \log n)$, a tradeoff between global and local optimizations over the adversarial 0/1 loss.
- 357, TITLE: Understanding and Improving Knowledge Graph Embedding for Entity Alignment
<https://proceedings.mlr.press/v162/guo22i.html>
AUTHORS: Lingbing Guo, Qiang Zhang, Zequn Sun, Mingyang Chen, Wei Hu, Huajun Chen
HIGHLIGHT: To fill the research gap, we define a typical paradigm abstracted from existing EEA methods and analyze how the embedding discrepancy between two potentially aligned entities is implicitly bounded by a predefined margin in the score function.
- 358, TITLE: NISPA: Neuro-Inspired Stability-Plasticity Adaptation for Continual Learning in Sparse Networks
<https://proceedings.mlr.press/v162/gurbuz22a.html>
AUTHORS: Mustafa B Gurbuz, Constantine Dovrolis
HIGHLIGHT: The main desiderata associated with CL are to maintain performance on older tasks, leverage the latter to improve learning of future tasks, and to introduce minimal overhead in the training process (for instance, to not require a growing model or retraining). We propose the Neuro-Inspired Stability-Plasticity Adaptation (NISPA) architecture that addresses these desiderata through a sparse neural network with fixed density.
- 359, TITLE: Active Learning on a Budget: Opposite Strategies Suit High and Low Budgets
<https://proceedings.mlr.press/v162/hacohen22a.html>
AUTHORS: Guy Hacohen, Avihu Dekel, Daphna Weinshall
HIGHLIGHT: Accordingly, we propose TypiClust – a deep active learning strategy suited for low budgets.
- 360, TITLE: You Only Cut Once: Boosting Data Augmentation with a Single Cut
<https://proceedings.mlr.press/v162/han22a.html>
AUTHORS: Junlin Han, Pengfei Fang, Weihao Li, Jie Hong, Mohammad Ali Armin, Ian Reid, Lars Petersson, Hongdong Li
HIGHLIGHT: We present You Only Cut Once (YOCO) for performing data augmentations.
- 361, TITLE: Scalable MCMC Sampling for Nonsymmetric Determinantal Point Processes
<https://proceedings.mlr.press/v162/han22b.html>
AUTHORS: Insu Han, Mike Gartrell, Elvis Dohmatob, Amin Karbasi
HIGHLIGHT: In this work, we develop a scalable MCMC sampling algorithm for k -NDPPs with low-rank kernels, thus enabling runtime that is sublinear in n .
- 362, TITLE: G-Mixup: Graph Data Augmentation for Graph Classification
<https://proceedings.mlr.press/v162/han22c.html>
AUTHORS: Xiaotian Han, Zhimeng Jiang, Ninghao Liu, Xia Hu
HIGHLIGHT: However, it is challenging to directly adopt Mixup to augment graph data because different graphs typically: 1) have different numbers of nodes; 2) are not readily aligned; and 3) have unique typologies in non-Euclidean space. To this end, we propose G-Mixup to augment graphs for graph classification by interpolating the generator (i.e., graphon) of different classes of graphs.
- 363, TITLE: Private Streaming SCO in ℓ_p geometry with Applications in High Dimensional Online Decision Making
<https://proceedings.mlr.press/v162/han22d.html>
AUTHORS: Yuxuan Han, Zhicong Liang, Zhipeng Liang, Yang Wang, Yuan Yao, Jiheng Zhang
HIGHLIGHT: We propose a private variant of the Frank-Wolfe algorithm with recursive gradients for variance reduction to update and reveal the parameters upon each data.

- 364, TITLE: Off-Policy Reinforcement Learning with Delayed Rewards
<https://proceedings.mlr.press/v162/han22e.html>
AUTHORS: Beining Han, Zhizhou Ren, Zuofan Wu, Yuan Zhou, Jian Peng
HIGHLIGHT: We study deep reinforcement learning (RL) algorithms with delayed rewards.
- 365, TITLE: Adversarial Attacks on Gaussian Process Bandits
<https://proceedings.mlr.press/v162/han22f.html>
AUTHORS: Eric Han, Jonathan Scarlett
HIGHLIGHT: Our goal is to understand adversarial attacks on GP bandits from theoretical and practical perspectives.
- 366, TITLE: Random Gegenbauer Features for Scalable Kernel Methods
<https://proceedings.mlr.press/v162/han22g.html>
AUTHORS: Insu Han, Amir Zandieh, Haim Avron
HIGHLIGHT: We propose efficient random features for approximating a new and rich class of kernel functions that we refer to as Generalized Zonal Kernels (GZK).
- 367, TITLE: Stochastic Reweighted Gradient Descent
<https://proceedings.mlr.press/v162/hanchi22a.html>
AUTHORS: Ayoub El Hanchi, David Stephens, Chris Maddison
HIGHLIGHT: In this work, we propose stochastic reweighted gradient descent (SRG), a stochastic gradient method based solely on importance sampling that can reduce the variance of the gradient estimator and improve on the asymptotic error of stochastic gradient descent (SGD) in the strongly convex and smooth case.
- 368, TITLE: Dual Perspective of Label-Specific Feature Learning for Multi-Label Classification
<https://proceedings.mlr.press/v162/hang22a.html>
AUTHORS: Jun-Yi Hang, Min-Ling Zhang
HIGHLIGHT: In this paper, we propose a dual perspective for label-specific feature learning, where label-specific discriminative properties are considered by identifying each label's own non-informative features and making the discrimination process immutable to variations of these features.
- 369, TITLE: Temporal Difference Learning for Model Predictive Control
<https://proceedings.mlr.press/v162/hansen22a.html>
AUTHORS: Nicklas A Hansen, Hao Su, Xiaolong Wang
HIGHLIGHT: In this work, we combine the strengths of model-free and model-based methods.
- 370, TITLE: Bisimulation Makes Analogies in Goal-Conditioned Reinforcement Learning
<https://proceedings.mlr.press/v162/hansen-estruch22a.html>
AUTHORS: Philippe Hansen-Estruch, Amy Zhang, Ashvin Nair, Patrick Yin, Sergey Levine
HIGHLIGHT: We propose a new form of state abstraction called goal-conditioned bisimulation that captures functional equivariance, allowing for the reuse of skills to achieve new goals.
- 371, TITLE: TURF: Two-Factor, Universal, Robust, Fast Distribution Learning Algorithm
<https://proceedings.mlr.press/v162/hao22a.html>
AUTHORS: Yi Hao, Ayush Jain, Alon Orlitsky, Vaishakh Ravindrakumar
HIGHLIGHT: We derive a near-linear-time and essentially sample-optimal estimator that establishes $c_{\{t,d\}}=2$ for all $(t,d) \neq (1,0)$.
- 372, TITLE: Contextual Information-Directed Sampling
<https://proceedings.mlr.press/v162/hao22b.html>
AUTHORS: Botao Hao, Tor Lattimore, Chao Qin
HIGHLIGHT: We investigate the IDS design through two contextual bandit problems: contextual bandits with graph feedback and sparse linear contextual bandits.
- 373, TITLE: GSmooth: Certified Robustness against Semantic Transformations via Generalized Randomized Smoothing
<https://proceedings.mlr.press/v162/hao22c.html>
AUTHORS: Zhongkai Hao, Chengyang Ying, Yinpeng Dong, Hang Su, Jian Song, Jun Zhu
HIGHLIGHT: However, existing methods are insufficient or unable to provably defend against semantic transformations, especially those without closed-form expressions (such as defocus blur and pixelate), which are more common in practice and often unrestricted. To fill up this gap, we propose generalized randomized smoothing (GSmooth), a unified theoretical framework for certifying robustness against general semantic transformations via a novel dimension augmentation strategy.
- 374, TITLE: Implicit Regularization with Polynomial Growth in Deep Tensor Factorization
<https://proceedings.mlr.press/v162/hariz22a.html>
AUTHORS: Kais Hariz, Hachem Kadri, Stephane Ayache, Maher Moakher, Thierry Artieres
HIGHLIGHT: We study the implicit regularization effects of deep learning in tensor factorization.
- 375, TITLE: Strategic Instrumental Variable Regression: Recovering Causal Relationships From Strategic Responses
<https://proceedings.mlr.press/v162/harris22a.html>
AUTHORS: Keegan Harris, Dung Daniel T Ngo, Logan Stapleton, Hoda Heidari, Steven Wu

HIGHLIGHT: Specifically, our work establishes a novel connection between strategic responses to ML models and instrumental variable (IV) regression by observing that the sequence of deployed models can be viewed as an instrument that affects agents' observable features but does not directly influence their outcomes.

376, **TITLE:** C*-algebra Net: A New Approach Generalizing Neural Network Parameters to C*-algebra
<https://proceedings.mlr.press/v162/hashimoto22a.html>

AUTHORS: Yuka Hashimoto, Zhao Wang, Tomoko Matsui

HIGHLIGHT: We propose a new framework that generalizes the parameters of neural network models to C^* -algebra-valued ones.

377, **TITLE:** General-purpose, long-context autoregressive modeling with Perceiver AR
<https://proceedings.mlr.press/v162/hawthorne22a.html>

AUTHORS: Curtis Hawthorne, Andrew Jaegle, Catalina Cangea, Sebastian Borgeaud, Charlie Nash, Mateusz Malinowski, Sander Dieleman, Oriol Vinyals, Matthew Botvinick, Ian Simon, Hannah Sheahan, Neil Zeghidour, Jean-Baptiste Alayrac, Joao Carreira, Jesse Engel

HIGHLIGHT: We develop Perceiver AR, an autoregressive, modality-agnostic architecture which uses cross-attention to map long-range inputs to a small number of latents while also maintaining end-to-end causal masking.

378, **TITLE:** On Distribution Shift in Learning-based Bug Detectors
<https://proceedings.mlr.press/v162/he22a.html>

AUTHORS: Jingxuan He, Luca Beurer-Kellner, Martin Vechev

HIGHLIGHT: In this work, we argue that this massive performance difference is caused by a distribution shift, i.e., a fundamental mismatch between the real bug distribution and the synthetic bug distribution used to train and evaluate the detectors.

379, **TITLE:** GNNRank: Learning Global Rankings from Pairwise Comparisons via Directed Graph Neural Networks
<https://proceedings.mlr.press/v162/he22b.html>

AUTHORS: Yixuan He, Quan Gan, David Wipf, Gesine D Reinert, Junchi Yan, Mihai Cucuringu

HIGHLIGHT: In this paper, we introduce neural networks into the ranking recovery problem by proposing the so-called GNNRank, a trainable GNN-based framework with digraph embedding.

380, **TITLE:** Exploring the Gap between Collapsed & Whitened Features in Self-Supervised Learning
<https://proceedings.mlr.press/v162/he22c.html>

AUTHORS: Bobby He, Mete Ozay

HIGHLIGHT: We identify power law behaviour in eigenvalue decay, parameterised by exponent $\beta=0$, as a spectrum that bridges between the collapsed & whitened feature extremes.

381, **TITLE:** Sparse Double Descent: Where Network Pruning Aggravates Overfitting
<https://proceedings.mlr.press/v162/he22d.html>

AUTHORS: Zheng He, Zeke Xie, Quanzhi Zhu, Zengchang Qin

HIGHLIGHT: In this paper, we have three main contributions. First, we report the novel sparse double descent phenomenon through extensive experiments.

382, **TITLE:** A Reduction from Linear Contextual Bandit Lower Bounds to Estimation Lower Bounds
<https://proceedings.mlr.press/v162/he22e.html>

AUTHORS: Jiahao He, Jiheng Zhang, Rachel Zhang

HIGHLIGHT: In this paper, we complete the reverse direction by establishing the necessity.

383, **TITLE:** HyperPrompt: Prompt-based Task-Conditioning of Transformers
<https://proceedings.mlr.press/v162/he22f.html>

AUTHORS: Yun He, Steven Zheng, Yi Tay, Jai Gupta, Yu Du, Vamsi Aribandi, Zhe Zhao, Yaguang Li, Zhao Chen, Donald Metzler, Heng-Tze Cheng, Ed H. Chi

HIGHLIGHT: Here, we explore the use of HyperNetworks to generate hyper-prompts: we propose HyperPrompt, a novel architecture for prompt-based task-conditioning of self-attention in Transformers.

384, **TITLE:** Label-Descriptive Patterns and Their Application to Characterizing Classification Errors
<https://proceedings.mlr.press/v162/hedderich22a.html>

AUTHORS: Michael A. Hedderich, Jonas Fischer, Dietrich Klakow, Jilles Vreeken

HIGHLIGHT: We propose to discover those feature-value combinations (i.e., patterns) that strongly correlate with correct resp.

385, **TITLE:** NOMU: Neural Optimization-based Model Uncertainty
<https://proceedings.mlr.press/v162/heiss22a.html>

AUTHORS: Jakob M Heiss, Jakob Weissteiner, Hanna S Wutte, Sven Seuken, Josef Teichmann

HIGHLIGHT: However, we find that established benchmarks often fail to reliably capture some of these desiderata, even those that are required by Bayesian theory. To address this, we introduce a new approach for capturing model uncertainty for NNs, which we call Neural Optimization-based Model Uncertainty (NOMU).

386, **TITLE:** Scaling Out-of-Distribution Detection for Real-World Settings
<https://proceedings.mlr.press/v162/hendrycks22a.html>

AUTHORS: Dan Hendrycks, Steven Basart, Mantas Mazeika, Andy Zou, Joseph Kwon, Mohammadreza Mostajabi, Jacob Steinhardt, Dawn Song

HIGHLIGHT: To set the stage for more realistic out-of-distribution detection, we depart from small-scale settings and explore large-scale multiclass and multi-label settings with high-resolution images and thousands of classes.

387, TITLE: Generalization Bounds using Lower Tail Exponents in Stochastic Optimizers

<https://proceedings.mlr.press/v162/hodgkinson22a.html>

AUTHORS: Liam Hodgkinson, Umut Simsekli, Rajiv Khanna, Michael Mahoney

HIGHLIGHT: While recent work has revealed connections between generalization and heavy-tailed behavior in stochastic optimization, they mainly relied on continuous-time approximations; and a rigorous treatment for the original discrete-time iterations is yet to be performed. To bridge this gap, we present novel bounds linking generalization to the lower tail exponent of the transition kernel associated with the optimizer around a local minimum, in both discrete- and continuous-time settings.

388, TITLE: Unsupervised Detection of Contextualized Embedding Bias with Application to Ideology

<https://proceedings.mlr.press/v162/hofmann22a.html>

AUTHORS: Valentin Hofmann, Janet Pierrehumbert, Hinrich Sch?tze

HIGHLIGHT: We propose a fully unsupervised method to detect bias in contextualized embeddings.

389, TITLE: Neural Laplace: Learning diverse classes of differential equations in the Laplace domain

<https://proceedings.mlr.press/v162/holt22a.html>

AUTHORS: Samuel I Holt, Zhaozhi Qian, Mihaela van der Schaar

HIGHLIGHT: In this work, we propose Neural Laplace, a unifying framework for learning diverse classes of DEs including all the aforementioned ones.

390, TITLE: Deep Hierarchy in Bandits

<https://proceedings.mlr.press/v162/hong22a.html>

AUTHORS: Joey Hong, Branislav Kveton, Sumeet Katariya, Manzil Zaheer, Mohammad Ghavamzadeh

HIGHLIGHT: Since the hierarchy can have multiple layers, we call it deep. We propose a hierarchical Thompson sampling algorithm (HierTS) for this problem and show how to implement it efficiently for Gaussian hierarchies.

391, TITLE: DAdaQuant: Doubly-adaptive quantization for communication-efficient Federated Learning

<https://proceedings.mlr.press/v162/hong22a.html>

AUTHORS: Robert H?nig, Yiren Zhao, Robert Mullins

HIGHLIGHT: We introduce DAdaQuant as a doubly-adaptive quantization algorithm that dynamically changes the quantization level across time and different clients.

392, TITLE: Equivariant Diffusion for Molecule Generation in 3D

<https://proceedings.mlr.press/v162/hooegeboom22a.html>

AUTHORS: Emiel Hooegeboom, Vi?ctor Garcia Satorras, Cl?ment Vignac, Max Welling

HIGHLIGHT: This work introduces a diffusion model for molecule generation in 3D that is equivariant to Euclidean transformations.

393, TITLE: Conditional GANs with Auxiliary Discriminative Classifier

<https://proceedings.mlr.press/v162/hou22a.html>

AUTHORS: Liang Hou, Qi Cao, Huawei Shen, Siyuan Pan, Xiaoshuang Li, Xueqi Cheng

HIGHLIGHT: The fundamental reason pointed out in this paper is that the classifier of AC-GAN is generator-agnostic, which therefore cannot provide informative guidance for the generator to approach the joint distribution, resulting in a minimization of the conditional entropy that decreases the intra-class diversity.

394, TITLE: AdAUC: End-to-end Adversarial AUC Optimization Against Long-tail Problems

<https://proceedings.mlr.press/v162/hou22b.html>

AUTHORS: Wenzheng Hou, Qianqian Xu, Zhiyong Yang, Shilong Bao, Yuan He, Qingming Huang

HIGHLIGHT: Under such a scenario, AUC is a much more reasonable metric than accuracy since it is insensitive toward class distribution. Motivated by this, we present an early trial to explore adversarial training methods to optimize AUC.

395, TITLE: Wide Bayesian neural networks have a simple weight posterior: theory and accelerated sampling

<https://proceedings.mlr.press/v162/hron22a.html>

AUTHORS: Jiri Hron, Roman Novak, Jeffrey Pennington, Jascha Sohl-Dickstein

HIGHLIGHT: We introduce reparameterisation, a data-dependent reparameterisation which transforms a Bayesian neural network (BNN) posterior to a distribution whose KL divergence to the BNN prior vanishes as layer widths grow.

396, TITLE: Learning inverse folding from millions of predicted structures

<https://proceedings.mlr.press/v162/hsu22a.html>

AUTHORS: Chloe Hsu, Robert Verkuil, Jason Liu, Zeming Lin, Brian Hie, Tom Sercu, Adam Lerer, Alexander Rives

HIGHLIGHT: We consider the problem of predicting a protein sequence from its backbone atom coordinates.

397, TITLE: Nearly Minimax Optimal Reinforcement Learning with Linear Function Approximation

<https://proceedings.mlr.press/v162/hu22a.html>

AUTHORS: Pihe Hu, Yu Chen, Longbo Huang

HIGHLIGHT: Specifically, we consider the episodic inhomogeneous linear Markov Decision Process (MDP), and propose a novel computation-efficient algorithm, LSVI-UCB⁺, which achieves an $\tilde{O}(\sqrt{Hd\sqrt{T}})$ regret bound where H is the episode length, d is the feature dimension, and T is the number of steps.

398, **TITLE:** Neuron Dependency Graphs: A Causal Abstraction of Neural Networks
<https://proceedings.mlr.press/v162/hu22b.html>

AUTHORS: Yaojie Hu, Jin Tian

HIGHLIGHT: We discover that neural networks exhibit approximate logical dependencies among neurons, and we introduce Neuron Dependency Graphs (NDG) that extract and present them as directed graphs.

399, **TITLE:** Policy Diagnosis via Measuring Role Diversity in Cooperative Multi-agent RL
<https://proceedings.mlr.press/v162/hu22c.html>

AUTHORS: Siyi Hu, Chuanlong Xie, Xiaodan Liang, Xiaojun Chang

HIGHLIGHT: In this study, we quantify the agent’s behavior difference and build its relationship with the policy performance via **Role Diversity**, a metric to measure the characteristics of MARL tasks.

400, **TITLE:** On the Role of Discount Factor in Offline Reinforcement Learning

<https://proceedings.mlr.press/v162/hu22d.html>

AUTHORS: Hao Hu, Yiqin Yang, Qianchuan Zhao, Chongjie Zhang

HIGHLIGHT: This paper examines two distinct effects of γ in offline RL with theoretical analysis, namely the regularization effect and the pessimism effect.

401, **TITLE:** Transformer Quality in Linear Time

<https://proceedings.mlr.press/v162/hua22a.html>

AUTHORS: Weizhe Hua, Zihang Dai, Hanxiao Liu, Quoc Le

HIGHLIGHT: We revisit the design choices in Transformers, and propose methods to address their weaknesses in handling long sequences.

402, **TITLE:** Language Models as Zero-Shot Planners: Extracting Actionable Knowledge for Embodied Agents

<https://proceedings.mlr.press/v162/huang22a.html>

AUTHORS: Wenlong Huang, Pieter Abbeel, Deepak Pathak, Igor Mordatch

HIGHLIGHT: In this paper, we investigate the possibility of grounding high-level tasks, expressed in natural language (e.g. “make breakfast”), to a chosen set of actionable steps (e.g. “open fridge”).

403, **TITLE:** Forward Operator Estimation in Generative Models with Kernel Transfer Operators

<https://proceedings.mlr.press/v162/huang22b.html>

AUTHORS: Zhichun Huang, Rudrasis Chakraborty, Vikas Singh

HIGHLIGHT: We propose a substantially cheaper (and simpler) forward operator estimation strategy based on adapting known results on kernel transfer operators.

404, **TITLE:** Adaptive Best-of-Both-Worlds Algorithm for Heavy-Tailed Multi-Armed Bandits

<https://proceedings.mlr.press/v162/huang22c.html>

AUTHORS: Jiatai Huang, Yan Dai, Longbo Huang

HIGHLIGHT: In this paper, we generalize the concept of heavy-tailed multi-armed bandits to adversarial environments, and develop robust best-of-both-worlds algorithms for heavy-tailed multi-armed bandits (MAB), where losses have α -th ($\alpha \leq 2$) moments bounded by σ^α , while the variances may not exist.

405, **TITLE:** Frustratingly Easy Transferability Estimation

<https://proceedings.mlr.press/v162/huang22d.html>

AUTHORS: Long-Kai Huang, Junzhou Huang, Yu Rong, Qiang Yang, Ying Wei

HIGHLIGHT: Existing estimation algorithms either require intensive training on target tasks or have difficulties in evaluating the transferability between layers. To this end, we propose a simple, efficient, and effective transferability measure named TransRate.

406, **TITLE:** Modality Competition: What Makes Joint Training of Multi-modal Network Fail in Deep Learning? (Provably)

<https://proceedings.mlr.press/v162/huang22e.html>

AUTHORS: Yu Huang, Junyang Lin, Chang Zhou, Hongxia Yang, Longbo Huang

HIGHLIGHT: Recently, it has been observed that the best uni-modal network outperforms the jointly trained multi-modal network across different combinations of modalities on various tasks, which is counter-intuitive since multiple signals would bring more information (Wang et al., 2020). This work provides a theoretical explanation for the emergence of such performance gap in neural networks for the prevalent joint training framework.

407, **TITLE:** Action-Sufficient State Representation Learning for Control with Structural Constraints

<https://proceedings.mlr.press/v162/huang22f.html>

AUTHORS: Biwei Huang, Chaochao Lu, Liu Leqi, Jose Miguel Hernandez-Lobato, Clark Glymour, Bernhard Schölkopf, Kun Zhang

HIGHLIGHT: In this paper, we focus on partially observable environments and propose to learn a minimal set of state representations that capture sufficient information for decision-making, termed Action-Sufficient state Representations (ASRs).

408, **TITLE:** 3DLinker: An E(3) Equivariant Variational Autoencoder for Molecular Linker Design

<https://proceedings.mlr.press/v162/huang22g.html>

AUTHORS: Yinan Huang, Xingang Peng, Jianzhu Ma, Muhan Zhang
HIGHLIGHT: In this work, we focus on a new type of drug design problem — generating a small “linker” to physically attach two independent molecules with their distinct functions.

409, TITLE: SDQ: Stochastic Differentiable Quantization with Mixed Precision
<https://proceedings.mlr.press/v162/huang22h.html>
AUTHORS: Xijie Huang, Zhiqiang Shen, Shichao Li, Zechun Liu, Hu Xianghong, Jeffrey Wicaksana, Eric Xing, Kwang-Ting Cheng
HIGHLIGHT: In this work, we present a novel Stochastic Differentiable Quantization (SDQ) method that can automatically learn the MPQ strategy in a more flexible and globally-optimized space with a smoother gradient approximation.

410, TITLE: Tackling Data Heterogeneity: A New Unified Framework for Decentralized SGD with Sample-induced Topology
<https://proceedings.mlr.press/v162/huang22i.html>
AUTHORS: Yan Huang, Ying Sun, Zehan Zhu, Changzhi Yan, Jinming Xu
HIGHLIGHT: We develop a general framework unifying several gradient-based stochastic optimization methods for empirical risk minimization problems both in centralized and distributed scenarios.

411, TITLE: Efficient Representation Learning via Adaptive Context Pooling
<https://proceedings.mlr.press/v162/huang22j.html>
AUTHORS: Chen Huang, Walter Talbott, Navdeep Jaitly, Joshua M Susskind
HIGHLIGHT: In doing so, they assume a fixed attention granularity defined by the individual tokens (e.g., text characters or image pixels), which may not be optimal for modeling complex dependencies at higher levels. In this paper, we propose ContextPool to address this problem by adapting the attention granularity for each token.

412, TITLE: On the Learning of Non-Autoregressive Transformers
<https://proceedings.mlr.press/v162/huang22k.html>
AUTHORS: Fei Huang, Tianhua Tao, Hao Zhou, Lei Li, Minlie Huang
HIGHLIGHT: In this paper, we present theoretical and empirical analyses to reveal the challenges of NAT learning and propose a unified perspective to understand existing successes.

413, TITLE: Going Deeper into Permutation-Sensitive Graph Neural Networks
<https://proceedings.mlr.press/v162/huang22l.html>
AUTHORS: Zhongyu Huang, Yingheng Wang, Chaozhuo Li, Huiguang He
HIGHLIGHT: In this work, we devise an efficient permutation-sensitive aggregation mechanism via permutation groups, capturing pairwise correlations between neighboring nodes.

414, TITLE: Directed Acyclic Transformer for Non-Autoregressive Machine Translation
<https://proceedings.mlr.press/v162/huang22m.html>
AUTHORS: Fei Huang, Hao Zhou, Yang Liu, Hang Li, Minlie Huang
HIGHLIGHT: In this paper, we propose Directed Acyclic Transformer (DA-Transformer), which represents the hidden states in a Directed Acyclic Graph (DAG), where each path of the DAG corresponds to a specific translation.

415, TITLE: Unsupervised Ground Metric Learning Using Wasserstein Singular Vectors
<https://proceedings.mlr.press/v162/huizing22a.html>
AUTHORS: Geert-Jan Huizing, Laura Cantini, Gabriel Peyr?
HIGHLIGHT: In this paper, we propose for the first time a canonical answer by simultaneously computing an OT distance between samples and between features of a dataset.

416, TITLE: Robust Kernel Density Estimation with Median-of-Means principle
<https://proceedings.mlr.press/v162/humbert22a.html>
AUTHORS: Pierre Humbert, Batiste Le Bars, Ludovic Minvielle
HIGHLIGHT: In this paper, we introduce a robust non-parametric density estimator combining the popular Kernel Density Estimation method and the Median-of-Means principle (MoM-KDE).

417, TITLE: A data-driven approach for learning to control computers
<https://proceedings.mlr.press/v162/humphreys22a.html>
AUTHORS: Peter C Humphreys, David Raposo, Tobias Pohlen, Gregory Thornton, Rachita Chhaparia, Alistair Muldal, Josh Abramson, Petko Georgiev, Adam Santoro, Timothy Lillicrap
HIGHLIGHT: Here we investigate the setting of computer control using keyboard and mouse, with goals specified via natural language.

418, TITLE: Proximal Denoiser for Convergent Plug-and-Play Optimization with Nonconvex Regularization
<https://proceedings.mlr.press/v162/hurault22a.html>
AUTHORS: Samuel Hurault, Arthur Leclaire, Nicolas Papadakis
HIGHLIGHT: Using such a denoiser guarantees the convergence of the PnP version of the Half-Quadratic-Splitting (PnP-HQS) iterative algorithm. In this paper, we show that this gradient denoiser can actually correspond to the proximal operator of another scalar function.

419, TITLE: Inverse Contextual Bandits: Learning How Behavior Evolves over Time

- <https://proceedings.mlr.press/v162/huyuk22a.html>
AUTHORS: Alihan H?y?k, Daniel Jarrett, Mihaela van der Schaar
HIGHLIGHT: To give an answer, we desire a policy learning method that provides interpretable representations of decision-making, in particular capturing an agent's non-stationary knowledge of the world, as well as operating in an offline manner.
- 420, TITLE: Datamodels: Understanding Predictions with Data and Data with Predictions
<https://proceedings.mlr.press/v162/ilyas22a.html>
AUTHORS: Andrew Ilyas, Sung Min Park, Logan Engstrom, Guillaume Leclerc, Aleksander Madry
HIGHLIGHT: We present a conceptual framework, datamodeling, for analyzing the behavior of a model class in terms of the training data.
- 421, TITLE: Parsimonious Learning-Augmented Caching
<https://proceedings.mlr.press/v162/im22a.html>
AUTHORS: Sungjin Im, Ravi Kumar, Aditya Petety, Manish Purohit
HIGHLIGHT: In this paper we introduce and study the setting in which the learning-augmented algorithm can utilize the predictions parsimoniously.
- 422, TITLE: Bayesian Optimization for Distributionally Robust Chance-constrained Problem
<https://proceedings.mlr.press/v162/inatsu22a.html>
AUTHORS: Yu Inatsu, Shion Takeno, Masayuki Karasuyama, Ichiro Takeuchi
HIGHLIGHT: In this study, we consider distributionally robust CC (DRCC) problem and propose a novel DRCC Bayesian optimization method for the case where the distribution of the environmental variables cannot be precisely specified.
- 423, TITLE: LeNSE: Learning To Navigate Subgraph Embeddings for Large-Scale Combinatorial Optimisation
<https://proceedings.mlr.press/v162/ireland22a.html>
AUTHORS: David Ireland, Giovanni Montana
HIGHLIGHT: We propose a low-complexity approach for identifying a (possibly much smaller) subgraph of the original graph where the heuristics can be run in reasonable time and with a high likelihood of finding a global near-optimal solution.
- 424, TITLE: The Dual Form of Neural Networks Revisited: Connecting Test Time Predictions to Training Patterns via Spotlights of Attention
<https://proceedings.mlr.press/v162/irie22a.html>
AUTHORS: Kazuki Irie, R?bert Csord?s, J?rgen Schmidhuber
HIGHLIGHT: However, this dual formulation offers a possibility of directly visualising how an NN makes use of training patterns at test time, by examining the corresponding attention weights. We conduct experiments on small scale supervised image classification tasks in single-task, multi-task, and continual learning settings, as well as language modelling, and discuss potentials and limits of this view for better understanding and interpreting how NNs exploit training patterns.
- 425, TITLE: A Modern Self-Referential Weight Matrix That Learns to Modify Itself
<https://proceedings.mlr.press/v162/irie22b.html>
AUTHORS: Kazuki Irie, Imanol Schlag, R?bert Csord?s, J?rgen Schmidhuber
HIGHLIGHT: We propose a scalable self-referential WM (SRWM) that learns to use outer products and the delta update rule to modify itself.
- 426, TITLE: Revisiting Online Submodular Minimization: Gap-Dependent Regret Bounds, Best of Both Worlds and Adversarial Robustness
<https://proceedings.mlr.press/v162/ito22a.html>
AUTHORS: Shinji Ito
HIGHLIGHT: In this paper, we consider online decision problems with submodular loss functions.
- 427, TITLE: Modeling Strong and Human-Like Gameplay with KL-Regularized Search
<https://proceedings.mlr.press/v162/jacob22a.html>
AUTHORS: Athul Paul Jacob, David J Wu, Gabriele Farina, Adam Lerer, Hengyuan Hu, Anton Bakhtin, Jacob Andreas, Noam Brown
HIGHLIGHT: We consider the task of accurately modeling strong human policies in multi-agent decision-making problems, given examples of human behavior.
- 428, TITLE: A deep convolutional neural network that is invariant to time rescaling
<https://proceedings.mlr.press/v162/jacques22a.html>
AUTHORS: Brandon G Jacques, Zoran Tiganj, Aakash Sarkar, Marc Howard, Per Sederberg
HIGHLIGHT: This paper presents a deep CNN (SITHCon) that uses a logarithmically compressed temporal representation at each level.
- 429, TITLE: Input Dependent Sparse Gaussian Processes
<https://proceedings.mlr.press/v162/jafraسته22a.html>
AUTHORS: Bahram Jafraسته, Carlos Villacampa-Calvo, Daniel Hernandez-Lobato
HIGHLIGHT: A limitation is, however, that in some tasks a large number of inducing points may be required to obtain good results. To alleviate this, we propose here to amortize the computation of the inducing points locations, as well as the parameters of $\$q\$$.

- 430, TITLE: Regret Minimization with Performative Feedback
<https://proceedings.mlr.press/v162/jagadeesan22a.html>
AUTHORS: Meena Jagadeesan, Tijana Zrnica, Celestine Mendler-Dünnlender
HIGHLIGHT: Our main contribution is regret bounds that scale only with the complexity of the distribution shifts and not that of the reward function.
- 431, TITLE: Biological Sequence Design with GFlowNets
<https://proceedings.mlr.press/v162/jain22a.html>
AUTHORS: Moksh Jain, Emmanuel Bengio, Alex Hernandez-Garcia, Jarrid Rector-Brooks, Bonaventure F. P. Dossou, Chanakya Ajit Ekbote, Jie Fu, Tianyu Zhang, Michael Kilgour, Dinghui Zhang, Lena Simine, Payel Das, Yoshua Bengio
HIGHLIGHT: In this work, we propose an active learning algorithm leveraging epistemic uncertainty estimation and the recently proposed GFlowNets as a generator of diverse candidate solutions, with the objective to obtain a diverse batch of useful (as defined by some utility function, for example, the predicted anti-microbial activity of a peptide) and informative candidates after each round.
- 432, TITLE: Combining Diverse Feature Priors
<https://proceedings.mlr.press/v162/jain22b.html>
AUTHORS: Saachi Jain, Dimitris Tsipras, Aleksander Madry
HIGHLIGHT: To improve model generalization, model designers often restrict the features that their models use, either implicitly or explicitly. In this work, we explore the design space of leveraging such feature priors by viewing them as distinct perspectives on the data.
- 433, TITLE: Training Your Sparse Neural Network Better with Any Mask
<https://proceedings.mlr.press/v162/jaiswal22a.html>
AUTHORS: Ajay Kumar Jaiswal, Haoyu Ma, Tianlong Chen, Ying Ding, Zhangyang Wang
HIGHLIGHT: Apart from the popular belief that only the quality of sparse masks matters for sparse training, in this paper we demonstrate an alternative opportunity: one can carefully customize the sparse training techniques to deviate from the default dense network training protocols, consisting of introducing “ghost” neurons and skip connections at the early stage of training, and strategically modifying the initialization as well as labels.
- 434, TITLE: Sequential Covariate Shift Detection Using Classifier Two-Sample Tests
<https://proceedings.mlr.press/v162/jang22a.html>
AUTHORS: Sooyong Jang, Sangdon Park, Insup Lee, Osbert Bastani
HIGHLIGHT: We consider the problem of detecting covariate shift, where the covariate distribution shifts but the conditional distribution of labels given covariates remains the same.
- 435, TITLE: Surrogate Likelihoods for Variational Annealed Importance Sampling
<https://proceedings.mlr.press/v162/jankowiak22a.html>
AUTHORS: Martin Jankowiak, Du Phan
HIGHLIGHT: However, supporting data subsampling in these hybrid methods can be a challenge, a shortcoming that we address by introducing a surrogate likelihood that can be learned jointly with other variational parameters.
- 436, TITLE: Planning with Diffusion for Flexible Behavior Synthesis
<https://proceedings.mlr.press/v162/janner22a.html>
AUTHORS: Michael Janner, Yilun Du, Joshua Tenenbaum, Sergey Levine
HIGHLIGHT: In this paper, we consider what it would look like to fold as much of the trajectory optimization pipeline as possible into the modeling problem, such that sampling from the model and planning with it become nearly identical.
- 437, TITLE: HyperImpute: Generalized Iterative Imputation with Automatic Model Selection
<https://proceedings.mlr.press/v162/jarrett22a.html>
AUTHORS: Daniel Jarrett, Bogdan C Cebere, Tennison Liu, Alicia Curth, Mihaela van der Schaar
HIGHLIGHT: In this work, we study an approach that marries the advantages of both: We propose *HyperImpute*, a generalized iterative imputation framework for adaptively and automatically configuring column-wise models and their hyperparameters.
- 438, TITLE: Mitigating Modality Collapse in Multimodal VAEs via Impartial Optimization
<https://proceedings.mlr.press/v162/javaloy22a.html>
AUTHORS: Adrian Javaloy, Maryam Meghdadi, Isabel Valera
HIGHLIGHT: We refer to this limitation as modality collapse. In this work, we argue that this effect is a consequence of conflicting gradients during multimodal VAE training.
- 439, TITLE: Towards understanding how momentum improves generalization in deep learning
<https://proceedings.mlr.press/v162/jelassi22a.html>
AUTHORS: Samy Jelassi, Yuezhi Li
HIGHLIGHT: In this paper, we adopt another perspective and first empirically show that gradient descent with momentum (GD+M) significantly improves generalization compared to gradient descent (GD) in some deep learning problems. From this observation, we formally study how momentum improves generalization.
- 440, TITLE: MASER: Multi-Agent Reinforcement Learning with Subgoals Generated from Experience Replay Buffer
<https://proceedings.mlr.press/v162/jeon22a.html>

- AUTHORS: Jeewon Jeon, Woojun Kim, Whiyung Jung, Youngchul Sung
HIGHLIGHT: In this paper, we consider cooperative multi-agent reinforcement learning (MARL) with sparse reward.
- 441, TITLE: An Exact Symbolic Reduction of Linear Smart Predict+Optimize to Mixed Integer Linear Programming
<https://proceedings.mlr.press/v162/jeong22a.html>
AUTHORS: Jihwan Jeong, Parth Jaggi, Andrew Butler, Scott Sanner
HIGHLIGHT: In this paper, we cast the SPO problem as a bi-level program and apply Symbolic Variable Elimination (SVE) to analytically solve the lower optimization.
- 442, TITLE: Agnostic Learnability of Halfspaces via Logistic Loss
<https://proceedings.mlr.press/v162/ji22a.html>
AUTHORS: Ziwei Ji, Kwangjun Ahn, Pranjal Awasthi, Satyen Kale, Stefani Karp
HIGHLIGHT: Previously, for a certain broad class of “well-behaved” distributions on the examples, Diakonikolas et al. (2020) proved an $\tilde{\Omega}(\text{OPT})$ lower bound, while Frei et al. (2021) proved an $\tilde{O}(\sqrt{\text{OPT}})$ upper bound, where OPT denotes the best zero-one/misclassification risk of a homogeneous halfspace. In this paper, we close this gap by constructing a well-behaved distribution such that the global minimizer of the logistic risk over this distribution only achieves $\Omega(\sqrt{\text{OPT}})$ misclassification risk, matching the upper bound in (Frei et al., 2021).
- 443, TITLE: Improving Policy Optimization with Generalist-Specialist Learning
<https://proceedings.mlr.press/v162/jia22a.html>
AUTHORS: Zhiwei Jia, Xuanlin Li, Zhan Ling, Shuang Liu, Yiran Wu, Hao Su
HIGHLIGHT: To have the best of both worlds, we propose a novel generalist-specialist training framework.
- 444, TITLE: Translatotron 2: High-quality direct speech-to-speech translation with voice preservation
<https://proceedings.mlr.press/v162/jia22b.html>
AUTHORS: Ye Jia, Michelle Tadmor Ramanovich, Tal Remez, Roi Pomerantz
HIGHLIGHT: We present Translatotron 2, a neural direct speech-to-speech translation model that can be trained end-to-end.
- 445, TITLE: Online Learning and Pricing with Reusable Resources: Linear Bandits with Sub-Exponential Rewards
<https://proceedings.mlr.press/v162/jia22c.html>
AUTHORS: Huiwen Jia, Cong Shi, Siqian Shen
HIGHLIGHT: We propose a rate-optimal online learning and pricing algorithm, termed Batch Linear Confidence Bound (BLinUCB), and prove that the cumulative regret is $\tilde{O}(d_f \sqrt{T})$.
- 446, TITLE: The Role of Deconfounding in Meta-learning
<https://proceedings.mlr.press/v162/jiang22a.html>
AUTHORS: Yinjie Jiang, Zhengyu Chen, Kun Kuang, Luotian Yuan, Xinhai Ye, Zhihua Wang, Fei Wu, Ying Wei
HIGHLIGHT: In this paper, we offer a novel causal perspective of meta-learning.
- 447, TITLE: Subspace Learning for Effective Meta-Learning
<https://proceedings.mlr.press/v162/jiang22b.html>
AUTHORS: Weisen Jiang, James Kwok, Yu Zhang
HIGHLIGHT: We propose an algorithm to learn the meta-parameters (i.e, subspace bases).
- 448, TITLE: Optimal Algorithms for Stochastic Multi-Level Compositional Optimization
<https://proceedings.mlr.press/v162/jiang22c.html>
AUTHORS: Wei Jiang, Bokun Wang, Yibo Wang, Lijun Zhang, Tianbao Yang
HIGHLIGHT: In this paper, we investigate the problem of stochastic multi-level compositional optimization, where the objective function is a composition of multiple smooth but possibly non-convex functions.
- 449, TITLE: Antibody-Antigen Docking and Design via Hierarchical Structure Refinement
<https://proceedings.mlr.press/v162/jin22a.html>
AUTHORS: Wengong Jin, Dr.Regina Barzilay, Tommi Jaakkola
HIGHLIGHT: In this paper, we propose a new model called Hierarchical Structure Refinement Network (HSRN) for paratope docking and design.
- 450, TITLE: Sharpened Quasi-Newton Methods: Faster Superlinear Rate and Larger Local Convergence Neighborhood
<https://proceedings.mlr.press/v162/jin22b.html>
AUTHORS: Quijiang Jin, Alec Koppel, Ketan Rajawat, Aryan Mokhtari
HIGHLIGHT: This is due to the fact that in Greedy-BFGS the Hessian is directly approximated and the Newton direction approximation may not be as accurate as the one for BFGS. In this paper, we close this gap and present a novel BFGS method that has the best of two worlds.
- 451, TITLE: The Power of Exploiter: Provable Multi-Agent RL in Large State Spaces
<https://proceedings.mlr.press/v162/jin22c.html>
AUTHORS: Chi Jin, Qinghua Liu, Tiancheng Yu
HIGHLIGHT: This paper considers two-player zero-sum Markov Games (MGs). We propose a new algorithm that can provably find the Nash equilibrium policy using a polynomial number of samples, for any MG with low multi-agent Bellman-Eluder dimension—a new complexity measure adapted from its single-agent version (Jin et al., 2021).

- 452, TITLE: Domain Adaptation for Time Series Forecasting via Attention Sharing
<https://proceedings.mlr.press/v162/jin22d.html>
AUTHORS: Xiaoyong Jin, Youngsuk Park, Danielle Maddix, Hao Wang, Yuyang Wang
HIGHLIGHT: This poses a challenge for typical forecasting problems in practice, where there is a limited number of time series or observations per time series, or both. To cope with this data scarcity issue, we propose a novel domain adaptation framework, Domain Adaptation Forecaster (DAF).
- 453, TITLE: Accelerated Federated Learning with Decoupled Adaptive Optimization
<https://proceedings.mlr.press/v162/jin22e.html>
AUTHORS: Jiayin Jin, Jiaxiang Ren, Yang Zhou, Lingjuan Lyu, Ji Liu, Dejing Dou
HIGHLIGHT: This work aims to develop novel adaptive optimization methods for FL from the perspective of dynamics of ordinary differential equations (ODEs).
- 454, TITLE: Supervised Off-Policy Ranking
<https://proceedings.mlr.press/v162/jin22f.html>
AUTHORS: Yue Jin, Yue Zhang, Tao Qin, Xudong Zhang, Jian Yuan, Houqiang Li, Tie-Yan Liu
HIGHLIGHT: We propose a method to solve SOPR, which learns a policy scoring model by minimizing a ranking loss of the training policies rather than estimating the precise policy performance.
- 455, TITLE: Input-agnostic Certified Group Fairness via Gaussian Parameter Smoothing
<https://proceedings.mlr.press/v162/jin22g.html>
AUTHORS: Jiayin Jin, Zeru Zhang, Yang Zhou, Lingfei Wu
HIGHLIGHT: This paper proposes an input-agnostic certified group fairness algorithm, FairSmooth, for improving the fairness of classification models while maintaining the remarkable prediction accuracy.
- 456, TITLE: Score-based Generative Modeling of Graphs via the System of Stochastic Differential Equations
<https://proceedings.mlr.press/v162/jo22a.html>
AUTHORS: Jaehyeong Jo, Seul Lee, Sung Ju Hwang
HIGHLIGHT: Yet, this is a challenging problem, and the previous graph generative methods either fail to capture the permutation-invariance property of graphs or cannot sufficiently model the complex dependency between nodes and edges, which is crucial for generating real-world graphs such as molecules. To overcome such limitations, we propose a novel score-based generative model for graphs with a continuous-time framework.
- 457, TITLE: Choosing Answers in Epsilon-Best-Answer Identification for Linear Bandits
<https://proceedings.mlr.press/v162/jourdan22a.html>
AUTHORS: Marc Jourdan, Romy Degenne
HIGHLIGHT: We demonstrate that picking the answer with highest mean does not allow an algorithm to reach asymptotic optimality in terms of expected sample complexity. Instead, a furthest answer should be identified.
- 458, TITLE: Robust Fine-Tuning of Deep Neural Networks with Hessian-based Generalization Guarantees
<https://proceedings.mlr.press/v162/ju22a.html>
AUTHORS: Haotian Ju, Dongyue Li, Hongyang R Zhang
HIGHLIGHT: We design an algorithm that incorporates consistent losses and distance-based regularization for fine-tuning.
- 459, TITLE: Robust alignment of cross-session recordings of neural population activity by behaviour via unsupervised domain adaptation
<https://proceedings.mlr.press/v162/jude22a.html>
AUTHORS: Justin Jude, Matthew Perich, Lee Miller, Matthias Hennig
HIGHLIGHT: On the other hand, evidence suggests that the latent dynamics underlying behaviour may be stable even over months and years. Based on this idea, we introduce a model capable of inferring behaviourally relevant latent dynamics from previously unseen data recorded from the same animal, without any need for decoder recalibration.
- 460, TITLE: On Measuring Causal Contributions via do-interventions
<https://proceedings.mlr.press/v162/jung22a.html>
AUTHORS: Yonghan Jung, Shiva Kasiviswanathan, Jin Tian, Dominik Janzing, Patrick Bloebaum, Elias Bareinboim
HIGHLIGHT: In this paper, we develop a principled method for quantifying causal contributions. First, we provide desiderata of properties axioms that causal contribution measures should satisfy and propose the do-Shapley values (inspired by do-interventions [Pearl, 2000]) as a unique method satisfying these properties.
- 461, TITLE: Efficient Approximate Inference for Stationary Kernel on Frequency Domain
<https://proceedings.mlr.press/v162/jung22b.html>
AUTHORS: Yohan Jung, Kyungwoo Song, Jinkyoo Park
HIGHLIGHT: However, despite its expressive power, training this kernel is typically difficult because scalability and overfitting issues often arise due to a large number of training parameters. To resolve these issues, we propose an approximate inference method for estimating the Spectral mixture kernel hyperparameters.
- 462, TITLE: Sketching Algorithms and Lower Bounds for Ridge Regression
<https://proceedings.mlr.press/v162/kacham22a.html>
AUTHORS: Praneeth Kacham, David Woodruff

HIGHLIGHT: We give a sketching-based iterative algorithm that computes a $1+\epsilon$ approximate solution for the ridge regression problem $\min_x \|Ax-b\|_2^2 + \lambda \|x\|_2^2$ where $A \in \mathbb{R}^{n \times d}$ with $d \geq n$.

463, **TITLE:** Flashlight: Enabling Innovation in Tools for Machine Learning

<https://proceedings.mlr.press/v162/kahn22a.html>

AUTHORS: Jacob D Kahn, Vineel Pratap, Tatiana Likhomanenko, Qiantong Xu, Awni Hannun, Jeff Cai, Paden Tomasello, Ann Lee, Edouard Grave, Gilad Avidov, Benoit Steiner, Vitaliy Liptchinsky, Gabriel Synnaeve, Ronan Collobert

HIGHLIGHT: We introduce Flashlight, an open-source library built to spur innovation in machine learning tools and systems by prioritizing open, modular, customizable internals and state-of-the-art, research-ready models and training setups across a variety of domains.

464, **TITLE:** Learning-based Optimisation of Particle Accelerators Under Partial Observability Without Real-World Training

<https://proceedings.mlr.press/v162/kaiser22a.html>

AUTHORS: Jan Kaiser, Oliver Stein, Annika Eichler

HIGHLIGHT: In this contribution, we demonstrate how to successfully apply RL to the optimisation of a highly complex real-world machine $\{-\}$ specifically a linear particle accelerator $\{-\}$ in an only partially observable setting and without requiring training on the real machine.

465, **TITLE:** Stochastic Deep Networks with Linear Competing Units for Model-Agnostic Meta-Learning

<https://proceedings.mlr.press/v162/kalais22a.html>

AUTHORS: Konstantinos Kalais, Sotirios Chatzis

HIGHLIGHT: Stochastic Deep Networks with Linear Competing Units for Model-Agnostic Meta-Learning

466, **TITLE:** Doubly Robust Distributionally Robust Off-Policy Evaluation and Learning

<https://proceedings.mlr.press/v162/kallus22a.html>

AUTHORS: Nathan Kallus, Xiaojie Mao, Kaiwen Wang, Zhengyuan Zhou

HIGHLIGHT: In this paper, we propose the first DR algorithms for DROPE/L with KL-divergence uncertainty sets.

467, **TITLE:** Improved Rates for Differentially Private Stochastic Convex Optimization with Heavy-Tailed Data

<https://proceedings.mlr.press/v162/kamath22a.html>

AUTHORS: Gautam Kamath, Xingtu Liu, Huanyu Zhang

HIGHLIGHT: We study stochastic convex optimization with heavy-tailed data under the constraint of differential privacy (DP).

468, **TITLE:** Comprehensive Analysis of Negative Sampling in Knowledge Graph Representation Learning

<https://proceedings.mlr.press/v162/kamigaito22a.html>

AUTHORS: Hidetaka Kamigaito, Katsuhiko Hayashi

HIGHLIGHT: Currently, empirical hyperparameter tuning addresses this problem at the cost of computational time. To solve this problem, we theoretically analyzed NS loss to assist hyperparameter tuning and understand the better use of the NS loss in KGE learning.

469, **TITLE:** Matching Learned Causal Effects of Neural Networks with Domain Priors

<https://proceedings.mlr.press/v162/kancheti22a.html>

AUTHORS: Sai Srinivas Kancheti, Abbavaram Gowtham Reddy, Vineeth N Balasubramanian, Amit Sharma

HIGHLIGHT: Therefore, we propose a regularization method that aligns the learned causal effects of a neural network with domain priors, including both direct and total causal effects.

470, **TITLE:** Deduplicating Training Data Mitigates Privacy Risks in Language Models

<https://proceedings.mlr.press/v162/kandpal22a.html>

AUTHORS: Nikhil Kandpal, Eric Wallace, Colin Raffel

HIGHLIGHT: Past work has shown that large language models are susceptible to privacy attacks, where adversaries generate sequences from a trained model and detect which sequences are memorized from the training set. In this work, we show that the success of these attacks is largely due to duplication in commonly used web-scraped training sets.

471, **TITLE:** Lyapunov Density Models: Constraining Distribution Shift in Learning-Based Control

<https://proceedings.mlr.press/v162/kang22a.html>

AUTHORS: Katie Kang, Paula Gradu, Jason J Choi, Michael Janner, Claire Tomlin, Sergey Levine

HIGHLIGHT: Can we combine these two concepts, producing learning-based control algorithms that constrain the system to in-distribution states using only in-distribution actions? In this paper, we propose to do this by combining concepts from Lyapunov stability and density estimation, introducing Lyapunov density models: a generalization of control Lyapunov functions and density models that provides guarantees about an agent's ability to stay in-distribution over its entire trajectory.

472, **TITLE:** Forget-free Continual Learning with Winning Subnetworks

<https://proceedings.mlr.press/v162/kang22b.html>

AUTHORS: Haeyong Kang, Rusty John Lloyd Mina, Sultan Rizky Hikmawan Madjid, Jaehong Yoon, Mark Hasegawa-Johnson, Sung Ju Hwang, Chang D. Yoo

HIGHLIGHT: Inspired by Lottery Ticket Hypothesis that competitive subnetworks exist within a dense network, we propose a continual learning method referred to as Winning SubNetworks (WSN), which sequentially learns and selects an optimal subnetwork for each task.

- 473, TITLE: Differentially Private Approximate Quantiles
<https://proceedings.mlr.press/v162/kaplan22a.html>
AUTHORS: Haim Kaplan, Shachar Schnapp, Uri Stemmer
HIGHLIGHT: In this work we study the problem of differentially private (DP) quantiles, in which given dataset X and quantiles $q_1, \dots, q_m \in [0,1]$, we want to output m quantile estimations which are as close as possible to the true quantiles and preserve DP.
- 474, TITLE: Simultaneous Graph Signal Clustering and Graph Learning
<https://proceedings.mlr.press/v162/karaaslanli22a.html>
AUTHORS: Abdullah Karaaslanli, Selin Aviyente
HIGHLIGHT: In this paper, we address the problem of learning multiple graphs from heterogeneous data by formulating an optimization problem for joint graph signal clustering and graph topology inference.
- 475, TITLE: Composing Partial Differential Equations with Physics-Aware Neural Networks
<https://proceedings.mlr.press/v162/karlbauer22a.html>
AUTHORS: Matthias Karlbauer, Timothy Praditia, Sebastian Otte, Sergey Oladyskhin, Wolfgang Nowak, Martin V. Butz
HIGHLIGHT: We introduce a compositional physics-aware Finite volume Neural Network (FINN) for learning spatiotemporal advection-diffusion processes.
- 476, TITLE: Meta-Learning Hypothesis Spaces for Sequential Decision-making
<https://proceedings.mlr.press/v162/kassraie22a.html>
AUTHORS: Parnian Kassraie, Jonas Rothfuss, Andreas Krause
HIGHLIGHT: In this work, we propose to meta-learn a kernel from offline data (Meta-KeL).
- 477, TITLE: FOCUS: Familiar Objects in Common and Uncommon Settings
<https://proceedings.mlr.press/v162/kattakinda22a.html>
AUTHORS: Priyatham Kattakinda, Soheil Feizi
HIGHLIGHT: In this paper, we introduce FOCUS (Familiar Objects in Common and Uncommon Settings), a dataset for stress-testing the generalization power of deep image classifiers.
- 478, TITLE: Training OOD Detectors in their Natural Habitats
<https://proceedings.mlr.press/v162/katz-samuels22a.html>
AUTHORS: Julian Katz-Samuels, Julia B Nakhleh, Robert Nowak, Yixuan Li
HIGHLIGHT: In this paper, we propose a novel framework that leverages wild mixture data—that naturally consists of both ID and OOD samples.
- 479, TITLE: Robustness Implies Generalization via Data-Dependent Generalization Bounds
<https://proceedings.mlr.press/v162/kawaguchi22a.html>
AUTHORS: Kenji Kawaguchi, Zhun Deng, Kyle Luh, Jiaoyang Huang
HIGHLIGHT: We present several examples, including ones for lasso and deep learning, in which our bounds are provably preferable.
- 480, TITLE: Generating Distributional Adversarial Examples to Evade Statistical Detectors
<https://proceedings.mlr.press/v162/kaya22a.html>
AUTHORS: Yigitcan Kaya, Muhammad Bilal Zafar, Sergul Aydore, Nathalie Rauschmayr, Krishnamurthy Kenthapadi
HIGHLIGHT: Due to the difficulties in designing adaptive attacks, however, recent work suggests that most detectors have incomplete evaluation. We aim to fill this gap by designing a generic adaptive attack against detectors: the 'statistical indistinguishability attack' (SIA).
- 481, TITLE: Secure Quantized Training for Deep Learning
<https://proceedings.mlr.press/v162/keller22a.html>
AUTHORS: Marcel Keller, Ke Sun
HIGHLIGHT: We implement training of neural networks in secure multi-party computation (MPC) using quantization commonly used in said setting.
- 482, TITLE: A Convergent and Dimension-Independent Min-Max Optimization Algorithm
<https://proceedings.mlr.press/v162/keswani22a.html>
AUTHORS: Vijay Keswani, Oren Mangoubi, Sushant Sachdeva, Nisheeth K. Vishnoi
HIGHLIGHT: We study a variant of a recently introduced min-max optimization framework where the max-player is constrained to update its parameters in a greedy manner until it reaches a first-order stationary point.
- 483, TITLE: Neural Network Poisson Models for Behavioural and Neural Spike Train Data
<https://proceedings.mlr.press/v162/khajehnejad22a.html>
AUTHORS: Moein Khajehnejad, Forough Habibollahi, Richard Nock, Ehsan Arabzadeh, Peter Dayan, Amir Dezfouli
HIGHLIGHT: Predominant modeling methods apply rather disjoint techniques to these scales; by contrast, we suggest an end-to-end model which exploits recent developments of flexible, but tractable, neural network point-process models to characterize dependencies between stimuli, actions, and neural data.
- 484, TITLE: Federated Reinforcement Learning: Linear Speedup Under Markovian Sampling
<https://proceedings.mlr.press/v162/khodadadian22a.html>

AUTHORS: Sajad Khodadadian, Pranay Sharma, Gauri Joshi, Siva Theja Maguluri
HIGHLIGHT: In this paper, we consider a federated reinforcement learning framework where multiple agents collaboratively learn a global model, without sharing their individual data and policies.

485, TITLE: Multi-Level Branched Regularization for Federated Learning
<https://proceedings.mlr.press/v162/kim22a.html>

AUTHORS: Jinkyu Kim, Geeho Kim, Bohyung Han
HIGHLIGHT: To alleviate the limitations, we propose a novel architectural regularization technique that constructs multiple auxiliary branches in each local model by grafting local and global subnetworks at several different levels and that learns the representations of the main pathway in the local model congruent to the auxiliary hybrid pathways via online knowledge distillation.

486, TITLE: Learning fair representation with a parametric integral probability metric
<https://proceedings.mlr.press/v162/kim22b.html>

AUTHORS: Dongha Kim, Kunwoong Kim, Insung Kong, Ilsang Ohn, Yongdai Kim
HIGHLIGHT: In this paper, we propose a new adversarial training scheme for LFR, where the integral probability metric (IPM) with a specific parametric family of discriminators is used.

487, TITLE: Dataset Condensation via Efficient Synthetic-Data Parameterization
<https://proceedings.mlr.press/v162/kim22c.html>

AUTHORS: Jang-Hyun Kim, Jinuk Kim, Seong Joon Oh, Sangdoon Yun, Hwanjun Song, Joonhyun Jeong, Jung-Woo Ha, Hyun Oh Song
HIGHLIGHT: To this end, we propose a novel condensation framework that generates multiple synthetic data with a limited storage budget via efficient parameterization considering data regularity.

488, TITLE: Guided-TTS: A Diffusion Model for Text-to-Speech via Classifier Guidance
<https://proceedings.mlr.press/v162/kim22d.html>

AUTHORS: Heeseung Kim, Sungwon Kim, Sungroh Yoon
HIGHLIGHT: We propose Guided-TTS, a high-quality text-to-speech (TTS) model that does not require any transcript of target speaker using classifier guidance.

489, TITLE: Variational On-the-Fly Personalization
<https://proceedings.mlr.press/v162/kim22e.html>

AUTHORS: Jangho Kim, Jun-Tae Lee, Simyung Chang, Nojun Kwak
HIGHLIGHT: In this paper, we propose a novel personalization method, Variational On-the-Fly Personalization.

490, TITLE: Fisher SAM: Information Geometry and Sharpness Aware Minimisation
<https://proceedings.mlr.press/v162/kim22f.html>

AUTHORS: Minyoung Kim, Da Li, Shell X Hu, Timothy Hospedales
HIGHLIGHT: In this paper we consider the information geometry of the model parameter space when defining the neighborhood, namely replacing SAM's Euclidean balls with ellipsoids induced by the Fisher information.

491, TITLE: ViT-NeT: Interpretable Vision Transformers with Neural Tree Decoder
<https://proceedings.mlr.press/v162/kim22g.html>

AUTHORS: Sangwon Kim, Jaeyeal Nam, Byoung Chul Ko
HIGHLIGHT: In this study, we propose a new ViT neural tree decoder (ViT-NeT).

492, TITLE: Sanity Simulations for Saliency Methods
<https://proceedings.mlr.press/v162/kim22h.html>

AUTHORS: Joon Sik Kim, Gregory Plumb, Ameet Talwalkar
HIGHLIGHT: In this work, we design a synthetic benchmarking framework, SMERF, that allows us to perform ground-truth-based evaluation while controlling the complexity of the model's reasoning.

493, TITLE: Soft Truncation: A Universal Training Technique of Score-based Diffusion Model for High Precision Score Estimation
<https://proceedings.mlr.press/v162/kim22i.html>

AUTHORS: Dongjun Kim, Seungjae Shin, Kyungwoo Song, Wanmo Kang, Il-Chul Moon
HIGHLIGHT: For successful training, therefore, we introduce Soft Truncation, a universally applicable training technique for diffusion models, that softens the fixed and static truncation hyperparameter into a random variable.

494, TITLE: Rotting Infinitely Many-Armed Bandits
<https://proceedings.mlr.press/v162/kim22j.html>

AUTHORS: Jung-Hun Kim, Milan Vojnovic, Se-Young Yun
HIGHLIGHT: We consider the infinitely many-armed bandit problem with rotting rewards, where the mean reward of an arm decreases at each pull of the arm according to an arbitrary trend with maximum rotting rate $\rho \in (0, 1)$.

495, TITLE: Accelerated Gradient Methods for Geodesically Convex Optimization: Tractable Algorithms and Convergence Analysis
<https://proceedings.mlr.press/v162/kim22k.html>

AUTHORS: Jungbin Kim, Insoon Yang

HIGHLIGHT: We propose computationally tractable accelerated first-order methods for Riemannian optimization, extending the Nesterov accelerated gradient (NAG) method.

496, **TITLE:** Generalizing to New Physical Systems via Context-Informed Dynamics Model

<https://proceedings.mlr.press/v162/kirchmeyer22a.html>

AUTHORS: Matthieu Kirchmeyer, Yuan Yin, Jeremie Dona, Nicolas Baskiotis, Alain Rakotomamonjy, Patrick Gallinari

HIGHLIGHT: Data-driven approaches to modeling physical systems fail to generalize to unseen systems that share the same general dynamics with the learning domain, but correspond to different physical contexts. We propose a new framework for this key problem, context-informed dynamics adaptation (CoDA), which takes into account the distributional shift across systems for fast and efficient adaptation to new dynamics.

497, **TITLE:** SoQal: Selective Oracle Questioning for Consistency Based Active Learning of Cardiac Signals

<https://proceedings.mlr.press/v162/kiyasseh22a.html>

AUTHORS: Dani Kiyasseh, Tingting Zhu, David A Clifton

HIGHLIGHT: One way to mitigate this burden is via active learning (AL) which involves the (a) acquisition and (b) annotation of informative unlabelled instances. Whereas previous work addresses either one of these elements independently, we propose an AL framework that addresses both.

498, **TITLE:** Curriculum Reinforcement Learning via Constrained Optimal Transport

<https://proceedings.mlr.press/v162/klink22a.html>

AUTHORS: Pascal Klink, Haoyi Yang, Carlo D'Eramo, Jan Peters, Joni Pajarinen

HIGHLIGHT: In this work, we focus on the idea of framing curricula as interpolations between task distributions, which has previously been shown to be a viable approach to CRL.

499, **TITLE:** Exploiting Redundancy: Separable Group Convolutional Networks on Lie Groups

<https://proceedings.mlr.press/v162/knigge22a.html>

AUTHORS: David M. Knigge, David W Romero, Erik J Bekkers

HIGHLIGHT: In this work, we investigate the properties of representations learned by regular G-CNNs, and show considerable parameter redundancy in group convolution kernels.

500, **TITLE:** Revisiting Contrastive Learning through the Lens of Neighborhood Component Analysis: an Integrated Framework

<https://proceedings.mlr.press/v162/ko22a.html>

AUTHORS: Ching-Yun Ko, Jeet Mohapatra, Sijia Liu, Pin-Yu Chen, Luca Daniel, Lily Weng

HIGHLIGHT: By investigating the connection between contrastive learning and neighborhood component analysis (NCA), we provide a novel stochastic nearest neighbor viewpoint of contrastive learning and subsequently propose a series of contrastive losses that outperform the existing ones.

501, **TITLE:** Transfer Learning In Differential Privacy's Hybrid-Model

<https://proceedings.mlr.press/v162/kohen22a.html>

AUTHORS: Refael Kohen, Or Sheffet

HIGHLIGHT: Here we study the problem of machine learning in the hybrid-model where the n individuals in the curator's dataset are drawn from a different distribution than the one of the general population (the local-agents).

502, **TITLE:** Markov Chain Monte Carlo for Continuous-Time Switching Dynamical Systems

<https://proceedings.mlr.press/v162/kohs22a.html>

AUTHORS: Lukas Kohs, Bastian Alt, Heinz Koepl

HIGHLIGHT: In this work, we propose a novel inference algorithm utilizing a Markov Chain Monte Carlo approach.

503, **TITLE:** Partial disentanglement for domain adaptation

<https://proceedings.mlr.press/v162/kong22a.html>

AUTHORS: Lingjing Kong, Shaoran Xie, Weiran Yao, Yujia Zheng, Guangyi Chen, Petar Stojanov, Victor Akinwande, Kun Zhang

HIGHLIGHT: Given the theoretical insights, we propose a practical domain adaptation framework, called iMSDA.

504, **TITLE:** Simultaneously Learning Stochastic and Adversarial Bandits with General Graph Feedback

<https://proceedings.mlr.press/v162/kong22b.html>

AUTHORS: Fang Kong, Yichi Zhou, Shuai Li

HIGHLIGHT: With a general feedback graph, the observation of an arm may not be available when this arm is pulled, which makes the exploration more expensive and the algorithms more challenging to perform optimally in both environments. In this work, we overcome this difficulty by a new trade-off mechanism with a carefully-designed proportion for exploration and exploitation.

505, **TITLE:** Adaptive Data Analysis with Correlated Observations

<https://proceedings.mlr.press/v162/kontorovich22a.html>

AUTHORS: Aryeh Kontorovich, Menachem Sadigurschi, Uri Stemmer

HIGHLIGHT: We embark on a systematic study of the possibilities of adaptive data analysis with correlated observations.

506, **TITLE:** Controlling Conditional Language Models without Catastrophic Forgetting

<https://proceedings.mlr.press/v162/korbak22a.html>

AUTHORS: Tomasz Korbak, Hady Elsahar, German Kruszewski, Marc Dymetman

- HIGHLIGHT:** In this paper, we extend DPG to conditional tasks by proposing Conditional DPG (CDPG).
- 507, **TITLE:** Batch Greenhorn Algorithm for Entropic-Regularized Multimarginal Optimal Transport: Linear Rate of Convergence and Iteration Complexity
<https://proceedings.mlr.press/v162/kostic22a.html>
AUTHORS: Vladimir R. Kostic, Saverio Salzo, Massimiliano Pontil
HIGHLIGHT: In this work we propose a batch multimarginal version of the Greenhorn algorithm for the entropic-regularized optimal transport problem.
- 508, **TITLE:** Certified Adversarial Robustness Under the Bounded Support Set
<https://proceedings.mlr.press/v162/kou22a.html>
AUTHORS: Yiwen Kou, Qinyuan Zheng, Yisen Wang
HIGHLIGHT: In this paper, we generalize the f -divergence-based framework to a Wasserstein-distance-based and total-variation-distance-based framework that is first able to analyze robustness properties of bounded support set smoothing measures both theoretically and experimentally.
- 509, **TITLE:** Exact Learning of Preference Structure: Single-peaked Preferences and Beyond
<https://proceedings.mlr.press/v162/kraiczy22a.html>
AUTHORS: Sonja Kraiczy, Edith Elkind
HIGHLIGHT: We consider the setting where the members of a society (voters) have preferences over candidates, and the candidates can be ordered on an axis so that the voters' preferences are single-peaked on this axis.
- 510, **TITLE:** Reconstructing Nonlinear Dynamical Systems from Multi-Modal Time Series
<https://proceedings.mlr.press/v162/kramer22a.html>
AUTHORS: Daniel Kramer, Philine L Bommer, Daniel Durstewitz, Carlo Tombolini, Georgia Koppe
HIGHLIGHT: Here we propose a general framework for multi-modal data integration for the purpose of nonlinear DS reconstruction and the analysis of cross-modal relations.
- 511, **TITLE:** Probabilistic ODE Solutions in Millions of Dimensions
<https://proceedings.mlr.press/v162/kramer22b.html>
AUTHORS: Nicholas Kr?mer, Nathanael Bosch, Jonathan Schmidt, Philipp Hennig
HIGHLIGHT: In this work, we explain the mathematical assumptions and detailed implementation schemes behind solving high-dimensional ODEs with a probabilistic numerical algorithm.
- 512, **TITLE:** Active Nearest Neighbor Regression Through Delaunay Refinement
<https://proceedings.mlr.press/v162/kravberg22a.html>
AUTHORS: Alexander Kravberg, Giovanni Luca Marchetti, Vladislav Polianskii, Anastasiia Varava, Florian T. Pokorny, Danica Kragic
HIGHLIGHT: We introduce an algorithm for active function approximation based on nearest neighbor regression.
- 513, **TITLE:** Functional Generalized Empirical Likelihood Estimation for Conditional Moment Restrictions
<https://proceedings.mlr.press/v162/kremer22a.html>
AUTHORS: Heiner Kremer, Jia-Jie Zhu, Krikamol Muandet, Bernhard Sch?lkopf
HIGHLIGHT: To benefit from recent developments in machine learning, we provide a functional reformulation of GEL in which arbitrary models can be leveraged. Motivated by a dual formulation of the resulting infinite dimensional optimization problem, we devise a practical method and explore its asymptotic properties.
- 514, **TITLE:** Calibrated and Sharp Uncertainties in Deep Learning via Density Estimation
<https://proceedings.mlr.press/v162/kuleshov22a.html>
AUTHORS: Volodymyr Kuleshov, Shachi Deshpande
HIGHLIGHT: We introduce a simple training procedure based on recalibration that yields calibrated models without sacrificing overall performance; unlike previous approaches, ours ensures the most general property of distribution calibration and applies to any model, including neural networks.
- 515, **TITLE:** ActiveHedge: Hedge meets Active Learning
<https://proceedings.mlr.press/v162/kumar22a.html>
AUTHORS: Bhuvish Kumar, Jacob D Abernethy, Venkatesh Saligrama
HIGHLIGHT: We consider the classical problem of multi-class prediction with expert advice, but with an active learning twist.
- 516, **TITLE:** Balancing Discriminability and Transferability for Source-Free Domain Adaptation
<https://proceedings.mlr.press/v162/kundu22a.html>
AUTHORS: Jogendra Nath Kundu, Akshay R Kulkarni, Suvaansh Bhambri, Deepesh Mehta, Shreyas Anand Kulkarni, Varun Jampani, Venkatesh Babu Radhakrishnan
HIGHLIGHT: Upon analyzing the hurdles from both theoretical and empirical standpoints, we derive novel insights to show that a mixup between original and corresponding translated generic samples enhances the discriminability-transferability trade-off while duly respecting the privacy-oriented source-free setting.
- 517, **TITLE:** Showing Your Offline Reinforcement Learning Work: Online Evaluation Budget Matters
<https://proceedings.mlr.press/v162/kurenkov22a.html>
AUTHORS: Vladislav Kurenkov, Sergey Kolesnikov

- HIGHLIGHT:** In this work, we argue for the importance of an online evaluation budget for a reliable comparison of deep offline RL algorithms.
- 518, **TITLE:** Equivariant Priors for compressed sensing with unknown orientation
<https://proceedings.mlr.press/v162/kuzina22a.html>
AUTHORS: Anna Kuzina, Kumar Pratik, Fabio Valerio Massoli, Arash Behboodi
HIGHLIGHT: Additionally, in many scenarios, the signal has an unknown orientation prior to measurements. To address such recovery problems, we propose using equivariant generative models as a prior, which encapsulate orientation information in their latent space.
- 519, **TITLE:** Coordinated Attacks against Contextual Bandits: Fundamental Limits and Defense Mechanisms
<https://proceedings.mlr.press/v162/kwon22a.html>
AUTHORS: Jeongyeol Kwon, Yonathan Efroni, Constantine Caramanis, Shie Mannor
HIGHLIGHT: Motivated by online recommendation systems, we propose the problem of finding the optimal policy in multitask contextual bandits when a small fraction α of tasks (users) are arbitrary and adversarial.
- 520, **TITLE:** Large Batch Experience Replay
<https://proceedings.mlr.press/v162/lahire22a.html>
AUTHORS: Thibault Lahire, Matthieu Geist, Emmanuel Rachelson
HIGHLIGHT: In this work, we cast the replay buffer sampling problem as an importance sampling one for estimating the gradient.
- 521, **TITLE:** FedScale: Benchmarking Model and System Performance of Federated Learning at Scale
<https://proceedings.mlr.press/v162/lai22a.html>
AUTHORS: Fan Lai, Yinwei Dai, Sanjay Singapuram, Jiachen Liu, Xiangfeng Zhu, Harsha Madhyastha, Mosharaf Chowdhury
HIGHLIGHT: We present FedScale, a federated learning (FL) benchmarking suite with realistic datasets and a scalable runtime to enable reproducible FL research.
- 522, **TITLE:** Smoothed Adaptive Weighting for Imbalanced Semi-Supervised Learning: Improve Reliability Against Unknown Distribution Data
<https://proceedings.mlr.press/v162/lai22b.html>
AUTHORS: Zhengfeng Lai, Chao Wang, Henry Gunawan, Sen-Ching S Cheung, Chen-Nee Chuah
HIGHLIGHT: Based on this study, we propose a self-adaptive algorithm, named Smoothed Adaptive Weighting (SAW).
- 523, **TITLE:** Functional Output Regression with Infimal Convolution: Exploring the Huber and ℓ_1 -insensitive Losses
<https://proceedings.mlr.press/v162/lambert22a.html>
AUTHORS: Alex Lambert, Dimitri Bouche, Zoltan Szabo, Florence D'Alch-Buc
HIGHLIGHT: We derive computationally tractable algorithms relying on duality to tackle the resulting tasks in the context of vector-valued reproducing kernel Hilbert spaces.
- 524, **TITLE:** Tell me why! Explanations support learning relational and causal structure
<https://proceedings.mlr.press/v162/lampinen22a.html>
AUTHORS: Andrew K Lampinen, Nicholas Roy, Ishita Dasgupta, Stephanie Cy Chan, Allison Tam, James Mcclelland, Chen Yan, Adam Santoro, Neil C Rabinowitz, Jane Wang, Felix Hill
HIGHLIGHT: Here, we show that language can play a similar role for deep RL agents in complex environments.
- 525, **TITLE:** Generative Cooperative Networks for Natural Language Generation
<https://proceedings.mlr.press/v162/lamprier22a.html>
AUTHORS: Sylvain Lamprier, Thomas Scialom, Antoine Chaffin, Vincent Claveau, Ewa Kijak, Jacopo Staiano, Benjamin Piwowarski
HIGHLIGHT: In this paper, we introduce Generative Cooperative Networks, in which the discriminator architecture is cooperatively used along with the generation policy to output samples of realistic texts for the task at hand.
- 526, **TITLE:** DSTAGNN: Dynamic Spatial-Temporal Aware Graph Neural Network for Traffic Flow Forecasting
<https://proceedings.mlr.press/v162/lan22a.html>
AUTHORS: Shiyong Lan, Yitong Ma, Weikang Huang, Wenwu Wang, Hongyu Yang, Pyang Li
HIGHLIGHT: This paper proposes a novel Dynamic Spatial-Temporal Aware Graph Neural Network (DSTAGNN) to model the complex spatial-temporal interaction in road network.
- 527, **TITLE:** Cooperative Online Learning in Stochastic and Adversarial MDPs
<https://proceedings.mlr.press/v162/lancewicki22a.html>
AUTHORS: Tal Lancewicki, Aviv Rosenberg, Yishay Mansour
HIGHLIGHT: We study cooperative online learning in stochastic and adversarial Markov decision process (MDP).
- 528, **TITLE:** PINs: Progressive Implicit Networks for Multi-Scale Neural Representations
<https://proceedings.mlr.press/v162/landgraf22a.html>
AUTHORS: Zoe Landgraf, Alexander Sorkine Hornung, Ricardo S Cabral

HIGHLIGHT: However, scenes with a wide frequency spectrum remain a challenge: choosing high frequencies for positional encoding introduces noise in low structure areas, while low frequencies results in poor fitting of detailed regions. To address this, we propose a progressive positional encoding, exposing a hierarchical MLP structure to incremental sets of frequency encodings.

529, **TITLE:** Co-training Improves Prompt-based Learning for Large Language Models

<https://proceedings.mlr.press/v162/lang22a.html>

AUTHORS: Hunter Lang, Monica N Agrawal, Yoon Kim, David Sontag

HIGHLIGHT: We demonstrate that co-training (Blum & Mitchell, 1998) can improve the performance of prompt-based learning by using unlabeled data.

530, **TITLE:** Goal Misgeneralization in Deep Reinforcement Learning

<https://proceedings.mlr.press/v162/langosco22a.html>

AUTHORS: Lauro Langosco Di Langosco, Jack Koch, Lee D Sharkey, Jacob Pfau, David Krueger

HIGHLIGHT: We study goal misgeneralization, a type of out-of-distribution robustness failure in reinforcement learning (RL).

531, **TITLE:** Marginal Tail-Adaptive Normalizing Flows

<https://proceedings.mlr.press/v162/laszkiwicz22a.html>

AUTHORS: Mike Laszkiewicz, Johannes Lederer, Asja Fischer

HIGHLIGHT: In this paper, we focus on improving the ability of normalizing flows to correctly capture the tail behavior and, thus, form more accurate models.

532, **TITLE:** Bregman Proximal Langevin Monte Carlo via Bregman-Moreau Envelopes

<https://proceedings.mlr.press/v162/lau22a.html>

AUTHORS: Tim Tsz-Kit Lau, Han Liu

HIGHLIGHT: We propose efficient Langevin Monte Carlo algorithms for sampling distributions with nonsmooth convex composite potentials, which is the sum of a continuously differentiable function and a possibly nonsmooth function.

533, **TITLE:** Scalable Deep Reinforcement Learning Algorithms for Mean Field Games

<https://proceedings.mlr.press/v162/lauriere22a.html>

AUTHORS: Mathieu Lauriere, Sarah Perrin, Sertan Girgin, Paul Muller, Ayush Jain, Theophile Cabannes, Georgios Piliouras, Julien Perolat, Romuald Elie, Olivier Pietquin, Matthieu Geist

HIGHLIGHT: This is far from being trivial in the case of non-linear function approximation that enjoy good generalization properties, e.g. neural networks. We propose two methods to address this shortcoming.

534, **TITLE:** Implicit Bias of Linear Equivariant Networks

<https://proceedings.mlr.press/v162/lawrence22a.html>

AUTHORS: Hannah Lawrence, Bobak Kiani, Kristian G Georgiev, Andrew K Dienes

HIGHLIGHT: In this context, we show that L-layer full-width linear G-CNNs trained via gradient descent for binary classification converge to solutions with low-rank Fourier matrix coefficients, regularized by the $2/L$ -Schatten matrix norm.

535, **TITLE:** Differentially Private Maximal Information Coefficients

<https://proceedings.mlr.press/v162/lazarsfeld22a.html>

AUTHORS: John Lazarsfeld, Aaron Johnson, Emmanuel Adeniran

HIGHLIGHT: As a solution, we present algorithms to approximate MIC in a way that provides differential privacy.

536, **TITLE:** Entropic Gromov-Wasserstein between Gaussian Distributions

<https://proceedings.mlr.press/v162/le22a.html>

AUTHORS: Khang Le, Dung Q Le, Huy Nguyen, Dat Do, Tung Pham, Nhat Ho

HIGHLIGHT: We study the entropic Gromov-Wasserstein and its unbalanced version between (unbalanced) Gaussian distributions with different dimensions.

537, **TITLE:** Neurocoder: General-Purpose Computation Using Stored Neural Programs

<https://proceedings.mlr.press/v162/le22b.html>

AUTHORS: Hung Le, Svetha Venkatesh

HIGHLIGHT: Here we design Neurocoder, a new class of general-purpose neural networks in which the neural network “codes” itself in a data-responsive way by composing relevant programs from a set of shareable, modular programs stored in external memory.

538, **TITLE:** Convergence of Policy Gradient for Entropy Regularized MDPs with Neural Network Approximation in the Mean-Field Regime

<https://proceedings.mlr.press/v162/leahy22a.html>

AUTHORS: James-Michael Leahy, Bekzhan Kerimkulov, David Siska, Lukasz Szpruch

HIGHLIGHT: We study the global convergence of policy gradient for infinite-horizon, continuous state and action space, and entropy-regularized Markov decision processes (MDPs).

539, **TITLE:** A Random Matrix Analysis of Data Stream Clustering: Coping With Limited Memory Resources

<https://proceedings.mlr.press/v162/lebeau22a.html>

AUTHORS: Hugo Lebeau, Romain Couillet, Florent Chatelain

HIGHLIGHT: This article introduces a random matrix framework for the analysis of clustering on high-dimensional data streams, a particularly relevant setting for a more sober processing of large amounts of data with limited memory and energy resources.

540, **TITLE:** Neural Tangent Kernel Analysis of Deep Narrow Neural Networks
<https://proceedings.mlr.press/v162/lee22a.html>

AUTHORS: Jongmin Lee, Joo Young Choi, Ernest K Ryu, Albert No

HIGHLIGHT: In this work, we present the first trainability guarantee of infinitely deep but narrow neural networks.

541, **TITLE:** Dataset Condensation with Contrastive Signals
<https://proceedings.mlr.press/v162/lee22b.html>

AUTHORS: Saehyung Lee, Sanghyuk Chun, Sangwon Jung, Sangdoon Yun, Sungroh Yoon

HIGHLIGHT: We attribute this to the lack of participation of the contrastive signals between the classes resulting from the class-wise gradient matching strategy. To address this problem, we propose Dataset Condensation with Contrastive signals (DCC) by modifying the loss function to enable the DC methods to effectively capture the differences between classes.

542, **TITLE:** Confidence Score for Source-Free Unsupervised Domain Adaptation

<https://proceedings.mlr.press/v162/lee22c.html>

AUTHORS: Jonghyun Lee, Dahuin Jung, Junho Yim, Sungroh Yoon

HIGHLIGHT: To differentiate between sample importance, in this study, we propose a novel sample-wise confidence score, the Joint Model-Data Structure (JMDS) score for SFUDA.

543, **TITLE:** A Statistical Manifold Framework for Point Cloud Data

<https://proceedings.mlr.press/v162/lee22d.html>

AUTHORS: Yonghyeon Lee, Seungyeon Kim, Jinwon Choi, Frank Park

HIGHLIGHT: A growing number of applications require a means of measuring not only distances between point clouds, but also angles, volumes, derivatives, and other more advanced concepts. To formulate and quantify these concepts in a coordinate-invariant way, we develop a Riemannian geometric framework for point cloud data.

544, **TITLE:** Low-Complexity Deep Convolutional Neural Networks on Fully Homomorphic Encryption Using Multiplexed Parallel Convolutions

<https://proceedings.mlr.press/v162/lee22e.html>

AUTHORS: Eunsang Lee, Joon-Woo Lee, Junghyun Lee, Young-Sik Kim, Yongjune Kim, Jong-Seon No, Woosuk Choi

HIGHLIGHT: To improve the performance, we first minimize total bootstrapping runtime using multiplexed parallel convolution that collects sparse output data for multiple channels compactly. We also propose the imaginary-removing bootstrapping to prevent the deep neural networks from catastrophic divergence during approximate ReLU operations.

545, **TITLE:** Statistical inference with implicit SGD: proximal Robbins-Monro vs. Polyak-Ruppert

<https://proceedings.mlr.press/v162/lee22f.html>

AUTHORS: Yoonhyung Lee, Sungdong Lee, Joong-Ho Won

HIGHLIGHT: In this paper, we conduct an in-depth analysis of the two modes of ISGD for smooth convex functions, namely proximal Robbins-Monro (proxRM) and proximal Polyak-Ruppert (proxPR) procedures, for their use in statistical inference on model parameters.

546, **TITLE:** Maslow's Hammer in Catastrophic Forgetting: Node Re-Use vs. Node Activation

<https://proceedings.mlr.press/v162/lee22g.html>

AUTHORS: Sebastian Lee, Stefano Sarao Mannelli, Claudia Clopath, Sebastian Goldt, Andrew Saxe

HIGHLIGHT: Surprisingly, the amount of forgetting does not increase with the dissimilarity between the learned tasks, but appears to be worst in an intermediate similarity regime. In this paper we theoretically analyse both a synthetic teacher-student framework and a real data setup to provide an explanation of this phenomenon that we name Maslow's Hammer hypothesis.

547, **TITLE:** Query-Efficient and Scalable Black-Box Adversarial Attacks on Discrete Sequential Data via Bayesian Optimization

<https://proceedings.mlr.press/v162/lee22h.html>

AUTHORS: Deokjae Lee, Seungyong Moon, Junhyeok Lee, Hyun Oh Song

HIGHLIGHT: We introduce block decomposition and history subsampling techniques to improve the scalability of Bayesian optimization when an input sequence becomes long.

548, **TITLE:** Least Squares Estimation using Sketched Data with Heteroskedastic Errors

<https://proceedings.mlr.press/v162/lee22i.html>

AUTHORS: Sokbae Lee, Serena Ng

HIGHLIGHT: This paper considers the case when the regression errors do not have constant variance and heteroskedasticity robust standard errors would normally be needed for test statistics to provide accurate inference.

549, **TITLE:** Why the Rich Get Richer? On the Balancedness of Random Partition Models

<https://proceedings.mlr.press/v162/lee22j.html>

AUTHORS: Changwoo J Lee, Huiyan Sang

HIGHLIGHT: We propose a principled way to compare the balancedness of random partition models, which gives a better understanding of what model works better and what doesn't for different applications.

- 550, TITLE: Model Selection in Batch Policy Optimization
<https://proceedings.mlr.press/v162/lee22k.html>
AUTHORS: Jonathan Lee, George Tucker, Ofir Nachum, Bo Dai
HIGHLIGHT: We study the problem of model selection in batch policy optimization: given a fixed, partial-feedback dataset and M model classes, learn a policy with performance that is competitive with the policy derived from the best model class.
- 551, TITLE: Supervised Learning with General Risk Functionals
<https://proceedings.mlr.press/v162/leqi22a.html>
AUTHORS: Liu Leqi, Audrey Huang, Zachary Lipton, Kamyar Azizzadenesheli
HIGHLIGHT: We establish the first uniform convergence results for estimating the CDF of the loss distribution, which yield uniform convergence guarantees that hold simultaneously both over a class of Hölder risk functionals and over a hypothesis class.
- 552, TITLE: Generalized Strategic Classification and the Case of Aligned Incentives
<https://proceedings.mlr.press/v162/levanon22a.html>
AUTHORS: Sagi Levanon, Nir Rosenfeld
HIGHLIGHT: In this work we argue for a broader perspective on what accounts for strategic user behavior, and propose and study a flexible model of generalized strategic classification.
- 553, TITLE: A Simple Unified Framework for High Dimensional Bandit Problems
<https://proceedings.mlr.press/v162/li22a.html>
AUTHORS: Wenjie Li, Adarsh Barik, Jean Honorio
HIGHLIGHT: Stochastic high dimensional bandit problems with low dimensional structures are useful in different applications such as online advertising and drug discovery. In this work, we propose a simple unified algorithm for such problems and present a general analysis framework for the regret upper bound of our algorithm.
- 554, TITLE: Robust Training of Neural Networks Using Scale Invariant Architectures
<https://proceedings.mlr.press/v162/li22b.html>
AUTHORS: Zhiyuan Li, Srinadh Bhojanapalli, Manzil Zaheer, Sashank Reddi, Sanjiv Kumar
HIGHLIGHT: However, the use of adaptivity not only comes at the cost of extra memory but also raises the fundamental question: can non-adaptive methods like SGD enjoy similar benefits? In this paper, we provide an affirmative answer to this question by proposing to achieve both robust and memory-efficient training via the following general recipe: (1) modify the architecture and make it scale invariant, (2) train with SGD and weight decay, and optionally (3) clip the global gradient norm proportional to weight norm multiplied by $\sqrt{\frac{2\lambda}{\eta}}$, where η is learning rate and λ is weight decay.
- 555, TITLE: Spatial-Channel Token Distillation for Vision MLPs
<https://proceedings.mlr.press/v162/li22c.html>
AUTHORS: Yanxi Li, Xinghao Chen, Minjing Dong, Yehui Tang, Yunhe Wang, Chang Xu
HIGHLIGHT: This work solves the problem from a novel knowledge distillation perspective. We propose a novel Spatial-channel Token Distillation (STD) method, which improves the information mixing in the two dimensions by introducing distillation tokens to each of them.
- 556, TITLE: An Analytical Update Rule for General Policy Optimization
<https://proceedings.mlr.press/v162/li22d.html>
AUTHORS: Hepeng Li, Nicholas Clavette, Haibo He
HIGHLIGHT: We present an analytical policy update rule that is independent of parametric function approximators.
- 557, TITLE: On Convergence of Gradient Descent Ascent: A Tight Local Analysis
<https://proceedings.mlr.press/v162/li22e.html>
AUTHORS: Haochuan Li, Farzan Farnia, Subhro Das, Ali Jadbabaie
HIGHLIGHT: While this stepsize ratio suggests a slow training of the min player, practical GAN algorithms typically adopt similar stepsizes for both variables, indicating a wide gap between theoretical and empirical results. In this paper, we aim to bridge this gap by analyzing the local convergence of general nonconvex-nonconcave minimax problems.
- 558, TITLE: On the Finite-Time Performance of the Knowledge Gradient Algorithm
<https://proceedings.mlr.press/v162/li22f.html>
AUTHORS: Yanwen Li, Siyang Gao
HIGHLIGHT: In this research, we present new theoretical results about the finite-time performance of the KG algorithm.
- 559, TITLE: Phasic Self-Imitative Reduction for Sparse-Reward Goal-Conditioned Reinforcement Learning
<https://proceedings.mlr.press/v162/li22g.html>
AUTHORS: Yunfei Li, Tian Gao, Jiaqi Yang, Huazhe Xu, Yi Wu
HIGHLIGHT: We propose a novel phasic solution by alternating online RL and offline SL for tackling sparse-reward goal-conditioned problems. In the online phase, we perform RL training and collect rollout data while in the offline phase, we perform SL on those successful trajectories from the dataset.
- 560, TITLE: G^2 CN: Graph Gaussian Convolution Networks with Concentrated Graph Filters
<https://proceedings.mlr.press/v162/li22h.html>
AUTHORS: Mingjie Li, Xiaojun Guo, Yifei Wang, Yisen Wang, Zhouchen Lin

HIGHLIGHT: Nevertheless, we notice that existing spectral analysis fails to explain why existing graph propagations with the same global tendency, such as low-pass or high-pass, still yield very different results. Motivated by this situation, we develop a new framework for spectral analysis in this paper called concentration analysis.

561, **TITLE:** Decomposing Temporal High-Order Interactions via Latent ODEs

<https://proceedings.mlr.press/v162/li22i.html>

AUTHORS: Shibo Li, Robert Kirby, Shandian Zhe

HIGHLIGHT: As a result, these methods might not be capable enough of capturing complex, fine-grained temporal dynamics or making accurate predictions for long-term interaction results. To overcome these limitations, we propose a novel Temporal High-order Interaction decomposition model based on Ordinary Differential Equations (THIS-ODE).

562, **TITLE:** Neural Inverse Transform Sampler

<https://proceedings.mlr.press/v162/li22j.html>

AUTHORS: Henry Li, Yuval Kluger

HIGHLIGHT: In this paper, we show that when modeling one-dimensional conditional densities with a neural network, \mathcal{Z} can be exactly and efficiently computed by letting the network represent the cumulative distribution function of a target density, and applying a generalized fundamental theorem of calculus.

563, **TITLE:** PLATINUM: Semi-Supervised Model Agnostic Meta-Learning using Submodular Mutual Information

<https://proceedings.mlr.press/v162/li22k.html>

AUTHORS: Changbin Li, Suraj Kothawade, Feng Chen, Rishabh Iyer

HIGHLIGHT: In this work, we propose PLATINUM (semi-supervised model agnostic meta learning using submodular mutual information), a novel semi-supervised model agnostic meta learning framework that uses the submodular mutual information (SMI) functions to boost the performance of FSC.

564, **TITLE:** Deconfounded Value Decomposition for Multi-Agent Reinforcement Learning

<https://proceedings.mlr.press/v162/li22l.html>

AUTHORS: Jiahui Li, Kun Kuang, Baoxiang Wang, Furui Liu, Long Chen, Changjie Fan, Fei Wu, Jun Xiao

HIGHLIGHT: In this paper, we investigate VD from a novel perspective of causal inference.

565, **TITLE:** C-MinHash: Improving Minwise Hashing with Circulant Permutation

<https://proceedings.mlr.press/v162/li22m.html>

AUTHORS: Xiaoyun Li, Ping Li

HIGHLIGHT: In this paper, we propose Circulant MinHash (C-MinHash) and provide the surprising theoretical results that using only two independent random permutations in a circulant manner leads to uniformly smaller Jaccard estimation variance than that of the classical MinHash with K independent permutations.

566, **TITLE:** BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation

<https://proceedings.mlr.press/v162/li22n.html>

AUTHORS: Junnan Li, Dongxu Li, Caiming Xiong, Steven Hoi

HIGHLIGHT: In this paper, we propose BLIP, a new VLP framework which transfers flexibly to both vision-language understanding and generation tasks.

567, **TITLE:** Restarted Nonconvex Accelerated Gradient Descent: No More Polylogarithmic Factor in the $\mathcal{SO}(e^{-7/4})$ Complexity

<https://proceedings.mlr.press/v162/li22o.html>

AUTHORS: Huan Li, Zhouchen Lin

HIGHLIGHT: This paper studies the accelerated gradient descent for general nonconvex problems under the gradient Lipschitz and Hessian Lipschitz assumptions.

568, **TITLE:** Achieving Fairness at No Utility Cost via Data Reweighting with Influence

<https://proceedings.mlr.press/v162/li22p.html>

AUTHORS: Peizhao Li, Hongfu Liu

HIGHLIGHT: In this paper, we focus on the pre-processing aspect for achieving fairness, and propose a data reweighting approach that only adjusts the weight for samples in the training phase.

569, **TITLE:** High Probability Guarantees for Nonconvex Stochastic Gradient Descent with Heavy Tails

<https://proceedings.mlr.press/v162/li22q.html>

AUTHORS: Shaojie Li, Yong Liu

HIGHLIGHT: In this paper, we develop high probability bounds for nonconvex SGD with a joint perspective of optimization and generalization performance.

570, **TITLE:** MetAug: Contrastive Learning via Meta Feature Augmentation

<https://proceedings.mlr.press/v162/li22r.html>

AUTHORS: Jiangmeng Li, Wenwen Qiang, Changwen Zheng, Bing Su, Hui Xiong

HIGHLIGHT: In response, we propose to directly augment the features in latent space, thereby learning discriminative representations without a large amount of input data.

571, **TITLE:** PMIC: Improving Multi-Agent Reinforcement Learning with Progressive Mutual Information Collaboration

<https://proceedings.mlr.press/v162/li22s.html>

- AUTHORS: Pengyi Li, Hongyao Tang, Tianpei Yang, Xiaotian Hao, Tong Sang, Yan Zheng, Jianye Hao, Matthew E. Taylor, Wenyuan Tao, Zhen Wang
HIGHLIGHT: However, we reveal sub-optimal collaborative behaviors also emerge with strong correlations, and simply maximizing the MI can, surprisingly, hinder the learning towards better collaboration. To address this issue, we propose a novel MARL framework, called Progressive Mutual Information Collaboration (PMIC), for more effective MI-driven collaboration.
- 572, TITLE: CerDEQ: Certifiable Deep Equilibrium Model
<https://proceedings.mlr.press/v162/li22t.html>
AUTHORS: Mingjie Li, Yisen Wang, Zhouchen Lin
HIGHLIGHT: In this work, we aim to tackle the problem of DEQ's certified training.
- 573, TITLE: Generalization Guarantee of Training Graph Convolutional Networks with Graph Topology Sampling
<https://proceedings.mlr.press/v162/li22u.html>
AUTHORS: Hongkang Li, Meng Wang, Sijia Liu, Pin-Yu Chen, Jinjun Xiong
HIGHLIGHT: To the best of our knowledge, this paper provides the first theoretical justification of graph topology sampling in training (up to) three-layer GCNs for semi-supervised node classification.
- 574, TITLE: Let Invariant Rationale Discovery Inspire Graph Contrastive Learning
<https://proceedings.mlr.press/v162/li22v.html>
AUTHORS: Sihang Li, Xiang Wang, An Zhang, Yingxin Wu, Xiangnan He, Tat-Seng Chua
HIGHLIGHT: Taking an invariance look at GCL, we argue that a high-performing augmentation should preserve the salient semantics of anchor graphs regarding instance-discrimination. To this end, we relate GCL with invariant rationale discovery, and propose a new framework, Rationale-aware Graph Contrastive Learning (RGCL).
- 575, TITLE: Difference Advantage Estimation for Multi-Agent Policy Gradients
<https://proceedings.mlr.press/v162/li22w.html>
AUTHORS: Yueheng Li, Guangming Xie, Zongqing Lu
HIGHLIGHT: In this paper, we investigate multi-agent credit assignment induced by reward shaping and provide a theoretical understanding in terms of its credit assignment and policy bias.
- 576, TITLE: Private Adaptive Optimization with Side information
<https://proceedings.mlr.press/v162/li22x.html>
AUTHORS: Tian Li, Manzil Zaheer, Sashank Reddi, Virginia Smith
HIGHLIGHT: To this end, we propose AdaDPS, a general framework that uses non-sensitive side information to precondition the gradients, allowing the effective use of adaptive methods in private settings.
- 577, TITLE: Permutation Search of Tensor Network Structures via Local Sampling
<https://proceedings.mlr.press/v162/li22y.html>
AUTHORS: Chao Li, Junhua Zeng, Zerui Tao, Qibin Zhao
HIGHLIGHT: In this paper, we consider a practical variant of TN-SS, dubbed TN permutation search (TN-PS), in which we search for good mappings from tensor modes onto TN vertices (core tensors) for compact TN representations.
- 578, TITLE: Hessian-Free High-Resolution Nesterov Acceleration For Sampling
<https://proceedings.mlr.press/v162/li22z.html>
AUTHORS: Ruilin Li, Hongyuan Zha, Molei Tao
HIGHLIGHT: Nesterov's Accelerated Gradient (NAG) for optimization has better performance than its continuous time limit (noiseless kinetic Langevin) when a finite step-size is employed (Shi et al., 2021). This work explores the sampling counterpart of this phenomenon and proposes a diffusion process, whose discretizations can yield accelerated gradient-based MCMC methods.
- 579, TITLE: Double Sampling Randomized Smoothing
<https://proceedings.mlr.press/v162/li22aa.html>
AUTHORS: Linyi Li, Jiawei Zhang, Tao Xie, Bo Li
HIGHLIGHT: We instantiate DSRS for a generalized family of Gaussian smoothing and propose an efficient and sound computing method based on customized dual optimization considering sampling error.
- 580, TITLE: HouseE: Knowledge Graph Embedding with Householder Parameterization
<https://proceedings.mlr.press/v162/li22ab.html>
AUTHORS: Rui Li, Jianan Zhao, Chaozhuo Li, Di He, Yiqi Wang, Yuming Liu, Hao Sun, Senzhang Wang, Weiwei Deng, Yanming Shen, Xing Xie, Qi Zhang
HIGHLIGHT: In this work, we propose a more powerful KGE framework named HouseE, which involves a novel parameterization based on two kinds of Householder transformations: (1) Householder rotations to achieve superior capacity of modeling relation patterns; (2) Householder projections to handle sophisticated relation mapping properties.
- 581, TITLE: Learning Multiscale Transformer Models for Sequence Generation
<https://proceedings.mlr.press/v162/li22ac.html>
AUTHORS: Bei Li, Tong Zheng, Yi Jing, Chengbo Jiao, Tong Xiao, Jingbo Zhu
HIGHLIGHT: We built a multiscale Transformer model by establishing relationships among scales based on word-boundary information and phrase-level prior knowledge.
- 582, TITLE: Finding Global Homophily in Graph Neural Networks When Meeting Heterophily

<https://proceedings.mlr.press/v162/li22ad.html>

AUTHORS: Xiang Li, Renyu Zhu, Yao Cheng, Caihua Shan, Siqiang Luo, Dongsheng Li, Weining Qian
HIGHLIGHT: Further, for other homophilous nodes excluded in the neighborhood, they are ignored for information aggregation. To address these problems, we propose two models GloGNN and GloGNN++, which generate a node's embedding by aggregating information from global nodes in the graph.

583, TITLE: Fat-Tailed Variational Inference with Anisotropic Tail Adaptive Flows

<https://proceedings.mlr.press/v162/liang22a.html>

AUTHORS: Feynman Liang, Michael Mahoney, Liam Hodgkinson

HIGHLIGHT: In doing so, we unveil a fundamental problem which plagues many existing flow-based methods: they can only model tail-isotropic distributions (i.e., distributions having the same tail parameter in every direction). To mitigate this and enable modeling of tail-anisotropic targets, we propose anisotropic tail-adaptive flows (ATAF).

584, TITLE: Exploring and Exploiting Hubness Priors for High-Quality GAN Latent Sampling

<https://proceedings.mlr.press/v162/liang22b.html>

AUTHORS: Yuanbang Liang, Jing Wu, Yu-Kun Lai, Yipeng Qin

HIGHLIGHT: In this paper, we propose a novel GAN latent sampling method by exploring and exploiting the hubness priors of GAN latent distributions.

585, TITLE: Reducing Variance in Temporal-Difference Value Estimation via Ensemble of Deep Networks

<https://proceedings.mlr.press/v162/liang22c.html>

AUTHORS: Litian Liang, Yaosheng Xu, Stephen McAleer, Dailin Hu, Alexander Ihler, Pieter Abbeel, Roy Fox

HIGHLIGHT: In this paper, we propose MeanQ, a simple ensemble method that estimates target values as ensemble means.

586, TITLE: TSPipe: Learn from Teacher Faster with Pipelines

<https://proceedings.mlr.press/v162/lim22a.html>

AUTHORS: Hwijoon Lim, Yechan Kim, Sukmin Yun, Jinwoo Shin, Dongsu Han

HIGHLIGHT: This paper presents TSPipe, a pipelined approach to accelerate the training process of any TS frameworks including KD and SSL.

587, TITLE: Order Constraints in Optimal Transport

<https://proceedings.mlr.press/v162/lim22b.html>

AUTHORS: Yu Chin Fabian Lim, Laura Wynter, Shiao Hong Lim

HIGHLIGHT: We introduce novel order constraints into the optimal transport formulation to allow for the incorporation of structure.

588, TITLE: Flow-Guided Sparse Transformer for Video Deblurring

<https://proceedings.mlr.press/v162/lin22a.html>

AUTHORS: Jing Lin, Yuanhao Cai, Xiaowan Hu, Haoqian Wang, Youliang Yan, Xueyi Zou, Henghui Ding, Yulun Zhang, Radu Timofte, Luc Van Gool

HIGHLIGHT: In this paper, we propose a novel framework, Flow-Guided Sparse Transformer (FGST), for video deblurring.

589, TITLE: Federated Learning with Positive and Unlabeled Data

<https://proceedings.mlr.press/v162/lin22b.html>

AUTHORS: Xinyang Lin, Hanting Chen, Yixing Xu, Chao Xu, Xiaolin Gui, Yiping Deng, Yunhe Wang

HIGHLIGHT: Therefore, existing PU learning methods can be hardly applied in this situation. To address this problem, we propose a novel framework, namely Federated learning with Positive and Unlabeled data (FedPU), to minimize the expected risk of multiple negative classes by leveraging the labeled data in other clients.

590, TITLE: Decentralized Online Convex Optimization in Networked Systems

<https://proceedings.mlr.press/v162/lin22c.html>

AUTHORS: Yiheng Lin, Judy Gan, Guannan Qu, Yash Kanoria, Adam Wierman

HIGHLIGHT: Our work proposes a novel online algorithm, Localized Predictive Control (LPC), which generalizes predictive control to multi-agent systems.

591, TITLE: Unsupervised Flow-Aligned Sequence-to-Sequence Learning for Video Restoration

<https://proceedings.mlr.press/v162/lin22d.html>

AUTHORS: Jing Lin, Xiaowan Hu, Yuanhao Cai, Haoqian Wang, Youliang Yan, Xueyi Zou, Yulun Zhang, Luc Van Gool

HIGHLIGHT: How to properly model the inter-frame relation within the video sequence is an important but unsolved challenge for video restoration (VR). In this work, we propose an unsupervised flow-aligned sequence-to-sequence model (S2SVR) to address this problem.

592, TITLE: Constrained Gradient Descent: A Powerful and Principled Evasion Attack Against Neural Networks

<https://proceedings.mlr.press/v162/lin22e.html>

AUTHORS: Weiran Lin, Keane Lucas, Lujo Bauer, Michael K. Reiter, Mahmood Sharif

HIGHLIGHT: We propose new, more efficient targeted white-box attacks against deep neural networks.

593, TITLE: Learning Augmented Binary Search Trees

<https://proceedings.mlr.press/v162/lin22f.html>

AUTHORS: Honghao Lin, Tian Luo, David Woodruff

HIGHLIGHT: Given recent advances in algorithms with predictions, we propose pairing treaps with machine advice to form a learning-augmented treap.

594, **TITLE:** Online Nonsubmodular Minimization with Delayed Costs: From Full Information to Bandit Feedback
<https://proceedings.mlr.press/v162/lin22g.html>

AUTHORS: Tianyi Lin, Aldo Pacchiano, Yaodong Yu, Michael Jordan

HIGHLIGHT: Motivated by applications to online learning in sparse estimation and Bayesian optimization, we consider the problem of online unconstrained nonsubmodular minimization with delayed costs in both full information and bandit feedback settings.

595, **TITLE:** Measuring the Effect of Training Data on Deep Learning Predictions via Randomized Experiments
<https://proceedings.mlr.press/v162/lin22h.html>

AUTHORS: Jinkun Lin, Anqi Zhang, Mathias L'cuycer, Jinyang Li, Aurojit Panda, Siddhartha Sen

HIGHLIGHT: We develop a new, principled algorithm for estimating the contribution of training data points to the behavior of a deep learning model, such as a specific prediction it makes.

596, **TITLE:** Interactively Learning Preference Constraints in Linear Bandits

<https://proceedings.mlr.press/v162/lindner22a.html>

AUTHORS: David Lindner, Sebastian Tschiatschek, Katja Hofmann, Andreas Krause

HIGHLIGHT: We formalize the challenge of interactively learning about these constraints as a novel linear bandit problem which we call constrained linear best-arm identification. To solve this problem, we propose the Adaptive Constraint Learning (ACOL) algorithm.

597, **TITLE:** Delayed Reinforcement Learning by Imitation

<https://proceedings.mlr.press/v162/liotet22a.html>

AUTHORS: Pierre Liotet, Davide Maran, Lorenzo Bisi, Marcello Restelli

HIGHLIGHT: When the agent's observations or interactions are delayed, classic reinforcement learning tools usually fail. In this paper, we propose a simple yet new and efficient solution to this problem.

598, **TITLE:** CITRIS: Causal Identifiability from Temporal Intervened Sequences

<https://proceedings.mlr.press/v162/lippe22a.html>

AUTHORS: Phillip Lippe, Sara Magliacane, Sindy L'we, Yuki M Asano, Taco Cohen, Stratis Gavves

HIGHLIGHT: In this paper, we propose CITRIS, a variational autoencoder framework that learns causal representations from temporal sequences of images in which underlying causal factors have possibly been intervened upon.

599, **TITLE:** StreamingQA: A Benchmark for Adaptation to New Knowledge over Time in Question Answering Models

<https://proceedings.mlr.press/v162/liska22a.html>

AUTHORS: Adam Liska, Tomas Kocisky, Elena Gribovskaya, Tayfun Terzi, Eren Sezener, Devang Agrawal, Cyprien De Masson D'Autume, Tim Scholtes, Manzil Zaheer, Susannah Young, Ellen Gilsenan-Mcmahon, Sophia Austin, Phil Blunsom, Angeliki Lazaridou

HIGHLIGHT: To study how semi-parametric QA models and their underlying parametric language models (LMs) adapt to evolving knowledge, we construct a new large-scale dataset, StreamingQA, with human written and generated questions asked on a given date, to be answered from 14 years of time-stamped news articles.

600, **TITLE:** Distributionally Robust $\mathcal{Q}\mathcal{S}$ -Learning

<https://proceedings.mlr.press/v162/liu22a.html>

AUTHORS: Zijian Liu, Qinxun Bai, Jose Blanchet, Perry Dong, Wei Xu, Zhengqing Zhou, Zhengyuan Zhou

HIGHLIGHT: In this paper, we propose a novel distributionally robust $\mathcal{Q}\mathcal{S}$ -learning algorithm that learns the best policy in the worst distributional perturbation of the environment.

601, **TITLE:** Constrained Variational Policy Optimization for Safe Reinforcement Learning

<https://proceedings.mlr.press/v162/liu22b.html>

AUTHORS: Zuxin Liu, Zhepeng Cen, Vladislav Isenbaev, Wei Liu, Steven Wu, Bo Li, Ding Zhao

HIGHLIGHT: We introduce a novel Expectation-Maximization approach to naturally incorporate constraints during the policy learning: 1) a provable optimal non-parametric variational distribution could be computed in closed form after a convex optimization (E-step); 2) the policy parameter is improved within the trust region based on the optimal variational distribution (M-step).

602, **TITLE:** Benefits of Overparameterized Convolutional Residual Networks: Function Approximation under Smoothness Constraint

<https://proceedings.mlr.press/v162/liu22c.html>

AUTHORS: Hao Liu, Minshuo Chen, Siawpeng Er, Wenjing Liao, Tong Zhang, Tuo Zhao

HIGHLIGHT: The neural network themselves, however, can be highly nonsmooth. To bridge this gap, we take convolutional residual networks (ConvResNets) as an example, and prove that large ConvResNets can not only approximate a target function in terms of function value, but also exhibit sufficient first-order smoothness.

603, **TITLE:** Boosting Graph Structure Learning with Dummy Nodes

<https://proceedings.mlr.press/v162/liu22d.html>

AUTHORS: Xin Liu, Jiayang Cheng, Yangqiu Song, Xin Jiang

HIGHLIGHT: In this paper, we use a particular dummy node connecting to all existing vertices without affecting original vertex and edge properties.

- 604, TITLE: Equivalence Analysis between Counterfactual Regret Minimization and Online Mirror Descent
<https://proceedings.mlr.press/v162/liu22e.html>
AUTHORS: Weiming Liu, Huacong Jiang, Bin Li, Houqiang Li
HIGHLIGHT: In this paper, we show that CFRs with Regret Matching and Regret Matching+ are equivalent to special cases of FTRL and OMD, respectively.
- 605, TITLE: Deep Probability Estimation
<https://proceedings.mlr.press/v162/liu22f.html>
AUTHORS: Sheng Liu, Aakash Kaku, Weicheng Zhu, Matan Leibovich, Sreyas Mohan, Boyang Yu, Haoxiang Huang, Laure Zanna, Narges Razavian, Jonathan Niles-Weed, Carlos Fernandez-Granda
HIGHLIGHT: We evaluate existing methods on the synthetic data as well as on three real-world probability estimation tasks, all of which involve inherent uncertainty: precipitation forecasting from radar images, predicting cancer patient survival from histopathology images, and predicting car crashes from dashcam videos. To address this, we build a synthetic dataset to study and compare different computable metrics.
- 606, TITLE: Gating Dropout: Communication-efficient Regularization for Sparsely Activated Transformers
<https://proceedings.mlr.press/v162/liu22g.html>
AUTHORS: Rui Liu, Young Jin Kim, Alexandre Muzio, Hany Hassan
HIGHLIGHT: In this paper, we propose Gating Dropout, which allows tokens to ignore the gating network and stay at their local machines, thus reducing the cross-machine communication.
- 607, TITLE: Simplex Neural Population Learning: Any-Mixture Bayes-Optimality in Symmetric Zero-sum Games
<https://proceedings.mlr.press/v162/liu22h.html>
AUTHORS: Siqi Liu, Marc Lanctot, Luke Marris, Nicolas Heess
HIGHLIGHT: In this paper, we propose simplex-NeuPL that satisfies two desiderata simultaneously: i) learning a population of strategically diverse basis policies, represented by a single conditional network; ii) using the same network, learn best-responses to any mixture over the simplex of basis policies.
- 608, TITLE: Rethinking Attention-Model Explainability through Faithfulness Violation Test
<https://proceedings.mlr.press/v162/liu22i.html>
AUTHORS: Yibing Liu, Haoliang Li, Yangyang Guo, Chenqi Kong, Jing Li, Shiqi Wang
HIGHLIGHT: However, in this paper, we find one critical limitation in attention explanations: weakness in identifying the polarity of feature impact.
- 609, TITLE: Optimization-Derived Learning with Essential Convergence Analysis of Training and Hyper-training
<https://proceedings.mlr.press/v162/liu22j.html>
AUTHORS: Risheng Liu, Xuan Liu, Shangzhi Zeng, Jin Zhang, Yixuan Zhang
HIGHLIGHT: In this work, we design a Generalized Krasnoselskii-Mann (GKM) scheme based on fixed-point iterations as our fundamental ODL module, which unifies existing ODL methods as special cases.
- 610, TITLE: Deep Neural Network Fusion via Graph Matching with Applications to Model Ensemble and Federated Learning
<https://proceedings.mlr.press/v162/liu22k.html>
AUTHORS: Chang Liu, Chenfei Lou, Runzhong Wang, Alan Yuhan Xi, Li Shen, Junchi Yan
HIGHLIGHT: For the rising problem scale and multi-model consistency issues, we propose an efficient graduated assignment-based model fusion method, dubbed GAMF, which iteratively updates the matchings in a consistency-maintaining manner.
- 611, TITLE: Welfare Maximization in Competitive Equilibrium: Reinforcement Learning for Markov Exchange Economy
<https://proceedings.mlr.press/v162/liu22l.html>
AUTHORS: Zhihan Liu, Miao Lu, Zhaoran Wang, Michael Jordan, Zhuoran Yang
HIGHLIGHT: Working in a setting in which the utility function and the system dynamics are both unknown, we propose to find the socially optimal policy and the CE from data via both online and offline variants of MARL.
- 612, TITLE: Generating 3D Molecules for Target Protein Binding
<https://proceedings.mlr.press/v162/liu22m.html>
AUTHORS: Meng Liu, Youzhi Luo, Kanji Uchino, Koji Maruhashi, Shuiwang Ji
HIGHLIGHT: A fundamental problem in drug discovery is to design molecules that bind to specific proteins. To tackle this problem using machine learning methods, here we propose a novel and effective framework, known as GraphBP, to generate 3D molecules that bind to given proteins by placing atoms of specific types and locations to the given binding site one by one.
- 613, TITLE: Communication-efficient Distributed Learning for Large Batch Optimization
<https://proceedings.mlr.press/v162/liu22n.html>
AUTHORS: Rui Liu, Barzan Mozafari
HIGHLIGHT: In this paper, we propose new gradient compression methods for large batch optimization, JointSpar and its variant JointSpar-LARS with layerwise adaptive learning rates, that jointly reduce both the computation and the communication cost.
- 614, TITLE: Adaptive Accelerated (Extra-)Gradient Methods with Variance Reduction
<https://proceedings.mlr.press/v162/liu22o.html>
AUTHORS: Zijian Liu, Ta Duy Nguyen, Alina Ene, Huy Nguyen

- HIGHLIGHT:** In this paper, we study the finite-sum convex optimization problem focusing on the general convex case.
- 615, **TITLE:** REvolveR: Continuous Evolutionary Models for Robot-to-robot Policy Transfer
<https://proceedings.mlr.press/v162/liu22p.html>
AUTHORS: Xingyu Liu, Deepak Pathak, Kris Kitani
HIGHLIGHT: In this work, we consider the problem of transferring a policy across two different robots with significantly different parameters such as kinematics and morphology.
- 616, **TITLE:** Kill a Bird with Two Stones: Closing the Convergence Gaps in Non-Strongly Convex Optimization by Directly Accelerated SVRG with Double Compensation and Snapshots
<https://proceedings.mlr.press/v162/liu22q.html>
AUTHORS: Yuanyuan Liu, Fanhua Shang, Weixin An, Hongying Liu, Zhouchen Lin
HIGHLIGHT: However, there are still some gaps between the oracle complexities and their lower bounds. To fill in these gaps, this paper proposes a novel Directly Accelerated stochastic Variance reduction (DAVIS) algorithm with two Snapshots for non-strongly convex (non-SC) unconstrained problems.
- 617, **TITLE:** Learning Markov Games with Adversarial Opponents: Efficient Algorithms and Fundamental Limits
<https://proceedings.mlr.press/v162/liu22r.html>
AUTHORS: Qinghua Liu, Yuanhao Wang, Chi Jin
HIGHLIGHT: To address this problem, this work studies no-regret learning in Markov games with adversarial opponents when competing against the best fixed policy in hindsight.
- 618, **TITLE:** Local Augmentation for Graph Neural Networks
<https://proceedings.mlr.press/v162/liu22s.html>
AUTHORS: Songtao Liu, Rex Ying, Hanze Dong, Lanqing Li, Tingyang Xu, Yu Rong, Peilin Zhao, Junzhou Huang, Dinghao Wu
HIGHLIGHT: However, it remains an open question whether the neighborhood information is adequately aggregated for learning representations of nodes with few neighbors. To address this, we propose a simple and efficient data augmentation strategy, local augmentation, to learn the distribution of the node representations of the neighbors conditioned on the central node's representation and enhance GNN's expressive power with generated features.
- 619, **TITLE:** Asking for Knowledge (AFK): Training RL Agents to Query External Knowledge Using Language
<https://proceedings.mlr.press/v162/liu22t.html>
AUTHORS: Iou-Jen Liu, Xingdi Yuan, Marc-Alexandre Côté, Pierre-Yves Oudeyer, Alexander Schwing
HIGHLIGHT: In order to study how agents can be taught to query external knowledge via language, we first introduce two new environments: the grid-world-based Q-BabyAI and the text-based Q-TextWorld. In addition to physical interactions, an agent can query an external knowledge source specialized for these environments to gather information. Second, we propose the 'Asking for Knowledge' (AFK) agent, which learns to generate language commands to query for meaningful knowledge that helps solve the tasks.
- 620, **TITLE:** Learning from Demonstration: Provably Efficient Adversarial Policy Imitation with Linear Function Approximation
<https://proceedings.mlr.press/v162/liu22u.html>
AUTHORS: Zhihan Liu, Yufeng Zhang, Zuyue Fu, Zhuoran Yang, Zhaoran Wang
HIGHLIGHT: In this paper, we study GAIL in both online and offline settings with linear function approximation, where both the transition and reward function are linear in the feature maps.
- 621, **TITLE:** GACT: Activation Compressed Training for Generic Network Architectures
<https://proceedings.mlr.press/v162/liu22v.html>
AUTHORS: Xiaoxuan Liu, Lianmin Zheng, Dequan Wang, Yukuo Cen, Weize Chen, Xu Han, Jianfei Chen, Zhiyuan Liu, Jie Tang, Joey Gonzalez, Michael Mahoney, Alvin Cheung
HIGHLIGHT: This paper presents GACT, an ACT framework to support a broad range of machine learning tasks for generic NN architectures with limited domain knowledge.
- 622, **TITLE:** Robust Training under Label Noise by Over-parameterization
<https://proceedings.mlr.press/v162/liu22w.html>
AUTHORS: Sheng Liu, Zhihui Zhu, Qing Qu, Chong You
HIGHLIGHT: In this work, we propose a principled approach for robust training of over-parameterized deep networks in classification tasks where a proportion of training labels are corrupted.
- 623, **TITLE:** Plan Your Target and Learn Your Skills: Transferable State-Only Imitation Learning via Decoupled Policy Optimization
<https://proceedings.mlr.press/v162/liu22x.html>
AUTHORS: Minghuan Liu, Zhengbang Zhu, Yuzheng Zhuang, Weinan Zhang, Jianye Hao, Yong Yu, Jun Wang
HIGHLIGHT: In this paper, we introduce Decoupled Policy Optimization (DePO), which explicitly decouples the policy as a high-level state planner and an inverse dynamics model.
- 624, **TITLE:** On the Impossibility of Learning to Cooperate with Adaptive Partner Strategies in Repeated Games
<https://proceedings.mlr.press/v162/loftin22a.html>
AUTHORS: Robert Loftin, Frans A Oliehoek

HIGHLIGHT: The goal of this work is to understand whether we can reliably learn to cooperate with other agents without such restrictive assumptions, which are unlikely to hold in real-world applications.

625, **TITLE:** AutoIP: A United Framework to Integrate Physics into Gaussian Processes

<https://proceedings.mlr.press/v162/long22a.html>

AUTHORS: Da Long, Zheng Wang, Aditi Krishnapriyan, Robert Kirby, Shandian Zhe, Michael Mahoney

HIGHLIGHT: In this work, we propose a simple, yet powerful and general framework {—} AutoIP, for Automatically Incorporating Physics {—} that can integrate all kinds of differential equations into Gaussian Processes (GPs) to enhance prediction accuracy and uncertainty quantification.

626, **TITLE:** Bayesian Model Selection, the Marginal Likelihood, and Generalization

<https://proceedings.mlr.press/v162/lotfi22a.html>

AUTHORS: Sanae Lotfi, Pavel Izmailov, Gregory Benton, Micah Goldblum, Andrew Gordon Wilson

HIGHLIGHT: We first revisit the appealing properties of the marginal likelihood for learning constraints and hypothesis testing. We then highlight the conceptual and practical issues in using the marginal likelihood as a proxy for generalization.

627, **TITLE:** Feature Learning and Signal Propagation in Deep Neural Networks

<https://proceedings.mlr.press/v162/lou22a.html>

AUTHORS: Yizhang Lou, Chris E Mingard, Soufiane Hayou

HIGHLIGHT: The curve of the alignment as a function of layer index (generally) exhibits an ascent-descent pattern where the maximum is reached for some hidden layer. In this work, we provide the first explanation for this phenomenon.

628, **TITLE:** Fluctuations, Bias, Variance & Ensemble of Learners: Exact Asymptotics for Convex Losses in High-Dimension

<https://proceedings.mlr.press/v162/loureiro22a.html>

AUTHORS: Bruno Loureiro, Cedric Gerbelot, Maria Refinetti, Gabriele Sicuro, Florent Krzakala

HIGHLIGHT: In this manuscript we develop a quantitative and rigorous theory for the study of fluctuations in an ensemble of generalised linear models trained on different, but correlated, features in high-dimensions.

629, **TITLE:** A Single-Loop Gradient Descent and Perturbed Ascent Algorithm for Nonconvex Functional Constrained Optimization

<https://proceedings.mlr.press/v162/lu22a.html>

AUTHORS: Songtao Lu

HIGHLIGHT: In this paper, we propose a novel gradient descent and perturbed ascent (GDPA) algorithm to solve a class of smooth nonconvex inequality constrained problems.

630, **TITLE:** Additive Gaussian Processes Revisited

<https://proceedings.mlr.press/v162/lu22b.html>

AUTHORS: Xiaoyu Lu, Alexis Boukouvalas, James Hensman

HIGHLIGHT: We propose the orthogonal additive kernel (OAK), which imposes an orthogonality constraint on the additive functions, enabling an identifiable, low-dimensional representation of the functional relationship.

631, **TITLE:** ModLaNets: Learning Generalisable Dynamics via Modularity and Physical Inductive Bias

<https://proceedings.mlr.press/v162/lu22c.html>

AUTHORS: Yupu Lu, Shijie Lin, Guanqi Chen, Jia Pan

HIGHLIGHT: Deep learning models are able to approximate one specific dynamical system but struggle at learning generalisable dynamics, where dynamical systems obey the same laws of physics but contain different numbers of elements (e.g., double- and triple-pendulum systems). To relieve this issue, we proposed the Modular Lagrangian Network (ModLaNet), a structural neural network framework with modularity and physical inductive bias.

632, **TITLE:** Model-Free Opponent Shaping

<https://proceedings.mlr.press/v162/lu22d.html>

AUTHORS: Christopher Lu, Timon Willi, Christian A Schroeder De Witt, Jakob Foerster

HIGHLIGHT: However, these methods are myopic since only a small number of steps can be anticipated, are asymmetric since they treat other agents as naive learners, and require the use of higher-order derivatives, which are calculated through white-box access to an opponent's differentiable learning algorithm. To address these issues, we propose Model-Free Opponent Shaping (M-FOS).

633, **TITLE:** Multi-slots Online Matching with High Entropy

<https://proceedings.mlr.press/v162/lu22e.html>

AUTHORS: Xingyu Lu, Qintong Wu, Wenliang Zhong

HIGHLIGHT: Particularly, the gradient computation and resource allocation are both challenging under this setting due to the absence of a closed-form solution. To overcome these obstacles, we develop a novel algorithm named Online subGradient descent for Multi-slots Allocation (OG-MA).

634, **TITLE:** Maximum Likelihood Training for Score-based Diffusion ODEs by High Order Denoising Score Matching

<https://proceedings.mlr.press/v162/lu22f.html>

AUTHORS: Cheng Lu, Kaiwen Zheng, Fan Bao, Jianfei Chen, Chongxuan Li, Jun Zhu

HIGHLIGHT: In this work, we prove that matching the first-order score is not sufficient to maximize the likelihood of the ODE, by showing a gap between the maximum likelihood and score matching objectives.

- 635, TITLE: Orchestra: Unsupervised Federated Learning via Globally Consistent Clustering
<https://proceedings.mlr.press/v162/lubana22a.html>
AUTHORS: Ekdeep Lubana, Chi Ian Tang, Fahim Kawsar, Robert Dick, Akhil Mathur
HIGHLIGHT: Prior work on this topic has focused on directly extending centralized self-supervised learning techniques, which are not designed to have the properties listed above. To address this situation, we propose Orchestra, a novel unsupervised federated learning technique that exploits the federation's hierarchy to orchestrate a distributed clustering task and enforce a globally consistent partitioning of clients' data into discriminable clusters.
- 636, TITLE: A Rigorous Study of Integrated Gradients Method and Extensions to Internal Neuron Attributions
<https://proceedings.mlr.press/v162/lundstrom22a.html>
AUTHORS: Daniel D Lundstrom, Tianjian Huang, Meisam Razaviyayn
HIGHLIGHT: This paper comments on fundamental aspects of IG and its applications/extensions: 1) We identify key differences between IG function spaces and the supporting literature's function spaces which problematize previous claims of IG uniqueness.
- 637, TITLE: BAMDT: Bayesian Additive Semi-Multivariate Decision Trees for Nonparametric Regression
<https://proceedings.mlr.press/v162/luo22a.html>
AUTHORS: Zhao Tang Luo, Huiyan Sang, Bani Mallick
HIGHLIGHT: In this paper, we develop a new class of Bayesian additive multivariate decision tree models that combine univariate split rules for handling possibly high dimensional features without known multivariate structures and novel multivariate split rules for features with multivariate structures in each weak learner.
- 638, TITLE: Disentangled Federated Learning for Tackling Attributes Skew via Invariant Aggregation and Diversity Transferring
<https://proceedings.mlr.press/v162/luo22b.html>
AUTHORS: Zhengquan Luo, Yunlong Wang, Zilei Wang, Zhenan Sun, Tieniu Tan
HIGHLIGHT: To cope with these, we proposed disentangled federated learning (DFL) to disentangle the domain-specific and cross-invariant attributes into two complementary branches, which are trained by the proposed alternating local-global optimization independently.
- 639, TITLE: Channel Importance Matters in Few-Shot Image Classification
<https://proceedings.mlr.press/v162/luo22c.html>
AUTHORS: Xu Luo, Jing Xu, Zenglin Xu
HIGHLIGHT: Understanding the difficulties posed by this task distribution shift is central to FSL. In this paper, we show that a simple channel-wise feature transformation may be the key to unraveling this secret from a channel perspective.
- 640, TITLE: Learning Dynamics and Generalization in Deep Reinforcement Learning
<https://proceedings.mlr.press/v162/lyle22a.html>
AUTHORS: Clare Lyle, Mark Rowland, Will Dabney, Marta Kwiatkowska, Yarin Gal
HIGHLIGHT: Solving a reinforcement learning (RL) problem poses two competing challenges: fitting a potentially discontinuous value function, and generalizing well to new observations. In this paper, we analyze the learning dynamics of temporal difference algorithms to gain novel insight into the tension between these two objectives.
- 641, TITLE: On Finite-Sample Identifiability of Contrastive Learning-Based Nonlinear Independent Component Analysis
<https://proceedings.mlr.press/v162/lyu22a.html>
AUTHORS: Qi Lyu, Xiao Fu
HIGHLIGHT: This work puts forth a finite-sample identifiability analysis of GCL-based nICA.
- 642, TITLE: Pessimism meets VCG: Learning Dynamic Mechanism Design via Offline Reinforcement Learning
<https://proceedings.mlr.press/v162/lyu22b.html>
AUTHORS: Boxiang Lyu, Zhaoran Wang, Mladen Kolar, Zhuoran Yang
HIGHLIGHT: To the best of our knowledge, our work provides the first offline RL algorithm for dynamic mechanism design without assuming uniform coverage.
- 643, TITLE: Versatile Offline Imitation from Observations and Examples via Regularized State-Occupancy Matching
<https://proceedings.mlr.press/v162/ma22a.html>
AUTHORS: Yecheng Ma, Andrew Shen, Dinesh Jayaraman, Osbert Bastani
HIGHLIGHT: We propose State Matching Offline Distribution Correction Estimation (SMODICE), a novel and versatile regression-based offline imitation learning algorithm derived via state-occupancy matching.
- 644, TITLE: Quantification and Analysis of Layer-wise and Pixel-wise Information Discarding
<https://proceedings.mlr.press/v162/ma22b.html>
AUTHORS: Haotian Ma, Hao Zhang, Fan Zhou, Yinqing Zhang, Quanshi Zhang
HIGHLIGHT: This paper presents a method to explain how the information of each input variable is gradually discarded during the forward propagation in a deep neural network (DNN), which provides new perspectives to explain DNNs.
- 645, TITLE: Interpretable Neural Networks with Frank-Wolfe: Sparse Relevance Maps and Relevance Orderings
<https://proceedings.mlr.press/v162/macdonald22a.html>
AUTHORS: Jan Macdonald, Mathieu E. Besançon, Sebastian Pokutta

- HIGHLIGHT: We study the effects of constrained optimization formulations and Frank-Wolfe algorithms for obtaining interpretable neural network predictions.
- 646, TITLE: A Tighter Analysis of Spectral Clustering, and Beyond
<https://proceedings.mlr.press/v162/macgregor22a.html>
AUTHORS: Peter Macgregor, He Sun
HIGHLIGHT: A Tighter Analysis of Spectral Clustering, and Beyond
- 647, TITLE: Zero-Shot Reward Specification via Grounded Natural Language
<https://proceedings.mlr.press/v162/mahmoudieh22a.html>
AUTHORS: Parsa Mahmoudieh, Deepak Pathak, Trevor Darrell
HIGHLIGHT: We use recent developments in building large-scale visuolanguage models like CLIP to devise a framework that generates the task reward signal just from goal text description and raw pixel observations which is then used to learn the task policy.
- 648, TITLE: Feature selection using e-values
<https://proceedings.mlr.press/v162/majumdar22a.html>
AUTHORS: Subhabrata Majumdar, Snigdhanu Chatterjee
HIGHLIGHT: In the context of supervised learning, we introduce the concept of e-value.
- 649, TITLE: SSL Enables Learning from Sparse Rewards in Image-Goal Navigation
<https://proceedings.mlr.press/v162/majumdar22b.html>
AUTHORS: Arjun Majumdar, Gunnar A Sigurdsson, Robinson Piramuthu, Jesse Thomason, Dhruv Batra, Gaurav S Sukhatme
HIGHLIGHT: We demonstrate that combining sparse rewards with self-supervised learning (SSL) not only makes them work, but also outperforms dense rewards, which is the first result of this kind.
- 650, TITLE: Knowledge-Grounded Self-Rationalization via Extractive and Natural Language Explanations
<https://proceedings.mlr.press/v162/majumder22a.html>
AUTHORS: Bodhisattwa Prasad Majumder, Oana Camburu, Thomas Lukasiewicz, Julian McAuley
HIGHLIGHT: However, current models that generate the best extractive rationales or NLEs often fall behind the state-of-the-art (SOTA) in terms of task performance. In this work, we bridge this gap by introducing REXC, a self-rationalizing framework that grounds its predictions and two complementary types of explanations (NLEs and extractive rationales) in background knowledge.
- 651, TITLE: Nonparametric Involutive Markov Chain Monte Carlo
<https://proceedings.mlr.press/v162/mak22a.html>
AUTHORS: Carol Mak, Fabian Zaiser, Luke Ong
HIGHLIGHT: We present the nonparametric involutive Markov chain Monte Carlo (NP-iMCMC) algorithm as a method for constructing MCMC inference algorithms for nonparametric models expressible in universal PPLs.
- 652, TITLE: Architecture Agnostic Federated Learning for Neural Networks
<https://proceedings.mlr.press/v162/makhija22a.html>
AUTHORS: Disha Makhija, Xing Han, Nhat Ho, Joydeep Ghosh
HIGHLIGHT: This work introduces a novel framework, Federated Heterogeneous Neural Networks (FedHeNN), that allows each client to build a personalised model without enforcing a common architecture across clients.
- 653, TITLE: Robustness in Multi-Objective Submodular Optimization: a Quantile Approach
<https://proceedings.mlr.press/v162/malherbe22a.html>
AUTHORS: Cedric Malherbe, Kevin Scaman
HIGHLIGHT: In this work, we propose to design and analyse novel algorithms for the robust allocation of submodular systems through lens of quantile maximization.
- 654, TITLE: More Efficient Sampling for Tensor Decomposition With Worst-Case Guarantees
<https://proceedings.mlr.press/v162/malik22a.html>
AUTHORS: Osman Asif Malik
HIGHLIGHT: In this paper, we propose sampling-based ALS methods for the CP and tensor ring decompositions whose cost does not have this exponential dependence, thereby significantly improving on the previous state-of-the-art.
- 655, TITLE: Unaligned Supervision for Automatic Music Transcription in The Wild
<https://proceedings.mlr.press/v162/maman22a.html>
AUTHORS: Ben Maman, Amit H Bermano
HIGHLIGHT: We introduce Note $\{EM\}$, a method for simultaneously training a transcriber and aligning the scores to their corresponding performances, in a fully-automated process.
- 656, TITLE: Decision-Focused Learning: Through the Lens of Learning to Rank
<https://proceedings.mlr.press/v162/mandi22a.html>
AUTHORS: Jayanta Mandi, Victor Bucarey, Maxime Mulamba Ke Tchomba, Tias Guns
HIGHLIGHT: We develop pointwise, pairwise and listwise ranking loss functions, which can be differentiated in closed form given a subset of solutions.
- 657, TITLE: Differentially Private Coordinate Descent for Composite Empirical Risk Minimization

- <https://proceedings.mlr.press/v162/mangold22a.html>
AUTHORS: Paul Mangold, Aurélien Bellet, Joseph Salmon, Marc Tommasi
HIGHLIGHT: In this paper, we propose Differentially Private proximal Coordinate Descent (DP-CD), a new method to solve composite DP-ERM problems.
- 658, TITLE: Refined Convergence Rates for Maximum Likelihood Estimation under Finite Mixture Models
<https://proceedings.mlr.press/v162/manole22a.html>
AUTHORS: Tudor A Manole, Nhat Ho
HIGHLIGHT: We revisit the classical problem of deriving convergence rates for the maximum likelihood estimator (MLE) in finite mixture models.
- 659, TITLE: On Improving Model-Free Algorithms for Decentralized Multi-Agent Reinforcement Learning
<https://proceedings.mlr.press/v162/mao22a.html>
AUTHORS: Weichao Mao, Lin Yang, Kaiqing Zhang, Tamer Basar
HIGHLIGHT: For learning Nash equilibria in Markov potential games, we propose an independent policy gradient algorithm with a decentralized momentum-based variance reduction technique.
- 660, TITLE: On the Effects of Artificial Data Modification
<https://proceedings.mlr.press/v162/marcu22a.html>
AUTHORS: Antonia Marcu, Adam Prugel-Bennett
HIGHLIGHT: Specifically, we show current shape bias identification methods and occlusion robustness measures are biased and propose a fairer alternative for the latter.
- 661, TITLE: Personalized Federated Learning through Local Memorization
<https://proceedings.mlr.press/v162/marfoq22a.html>
AUTHORS: Othmane Marfoq, Giovanni Neglia, Richard Vidal, Laetitia Kameni
HIGHLIGHT: In this work, we exploit the ability of deep neural networks to extract high quality vectorial representations (embeddings) from non-tabular data, e.g., images and text, to propose a personalization mechanism based on local memorization.
- 662, TITLE: Nested Bandits
<https://proceedings.mlr.press/v162/martin22a.html>
AUTHORS: Matthieu Martin, Panayotis Mertikopoulos, Thibaud Rahier, Houssam Zenati
HIGHLIGHT: In this setting, optimal algorithms based on the exponential weights blueprint (like Hedge, EXP3, and their variants) may incur significant regret because they tend to spend excessive amounts of time exploring irrelevant alternatives with similar, suboptimal costs. To account for this, we propose a nested exponential weights (NEW) algorithm that performs a layered exploration of the learner’s set of alternatives based on a nested, step-by-step selection method.
- 663, TITLE: Closed-Form Diffeomorphic Transformations for Time Series Alignment
<https://proceedings.mlr.press/v162/martinez22a.html>
AUTHORS: Iñigo Martínez, Elisabeth Viles, Igor G. Olaizola
HIGHLIGHT: In this work, we present a closed-form expression for the ODE solution and its gradient under continuous piecewise-affine (CPA) velocity functions.
- 664, TITLE: SPECTRE: Spectral Conditioning Helps to Overcome the Expressivity Limits of One-shot Graph Generators
<https://proceedings.mlr.press/v162/martinkus22a.html>
AUTHORS: Karolis Martinkus, Andreas Loukas, Nathanaël Perraudin, Roger Wattenhofer
HIGHLIGHT: We approach the graph generation problem from a spectral perspective by first generating the dominant parts of the graph Laplacian spectrum and then building a graph matching these eigenvalues and eigenvectors.
- 665, TITLE: Modular Conformal Calibration
<https://proceedings.mlr.press/v162/marx22a.html>
AUTHORS: Charles Marx, Shengjia Zhao, Willie Neiswanger, Stefano Ermon
HIGHLIGHT: We introduce a versatile class of algorithms for recalibration in regression that we call modular conformal calibration (MCC).
- 666, TITLE: Continual Repeated Annealed Flow Transport Monte Carlo
<https://proceedings.mlr.press/v162/matthews22a.html>
AUTHORS: Alex Matthews, Michael Arbel, Danilo Jimenez Rezende, Arnaud Doucet
HIGHLIGHT: We propose Continual Repeated Annealed Flow Transport Monte Carlo (CRAFT), a method that combines a sequential Monte Carlo (SMC) sampler (itself a generalization of Annealed Importance Sampling) with variational inference using normalizing flows.
- 667, TITLE: How to Stay Curious while avoiding Noisy TVs using Aleatoric Uncertainty Estimation
<https://proceedings.mlr.press/v162/mavor-parker22a.html>
AUTHORS: Augustine Mavor-Parker, Kimberly Young, Caswell Barry, Lewis Griffin
HIGHLIGHT: In an attempt to make exploring agents robust to Noisy TVs, we present a simple solution: aleatoric mapping agents (AMAs).
- 668, TITLE: How to Steer Your Adversary: Targeted and Efficient Model Stealing Defenses with Gradient Redirection
<https://proceedings.mlr.press/v162/mazeika22a.html>

- AUTHORS: Mantas Mazeika, Bo Li, David Forsyth
HIGHLIGHT: However, existing defenses have poor performance in practice, either requiring enormous computational overheads or severe utility trade-offs. To meet these challenges, we present a new approach to model stealing defenses called gradient redirection.
- 669, TITLE: Quant-BnB: A Scalable Branch-and-Bound Method for Optimal Decision Trees with Continuous Features
<https://proceedings.mlr.press/v162/mazumder22a.html>
AUTHORS: Rahul Mazumder, Xiang Meng, Haoyue Wang
HIGHLIGHT: In this paper, we consider the problem of learning optimal decision trees, a combinatorial optimization problem that is challenging to solve at scale.
- 670, TITLE: Optimizing Tensor Network Contraction Using Reinforcement Learning
<https://proceedings.mlr.press/v162/meirom22a.html>
AUTHORS: Eli Meirom, Haggai Maron, Shie Mannor, Gal Chechik
HIGHLIGHT: We propose a Reinforcement Learning (RL) approach combined with Graph Neural Networks (GNN) to address the contraction ordering problem.
- 671, TITLE: Causal Transformer for Estimating Counterfactual Outcomes
<https://proceedings.mlr.press/v162/melnychuk22a.html>
AUTHORS: Valentyn Melnychuk, Dennis Frauen, Stefan Feuerriegel
HIGHLIGHT: In this paper, we develop a novel Causal Transformer for estimating counterfactual outcomes over time.
- 672, TITLE: Steerable 3D Spherical Neurons
<https://proceedings.mlr.press/v162/melnyk22a.html>
AUTHORS: Pavlo Melnyk, Michael Felsberg, M?rten Wadenb?ck
HIGHLIGHT: In our work, we propose a steerable feed-forward learning-based approach that consists of neurons with spherical decision surfaces and operates on point clouds.
- 673, TITLE: Transformers are Meta-Reinforcement Learners
<https://proceedings.mlr.press/v162/melo22a.html>
AUTHORS: Luckeciano C Melo
HIGHLIGHT: In this work, we present TrMRL (Transformers for Meta-Reinforcement Learning), a meta-RL agent that mimics the memory reinstatement mechanism using the transformer architecture.
- 674, TITLE: ButterflyFlow: Building Invertible Layers with Butterfly Matrices
<https://proceedings.mlr.press/v162/meng22a.html>
AUTHORS: Chenlin Meng, Linqi Zhou, Kristy Choi, Tri Dao, Stefano Ermon
HIGHLIGHT: We propose a new family of invertible linear layers based on butterfly layers, which are known to theoretically capture complex linear structures including permutations and periodicity, yet can be inverted efficiently.
- 675, TITLE: In defense of dual-encoders for neural ranking
<https://proceedings.mlr.press/v162/menon22a.html>
AUTHORS: Aditya Menon, Sadeep Jayasumana, Ankit Singh Rawat, Seungyeon Kim, Sashank Reddi, Sanjiv Kumar
HIGHLIGHT: However, a more fundamental question remains less explored: does this performance gap reflect an inherent limitation in the capacity of DE models, or a limitation in the training of such models? And does such an understanding suggest a principled means of improving DE models? In this paper, we study these questions, with three contributions.
- 676, TITLE: Equivariant Quantum Graph Circuits
<https://proceedings.mlr.press/v162/mernyei22a.html>
AUTHORS: Peter Mernyei, Konstantinos Meichanetzidis, Ismail Ilkan Ceylan
HIGHLIGHT: We investigate quantum circuits for graph representation learning, and propose equivariant quantum graph circuits (EQGCs), as a class of parameterized quantum circuits with strong relational inductive bias for learning over graph-structured data.
- 677, TITLE: Stochastic Rising Bandits
<https://proceedings.mlr.press/v162/metelli22a.html>
AUTHORS: Alberto Maria Metelli, Francesco Trov?, Matteo Pirola, Marcello Restelli
HIGHLIGHT: We design an algorithm for the rested case (R-ed-UCB) and one for the restless case (R-less-UCB), providing a regret bound depending on the properties of the instance and, under certain circumstances, of $\mathcal{O}(\frac{2}{3})$.
- 678, TITLE: Minimizing Control for Credit Assignment with Strong Feedback
<https://proceedings.mlr.press/v162/meulemans22a.html>
AUTHORS: Alexander Meulemans, Matilde Tristany Farinha, Maria R. Cervera, Jo?o Sacramento, Benjamin F. Grewe
HIGHLIGHT: Overall, our work presents a fundamentally novel view of learning as control minimization, while sidestepping biologically unrealistic assumptions.
- 679, TITLE: A Dynamical System Perspective for Lipschitz Neural Networks
<https://proceedings.mlr.press/v162/meunier22a.html>
AUTHORS: Laurent Meunier, Blaise J Delattre, Alexandre Araujo, Alexandre Allauzen

- HIGHLIGHT:** In this paper, we tackle the problem of building \mathcal{L}_1 -Lipschitz Neural Networks.
- 680, **TITLE:** Distribution Regression with Sliced Wasserstein Kernels
<https://proceedings.mlr.press/v162/meunier22b.html>
AUTHORS: Dimitri Meunier, Massimiliano Pontil, Carlo Ciliberto
HIGHLIGHT: In this work, we propose an OT-based estimator for distribution regression.
- 681, **TITLE:** Interpretable and Generalizable Graph Learning via Stochastic Attention Mechanism
<https://proceedings.mlr.press/v162/miao22a.html>
AUTHORS: Siqi Miao, Mia Liu, Pan Li
HIGHLIGHT: However, those post-hoc methods often fail to provide stable interpretation and may extract features that are spuriously correlated with the task. In this work, we address these issues by proposing Graph Stochastic Attention (GSAT).
- 682, **TITLE:** Modeling Structure with Undirected Neural Networks
<https://proceedings.mlr.press/v162/mihaylova22a.html>
AUTHORS: Tsvetomila Mihaylova, Vlad Niculae, Andre Martins
HIGHLIGHT: In this paper, we combine the representational strengths of factor graphs and of neural networks, proposing undirected neural networks (UNNs): a flexible framework for specifying computations that can be performed in any order.
- 683, **TITLE:** Universal Hopfield Networks: A General Framework for Single-Shot Associative Memory Models
<https://proceedings.mlr.press/v162/millidge22a.html>
AUTHORS: Beren Millidge, Tommaso Salvatori, Yuhang Song, Thomas Lukasiewicz, Rafal Bogacz
HIGHLIGHT: These include the classical Hopfield networks (HNs), sparse distributed memories (SDMs), and more recently the modern continuous Hopfield networks (MCHNs), which possess close links with self-attention in machine learning. In this paper, we propose a general framework for understanding the operation of such memory networks as a sequence of three operations: similarity, separation, and projection.
- 684, **TITLE:** Learning Stochastic Shortest Path with Linear Function Approximation
<https://proceedings.mlr.press/v162/min22a.html>
AUTHORS: Yifei Min, Jiafan He, Tianhao Wang, Quanquan Gu
HIGHLIGHT: We propose a novel algorithm with Hoeffding-type confidence sets for learning the linear mixture SSP, which can attain an $\tilde{O}(B_{\star}^{1.5} \sqrt{K/c_{\min}})$ regret.
- 685, **TITLE:** Prioritized Training on Points that are Learnable, Worth Learning, and not yet Learnt
<https://proceedings.mlr.press/v162/mindermann22a.html>
AUTHORS: Soren Mindermann, Jan M Brauner, Muhammed T Razzak, Mrinank Sharma, Andreas Kirsch, Winnie Xu, Benedikt H?ltgen, Aidan N Gomez, Adrien Morisot, Sebastian Farquhar, Yarin Gal
HIGHLIGHT: To accelerate training, we introduce Reducible Holdout Loss Selection (RHO-LOSS), a simple but principled technique which selects approximately those points for training that most reduce the model's generalization loss.
- 686, **TITLE:** POEM: Out-of-Distribution Detection with Posterior Sampling
<https://proceedings.mlr.press/v162/ming22a.html>
AUTHORS: Yifei Ming, Ying Fan, Yixuan Li
HIGHLIGHT: In this work, we propose a novel posterior sampling based outlier mining framework, POEM, which facilitates efficient use of outlier data and promotes learning a compact decision boundary between ID and OOD data for improved detection.
- 687, **TITLE:** A Simple Reward-free Approach to Constrained Reinforcement Learning
<https://proceedings.mlr.press/v162/miryoosefi22a.html>
AUTHORS: Sobhan Miryoosefi, Chi Jin
HIGHLIGHT: This paper bridges reward-free RL and constrained RL. Particularly, we propose a simple meta-algorithm such that given any reward-free RL oracle, the approachability and constrained RL problems can be directly solved with negligible overheads in sample complexity.
- 688, **TITLE:** Wide Neural Networks Forget Less Catastrophically
<https://proceedings.mlr.press/v162/mirzadeh22a.html>
AUTHORS: Seyed Iman Mirzadeh, Arslan Chaudhry, Dong Yin, Huiyi Hu, Razvan Pascanu, Dilan Gorur, Mehrdad Farajtabar
HIGHLIGHT: While the recent progress in continual learning literature is encouraging, our understanding of what properties of neural networks contribute to catastrophic forgetting is still limited. To address this, instead of focusing on continual learning algorithms, in this work, we focus on the model itself and study the impact of "width" of the neural network architecture on catastrophic forgetting, and show that width has a surprisingly significant effect on forgetting.
- 689, **TITLE:** Proximal and Federated Random Reshuffling
<https://proceedings.mlr.press/v162/mishchenko22a.html>
AUTHORS: Konstantin Mishchenko, Ahmed Khaled, Peter Richtarik
HIGHLIGHT: We propose two new algorithms: Proximal and Federated Random Reshuffling (ProxRR and FedRR).
- 690, **TITLE:** ProxSkip: Yes! Local Gradient Steps Provably Lead to Communication Acceleration! Finally!
<https://proceedings.mlr.press/v162/mishchenko22b.html>
AUTHORS: Konstantin Mishchenko, Grigory Malinovsky, Sebastian Stich, Peter Richtarik

HIGHLIGHT: In this work we are specifically interested in the regime in which the evaluation of prox is costly relative to the evaluation of the gradient, which is the case in many applications.

691, **TITLE:** Fast Convex Optimization for Two-Layer ReLU Networks: Equivalent Model Classes and Cone Decompositions

<https://proceedings.mlr.press/v162/mishkin22a.html>

AUTHORS: Aaron Mishkin, Arda Sahiner, Mert Pilanci

HIGHLIGHT: We develop fast algorithms and robust software for convex optimization of two-layer neural networks with ReLU activation functions.

692, **TITLE:** Memory-Based Model Editing at Scale

<https://proceedings.mlr.press/v162/mitchell22a.html>

AUTHORS: Eric Mitchell, Charles Lin, Antoine Bosselut, Christopher D Manning, Chelsea Finn

HIGHLIGHT: As a higher-capacity alternative, we propose Semi-Parametric Editing with a Retrieval-Augmented Counterfactual Model (SERAC), which stores edits in an explicit memory and learns to reason over them to modulate the base model's predictions as needed.

693, **TITLE:** Invariant Ancestry Search

<https://proceedings.mlr.press/v162/mogensen22a.html>

AUTHORS: Phillip B Mogensen, Nikolaj Thams, Jonas Peters

HIGHLIGHT: We introduce the concept of minimal invariance and propose invariant ancestry search (IAS).

694, **TITLE:** Differentially Private Community Detection for Stochastic Block Models

<https://proceedings.mlr.press/v162/mohamed22a.html>

AUTHORS: Mohamed S Mohamed, Dung Nguyen, Anil Vullikanti, Ravi Tandon

HIGHLIGHT: In this paper, we study the community detection problem while preserving the privacy of the individual connections between the vertices.

695, **TITLE:** A Multi-objective / Multi-task Learning Framework Induced by Pareto Stationarity

<https://proceedings.mlr.press/v162/momma22a.html>

AUTHORS: Michinari Momma, Chaosheng Dong, Jia Liu

HIGHLIGHT: In this paper, we develop a novel and generic framework to discover a PO solution with multiple forms of preferences.

696, **TITLE:** EqR: Equivariant Representations for Data-Efficient Reinforcement Learning

<https://proceedings.mlr.press/v162/mondal22a.html>

AUTHORS: Arnab Kumar Mondal, Vineet Jain, Kaleem Siddiqi, Siamak Ravanbakhsh

HIGHLIGHT: In particular, we propose new mechanisms for learning representations that are equivariant to both the agent's action, as well as symmetry transformations of the state-action pairs.

697, **TITLE:** Feature and Parameter Selection in Stochastic Linear Bandits

<https://proceedings.mlr.press/v162/moradipari22a.html>

AUTHORS: Ahmadreza Moradipari, Berkay Turan, Yasin Abbasi-Yadkori, Mahnoosh Alizadeh, Mohammad Ghavamzadeh

HIGHLIGHT: We study two model selection settings in stochastic linear bandits (LB).

698, **TITLE:** Power-Law Escape Rate of SGD

<https://proceedings.mlr.press/v162/mori22a.html>

AUTHORS: Takashi Mori, Liu Ziyin, Kangqiao Liu, Masahito Ueda

HIGHLIGHT: Stochastic gradient descent (SGD) undergoes complicated multiplicative noise for the mean-square loss. We use this property of SGD noise to derive a stochastic differential equation (SDE) with simpler additive noise by performing a random time change.

699, **TITLE:** Rethinking Fano's Inequality in Ensemble Learning

<https://proceedings.mlr.press/v162/morishita22a.html>

AUTHORS: Terufumi Morishita, Gaku Morio, Shota Horiguchi, Hiroaki Ozaki, Nobuo Nukaga

HIGHLIGHT: We propose a fundamental theory on ensemble learning that evaluates a given ensemble system by a well-grounded set of metrics.

700, **TITLE:** SpeqNets: Sparsity-aware permutation-equivariant graph networks

<https://proceedings.mlr.press/v162/morris22a.html>

AUTHORS: Christopher Morris, Gaurav Rattan, Sandra Kiefer, Siamak Ravanbakhsh

HIGHLIGHT: By introducing new heuristics for the graph isomorphism problem, we devise a class of universal, permutation-equivariant graph networks, which, unlike previous architectures, offer a fine-grained control between expressivity and scalability and adapt to the sparsity of the graph.

701, **TITLE:** CtrlFormer: Learning Transferable State Representation for Visual Control via Transformer

<https://proceedings.mlr.press/v162/mu22a.html>

AUTHORS: Yao Mark Mu, Shoufa Chen, Mingyu Ding, Jianyu Chen, Runjian Chen, Ping Luo

HIGHLIGHT: However, porting Transformer to sample-efficient visual control remains a challenging and unsolved problem. To this end, we propose a novel Control Transformer (CtrlFormer), possessing many appealing benefits that prior arts do not have.

- 702, TITLE: Generalized Beliefs for Cooperative AI
<https://proceedings.mlr.press/v162/muglich22a.html>
AUTHORS: Darius Muglich, Luisa M Zintgraf, Christian A Schroeder De Witt, Shimon Whiteson, Jakob Foerster
HIGHLIGHT: To address this, recent approaches rely on encoding symmetry and convention-awareness into policy training, but these require strong environmental assumptions and can complicate policy training. To overcome this, we propose moving the learning of conventions to the belief space.
- 703, TITLE: Bounding the Width of Neural Networks via Coupled Initialization A Worst Case Analysis
<https://proceedings.mlr.press/v162/munteanu22a.html>
AUTHORS: Alexander Munteanu, Simon Omlor, Zhao Song, David Woodruff
HIGHLIGHT: We observe that by instead initializing the weights into independent pairs, where each pair consists of two identical Gaussian vectors, we can significantly improve the convergence analysis.
- 704, TITLE: Constants Matter: The Performance Gains of Active Learning
<https://proceedings.mlr.press/v162/musmann22a.html>
AUTHORS: Stephen O Mussmann, Sanjoy Dasgupta
HIGHLIGHT: In this work, we show through upper and lower bounds, that for a simple benign setting of well-specified logistic regression on a uniform distribution over a sphere, the expected excess error of both active learning and random sampling have the same inverse proportional dependence on the number of samples.
- 705, TITLE: On the Generalization Analysis of Adversarial Learning
<https://proceedings.mlr.press/v162/mustafa22a.html>
AUTHORS: Waleed Mustafa, Yunwen Lei, Marius Kloft
HIGHLIGHT: In this paper, we study the generalization properties of adversarial learning.
- 706, TITLE: Universal and data-adaptive algorithms for model selection in linear contextual bandits
<https://proceedings.mlr.press/v162/muthukumar22a.html>
AUTHORS: Vidya K Muthukumar, Akshay Krishnamurthy
HIGHLIGHT: In this paper, we introduce new algorithms that a) explore in a data-adaptive manner, and b) provide model selection guarantees of the form $O(d^{\alpha} T^{1-\alpha})$ with no feature diversity conditions whatsoever, where d denotes the dimension of the linear model and T denotes the total number of rounds.
- 707, TITLE: The Importance of Non-Markovianity in Maximum State Entropy Exploration
<https://proceedings.mlr.press/v162/mutti22a.html>
AUTHORS: Mirco Mutti, Riccardo De Santi, Marcello Restelli
HIGHLIGHT: In this paper, we argue that non-Markovianity is instead paramount for maximum state entropy exploration in a finite-sample regime.
- 708, TITLE: PAC-Net: A Model Pruning Approach to Inductive Transfer Learning
<https://proceedings.mlr.press/v162/myung22a.html>
AUTHORS: Sanghoon Myung, In Huh, Wonik Jang, Jae Myung Choe, Jisu Ryu, Daesin Kim, Kee-Eung Kim, Changwook Jeong
HIGHLIGHT: In this paper, we propose PAC-Net, a simple yet effective approach for transfer learning based on pruning.
- 709, TITLE: AutoSNN: Towards Energy-Efficient Spiking Neural Networks
<https://proceedings.mlr.press/v162/na22a.html>
AUTHORS: Byunggook Na, Jisoo Mok, Seongsik Park, Dongjin Lee, Hyeokjun Choe, Sungroh Yoon
HIGHLIGHT: To further improve the accuracy and reduce the spikes generated by SNNs, we propose a spike-aware neural architecture search framework called AutoSNN.
- 710, TITLE: Implicit Bias of the Step Size in Linear Diagonal Neural Networks
<https://proceedings.mlr.press/v162/nacson22a.html>
AUTHORS: Mor Shpigel Nacson, Kavya Ravichandran, Nathan Srebro, Daniel Soudry
HIGHLIGHT: Focusing on diagonal linear networks as a model for understanding the implicit bias in underdetermined models, we show how the gradient descent step size can have a large qualitative effect on the implicit bias, and thus on generalization ability.
- 711, TITLE: DNNR: Differential Nearest Neighbors Regression
<https://proceedings.mlr.press/v162/nader22a.html>
AUTHORS: Youssef Nader, Leon Sixt, Tim Landgraf
HIGHLIGHT: We propose a novel method called Differential Nearest Neighbors Regression (DNNR) that addresses both issues simultaneously: during training, DNNR estimates local gradients to scale the features; during inference, it performs an n -th order Taylor approximation using estimated gradients.
- 712, TITLE: Overcoming Oscillations in Quantization-Aware Training
<https://proceedings.mlr.press/v162/nagel22a.html>
AUTHORS: Markus Nagel, Marios Fournarakis, Yelysei Bondarenko, Tijmen Blankevoort

HIGHLIGHT: In this paper, we delve deeper into the phenomenon of weight oscillations and show that it can lead to a significant accuracy degradation due to wrongly estimated batch-normalization statistics during inference and increased noise during training.

713, **TITLE:** Strategic Representation

<https://proceedings.mlr.press/v162/nair22a.html>

AUTHORS: Vineet Nair, Ganesh Ghalme, Inbal Talgam-Cohen, Nir Rosenfeld

HIGHLIGHT: How can a user make good choices based on strategic representations? We formalize this as a learning problem, and pursue algorithms for decision-making that are robust to manipulation.

714, **TITLE:** Improving Ensemble Distillation With Weight Averaging and Diversifying Perturbation

<https://proceedings.mlr.press/v162/nam22a.html>

AUTHORS: Giung Nam, Hyungi Lee, Byeongho Heo, Juho Lee

HIGHLIGHT: In this paper, we propose a weight averaging technique where a student with multiple subnetworks is trained to absorb the functional diversity of ensemble teachers, but then those subnetworks are properly averaged for inference, giving a single student network with no additional inference cost.

715, **TITLE:** Measuring Representational Robustness of Neural Networks Through Shared Invariances

<https://proceedings.mlr.press/v162/nanda22a.html>

AUTHORS: Vedant Nanda, Till Speicher, Camila Kolling, John P Dickerson, Krishna Gummadi, Adrian Weller

HIGHLIGHT: We propose a measure called `\stir`, which faithfully captures the extent to which two NNs share invariances.

716, **TITLE:** Tight and Robust Private Mean Estimation with Few Users

<https://proceedings.mlr.press/v162/narayanan22a.html>

AUTHORS: Shyam Narayanan, Vahab Mirrokni, Hossein Esfandiari

HIGHLIGHT: In this work, we study high-dimensional mean estimation under user-level differential privacy, and design an (ϵ, δ) -differentially private mechanism using as few users as possible.

717, **TITLE:** Fast Aquatic Swimmer Optimization with Differentiable Projective Dynamics and Neural Network Hydrodynamic Models

<https://proceedings.mlr.press/v162/nava22a.html>

AUTHORS: Elvis Nava, John Z Zhang, Mike Yan Michelis, Tao Du, Pingchuan Ma, Benjamin F. Grewe, Wojciech Matusik, Robert Kevin Katzschmann

HIGHLIGHT: Optimizing robotic swimmer design within such a system generally involves cumbersome, gradient-free procedures on top of the already costly simulation. To address this challenge we present a novel, fully differentiable hybrid approach to FSI that combines a 2D direct numerical simulation for the deformable solid structure of the swimmer and a physics-constrained neural network surrogate to capture hydrodynamic effects of the fluid.

718, **TITLE:** Multi-Task Learning as a Bargaining Game

<https://proceedings.mlr.press/v162/navon22a.html>

AUTHORS: Aviv Navon, Aviv Shamsian, Idan Achituve, Haggai Maron, Kenji Kawaguchi, Gal Chechik, Ethan Fetaya

HIGHLIGHT: In this paper, we propose viewing the gradients combination step as a bargaining game, where tasks negotiate to reach an agreement on a joint direction of parameter update.

719, **TITLE:** Variational Inference for Infinitely Deep Neural Networks

<https://proceedings.mlr.press/v162/nazaret22a.html>

AUTHORS: Achille Nazaret, David Blei

HIGHLIGHT: We introduce the unbounded depth neural network (UDN), an infinitely deep probabilistic model that adapts its complexity to the training data.

720, **TITLE:** Stable Conformal Prediction Sets

<https://proceedings.mlr.press/v162/ndiaye22a.html>

AUTHORS: Eugene Ndiaye

HIGHLIGHT: In this paper, we combine CP techniques with classical algorithmic stability bounds to derive a prediction set computable with a single model fit.

721, **TITLE:** Discovering Generalizable Spatial Goal Representations via Graph-based Active Reward Learning

<https://proceedings.mlr.press/v162/netanyahu22a.html>

AUTHORS: Aviv Netanyahu, Tianmin Shu, Joshua Tenenbaum, Pulkit Agrawal

HIGHLIGHT: In this work, we consider one-shot imitation learning for object rearrangement tasks, where an AI agent needs to watch a single expert demonstration and learn to perform the same task in different environments.

722, **TITLE:** Sublinear-Time Clustering Oracle for Signed Graphs

<https://proceedings.mlr.press/v162/neumann22a.html>

AUTHORS: Stefan Neumann, Pan Peng

HIGHLIGHT: We provide a local clustering oracle for signed graphs with such a clear community structure, that can answer membership queries, i.e., "Given a vertex v , which community does v belong to?", in sublinear time by reading only a small portion of the graph.

723, **TITLE:** Improved Regret for Differentially Private Exploration in Linear MDP

<https://proceedings.mlr.press/v162/ngo22a.html>
AUTHORS: Dung Daniel T Ngo, Giuseppe Vietri, Steven Wu
HIGHLIGHT: We provide a private algorithm with an improved regret rate with an optimal dependence of $O(\sqrt{K})$ on the number of episodes.

724, TITLE: A Framework for Learning to Request Rich and Contextually Useful Information from Humans
<https://proceedings.mlr.press/v162/nguyen22a.html>
AUTHORS: Khanh X Nguyen, Yonatan Bisk, Hal Daum? Iii
HIGHLIGHT: We present a general interactive framework that enables an agent to request and interpret rich, contextually useful information from an assistant that has knowledge about the task and the environment.

725, TITLE: Transformer Neural Processes: Uncertainty-Aware Meta Learning Via Sequence Modeling
<https://proceedings.mlr.press/v162/nguyen22b.html>
AUTHORS: Tung Nguyen, Aditya Grover
HIGHLIGHT: We propose Transformer Neural Processes (TNPs), a new member of the NP family that casts uncertainty-aware meta learning as a sequence modeling problem.

726, TITLE: Improving Transformers with Probabilistic Attention Keys
<https://proceedings.mlr.press/v162/nguyen22c.html>
AUTHORS: Tam Minh Nguyen, Tan Minh Nguyen, Dung D. D. Le, Duy Khuong Nguyen, Viet-Anh Tran, Richard Baraniuk, Nhat Ho, Stanley Osher
HIGHLIGHT: It has been observed that for many applications, those attention heads learn redundant embedding, and most of them can be removed without degrading the performance of the model. Inspired by this observation, we propose Transformer with a Mixture of Gaussian Keys (Transformer-MGK), a novel transformer architecture that replaces redundant heads in transformers with a mixture of keys at each head.

727, TITLE: On Transportation of Mini-batches: A Hierarchical Approach
<https://proceedings.mlr.press/v162/nguyen22d.html>
AUTHORS: Khai Nguyen, Dang Nguyen, Quoc Dinh Nguyen, Tung Pham, Hung Bui, Dinh Phung, Trung Le, Nhat Ho
HIGHLIGHT: Moreover, the m-OT does not approximate a proper metric between probability measures since the identity property is not satisfied. To address these problems, we propose a novel mini-batch scheme for optimal transport, named Batch of Mini-batches Optimal Transport (BoMb-OT), that finds the optimal coupling between mini-batches and it can be seen as an approximation to a well-defined distance on the space of probability measures.

728, TITLE: Improving Mini-batch Optimal Transport via Partial Transportation
<https://proceedings.mlr.press/v162/nguyen22e.html>
AUTHORS: Khai Nguyen, Dang Nguyen, The-Anh Vu-Le, Tung Pham, Nhat Ho
HIGHLIGHT: Motivated by the misspecified mappings issue, we propose a novel mini-batch method by using partial optimal transport (POT) between mini-batch empirical measures, which we refer to as mini-batch partial optimal transport (m-POT).

729, TITLE: Recurrent Model-Free RL Can Be a Strong Baseline for Many POMDPs
<https://proceedings.mlr.press/v162/ni22a.html>
AUTHORS: Tianwei Ni, Benjamin Eysenbach, Ruslan Salakhutdinov
HIGHLIGHT: We find that careful architecture and hyperparameter decisions can often yield a recurrent model-free implementation that performs on par with (and occasionally substantially better than) more sophisticated recent techniques.

730, TITLE: Optimal Estimation of Policy Gradient via Double Fitted Iteration
<https://proceedings.mlr.press/v162/ni22b.html>
AUTHORS: Chengzhuo Ni, Ruiqi Zhang, Xiang Ji, Xuezhou Zhang, Mengdi Wang
HIGHLIGHT: In this paper, we propose the double Fitted PG estimation (FPG) algorithm.

731, TITLE: GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models
<https://proceedings.mlr.press/v162/nichol22a.html>
AUTHORS: Alexander Quinn Nichol, Prafulla Dhariwal, Aditya Ramesh, Pranav Shyam, Pamela Mishkin, Bob McGrew, Ilya Sutskever, Mark Chen
HIGHLIGHT: We explore diffusion models for the problem of text-conditional image synthesis and compare two different guidance strategies: CLIP guidance and classifier-free guidance.

732, TITLE: Diffusion Models for Adversarial Purification
<https://proceedings.mlr.press/v162/nie22a.html>
AUTHORS: Weili Nie, Brandon Guo, Yujia Huang, Chaowei Xiao, Arash Vahdat, Animashree Anandkumar
HIGHLIGHT: In this work, we propose DiffPure that uses diffusion models for adversarial purification: Given an adversarial example, we first diffuse it with a small amount of noise following a forward diffusion process, and then recover the clean image through a reverse generative process.

733, TITLE: The Primacy Bias in Deep Reinforcement Learning
<https://proceedings.mlr.press/v162/nikishin22a.html>
AUTHORS: Evgenii Nikishin, Max Schwarzer, Pierluca D'Oro, Pierre-Luc Bacon, Aaron Courville
HIGHLIGHT: This work identifies a common flaw of deep reinforcement learning (RL) algorithms: a tendency to rely on early interactions and ignore useful evidence encountered later.

- 734, TITLE: Causal Conceptions of Fairness and their Consequences
<https://proceedings.mlr.press/v162/nilforoshan22a.html>
AUTHORS: Hamed Nilforoshan, Johann D Gaebler, Ravi Shroff, Sharad Goel
HIGHLIGHT: Here, we first assemble and categorize popular causal definitions of algorithmic fairness into two broad families: (1) those that constrain the effects of decisions on counterfactual disparities; and (2) those that constrain the effects of legally protected characteristics, like race and gender, on decisions. We then show, analytically and empirically, that both families of definitions almost always—in a measure theoretic sense—result in strongly Pareto dominated decision policies, meaning there is an alternative, unconstrained policy favored by every stakeholder with preferences drawn from a large, natural class.
- 735, TITLE: Efficient Test-Time Model Adaptation without Forgetting
<https://proceedings.mlr.press/v162/niu22a.html>
AUTHORS: Shuaicheng Niu, Jiayang Wu, Yifan Zhang, Yaofu Chen, Shijian Zheng, Peilin Zhao, Mingkui Tan
HIGHLIGHT: To address these issues, we propose an efficient anti-forgetting test-time adaptation (EATA) method.
- 736, TITLE: Generative Trees: Adversarial and Copycat
<https://proceedings.mlr.press/v162/nock22a.html>
AUTHORS: Richard Nock, Mathieu Guillaume-Bert
HIGHLIGHT: This paper proposes a new path forward for the generation of tabular data, exploiting decades-old understanding of the supervised task's best components for DT induction, from losses (properness), models (tree-based) to algorithms (boosting).
- 737, TITLE: Path-Aware and Structure-Preserving Generation of Synthetically Accessible Molecules
<https://proceedings.mlr.press/v162/noh22a.html>
AUTHORS: Juhwan Noh, Dae-Woong Jeong, Kiyoun Kim, Sehui Han, Moontae Lee, Honglak Lee, Yousung Jung
HIGHLIGHT: To design synthetically accessible molecules that preserve main structural motifs of target molecules, we propose a reaction-embedded and structure-conditioned variational autoencoder.
- 738, TITLE: Utilizing Expert Features for Contrastive Learning of Time-Series Representations
<https://proceedings.mlr.press/v162/nonnenmacher22a.html>
AUTHORS: Manuel T Nonnenmacher, Lukas Oldenburg, Ingo Steinwart, David Reeb
HIGHLIGHT: We present an approach that incorporates expert knowledge for time-series representation learning.
- 739, TITLE: Tranception: Protein Fitness Prediction with Autoregressive Transformers and Inference-time Retrieval
<https://proceedings.mlr.press/v162/notin22a.html>
AUTHORS: Pascal Notin, Mafalda Dias, Jonathan Frazer, Javier Marchena Hurtado, Aidan N Gomez, Debora Marks, Yarin Gal
HIGHLIGHT: We introduce Tranception, a novel transformer architecture leveraging autoregressive predictions and retrieval of homologous sequences at inference to achieve state-of-the-art fitness prediction performance.
- 740, TITLE: Fast Finite Width Neural Tangent Kernel
<https://proceedings.mlr.press/v162/novak22a.html>
AUTHORS: Roman Novak, Jascha Sohl-Dickstein, Samuel S Schoenholz
HIGHLIGHT: We perform the first in-depth analysis of the compute and memory requirements for NTK computation in finite width networks.
- 741, TITLE: Multicoated Supermasks Enhance Hidden Networks
<https://proceedings.mlr.press/v162/okoshi22a.html>
AUTHORS: Yasuyuki Okoshi, Angel Lopez Garcia-Arias, Kazutoshi Hirose, Kota Ando, Kazushi Kawamura, Thiem Van Chu, Masato Motomura, Jaehoon Yu
HIGHLIGHT: We show that the supermask stops improving even though gradients are not zero, thus underutilizing backpropagated information. To address this we propose a method that extends Hidden Networks by training an overlay of multiple hierarchical supermasks $\{—\}$ a multicoated supermask.
- 742, TITLE: Generalized Leverage Scores: Geometric Interpretation and Applications
<https://proceedings.mlr.press/v162/ordozgoiti22a.html>
AUTHORS: Bruno Ordozgoiti, Antonis Matakos, Aristides Gionis
HIGHLIGHT: In this paper we extend the definition of leverage scores to relate the columns of a matrix to arbitrary subsets of singular vectors.
- 743, TITLE: Practical Almost-Linear-Time Approximation Algorithms for Hybrid and Overlapping Graph Clustering
<https://proceedings.mlr.press/v162/orecchia22a.html>
AUTHORS: Lorenzo Orecchia, Konstantinos Ameranis, Charalampos Tsourakakis, Kunal Talwar
HIGHLIGHT: In this work, we introduce a frame-work based on two novel clustering objectives, which naturally extend the well-studied notion of conductance to clusters with hybrid vertex-and edge-boundary structure.
- 744, TITLE: Anticorrelated Noise Injection for Improved Generalization
<https://proceedings.mlr.press/v162/orvieto22a.html>
AUTHORS: Antonio Orvieto, Hans Kersting, Frank Proske, Francis Bach, Aurelien Lucchi
HIGHLIGHT: In this paper, we zoom in on the problem of correlating the perturbations of consecutive PGD steps.

- 745, TITLE: Scalable Deep Gaussian Markov Random Fields for General Graphs
<https://proceedings.mlr.press/v162/oskarsson22a.html>
AUTHORS: Joel Oskarsson, Per Sid?n, Fredrik Lindsten
HIGHLIGHT: We propose a flexible GMRF model for general graphs built on the multi-layer structure of Deep GMRFs, originally proposed for lattice graphs only.
- 746, TITLE: Zero-shot AutoML with Pretrained Models
<https://proceedings.mlr.press/v162/ozturk22a.html>
AUTHORS: Ekrem ?zt?rk, Fabio Ferreira, Hadi Jomaa, Lars Schmidt-Thieme, Josif Grabocka, Frank Hutter
HIGHLIGHT: Given a new dataset D and a low compute budget, how should we choose a pre-trained model to fine-tune to D, and set the fine-tuning hyperparameters without risking overfitting, particularly if D is small? Here, we extend automated machine learning (AutoML) to best make these choices.
- 747, TITLE: History Compression via Language Models in Reinforcement Learning
<https://proceedings.mlr.press/v162/paischer22a.html>
AUTHORS: Fabian Paischer, Thomas Adler, Vihang Patil, Angela Bitto-Nemling, Markus Holzleitner, Sebastian Lehner, Hamid Eghbal-Zadeh, Sepp Hochreiter
HIGHLIGHT: We propose to utilize a frozen Pretrained Language Transformer (PLT) for history representation and compression to improve sample efficiency.
- 748, TITLE: A Study on the Ramanujan Graph Property of Winning Lottery Tickets
<https://proceedings.mlr.press/v162/pal22a.html>
AUTHORS: Bithika Pal, Arindam Biswas, Sudeshna Kolay, Pabitra Mitra, Biswajit Basu
HIGHLIGHT: We analyze the Ramanujan graph property of such bipartite layers in terms of their spectral characteristics using the Cheeger’s inequality for irregular graphs.
- 749, TITLE: On Learning Mixture of Linear Regressions in the Non-Realizable Setting
<https://proceedings.mlr.press/v162/pal22b.html>
AUTHORS: Soumyabrata Pal, Arya Mazumdar, Rajat Sen, Avishek Ghosh
HIGHLIGHT: In this paper we show that a version of the popular expectation minimization (EM) algorithm finds out the best fit lines in a dataset even when a realizable model is not assumed, under some regularity conditions on the dataset and the initial points, and thereby provides a solution for the ERM.
- 750, TITLE: Plan Better Amid Conservatism: Offline Multi-Agent Reinforcement Learning with Actor Rectification
<https://proceedings.mlr.press/v162/pan22a.html>
AUTHORS: Ling Pan, Longbo Huang, Tengyu Ma, Huazhe Xu
HIGHLIGHT: Multiple agents exacerbate the problem severely, since the suboptimal policy by any agent can lead to uncoordinated global failure. Following this intuition, we propose a simple yet effective method, Offline Multi-Agent RL with Actor Rectification (OMAR), which combines the first-order policy gradients and zeroth-order optimization methods to better optimize the conservative value functions over the actor parameters.
- 751, TITLE: A Unified Weight Initialization Paradigm for Tensorial Convolutional Neural Networks
<https://proceedings.mlr.press/v162/pan22b.html>
AUTHORS: Yu Pan, Zeyong Su, Ao Liu, Wang Jingquan, Nannan Li, Zenglin Xu
HIGHLIGHT: Meanwhile, although there are ad-hoc approaches for specific architectures (e.g., Tensor Ring Nets), they are not applicable to TCNNs with other tensor decomposition methods (e.g., CP or Tucker decomposition). To address this problem, we propose a universal weight initialization paradigm, which generalizes Xavier and Kaiming methods and can be widely applicable to arbitrary TCNNs.
- 752, TITLE: Robustness and Accuracy Could Be Reconcilable by (Proper) Definition
<https://proceedings.mlr.press/v162/pang22a.html>
AUTHORS: Tianyu Pang, Min Lin, Xiao Yang, Jun Zhu, Shuicheng Yan
HIGHLIGHT: Thus, we dig for the origin of this trade-off in adversarial training and find that it may stem from the improperly defined robust error, which imposes an inductive bias of local invariance — an overcorrection towards smoothness.
- 753, TITLE: Towards Coherent and Consistent Use of Entities in Narrative Generation
<https://proceedings.mlr.press/v162/papalampidi22a.html>
AUTHORS: Pinelopi Papalampidi, Kris Cao, Tomas Kocisky
HIGHLIGHT: In this work, we focus on the end task of narrative generation and systematically analyse the long-range entity coherence and consistency in generated stories. First, we propose a set of automatic metrics for measuring model performance in terms of entity usage.
- 754, TITLE: Constrained Discrete Black-Box Optimization using Mixed-Integer Programming
<https://proceedings.mlr.press/v162/papalexopoulos22a.html>
AUTHORS: Theodore P Papalexopoulos, Christian Tjandraatmadja, Ross Anderson, Juan Pablo Vielma, David Belanger
HIGHLIGHT: In response, we propose NN+MILP, a general discrete MBO framework using piecewise-linear neural networks as surrogate models and mixed-integer linear programming (MILP) to optimize the acquisition function.
- 755, TITLE: A Theoretical Comparison of Graph Neural Network Extensions
<https://proceedings.mlr.press/v162/papp22a.html>

AUTHORS: P?l Andr?s Papp, Roger Wattenhofer
HIGHLIGHT: We study and compare different Graph Neural Network extensions that increase the expressive power of GNNs beyond the Weisfeiler-Leman test.

756, TITLE: Validating Causal Inference Methods
<https://proceedings.mlr.press/v162/parikh22a.html>
AUTHORS: Harsh Parikh, Carlos Varjao, Louise Xu, Eric Tchetgen Tchetgen
HIGHLIGHT: Our work introduces a deep generative model-based framework, Credence, to validate causal inference methods.

757, TITLE: The Unsurprising Effectiveness of Pre-Trained Vision Models for Control
<https://proceedings.mlr.press/v162/parisi22a.html>
AUTHORS: Simone Parisi, Aravind Rajeswaran, Senthil Purushwalkam, Abhinav Gupta
HIGHLIGHT: In this context, we revisit and study the role of pre-trained visual representations for control, and in particular representations trained on large-scale computer vision datasets.

758, TITLE: Learning Symmetric Embeddings for Equivariant World Models
<https://proceedings.mlr.press/v162/park22a.html>
AUTHORS: Jung Yeon Park, Ondrej Biza, Linfeng Zhao, Jan-Willem Van De Meent, Robin Walters
HIGHLIGHT: However, characterizing how transformations act on input data is often difficult, limiting the applicability of equivariant models. We propose learning symmetric embedding networks (SEs) that encode an input space (e.g. images), where we do not know the effect of transformations (e.g. rotations), to a feature space that transforms in a known manner under these operations.

759, TITLE: Blurs Behave Like Ensembles: Spatial Smoothings to Improve Accuracy, Uncertainty, and Robustness
<https://proceedings.mlr.press/v162/park22b.html>
AUTHORS: Namuk Park, Songkuk Kim
HIGHLIGHT: BNNs require a large number of predictions to produce reliable results, leading to a significant increase in computational cost. To alleviate this issue, we propose spatial smoothing, a method that ensembles neighboring feature map points of convolutional neural networks.

760, TITLE: Exact Optimal Accelerated Complexity for Fixed-Point Iterations
<https://proceedings.mlr.press/v162/park22c.html>
AUTHORS: Jisun Park, Ernest K Ryu
HIGHLIGHT: This work presents an acceleration mechanism for fixed-point iterations with nonexpansive operators, contractive operators, and nonexpansive operators satisfying a Hölder-type growth condition.

761, TITLE: Kernel Methods for Radial Transformed Compositional Data with Many Zeros
<https://proceedings.mlr.press/v162/park22d.html>
AUTHORS: Junyoung Park, Changwon Yoon, Cheolwoo Park, Jeongyoun Ahn
HIGHLIGHT: We propose a radial transformation that does not require zero substitutions and more importantly results in essential equivalence between domains before and after the transformation.

762, TITLE: Evolving Curricula with Regret-Based Environment Design
<https://proceedings.mlr.press/v162/parker-holder22a.html>
AUTHORS: Jack Parker-Holder, Minqi Jiang, Michael Dennis, Mikayel Samvelyan, Jakob Foerster, Edward Grefenstette, Tim Rockt?schel
HIGHLIGHT: This work proposes harnessing the power of evolution in a principled, regret-based curriculum.

763, TITLE: Neural Language Models are not Born Equal to Fit Brain Data, but Training Helps
<https://proceedings.mlr.press/v162/pasquiou22a.html>
AUTHORS: Alexandre Pasquiou, Yair Lakretz, John T Hale, Bertrand Thirion, Christophe Pallier
HIGHLIGHT: Here, we make first steps in this direction and examine the impact of test loss, training corpus and model architecture (comparing GloVe, LSTM, GPT-2 and BERT), on the prediction of functional Magnetic Resonance Imaging time-courses of participants listening to an audiobook.

764, TITLE: A new similarity measure for covariate shift with applications to nonparametric regression
<https://proceedings.mlr.press/v162/pathak22a.html>
AUTHORS: Reese Pathak, Cong Ma, Martin Wainwright
HIGHLIGHT: We introduce a new measure of distribution mismatch between the source and target distributions using the integrated ratio of probabilities of balls at a given radius.

765, TITLE: Align-RUDDER: Learning From Few Demonstrations by Reward Redistribution
<https://proceedings.mlr.press/v162/patil22a.html>
AUTHORS: Vihang Patil, Markus Hofmarcher, Marius-Constantin Dinu, Matthias Dorfer, Patrick M Blies, Johannes Brandstetter, Jos? Arjona-Medina, Sepp Hochreiter
HIGHLIGHT: In this work, we introduce Align-RUDDER, which utilizes a profile model for reward redistribution that is obtained from multiple sequence alignment of demonstrations.

766, TITLE: POET: Training Neural Networks on Tiny Devices with Integrated Rematerialization and Paging
<https://proceedings.mlr.press/v162/patil22b.html>

- AUTHORS: Shishir G. Patil, Paras Jain, Prabal Dutta, Ion Stoica, Joseph Gonzalez
HIGHLIGHT: We present POET, an algorithm to enable training large neural networks on memory-scarce battery-operated edge devices.
- 767, TITLE: Learning to Cut by Looking Ahead: Cutting Plane Selection via Imitation Learning
<https://proceedings.mlr.press/v162/paulus22a.html>
AUTHORS: Max B Paulus, Giulia Zarpellon, Andreas Krause, Laurent Charlin, Chris Maddison
HIGHLIGHT: In response, we propose a new neural architecture (NeuralCut) for imitation learning on the lookahead expert.
- 768, TITLE: Neural Network Pruning Denoises the Features and Makes Local Connectivity Emerge in Visual Tasks
<https://proceedings.mlr.press/v162/pellegrini22a.html>
AUTHORS: Franco Pellegrini, Giulio Biroli
HIGHLIGHT: Here, we characterize the inductive bias that pruning imprints in such "winning lottery tickets": focusing on visual tasks, we analyze the architecture resulting from iterative magnitude pruning of a simple fully connected network.
- 769, TITLE: Branchformer: Parallel MLP-Attention Architectures to Capture Local and Global Context for Speech Recognition and Understanding
<https://proceedings.mlr.press/v162/peng22a.html>
AUTHORS: Yifan Peng, Siddharth Dalmia, Ian Lane, Shinji Watanabe
HIGHLIGHT: It combines the benefits of extracting local dependencies using convolutions and global dependencies using self-attention. Inspired by this, we propose a more flexible, interpretable and customizable encoder alternative, Branchformer, with parallel branches for modeling various ranged dependencies in end-to-end speech processing.
- 770, TITLE: Pocket2Mol: Efficient Molecular Sampling Based on 3D Protein Pockets
<https://proceedings.mlr.press/v162/peng22b.html>
AUTHORS: Xingang Peng, Shitong Luo, Jiaqi Guan, Qi Xie, Jian Peng, Jianzhu Ma
HIGHLIGHT: To address the challenge, we develop an E(3)-equivariant generative network composed of two modules: 1) a new graph neural network capturing both spatial and bonding relationships between atoms of the binding pockets and 2) a new efficient algorithm which samples new drug candidates conditioned on the pocket representations from a tractable distribution without relying on MCMC.
- 771, TITLE: Differentiable Top-k Classification Learning
<https://proceedings.mlr.press/v162/petersen22a.html>
AUTHORS: Felix Petersen, Hilde Kuehne, Christian Borgelt, Oliver Deussen
HIGHLIGHT: Leveraging recent advances in differentiable sorting and ranking, we propose a family of differentiable top-k cross-entropy classification losses.
- 772, TITLE: Multi-scale Feature Learning Dynamics: Insights for Double Descent
<https://proceedings.mlr.press/v162/pezeshki22a.html>
AUTHORS: Mohammad Pezeshki, Amartya Mitra, Yoshua Bengio, Guillaume Lajoie
HIGHLIGHT: In this work, we investigate the origins of the less studied epoch-wise double descent in which the test error undergoes two non-monotonous transitions, or descents as the training time increases.
- 773, TITLE: A Differential Entropy Estimator for Training Neural Networks
<https://proceedings.mlr.press/v162/pichler22a.html>
AUTHORS: Georg Pichler, Pierre Jean A. Colombo, Malik Boudiaf, G?nther Koliander, Pablo Piantanida
HIGHLIGHT: To address shortcomings in previously proposed estimators for DE, here we introduce KNIFE, a fully parameterized, differentiable kernel-based estimator of DE.
- 774, TITLE: Federated Learning with Partial Model Personalization
<https://proceedings.mlr.press/v162/pillutla22a.html>
AUTHORS: Krishna Pillutla, Kshitiz Malik, Abdel-Rahman Mohamed, Mike Rabbat, Maziar Sanjabi, Lin Xiao
HIGHLIGHT: We consider two federated learning algorithms for training partially personalized models, where the shared and personal parameters are updated either simultaneously or alternately on the devices.
- 775, TITLE: Deep Networks on Toroids: Removing Symmetries Reveals the Structure of Flat Regions in the Landscape Geometry
<https://proceedings.mlr.press/v162/pittorino22a.html>
AUTHORS: Fabrizio Pittorino, Antonio Ferraro, Gabriele Perugini, Christoph Feinauer, Carlo Baldassi, Riccardo Zecchina
HIGHLIGHT: Grouping classifiers into equivalence classes, we develop a standardized parameterization in which all symmetries are removed, resulting in a toroidal topology. On this space, we explore the error landscape rather than the loss.
- 776, TITLE: Geometric Multimodal Contrastive Representation Learning
<https://proceedings.mlr.press/v162/poklukar22a.html>
AUTHORS: Petra Poklukar, Miguel Vasco, Hang Yin, Francisco S. Melo, Ana Paiva, Danica Kragic
HIGHLIGHT: To address it, we present a novel Geometric Multimodal Contrastive (GMC) representation learning method consisting of two main components: i) a two-level architecture consisting of modality-specific base encoders, allowing to process an arbitrary number of modalities to an intermediate representation of fixed dimensionality, and a shared projection head, mapping the intermediate representations to a latent representation space; ii) a multimodal contrastive loss function that encourages the geometric alignment of the learned representations.

- 777, TITLE: Constrained Offline Policy Optimization
<https://proceedings.mlr.press/v162/polosky22a.html>
AUTHORS: Nicholas Polosky, Bruno C. Da Silva, Madalina Fiterau, Jithin Jagannath
HIGHLIGHT: In this work we introduce Constrained Offline Policy Optimization (COPO), an offline policy optimization algorithm for learning in MDPs with cost constraints.
- 778, TITLE: Offline Meta-Reinforcement Learning with Online Self-Supervision
<https://proceedings.mlr.press/v162/pong22a.html>
AUTHORS: Vitchyr H Pong, Ashvin V Nair, Laura M Smith, Catherine Huang, Sergey Levine
HIGHLIGHT: We propose a hybrid offline meta-RL algorithm, which uses offline data with rewards to meta-train an adaptive policy, and then collects additional unsupervised online data, without any reward labels to bridge this distribution shift.
- 779, TITLE: Debiasser Beware: Pitfalls of Centering Regularized Transport Maps
<https://proceedings.mlr.press/v162/pooladian22a.html>
AUTHORS: Aram-Alexandre Pooladian, Marco Cuturi, Jonathan Niles-Weed
HIGHLIGHT: However, and perhaps surprisingly, we present a few cases in which debiasing is provably detrimental in a statistical sense, notably when the regularization strength is large or the number of samples is small.
- 780, TITLE: Adaptive Second Order Coresets for Data-efficient Machine Learning
<https://proceedings.mlr.press/v162/pooladzandi22a.html>
AUTHORS: Omead Pooladzandi, David Davini, Baharan Mirzasoleiman
HIGHLIGHT: We propose AdaCore, a method that leverages the geometry of the data to extract subsets of the training examples for efficient machine learning.
- 781, TITLE: On the Practicality of Deterministic Epistemic Uncertainty
<https://proceedings.mlr.press/v162/postels22a.html>
AUTHORS: Janis Postels, Mattia Seg?, Tao Sun, Luca Daniel Sieber, Luc Van Gool, Fisher Yu, Federico Tombari
HIGHLIGHT: To this end, we first provide a taxonomy of DUMs, and evaluate their calibration under continuous distributional shifts. Then, we extend them to semantic segmentation.
- 782, TITLE: A Simple Guard for Learned Optimizers
<https://proceedings.mlr.press/v162/premont-schwarz22a.html>
AUTHORS: Isabeau Pr?mont-Schwarz, Jaroslav Vi?tku, Jan Feyereisl
HIGHLIGHT: We propose a new class of Safeguarded L2O, called Loss-Guarded L2O (LGL2O), which is both conceptually simpler and computationally less expensive.
- 783, TITLE: Hardness and Algorithms for Robust and Sparse Optimization
<https://proceedings.mlr.press/v162/price22a.html>
AUTHORS: Eric Price, Sandeep Silwal, Samson Zhou
HIGHLIGHT: We explore algorithms and limitations for sparse optimization problems such as sparse linear regression and robust linear regression.
- 784, TITLE: Nonlinear Feature Diffusion on Hypergraphs
<https://proceedings.mlr.press/v162/prokopchik22a.html>
AUTHORS: Konstantin Prokopchik, Austin R Benson, Francesco Tudisco
HIGHLIGHT: Here, we develop a nonlinear diffusion process on hypergraphs that spreads both features and labels following the hypergraph structure.
- 785, TITLE: Universal Joint Approximation of Manifolds and Densities by Simple Injective Flows
<https://proceedings.mlr.press/v162/puthawala22a.html>
AUTHORS: Michael Puthawala, Matti Lassas, Ivan Dokmanic, Maarten De Hoop
HIGHLIGHT: We study approximation of probability measures supported on n -dimensional manifolds embedded in \mathbb{R}^m by injective flows—neural networks composed of invertible flows and injective layers.
- 786, TITLE: The Teaching Dimension of Regularized Kernel Learners
<https://proceedings.mlr.press/v162/qian22a.html>
AUTHORS: Hong Qian, Xu-Hui Liu, Chen-Xi Su, Aimin Zhou, Yang Yu
HIGHLIGHT: Inspired by the fact that regularization can reduce the learning complexity in machine learning, a natural question is whether the similar fact happens in machine teaching. To answer this essential question, this paper proposes a unified theoretical framework termed STARKE to analyze the TD of regularized kernel learners.
- 787, TITLE: ContentVec: An Improved Self-Supervised Speech Representation by Disentangling Speakers
<https://proceedings.mlr.press/v162/qian22b.html>
AUTHORS: Kaizhi Qian, Yang Zhang, Heting Gao, Junrui Ni, Cheng-I Lai, David Cox, Mark Hasegawa-Johnson, Shiyu Chang
HIGHLIGHT: In this paper, we propose a new SSL method that can achieve speaker disentanglement without severe loss of content.
- 788, TITLE: Interventional Contrastive Learning with Meta Semantic Regularizer

- <https://proceedings.mlr.press/v162/qiang22a.html>
AUTHORS: Wenwen Qiang, Jiangmeng Li, Changwen Zheng, Bing Su, Hui Xiong
HIGHLIGHT: We propose a backdoor adjustment-based regularization method, namely Interventional Contrastive Learning with Meta Semantic Regularizer (ICL-MSR), to perform causal intervention towards the proposed SCM.
- 789, TITLE: Sample-Efficient Reinforcement Learning with $\log(\log(T))$ Switching Cost
<https://proceedings.mlr.press/v162/qiao22a.html>
AUTHORS: Dan Qiao, Ming Yin, Ming Min, Yu-Xiang Wang
HIGHLIGHT: In this paper, we propose a new algorithm based on stage-wise exploration and adaptive policy elimination that achieves a regret of $\tilde{O}(\sqrt{H^4 S^2 AT})$ while requiring a switching cost of $O(HSA \log \log T)$.
- 790, TITLE: Generalizing to Evolving Domains with Latent Structure-Aware Sequential Autoencoder
<https://proceedings.mlr.press/v162/qin22a.html>
AUTHORS: Tiexin Qin, Shiqi Wang, Haoliang Li
HIGHLIGHT: In this paper, we formulate the aforementioned setting as the problem of evolving domain generalization.
- 791, TITLE: Graph Neural Architecture Search Under Distribution Shifts
<https://proceedings.mlr.press/v162/qin22b.html>
AUTHORS: Yijian Qin, Xin Wang, Ziwei Zhang, Pengtao Xie, Wenwu Zhu
HIGHLIGHT: However, when there is a distribution shift between training and testing graphs, the existing approaches fail to deal with the problem of adapting to unknown test graph structures since they only search for a fixed architecture for all graphs. To solve this problem, we propose a novel GRACES model which is able to generalize under distribution shifts through tailoring a customized GNN architecture suitable for each graph instance with unknown distribution.
- 792, TITLE: Spectral Representation of Robustness Measures for Optimization Under Input Uncertainty
<https://proceedings.mlr.press/v162/qing22a.html>
AUTHORS: Jixiang Qing, Tom Dhaene, Ivo Couckuyt
HIGHLIGHT: In this paper, we propose a Spectral Representation of Robustness Measures based on the GP's spectral representation, i.e., an analytical approach to approximately infer both robustness measures for normal and uniform input uncertainty distributions.
- 793, TITLE: Large-scale Stochastic Optimization of NDCG Surrogates for Deep Learning with Provable Convergence
<https://proceedings.mlr.press/v162/qui22a.html>
AUTHORS: Zi-Hao Qiu, Quanqi Hu, Yongjian Zhong, Lijun Zhang, Tianbao Yang
HIGHLIGHT: In this paper, we propose a principled approach to optimize NDCG and its top- K variant.
- 794, TITLE: Latent Outlier Exposure for Anomaly Detection with Contaminated Data
<https://proceedings.mlr.press/v162/qui22b.html>
AUTHORS: Chen Qiu, Aodong Li, Marius Kloft, Maja Rudolph, Stephan Mandt
HIGHLIGHT: We propose a strategy for training an anomaly detector in the presence of unlabeled anomalies that is compatible with a broad class of models.
- 795, TITLE: Contrastive UCB: Provably Efficient Contrastive Self-Supervised Learning in Online Reinforcement Learning
<https://proceedings.mlr.press/v162/qui22c.html>
AUTHORS: Shuang Qiu, Lingxiao Wang, Chenjia Bai, Zhuoran Yang, Zhaoran Wang
HIGHLIGHT: To narrow such a gap, we study contrastive-learning empowered RL for a class of Markov decision processes (MDPs) and Markov games (MGs) with low-rank transitions.
- 796, TITLE: Fast and Provable Nonconvex Tensor RPCA
<https://proceedings.mlr.press/v162/qui22d.html>
AUTHORS: Haiquan Qiu, Yao Wang, Shaojie Tang, Deyu Meng, Quanming Yao
HIGHLIGHT: In this paper, we study nonconvex tensor robust principal component analysis (RPCA) based on the StS-SVD .
- 797, TITLE: Generalized Federated Learning via Sharpness Aware Minimization
<https://proceedings.mlr.press/v162/qu22a.html>
AUTHORS: Zhe Qu, Xingyu Li, Rui Duan, Yao Liu, Bo Tang, Zhuo Lu
HIGHLIGHT: Therefore, in this paper, we revisit the solutions to the distribution shift problem in FL with a focus on local learning generality.
- 798, TITLE: Particle Transformer for Jet Tagging
<https://proceedings.mlr.press/v162/qu22b.html>
AUTHORS: Huilin Qu, Congqiao Li, Sitian Qian
HIGHLIGHT: In this work, we present JetClass, a new comprehensive dataset for jet tagging.
- 799, TITLE: Winning the Lottery Ahead of Time: Efficient Early Network Pruning
<https://proceedings.mlr.press/v162/rachwan22a.html>
AUTHORS: John Rachwan, Daniel Z'gner, Bertrand Charpentier, Simon Geisler, Morgane Ayle, Stephan G'nnemann
HIGHLIGHT: We propose Early Compression via Gradient Flow Preservation (EarlyCroP), which efficiently extracts state-of-the-art sparse models before or early in training addressing challenge (1), and can be applied in a structured manner addressing challenge (2).

- 800, TITLE: Convergence of Uncertainty Sampling for Active Learning
<https://proceedings.mlr.press/v162/raj22a.html>
AUTHORS: Anant Raj, Francis Bach
HIGHLIGHT: In this work, we propose an efficient uncertainty estimator for binary classification which we also extend to multiple classes, and provide a non-asymptotic rate of convergence for our uncertainty sampling based active learning algorithm in both cases under no-noise conditions (i.e., linearly separable data).
- 801, TITLE: DeepSpeed-MoE: Advancing Mixture-of-Experts Inference and Training to Power Next-Generation AI Scale
<https://proceedings.mlr.press/v162/rajbhandari22a.html>
AUTHORS: Samyam Rajbhandari, Conglong Li, Zhewei Yao, Minjia Zhang, Reza Yazdani Aminabadi, Ammar Ahmad Awan, Jeff Rasley, Yuxiong He
HIGHLIGHT: However, due to the much larger model size and unique architecture, how to provide fast MoE model inference remains challenging and unsolved, limiting their practical usage. To tackle this, we present DeepSpeed-MoE, an end-to-end MoE training and inference solution, including novel MoE architecture designs and model compression techniques that reduce MoE model size by up to 3.7x, and a highly optimized inference system that provides 7.3x better latency and cost compared to existing MoE inference solutions.
- 802, TITLE: Fishr: Invariant Gradient Variances for Out-of-Distribution Generalization
<https://proceedings.mlr.press/v162/rame22a.html>
AUTHORS: Alexandre Rame, Corentin Dancette, Matthieu Cord
HIGHLIGHT: In this paper, we introduce a new regularization - named Fishr - that enforces domain invariance in the space of the gradients of the loss: specifically, the domain-level variances of gradients are matched across training domains.
- 803, TITLE: A Closer Look at Smoothness in Domain Adversarial Training
<https://proceedings.mlr.press/v162/rangwani22a.html>
AUTHORS: Harsh Rangwani, Sumukh K Aithal, Mayank Mishra, Arihant Jain, Venkatesh Babu Radhakrishnan
HIGHLIGHT: In this work, we analyze the effect of smoothness enhancing formulations on domain adversarial training, the objective of which is a combination of task loss (eg.
- 804, TITLE: Linear Adversarial Concept Erasure
<https://proceedings.mlr.press/v162/ravfogel22a.html>
AUTHORS: Shauli Ravfogel, Michael Twiton, Yoav Goldberg, Ryan D Cotterell
HIGHLIGHT: In this work, we formulate the problem of identifying a linear subspace that corresponds to a given concept, and removing it from the representation.
- 805, TITLE: Implicit Regularization in Hierarchical Tensor Factorization and Deep Convolutional Neural Networks
<https://proceedings.mlr.press/v162/razin22a.html>
AUTHORS: Noam Razin, Asaf Maman, Nadav Cohen
HIGHLIGHT: Through a dynamical systems lens, we overcome challenges associated with hierarchy, and establish implicit regularization towards low hierarchical tensor rank.
- 806, TITLE: One-Pass Algorithms for MAP Inference of Nonsymmetric Determinantal Point Processes
<https://proceedings.mlr.press/v162/reddy22a.html>
AUTHORS: Aravind Reddy, Ryan A. Rossi, Zhao Song, Anup Rao, Tung Mai, Nedim Lipka, Gang Wu, Eunyee Koh, Nesreen Ahmed
HIGHLIGHT: In this paper, we initiate the study of one-pass algorithms for solving the maximum-a-posteriori (MAP) inference problem for Non-symmetric Determinantal Point Processes (NDPPs).
- 807, TITLE: Universality of Winning Tickets: A Renormalization Group Perspective
<https://proceedings.mlr.press/v162/redman22a.html>
AUTHORS: William T Redman, Tianlong Chen, Zhangyang Wang, Akshunna S. Dogra
HIGHLIGHT: We find that iterative magnitude pruning, the principal algorithm used for discovering winning tickets, is a renormalization group scheme, and can be viewed as inducing a flow in parameter space.
- 808, TITLE: The dynamics of representation learning in shallow, non-linear autoencoders
<https://proceedings.mlr.press/v162/refinetti22a.html>
AUTHORS: Maria Refinetti, Sebastian Goldt
HIGHLIGHT: Here, we study the dynamics of feature learning in non-linear, shallow autoencoders.
- 809, TITLE: Proximal Exploration for Model-guided Protein Sequence Design
<https://proceedings.mlr.press/v162/ren22a.html>
AUTHORS: Zhizhou Ren, Jiahan Li, Fan Ding, Yuan Zhou, Jianzhu Ma, Jian Peng
HIGHLIGHT: In this paper, we study the exploration mechanism of model-guided sequence design.
- 810, TITLE: Towards Theoretical Analysis of Transformation Complexity of ReLU DNNs
<https://proceedings.mlr.press/v162/ren22b.html>
AUTHORS: Jie Ren, Mingjie Li, Meng Zhou, Shih-Han Chan, Quanshi Zhang
HIGHLIGHT: This paper aims to theoretically analyze the complexity of feature transformations encoded in piecewise linear DNNs with ReLU layers.

- 811, TITLE: Benchmarking and Analyzing Point Cloud Classification under Corruptions
<https://proceedings.mlr.press/v162/ren22c.html>
AUTHORS: Jiawei Ren, Liang Pan, Ziwei Liu
HIGHLIGHT: In this work, we aim to rigorously benchmark and analyze point cloud classification under corruptions.
- 812, TITLE: A Unified View on PAC-Bayes Bounds for Meta-Learning
<https://proceedings.mlr.press/v162/rezazadeh22a.html>
AUTHORS: Arezou Rezazadeh
HIGHLIGHT: In this paper, by upper bounding arbitrary convex functions, which link the expected and empirical losses at the environment and also per-task levels, we obtain new PAC-Bayes bounds.
- 813, TITLE: 3PC: Three Point Compressors for Communication-Efficient Distributed Training and a Better Theory for Lazy Aggregation
<https://proceedings.mlr.press/v162/richtarik22a.html>
AUTHORS: Peter Richtarik, Igor Sokolov, Elnur Gasanov, Ilyas Fatkhullin, Zhize Li, Eduard Gorbunov
HIGHLIGHT: We propose and study a new class of gradient compressors for communication-efficient training—three point compressors (3PC)—as well as efficient distributed nonconvex optimization algorithms that can take advantage of them.
- 814, TITLE: Robust SDE-Based Variational Formulations for Solving Linear PDEs via Deep Learning
<https://proceedings.mlr.press/v162/richter22a.html>
AUTHORS: Lorenz Richter, Julius Berner
HIGHLIGHT: In this article, we rigorously investigate corresponding numerical aspects that appear in the context of linear Kolmogorov PDEs.
- 815, TITLE: Probabilistically Robust Learning: Balancing Average and Worst-case Performance
<https://proceedings.mlr.press/v162/robey22a.html>
AUTHORS: Alexander Robey, Luiz Chamon, George J. Pappas, Hamed Hassani
HIGHLIGHT: To this end, in this paper we propose a framework called probabilistic robustness that bridges the gap between the accurate, yet brittle average case and the robust, yet conservative worst case by enforcing robustness to most rather than to all perturbations.
- 816, TITLE: LyaNet: A Lyapunov Framework for Training Neural ODEs
<https://proceedings.mlr.press/v162/rodriguez22a.html>
AUTHORS: Ivan Dario Jimenez Rodriguez, Aaron Ames, Yisong Yue
HIGHLIGHT: We propose a method for training ordinary differential equations by using a control-theoretic Lyapunov condition for stability.
- 817, TITLE: Short-Term Plasticity Neurons Learning to Learn and Forget
<https://proceedings.mlr.press/v162/rodriguez22b.html>
AUTHORS: Hector Garcia Rodriguez, Qinghai Guo, Timoleon Moraitis
HIGHLIGHT: Here we present a new type of recurrent neural unit, the STP Neuron (STPN), which indeed turns out strikingly powerful.
- 818, TITLE: Function-space Inference with Sparse Implicit Processes
<https://proceedings.mlr.press/v162/rodri-guez-santana22a.html>
AUTHORS: Simon Rodri?guez-Santana, Bryan Zaldivar, Daniel Hernandez-Lobato
HIGHLIGHT: Existing methods that can tune the prior IP result in a Gaussian predictive distribution, which fails to capture important data patterns. By contrast, methods producing flexible predictive distributions by using another IP to approximate the posterior process cannot tune the prior IP to the observed data. We propose here the first method that can accomplish both goals.
- 819, TITLE: Score Matching Enables Causal Discovery of Nonlinear Additive Noise Models
<https://proceedings.mlr.press/v162/rolland22a.html>
AUTHORS: Paul Rolland, Volkan Cevher, Matth?us Kleindessner, Chris Russell, Dominik Janzing, Bernhard Sch?lkopf, Francesco Locatello
HIGHLIGHT: This paper demonstrates how to recover causal graphs from the score of the data distribution in non-linear additive (Gaussian) noise models.
- 820, TITLE: Dual Decomposition of Convex Optimization Layers for Consistent Attention in Medical Images
<https://proceedings.mlr.press/v162/ron22a.html>
AUTHORS: Tom Ron, Tamir Hazan
HIGHLIGHT: We propose a multi-layer attention mechanism that enforces consistent interpretations between attended convolutional layers using convex optimization.
- 821, TITLE: A Consistent and Efficient Evaluation Strategy for Attribution Methods
<https://proceedings.mlr.press/v162/rong22a.html>
AUTHORS: Yao Rong, Tobias Leemann, Vadim Borisov, Gjergji Kasneci, Enkelejda Kasneci
HIGHLIGHT: In this work, we present an information-theoretic analysis of evaluation strategies based on pixel perturbations.
- 822, TITLE: Efficiently Learning the Topology and Behavior of a Networked Dynamical System Via Active Queries

<https://proceedings.mlr.press/v162/rosenkrantz22a.html>

AUTHORS: Daniel J Rosenkrantz, Abhijin Adiga, Madhav Marathe, Zirou Qiu, S S Ravi, Richard Stearns, Anil Vullikanti
HIGHLIGHT: We present algorithms to learn the topology and the behavior under both batch and adaptive query models for several classes of dynamical systems.

823, TITLE: Learning to Infer Structures of Network Games

<https://proceedings.mlr.press/v162/rossi22a.html>

AUTHORS: Emanuele Rossi, Federico Monti, Yan Leng, Michael Bronstein, Xiaowen Dong

HIGHLIGHT: We adopt a transformer-like architecture which correctly accounts for the symmetries of the problem and learns a mapping from the equilibrium actions to the network structure of the game without explicit knowledge of the utility function.

824, TITLE: Direct Behavior Specification via Constrained Reinforcement Learning

<https://proceedings.mlr.press/v162/roy22a.html>

AUTHORS: Julien Roy, Roger Girgis, Joshua Romoff, Pierre-Luc Bacon, Chris J Pal

HIGHLIGHT: In this work, we argue that constrained RL, which has almost exclusively been used for safe RL, also has the potential to significantly reduce the amount of work spent for reward specification in applied RL projects.

825, TITLE: Constraint-based graph network simulator

<https://proceedings.mlr.press/v162/rubanova22a.html>

AUTHORS: Yulia Rubanova, Alvaro Sanchez-Gonzalez, Tobias Pfaff, Peter Battaglia

HIGHLIGHT: Here we present a framework for constraint-based learned simulation, where a scalar constraint function is implemented as a graph neural network, and future predictions are computed by solving the optimization problem defined by the learned constraint.

826, TITLE: Continual Learning via Sequential Function-Space Variational Inference

<https://proceedings.mlr.press/v162/rudner22a.html>

AUTHORS: Tim G. J. Rudner, Freddie Bickford Smith, Qixuan Feng, Yee Whye Teh, Yarin Gal

HIGHLIGHT: Addressing the drawbacks of existing techniques, we propose an optimization objective derived by formulating continual learning as sequential function-space variational inference.

827, TITLE: Graph-Coupled Oscillator Networks

<https://proceedings.mlr.press/v162/rusch22a.html>

AUTHORS: T. Konstantin Rusch, Ben Chamberlain, James Rowbottom, Siddhartha Mishra, Michael Bronstein

HIGHLIGHT: We propose Graph-Coupled Oscillator Networks (GraphCON), a novel framework for deep learning on graphs.

828, TITLE: Hindering Adversarial Attacks with Implicit Neural Representations

<https://proceedings.mlr.press/v162/rusu22a.html>

AUTHORS: Andrei A Rusu, Dan Andrei Calian, Sven Gowal, Raia Hadsell

HIGHLIGHT: We introduce the Lossy Implicit Network Activation Coding (LINAC) defence, an input transformation which successfully hinders several common adversarial attacks on CIFAR-10 classifiers for perturbations up to $8/255$ in Linf norm and 0.5 in L2 norm.

829, TITLE: Exploiting Independent Instruments: Identification and Distribution Generalization

<https://proceedings.mlr.press/v162/saengkyongam22a.html>

AUTHORS: Sorawit Saengkyongam, Leonard Henckel, Niklas Pfister, Jonas Peters

HIGHLIGHT: We connect to the existing literature in econometrics and provide a practical method called HSIC-X for exploiting independence that can be combined with any gradient-based learning procedure.

830, TITLE: FedNL: Making Newton-Type Methods Applicable to Federated Learning

<https://proceedings.mlr.press/v162/safaryan22a.html>

AUTHORS: Mher Safaryan, Rustem Islamov, Xun Qian, Peter Richtarik

HIGHLIGHT: Inspired by recent work of Islamov et al (2021), we propose a family of Federated Newton Learn ($\backslash\text{algname}\{\text{FedNL}\}$) methods, which we believe is a marked step in the direction of making second-order methods applicable to FL.

831, TITLE: Versatile Dueling Bandits: Best-of-both World Analyses for Learning from Relative Preferences

<https://proceedings.mlr.press/v162/saha22a.html>

AUTHORS: Aadirupa Saha, Pierre Gaillard

HIGHLIGHT: We study the problem of K-armed dueling bandit for both stochastic and adversarial environments, where the goal of the learner is to aggregate information through relative preferences of pair of decision points queried in an online sequential manner.

832, TITLE: Optimal and Efficient Dynamic Regret Algorithms for Non-Stationary Dueling Bandits

<https://proceedings.mlr.press/v162/saha22b.html>

AUTHORS: Aadirupa Saha, Shubham Gupta

HIGHLIGHT: We study the problem of dynamic regret minimization in SK -armed Dueling Bandits under non-stationary or time-varying preferences.

833, TITLE: Unraveling Attention via Convex Duality: Analysis and Interpretations of Vision Transformers

<https://proceedings.mlr.press/v162/sahiner22a.html>

AUTHORS: Arda Sahiner, Tolga Ergen, Batu Ozturkler, John Pauly, Morteza Mardani, Mert Pilanci

- HIGHLIGHT:** However, the underpinning inductive bias of attention is not well understood. To address this issue, this paper analyzes attention through the lens of convex duality.
- 834, **TITLE:** Off-Policy Evaluation for Large Action Spaces via Embeddings
<https://proceedings.mlr.press/v162/saito22a.html>
AUTHORS: Yuta Saito, Thorsten Joachims
HIGHLIGHT: This foils the use of OPE in many applications from recommender systems to language models. To overcome this issue, we propose a new OPE estimator that leverages marginalized importance weights when action embeddings provide structure in the action space.
- 835, **TITLE:** Optimal Clipping and Magnitude-aware Differentiation for Improved Quantization-aware Training
<https://proceedings.mlr.press/v162/sakr22a.html>
AUTHORS: Charbel Sakr, Steve Dai, Rangha Venkatesan, Brian Zimmer, William Dally, Brucec Khailany
HIGHLIGHT: We propose Optimally Clipped Tensors And Vectors (OCTAV), a recursive algorithm to determine MSE-optimal clipping scalars.
- 836, **TITLE:** A Convergence Theory for SVGD in the Population Limit under Talagrand's Inequality T1
<https://proceedings.mlr.press/v162/salim22a.html>
AUTHORS: Adil Salim, Lukang Sun, Peter Richtarik
HIGHLIGHT: We study the convergence of SVGD in the population limit, (i.e., with an infinite number of particles) to sample from a non-logconcave target distribution satisfying Talagrand's inequality T1.
- 837, **TITLE:** FITNESS: (Fine Tune on New and Similar Samples) to detect anomalies in streams with drift and outliers
<https://proceedings.mlr.press/v162/sankararaman22a.html>
AUTHORS: Abishek Sankararaman, Balakrishnan Narayanaswamy, Vikramank Y Singh, Zhao Song
HIGHLIGHT: We propose FITNESS (Fine Tune on New and Similar Samples), a flexible framework for detecting anomalies on data streams.
- 838, **TITLE:** The Algebraic Path Problem for Graph Metrics
<https://proceedings.mlr.press/v162/sanmarti-n22a.html>
AUTHORS: Enrique Fita Sanmarti, Sebastian Damrich, Fred Hamprecht
HIGHLIGHT: We here clarify, for the first time, the relation between the potential distance and the log-semiring.
- 839, **TITLE:** LSB: Local Self-Balancing MCMC in Discrete Spaces
<https://proceedings.mlr.press/v162/sansone22a.html>
AUTHORS: Emanuele Sansone
HIGHLIGHT: We present the Local Self-Balancing sampler (LSB), a local Markov Chain Monte Carlo (MCMC) method for sampling in purely discrete domains, which is able to autonomously adapt to the target distribution and to reduce the number of target evaluations required to converge.
- 840, **TITLE:** PoF: Post-Training of Feature Extractor for Improving Generalization
<https://proceedings.mlr.press/v162/sato22a.html>
AUTHORS: Ikuro Sato, Yamada Ryota, Masayuki Tanaka, Nakamasa Inoue, Rei Kawakami
HIGHLIGHT: We developed a training algorithm called PoF: Post-Training of Feature Extractor that updates the feature extractor part of an already-trained deep model to search a flatter minimum.
- 841, **TITLE:** Re-evaluating Word Mover's Distance
<https://proceedings.mlr.press/v162/sato22b.html>
AUTHORS: Ryoma Sato, Makoto Yamada, Hisashi Kashima
HIGHLIGHT: In this paper, we point out that the evaluation in the original study could be misleading.
- 842, **TITLE:** Understanding Contrastive Learning Requires Incorporating Inductive Biases
<https://proceedings.mlr.press/v162/saunshi22a.html>
AUTHORS: Nikunj Saunshi, Jordan Ash, Surbhi Goel, Dipendra Misra, Cyril Zhang, Sanjeev Arora, Sham Kakade, Akshay Krishnamurthy
HIGHLIGHT: Theoretical analysis is presented for the class of linear representations, where incorporating inductive biases of the function class allows contrastive learning to work with less stringent conditions compared to prior analyses.
- 843, **TITLE:** The Neural Race Reduction: Dynamics of Abstraction in Gated Networks
<https://proceedings.mlr.press/v162/saxe22a.html>
AUTHORS: Andrew Saxe, Shagun Sodhani, Sam Jay Lewallen
HIGHLIGHT: In this work, we begin to address this gap by introducing the Gated Deep Linear Network framework that schematizes how pathways of information flow impact learning dynamics within an architecture.
- 844, **TITLE:** Convergence Rates of Non-Convex Stochastic Gradient Descent Under a Generic Lojasiewicz Condition and Local Smoothness
<https://proceedings.mlr.press/v162/scaman22a.html>
AUTHORS: Kevin Scaman, Cedric Malherbe, Ludovic Dos Santos

HIGHLIGHT: In this work, we propose to extend these results by analyzing stochastic gradient descent under more generic Lojasiewicz conditions that are applicable to any convex loss function, thus extending the current theory to a larger panel of losses commonly used in practice such as cross-entropy.

845, **TITLE:** An Asymptotic Test for Conditional Independence using Analytic Kernel Embeddings

<https://proceedings.mlr.press/v162/scetbon22a.html>

AUTHORS: Meyer Scetbon, Laurent Meunier, Yaniv Romano

HIGHLIGHT: We propose a new conditional dependence measure and a statistical test for conditional independence.

846, **TITLE:** Linear-Time Gromov Wasserstein Distances using Low Rank Couplings and Costs

<https://proceedings.mlr.press/v162/scetbon22b.html>

AUTHORS: Meyer Scetbon, Gabriel Peyr?, Marco Cuturi

HIGHLIGHT: We show in this work how a recent variant of the OT problem that restricts the set of admissible couplings to those having a low-rank factorization is remarkably well suited to the resolution of GW: when applied to GW, we show that this approach is not only able to compute a stationary point of the GW problem in time $\mathcal{O}(n^2)$, but also uniquely positioned to benefit from the knowledge that the initial cost matrices are low-rank, to yield a linear time $\mathcal{O}(n)$ GW approximation.

847, **TITLE:** Streaming Inference for Infinite Feature Models

<https://proceedings.mlr.press/v162/schaeffer22a.html>

AUTHORS: Rylan Schaeffer, Yilun Du, Gabrielle K Liu, Ila Fiete

HIGHLIGHT: In this work, we make feature models significantly more applicable to streaming data by imbuing them with the ability to create new features, online, in a probabilistic and principled manner.

848, **TITLE:** Modeling Irregular Time Series with Continuous Recurrent Units

<https://proceedings.mlr.press/v162/schirmer22a.html>

AUTHORS: Mona Schirmer, Mazin Eltayeb, Stefan Lessmann, Maja Rudolph

HIGHLIGHT: However, in many datasets (e.g. medical records) observation times are irregular and can carry important information. To address this challenge, we propose continuous recurrent units (CRUs) $\{-\}$ a neural architecture that can naturally handle irregular intervals between observations.

849, **TITLE:** Structure Preserving Neural Networks: A Case Study in the Entropy Closure of the Boltzmann Equation

<https://proceedings.mlr.press/v162/schotthofer22a.html>

AUTHORS: Steffen Schotthofer, Tianbai Xiao, Martin Frank, Cory Hauck

HIGHLIGHT: In this paper, we explore applications of deep learning in statistical physics.

850, **TITLE:** Improving Robustness against Real-World and Worst-Case Distribution Shifts through Decision Region Quantification

<https://proceedings.mlr.press/v162/schwinn22a.html>

AUTHORS: Leo Schwinn, Leon Bungert, An Nguyen, Ren? Raab, Falk Pulsmeier, Doina Precup, Bjoern Eskofier, Dario Zanca

HIGHLIGHT: In this work, we propose the Decision Region Quantification (DRQ) algorithm to improve the robustness of any differentiable pre-trained model against both real-world and worst-case distribution shifts in the data.

851, **TITLE:** Symmetric Machine Theory of Mind

<https://proceedings.mlr.press/v162/sclar22a.html>

AUTHORS: Melanie Sclar, Graham Neubig, Yonatan Bisk

HIGHLIGHT: In contrast, we propose to model machine theory of mind in a more general symmetric scenario.

852, **TITLE:** Data-SUITE: Data-centric identification of in-distribution incongruous examples

<https://proceedings.mlr.press/v162/seedat22a.html>

AUTHORS: Nabeel Seedat, Jonathan Crabb?, Mihaela van der Schaar

HIGHLIGHT: To this end, we propose a paradigm shift with Data-SUITE: a data-centric AI framework to identify these regions, independent of a task-specific model.

853, **TITLE:** Continuous-Time Modeling of Counterfactual Outcomes Using Neural Controlled Differential Equations

<https://proceedings.mlr.press/v162/seedat22b.html>

AUTHORS: Nabeel Seedat, Fergus Imrie, Alexis Bellot, Zhaozhi Qian, Mihaela van der Schaar

HIGHLIGHT: To handle arbitrary observation patterns, we interpret the data as samples from an underlying continuous-time process and propose to model its latent trajectory explicitly using the mathematics of controlled differential equations.

854, **TITLE:** Neural Tangent Kernel Beyond the Infinite-Width Limit: Effects of Depth and Initialization

<https://proceedings.mlr.press/v162/seleznova22a.html>

AUTHORS: Mariia Seleznova, Gitta Kutyniok

HIGHLIGHT: In this paper, we study the NTK of fully-connected ReLU networks with depth comparable to width.

855, **TITLE:** Reinforcement Learning with Action-Free Pre-Training from Videos

<https://proceedings.mlr.press/v162/seo22a.html>

AUTHORS: Younggyo Seo, Kimin Lee, Stephen L James, Pieter Abbeel

HIGHLIGHT: To this end, we introduce a framework that learns representations useful for understanding the dynamics via generative pre-training on videos.

- 856, TITLE: Efficient Model-based Multi-agent Reinforcement Learning via Optimistic Equilibrium Computation
<https://proceedings.mlr.press/v162/sessa22a.html>
AUTHORS: Pier Giuseppe Sessa, Maryam Kamgarpour, Andreas Krause
HIGHLIGHT: We propose H-MARL (Hallucinated Multi-Agent Reinforcement Learning), a novel sample-efficient algorithm that can efficiently balance exploration, i.e., learning about the environment, and exploitation, i.e., achieve good equilibrium performance in the underlying general-sum Markov game.
- 857, TITLE: Selective Regression under Fairness Criteria
<https://proceedings.mlr.press/v162/shah22a.html>
AUTHORS: Abhin Shah, Yuheng Bu, Joshua K Lee, Subhro Das, Rameswar Panda, Prasanna Sattigeri, Gregory W Wornell
HIGHLIGHT: However, as we show, in some cases, the performance of a minority subgroup can decrease while we reduce the coverage, and thus selective regression can magnify disparities between different sensitive subgroups. Motivated by these disparities, we propose new fairness criteria for selective regression requiring the performance of every subgroup to improve with a decrease in coverage.
- 858, TITLE: Utility Theory for Sequential Decision Making
<https://proceedings.mlr.press/v162/shakerinava22a.html>
AUTHORS: Mehran Shakerinava, Siamak Ravanbakhsh
HIGHLIGHT: The von Neumann-Morgenstern (VNM) utility theorem shows that under certain axioms of rationality, decision-making is reduced to maximizing the expectation of some utility function. We extend these axioms to increasingly structured sequential decision making settings and identify the structure of the corresponding utility functions.
- 859, TITLE: Translating Robot Skills: Learning Unsupervised Skill Correspondences Across Robots
<https://proceedings.mlr.press/v162/shankar22a.html>
AUTHORS: Tanmay Shankar, Yixin Lin, Aravind Rajeswaran, Vikash Kumar, Stuart Anderson, Jean Oh
HIGHLIGHT: In this paper, we explore how we can endow robots with the ability to learn correspondences between their own skills, and those of morphologically different robots in different domains, in an entirely unsupervised manner.
- 860, TITLE: A State-Distribution Matching Approach to Non-Episodic Reinforcement Learning
<https://proceedings.mlr.press/v162/sharma22a.html>
AUTHORS: Archit Sharma, Rehaan Ahmad, Chelsea Finn
HIGHLIGHT: Assuming access to a few demonstrations, we propose a new method, MEDAL, that trains the backward policy to match the state distribution in the provided demonstrations.
- 861, TITLE: Content Addressable Memory Without Catastrophic Forgetting by Heteroassociation with a Fixed Scaffold
<https://proceedings.mlr.press/v162/sharma22b.html>
AUTHORS: Sugandha Sharma, Sarthak Chandra, Ila Fiete
HIGHLIGHT: We propose a novel CAM architecture, Memory Scaffold with Heteroassociation (MESH), that factorizes the problems of internal attractor dynamics and association with external content to generate a CAM continuum without a memory cliff: Small numbers of patterns are stored with complete information recovery matching standard CAMs, while inserting more patterns still results in partial recall of every pattern, with a graceful trade-off between pattern number and pattern richness.
- 862, TITLE: Federated Minimax Optimization: Improved Convergence Analyses and Algorithms
<https://proceedings.mlr.press/v162/sharma22c.html>
AUTHORS: Pranay Sharma, Rohan Panda, Gauri Joshi, Pramod Varshney
HIGHLIGHT: In this paper, we consider nonconvex minimax optimization, which is gaining prominence in many modern machine learning applications, such as GANs.
- 863, TITLE: DNS: Determinantal Point Process Based Neural Network Sampler for Ensemble Reinforcement Learning
<https://proceedings.mlr.press/v162/sheikh22a.html>
AUTHORS: Hassam Sheikh, Kizza Frisbee, Mariano Phielipp
HIGHLIGHT: In this paper, we propose DNS: a Determinantal Point Process based Neural Network Sampler that specifically uses k-DPP to sample a subset of neural networks for backpropagation at every training step thus significantly reducing the training time and computation cost.
- 864, TITLE: Instance Dependent Regret Analysis of Kernelized Bandits
<https://proceedings.mlr.press/v162/shekhar22a.html>
AUTHORS: Shubhanshu Shekhar, Tara Javidi
HIGHLIGHT: We study the problem of designing an adaptive strategy for querying a noisy zeroth-order-oracle to efficiently learn about the optimizer of an unknown function f^* .
- 865, TITLE: Data Augmentation as Feature Manipulation
<https://proceedings.mlr.press/v162/shen22a.html>
AUTHORS: Ruoqi Shen, Sebastien Bubeck, Suriya Gunasekar
HIGHLIGHT: In this work we consider another angle, and we study the effect of data augmentation on the dynamic of the learning process.
- 866, TITLE: Metric-Fair Active Learning
<https://proceedings.mlr.press/v162/shen22b.html>

- AUTHORS: Jie Shen, Nan Cui, Jing Wang
HIGHLIGHT: In this paper, we henceforth study metric-fair active learning of homogeneous halfspaces, and show that under the distribution-dependent PAC learning model, fairness and label efficiency can be achieved simultaneously.
- 867, TITLE: PDO-s3DCNNs: Partial Differential Operator Based Steerable 3D CNNs
<https://proceedings.mlr.press/v162/shen22c.html>
AUTHORS: Zhengyang Shen, Tao Hong, Qi She, Jinwen Ma, Zhouchen Lin
HIGHLIGHT: In this work, we employ partial differential operators (PDOs) to model 3D filters, and derive general steerable 3D CNNs, which are called PDO-s3DCNNs.
- 868, TITLE: Connect, Not Collapse: Explaining Contrastive Learning for Unsupervised Domain Adaptation
<https://proceedings.mlr.press/v162/shen22d.html>
AUTHORS: Kendrick Shen, Robbie M Jones, Ananya Kumar, Sang Michael Xie, Jeff Z. Haochen, Tengyu Ma, Percy Liang
HIGHLIGHT: In this paper, we show that contrastive pre-training, which learns features on unlabeled source and target data and then fine-tunes on labeled source data, is competitive with strong UDA methods.
- 869, TITLE: Constrained Optimization with Dynamic Bound-scaling for Effective NLP Backdoor Defense
<https://proceedings.mlr.press/v162/shen22e.html>
AUTHORS: Guangyu Shen, Yingqi Liu, Guanhong Tao, Qiuling Xu, Zhuo Zhang, Shengwei An, Shiqing Ma, Xiangyu Zhang
HIGHLIGHT: In this paper, we develop a novel optimization method for NLP backdoor inversion.
- 870, TITLE: Staged Training for Transformer Language Models
<https://proceedings.mlr.press/v162/shen22f.html>
AUTHORS: Sheng Shen, Pete Walsh, Kurt Keutzer, Jesse Dodge, Matthew Peters, Iz Beltagy
HIGHLIGHT: As an alternative, we consider a staged training setup that begins with a small model and incrementally increases the amount of compute used for training by applying a "growth operator" to increase the model depth and width.
- 871, TITLE: Deep Network Approximation in Terms of Intrinsic Parameters
<https://proceedings.mlr.press/v162/shen22g.html>
AUTHORS: Zuowei Shen, Haizhao Yang, Shijun Zhang
HIGHLIGHT: From an approximation perspective, this paper shows that the number of parameters that need to be learned can be significantly smaller than people typically expect.
- 872, TITLE: Gradient-Free Method for Heavily Constrained Nonconvex Optimization
<https://proceedings.mlr.press/v162/shi22a.html>
AUTHORS: Wanli Shi, Hongchang Gao, Bin Gu
HIGHLIGHT: In this paper, to solve the nonconvex problem with a large number of white/black-box constraints, we proposed a doubly stochastic zeroth-order gradient method (DSZOG) with momentum method and adaptive step size.
- 873, TITLE: Global Optimization of K-Center Clustering
<https://proceedings.mlr.press/v162/shi22b.html>
AUTHORS: Mingfei Shi, Kaixun Hua, Jiayang Ren, Yankai Cao
HIGHLIGHT: This work provides a practical global optimization algorithm for this task based on a reduced-space spatial branch and bound scheme.
- 874, TITLE: Pessimistic Q-Learning for Offline Reinforcement Learning: Towards Optimal Sample Complexity
<https://proceedings.mlr.press/v162/shi22c.html>
AUTHORS: Laixi Shi, Gen Li, Yuting Wei, Yuxin Chen, Yuejie Chi
HIGHLIGHT: Altogether, this work highlights the efficiency of model-free algorithms in offline RL when used in conjunction with pessimism and variance reduction.
- 875, TITLE: Adversarial Masking for Self-Supervised Learning
<https://proceedings.mlr.press/v162/shi22d.html>
AUTHORS: Yuge Shi, N Siddharth, Philip Torr, Adam R Kosiorek
HIGHLIGHT: We propose ADIOS, a masked image model (MIM) framework for self-supervised learning, which simultaneously learns a masking function and an image encoder using an adversarial objective.
- 876, TITLE: Visual Attention Emerges from Recurrent Sparse Reconstruction
<https://proceedings.mlr.press/v162/shi22e.html>
AUTHORS: Baifeng Shi, Yale Song, Neel Joshi, Trevor Darrell, Xin Wang
HIGHLIGHT: We present VARS, Visual Attention from Recurrent Sparse reconstruction, a new attention formulation built on two prominent features of the human visual attention mechanism: recurrency and sparsity.
- 877, TITLE: A Minimax Learning Approach to Off-Policy Evaluation in Confounded Partially Observable Markov Decision Processes
<https://proceedings.mlr.press/v162/shi22f.html>
AUTHORS: Chengchun Shi, Masatoshi Uehara, Jiawei Huang, Nan Jiang
HIGHLIGHT: We consider off-policy evaluation (OPE) in Partially Observable Markov Decision Processes (POMDPs), where the evaluation policy depends only on observable variables and the behavior policy depends on unobservable latent variables.

- 878, TITLE: Robust Group Synchronization via Quadratic Programming
<https://proceedings.mlr.press/v162/shi22g.html>
AUTHORS: Yunpeng Shi, Cole M Wyeth, Gilad Lerman
HIGHLIGHT: We propose a novel quadratic programming formulation for estimating the corruption levels in group synchronization, and use these estimates to solve this problem.
- 879, TITLE: Log-Euclidean Signatures for Intrinsic Distances Between Unaligned Datasets
<https://proceedings.mlr.press/v162/shnitzer22a.html>
AUTHORS: Tal Shnitzer, Mikhail Yurochkin, Kristjan Greenewald, Justin M Solomon
HIGHLIGHT: Instead, we exploit the Riemannian geometry of SPD matrices to compare these operators and define a new theoretically-motivated distance based on a lower bound of the log-Euclidean metric.
- 880, TITLE: Scalable Computation of Causal Bounds
<https://proceedings.mlr.press/v162/shridharan22a.html>
AUTHORS: Madhumitha Shridharan, Garud Iyengar
HIGHLIGHT: We consider the problem of computing bounds for causal inference problems with unobserved confounders, where identifiability does not hold.
- 881, TITLE: Bit Prioritization in Variational Autoencoders via Progressive Coding
<https://proceedings.mlr.press/v162/shu22a.html>
AUTHORS: Rui Shu, Stefano Ermon
HIGHLIGHT: In this work, we treat image synthesis itself as a hierarchical representation learning problem and regularize an HVAE toward representations that improve the model's image synthesis performance.
- 882, TITLE: Fair Representation Learning through Implicit Path Alignment
<https://proceedings.mlr.press/v162/shui22a.html>
AUTHORS: Changjian Shui, Qi Chen, Jiaqi Li, Boyu Wang, Christian Gagn?
HIGHLIGHT: Besides, to avoid the high computational and memory cost of differentiating in the inner-loop of bi-level objective, we propose an implicit path alignment algorithm, which only relies on the solution of inner optimization and the implicit differentiation rather than the exact optimization path.
- 883, TITLE: Faster Algorithms for Learning Convex Functions
<https://proceedings.mlr.press/v162/siahkamari22a.html>
AUTHORS: Ali Siahkamari, Durmus Alp Emre Acar, Christopher Liao, Kelly L Geyer, Venkatesh Saligrama, Brian Kulis
HIGHLIGHT: In this paper, we develop and analyze an approach for solving a broad range of convex function learning problems that is faster than state-of-the-art approaches.
- 884, TITLE: Coin Flipping Neural Networks
<https://proceedings.mlr.press/v162/sieradzki22a.html>
AUTHORS: Yuval Sieradzki, Nitzan Hodos, Gal Yehuda, Assaf Schuster
HIGHLIGHT: We show that neural networks with access to randomness can outperform deterministic networks by using amplification.
- 885, TITLE: Reverse Engineering the Neural Tangent Kernel
<https://proceedings.mlr.press/v162/simon22a.html>
AUTHORS: James Benjamin Simon, Sajant Anand, Mike Deweese
HIGHLIGHT: As a paradigm for principled neural architecture design, we propose the translation of high-performing kernels, which are better-understood and amenable to first-principles design, into equivalent network architectures, which have superior efficiency, flexibility, and feature learning.
- 886, TITLE: Demystifying the Adversarial Robustness of Random Transformation Defenses
<https://proceedings.mlr.press/v162/sitawarin22a.html>
AUTHORS: Chawin Sitawarin, Zachary J Golan-Strieb, David Wagner
HIGHLIGHT: First, we show that the BPDA attack (Athalye et al., 2018a) used in BaRT's evaluation is ineffective and likely overestimates its robustness. We then attempt to construct the strongest possible RT defense through the informed selection of transformations and Bayesian optimization for tuning their parameters.
- 887, TITLE: Smoothed Adversarial Linear Contextual Bandits with Knapsacks
<https://proceedings.mlr.press/v162/sivakumar22a.html>
AUTHORS: Vidyashankar Sivakumar, Shiliang Zuo, Arindam Banerjee
HIGHLIGHT: We present algorithms and characterize regret for LinCBwK in the smoothed setting where base context vectors are assumed to be perturbed by Gaussian noise.
- 888, TITLE: GenLabel: Mixup Relabeling using Generative Models
<https://proceedings.mlr.press/v162/sohn22a.html>
AUTHORS: Jy-Yong Sohn, Liang Shang, Hongxu Chen, Jackyun Moon, Dimitris Papailiopoulos, Kangwook Lee
HIGHLIGHT: In this paper, we first identify the main causes of this phenomenon by theoretically and empirically analyzing the mixup algorithm. To resolve this, we propose GenLabel, a simple yet effective relabeling algorithm designed for mixup.

- 889, TITLE: Communicating via Markov Decision Processes
<https://proceedings.mlr.press/v162/sokota22a.html>
AUTHORS: Samuel Sokota, Christian A Schroeder De Witt, Maximilian Igl, Luisa M Zintgraf, Philip Torr, Martin Strohmeier, Zico Kolter, Shimon Whiteson, Jakob Foerster
HIGHLIGHT: We consider the problem of communicating exogenous information by means of Markov decision process trajectories.
- 890, TITLE: The Multivariate Community Hawkes Model for Dependent Relational Events in Continuous-time Networks
<https://proceedings.mlr.press/v162/soliman22a.html>
AUTHORS: Hadeel Soliman, Lingfei Zhao, Zhipeng Huang, Subhadeep Paul, Kevin S Xu
HIGHLIGHT: We propose the multivariate community Hawkes (MULCH) model, an extremely flexible community-based model for continuous-time networks that introduces dependence between node pairs using structured multivariate Hawkes processes.
- 891, TITLE: Disentangling Sources of Risk for Distributional Multi-Agent Reinforcement Learning
<https://proceedings.mlr.press/v162/son22a.html>
AUTHORS: Kyunghwan Son, Junsu Kim, Sungsoo Ahn, Roben D Delos Reyes, Yung Yi, Jinwoo Shin
HIGHLIGHT: To this end, we propose Disentangled RLsk-sensitive Multi-Agent reinforcement learning (DRIMA) to separately access the risk sources.
- 892, TITLE: TAM: Topology-Aware Margin Loss for Class-Imbalanced Node Classification
<https://proceedings.mlr.press/v162/song22a.html>
AUTHORS: Jaeyun Song, Joonhyung Park, Eunho Yang
HIGHLIGHT: In addition, in order to handle this issue, we propose Topology-Aware Margin (TAM) to reflect local topology on the learning objective.
- 893, TITLE: A General Recipe for Likelihood-free Bayesian Optimization
<https://proceedings.mlr.press/v162/song22b.html>
AUTHORS: Jiaming Song, Lantao Yu, Willie Neiswanger, Stefano Ermon
HIGHLIGHT: To extend BO to a broader class of models and utilities, we propose likelihood-free BO (LFBO), an approach based on likelihood-free inference.
- 894, TITLE: Fully-Connected Network on Noncompact Symmetric Space and Ridgelet Transform based on Helgason-Fourier Analysis
<https://proceedings.mlr.press/v162/sonoda22a.html>
AUTHORS: Sho Sonoda, Isao Ishikawa, Masahiro Ikeda
HIGHLIGHT: Based on the well-established framework of the Helgason-Fourier transform on the noncompact symmetric space, we present a fully-connected network and its associated ridgelet transform on the noncompact symmetric space, covering the hyperbolic neural network (HNN) and the SPDNet as special cases.
- 895, TITLE: Saute RL: Almost Surely Safe Reinforcement Learning Using State Augmentation
<https://proceedings.mlr.press/v162/sootla22a.html>
AUTHORS: Aivar Sootla, Alexander I Cowen-Rivers, Taher Jafferjee, Ziyang Wang, David H Mguni, Jun Wang, Haitham Ammar
HIGHLIGHT: For example, plane landing and take-off should ideally occur with probability one. We address the problem by introducing Safety Augmented (Saute) Markov Decision Processes (MDPs), where the safety constraints are eliminated by augmenting them into the state-space and reshaping the objective.
- 896, TITLE: Lightweight Projective Derivative Codes for Compressed Asynchronous Gradient Descent
<https://proceedings.mlr.press/v162/soto22a.html>
AUTHORS: Pedro J Soto, Ilia Ilmer, Haibin Guan, Jun Li
HIGHLIGHT: This paper proposes a novel algorithm that encodes the partial derivatives themselves and furthermore optimizes the codes by performing lossy compression on the derivative codewords by maximizing the information contained in the codewords while minimizing the information between the codewords.
- 897, TITLE: Accelerating Bayesian Optimization for Biological Sequence Design with Denoising Autoencoders
<https://proceedings.mlr.press/v162/stanton22a.html>
AUTHORS: Samuel Stanton, Wesley Maddox, Nate Gruver, Phillip Maffettone, Emily Delaney, Peyton Greenside, Andrew Gordon Wilson
HIGHLIGHT: We develop a new approach (LaMBO) which jointly trains a denoising autoencoder with a discriminative multi-task Gaussian process head, allowing gradient-based optimization of multi-objective acquisition functions in the latent space of the autoencoder.
- 898, TITLE: 3D Infomax improves GNNs for Molecular Property Prediction
<https://proceedings.mlr.press/v162/stark22a.html>
AUTHORS: Hannes Stark, Dominique Beaini, Gabriele Corso, Prudencio Tossou, Christian Dallago, Stephan G?nnemann, Pietro Li?
HIGHLIGHT: Although the 3D molecular graph structure is necessary for models to achieve strong performance on many tasks, it is infeasible to obtain 3D structures at the scale required by many real-world applications. To tackle this issue, we propose to use existing 3D molecular datasets to pre-train a model to reason about the geometry of molecules given only their 2D molecular graphs.

- 899, TITLE: EquiBind: Geometric Deep Learning for Drug Binding Structure Prediction
<https://proceedings.mlr.press/v162/stark22b.html>
AUTHORS: Hannes Stark, Octavian Ganeva, Lagnajit Pattanaik, Dr.Regina Barzilay, Tommi Jaakkola
HIGHLIGHT: We challenge this paradigm with EquiBind, an SE(3)-equivariant geometric deep learning model performing direct-shot prediction of both i) the receptor binding location (blind docking) and ii) the ligand's bound pose and orientation.
- 900, TITLE: Plug & Play Attacks: Towards Robust and Flexible Model Inversion Attacks
<https://proceedings.mlr.press/v162/struppek22a.html>
AUTHORS: Lukas Struppek, Dominik Hintersdorf, Antonio De Almeida Correia, Antonia Adler, Kristian Kersting
HIGHLIGHT: To overcome these drawbacks, we present Plug & Play Attacks, which relax the dependency between the target model and image prior, and enable the use of a single GAN to attack a wide range of targets, requiring only minor adjustments to the attack.
- 901, TITLE: Scaling-up Diverse Orthogonal Convolutional Networks by a Paraunitary Framework
<https://proceedings.mlr.press/v162/su22a.html>
AUTHORS: Jiahao Su, Wonmin Byeon, Furong Huang
HIGHLIGHT: In this work, we propose a theoretical framework that establishes the equivalence between diverse orthogonal convolutional layers in the spatial domain and the paraunitary systems in the spectral domain.
- 902, TITLE: Divergence-Regularized Multi-Agent Actor-Critic
<https://proceedings.mlr.press/v162/su22b.html>
AUTHORS: Kefan Su, Zongqing Lu
HIGHLIGHT: In this paper, we investigate divergence regularization in cooperative MARL and propose a novel off-policy cooperative MARL framework, divergence-regularized multi-agent actor-critic (DMAC).
- 903, TITLE: Influence-Augmented Local Simulators: a Scalable Solution for Fast Deep RL in Large Networked Systems
<https://proceedings.mlr.press/v162/suau22a.html>
AUTHORS: Miguel Suau, Jinke He, Matthijs T. J. Spaan, Frans Oliehoek
HIGHLIGHT: In this paper, we study how to build lightweight simulators of complicated systems that can run sufficiently fast for deep RL to be applicable.
- 904, TITLE: Improved StyleGAN-v2 based Inversion for Out-of-Distribution Images
<https://proceedings.mlr.press/v162/subramanyam22a.html>
AUTHORS: Rakshith Subramanyam, Vivek Narayanaswamy, Mark Naufel, Andreas Spanias, Jayaraman J. Thiagarajan
HIGHLIGHT: In this paper, we propose SPHInX (StyleGAN with Projection Heads for Inverting X), an approach for accurately embedding OOD images onto the StyleGAN latent space.
- 905, TITLE: Continuous-Time Analysis of Accelerated Gradient Methods via Conservation Laws in Dilated Coordinate Systems
<https://proceedings.mlr.press/v162/suh22a.html>
AUTHORS: Jaewook J Suh, Gyumin Roh, Ernest K Ryu
HIGHLIGHT: We analyze continuous-time models of accelerated gradient methods through deriving conservation laws in dilated coordinate systems.
- 906, TITLE: Do Differentiable Simulators Give Better Policy Gradients?
<https://proceedings.mlr.press/v162/suh22b.html>
AUTHORS: Hyung Ju Suh, Max Simchowitz, Kaiqing Zhang, Russ Tedrake
HIGHLIGHT: We show that characteristics of certain physical systems, such as stiffness or discontinuities, may compromise the efficacy of the first-order estimator, and analyze this phenomenon through the lens of bias and variance.
- 907, TITLE: Intriguing Properties of Input-Dependent Randomized Smoothing
<https://proceedings.mlr.press/v162/sukeni-k22a.html>
AUTHORS: Peter Šukenič, Aleksei Kuvshinov, Stephan G. Schömann
HIGHLIGHT: We present one concrete design of the smoothing variance function and test it on CIFAR10 and MNIST.
- 908, TITLE: Cliff Diving: Exploring Reward Surfaces in Reinforcement Learning Environments
<https://proceedings.mlr.press/v162/sullivan22a.html>
AUTHORS: Ryan Sullivan, Jordan K Terry, Benjamin Black, John P Dickerson
HIGHLIGHT: This work presents reward surfaces and related visualizations of 27 of the most widely used reinforcement learning environments in Gym for the first time.
- 909, TITLE: AGNAS: Attention-Guided Micro and Macro-Architecture Search
<https://proceedings.mlr.press/v162/sun22a.html>
AUTHORS: Zihao Sun, Yu Hu, Shun Lu, Longxing Yang, Jilin Mei, Yinhe Han, Xiaowei Li
HIGHLIGHT: To address the two issues, we propose a new search paradigm, that is, leverage the attention mechanism to guide the micro- and macro-architecture search, namely AGNAS.
- 910, TITLE: Adaptive Random Walk Gradient Descent for Decentralized Optimization
<https://proceedings.mlr.press/v162/sun22b.html>

- AUTHORS: Tao Sun, Dongsheng Li, Bao Wang
HIGHLIGHT: In this paper, we study the adaptive step size random walk gradient descent with momentum for decentralized optimization, in which the training samples are drawn dependently with each other.
- 911, TITLE: MAE-DET: Revisiting Maximum Entropy Principle in Zero-Shot NAS for Efficient Object Detection
<https://proceedings.mlr.press/v162/sun22c.html>
AUTHORS: Zhenhong Sun, Ming Lin, Xiuyu Sun, Zhiyu Tan, Hao Li, Rong Jin
HIGHLIGHT: However, existing NAS methods for object detection require hundreds to thousands of GPU hours of searching, making them impractical in fast-paced research and development. In this work, we propose a novel zero-shot NAS method to address this issue.
- 912, TITLE: Out-of-Distribution Detection with Deep Nearest Neighbors
<https://proceedings.mlr.press/v162/sun22d.html>
AUTHORS: Yiyu Sun, Yifei Ming, Xiaojin Zhu, Yixuan Li
HIGHLIGHT: In this paper, we explore the efficacy of non-parametric nearest-neighbor distance for OOD detection, which has been largely overlooked in the literature.
- 913, TITLE: Black-Box Tuning for Language-Model-as-a-Service
<https://proceedings.mlr.press/v162/sun22e.html>
AUTHORS: Tianxiang Sun, Yunfan Shao, Hong Qian, Xuanjing Huang, Xipeng Qiu
HIGHLIGHT: This paper proposes the black-box tuning framework to optimize the continuous prompt prepended to the input text via derivative-free optimization.
- 914, TITLE: Correlated Quantization for Distributed Mean Estimation and Optimization
<https://proceedings.mlr.press/v162/suresh22a.html>
AUTHORS: Ananda Theertha Suresh, Ziteng Sun, Jae Ro, Felix Yu
HIGHLIGHT: We propose a correlated quantization protocol whose error guarantee depends on the deviation of data points instead of their absolute range.
- 915, TITLE: Causal Imitation Learning under Temporally Correlated Noise
<https://proceedings.mlr.press/v162/swamy22a.html>
AUTHORS: Gokul Swamy, Sanjiban Choudhury, Drew Bagnell, Steven Wu
HIGHLIGHT: In particular, we present two techniques, one of a generative-modeling flavor (DoubIL) that can utilize access to a simulator, and one of a game-theoretic flavor (ResiduLL) that can be run entirely offline.
- 916, TITLE: Being Properly Improper
<https://proceedings.mlr.press/v162/sypherd22a.html>
AUTHORS: Tyler Sypherd, Richard Nock, Lalitha Sankar
HIGHLIGHT: Our chief theoretical contribution is a generalization of the properness framework with a notion called twist-properness, which delineates loss functions with the ability to “untwist”; the twisted posterior into the clean posterior.
- 917, TITLE: Distributionally-Aware Kernelized Bandit Problems for Risk Aversion
<https://proceedings.mlr.press/v162/takemori22a.html>
AUTHORS: Sho Takemori
HIGHLIGHT: To address the issues, in this paper, we model environments using a family of the output distributions (or more precisely, probability kernel) and Kernel Mean Embeddings (KME), and provide novel UCB-type algorithms for CVaR and MV.
- 918, TITLE: Sequential and Parallel Constrained Max-value Entropy Search via Information Lower Bound
<https://proceedings.mlr.press/v162/takeno22a.html>
AUTHORS: Shion Takeno, Tomoyuki Tamura, Kazuki Shitara, Masayuki Karasuyama
HIGHLIGHT: In this paper, we propose a novel variant of MES for constrained problems, called Constrained MES via Information lower BOund (CMES-IBO), that is based on a Monte Carlo (MC) estimator of a lower bound of a mutual information (MI).
- 919, TITLE: SQ-VAE: Variational Bayes on Discrete Representation with Self-annealed Stochastic Quantization
<https://proceedings.mlr.press/v162/takida22a.html>
AUTHORS: Yuhta Takida, Takashi Shibuya, Weihsiang Liao, Chieh-Hsin Lai, Junki Ohmura, Toshimitsu Uesaka, Naoki Murata, Shusuke Takahashi, Toshiyuki Kumakura, Yuki Mitsufuji
HIGHLIGHT: In this paper, we propose a new training scheme that extends the standard VAE via novel stochastic dequantization and quantization, called stochastically quantized variational autoencoder (SQ-VAE).
- 920, TITLE: A Tree-based Model Averaging Approach for Personalized Treatment Effect Estimation from Heterogeneous Data Sources
<https://proceedings.mlr.press/v162/tan22a.html>
AUTHORS: Xiaoqing Tan, Chung-Chou H. Chang, Ling Zhou, Lu Tang
HIGHLIGHT: We propose a tree-based model averaging approach to improve the estimation accuracy of conditional average treatment effects (CATE) at a target site by leveraging models derived from other potentially heterogeneous sites, without them sharing subject-level data.
- 921, TITLE: N-Penetrate: Active Learning of Neural Collision Handler for Complex 3D Mesh Deformations

- <https://proceedings.mlr.press/v162/tan22b.html>
AUTHORS: Qingyang Tan, Zherong Pan, Breannan Smith, Takaaki Shiratori, Dinesh Manocha
HIGHLIGHT: We present a robust learning algorithm to detect and handle collisions in 3D deforming meshes.
- 922, TITLE: Biased Gradient Estimate with Drastic Variance Reduction for Meta Reinforcement Learning
<https://proceedings.mlr.press/v162/tang22a.html>
AUTHORS: Yunhao Tang
HIGHLIGHT: Critically, biased gradient estimates are almost always implemented in practice, whereas prior theory on meta-RL only establishes convergence under unbiased gradient estimates. In this work, we investigate such a discrepancy.
- 923, TITLE: Rethinking Graph Neural Networks for Anomaly Detection
<https://proceedings.mlr.press/v162/tang22b.html>
AUTHORS: Jianheng Tang, Jiajin Li, Ziqi Gao, Jia Li
HIGHLIGHT: Our crucial observation is the existence of anomalies will lead to the ‘right-shift’ phenomenon, that is, the spectral energy distribution concentrates less on low frequencies and more on high frequencies.
- 924, TITLE: Deep Safe Incomplete Multi-view Clustering: Theorem and Algorithm
<https://proceedings.mlr.press/v162/tang22c.html>
AUTHORS: Huayi Tang, Yong Liu
HIGHLIGHT: Although jointly imputing incomplete samples and conducting clustering has been shown to achieve promising performance, learning from both complete and incomplete data may be worse than learning only from complete data, particularly when imputed views are semantic inconsistent with missing views. To address this issue, we propose a novel framework to reduce the clustering performance degradation risk from semantic inconsistent imputed views.
- 925, TITLE: Virtual Homogeneity Learning: Defending against Data Heterogeneity in Federated Learning
<https://proceedings.mlr.press/v162/tang22d.html>
AUTHORS: Zhenheng Tang, Yonggang Zhang, Shaohuai Shi, Xin He, Bo Han, Xiaowen Chu
HIGHLIGHT: We propose a different approach named virtual homogeneity learning (VHL) to directly “rectify” the data heterogeneity.
- 926, TITLE: Cross-Space Active Learning on Graph Convolutional Networks
<https://proceedings.mlr.press/v162/tao22a.html>
AUTHORS: Yufei Tao, Hao Wu, Shiyuan Deng
HIGHLIGHT: Our study covers both budget algorithms which terminate after a designated number of label requests, and verifiable algorithms which terminate only after having found an accurate hypothesis.
- 927, TITLE: FedNest: Federated Bilevel, Minimax, and Compositional Optimization
<https://proceedings.mlr.press/v162/tarzanagh22a.html>
AUTHORS: Davoud Ataee Tarzanagh, Mingchen Li, Christos Thrampoulidis, Samet Oymak
HIGHLIGHT: In this work, we propose FedNest: A federated alternating stochastic gradient method to address general nested problems.
- 928, TITLE: Efficient Distributionally Robust Bayesian Optimization with Worst-case Sensitivity
<https://proceedings.mlr.press/v162/tay22a.html>
AUTHORS: Sebastian Shenghong Tay, Chuan Sheng Foo, Urano Daisuke, Richalynn Leong, Bryan Kian Hsiang Low
HIGHLIGHT: We provide a regret bound for our novel DRBO algorithm with the fast approximation, and empirically show it is competitive with that using the exact worst-case expected value while incurring significantly less computation time.
- 929, TITLE: LIDL: Local Intrinsic Dimension Estimation Using Approximate Likelihood
<https://proceedings.mlr.press/v162/tempczyk22a.html>
AUTHORS: Piotr Tempczyk, Rafal Michaluk, Lukasz Garncarek, Przemyslaw Spurek, Jacek Tabor, Adam Golinski
HIGHLIGHT: We attempt to address that challenge by proposing a novel approach to the problem: Local Intrinsic Dimension estimation using approximate Likelihood (LIDL).
- 930, TITLE: LCA-Nets: Lateral Competition Improves Robustness Against Corruption and Attack
<https://proceedings.mlr.press/v162/teti22a.html>
AUTHORS: Michael Teti, Garrett Kenyon, Ben Migliori, Juston Moore
HIGHLIGHT: Motivated by evidence that neural activity in V1 is sparse, we develop a class of hybrid CNNs, called LCA-Nets, which feature a frontend that performs sparse coding via local lateral competition.
- 931, TITLE: Reverse Engineering ℓ_p attacks: A block-sparse optimization approach with recovery guarantees
<https://proceedings.mlr.press/v162/thaker22a.html>
AUTHORS: Darshan Thaker, Paris Giampouras, Rene Vidal
HIGHLIGHT: Specifically, given an attacked signal, we study conditions under which one can determine the type of attack (ℓ_1 , ℓ_2 or ℓ_∞) and recover the clean signal.
- 932, TITLE: Generalised Policy Improvement with Geometric Policy Composition
<https://proceedings.mlr.press/v162/thakoor22a.html>
AUTHORS: Shantanu Thakoor, Mark Rowland, Diana Borsa, Will Dabney, Remi Munos, Andre Barreto

- HIGHLIGHT:** We introduce a method for policy improvement that interpolates between the greedy approach of value-based reinforcement learning (RL) and the full planning approach typical of model-based RL.
- 933, **TITLE:** Algorithms for the Communication of Samples
<https://proceedings.mlr.press/v162/theis22a.html>
AUTHORS: Lucas Theis, Noureldin Y Ahmed
HIGHLIGHT: Here we propose two new coding schemes with practical advantages over existing approaches.
- 934, **TITLE:** Consistent Polyhedral Surrogates for Top-k Classification and Variants
<https://proceedings.mlr.press/v162/thilagar22a.html>
AUTHORS: Anish Thilagar, Rafael Frongillo, Jessica J Finocchiaro, Emma Goodwill
HIGHLIGHT: We leverage this analysis to derive constraints on the conditional label distributions under which these proposed surrogates become consistent for top- k .
- 935, **TITLE:** On the Finite-Time Complexity and Practical Computation of Approximate Stationarity Concepts of Lipschitz Functions
<https://proceedings.mlr.press/v162/tian22a.html>
AUTHORS: Lai Tian, Kaiwen Zhou, Anthony Man-Cho So
HIGHLIGHT: Complement to these developments, in this paper, we isolate a new class of functions that could be Clarke irregular (and thus not weakly convex anymore) and show that our new algorithmic scheme can compute NAS points for functions in that class within finite time.
- 936, **TITLE:** From Dirichlet to Rubin: Optimistic Exploration in RL without Bonuses
<https://proceedings.mlr.press/v162/tiapkin22a.html>
AUTHORS: Daniil Tiapkin, Denis Belomestny, Eric Moulines, Alexey Naumov, Sergey Samsonov, Yunhao Tang, Michal Valko, Pierre Menard
HIGHLIGHT: We propose the Bayes-UCBVI algorithm for reinforcement learning in tabular, stage-dependent, episodic Markov decision process: a natural extension of the Bayes-UCB algorithm by Kaufmann et al. 2012 for multi-armed bandits.
- 937, **TITLE:** Nonparametric Sparse Tensor Factorization with Hierarchical Gamma Processes
<https://proceedings.mlr.press/v162/tillinghast22a.html>
AUTHORS: Conor Tillinghast, Zheng Wang, Shandian Zhe
HIGHLIGHT: We propose a nonparametric factorization approach for sparsely observed tensors.
- 938, **TITLE:** Deciphering Lasso-based Classification Through a Large Dimensional Analysis of the Iterative Soft-Thresholding Algorithm
<https://proceedings.mlr.press/v162/tiomoko22a.html>
AUTHORS: Malik Tiomoko, Ekkehard Schnoor, Mohamed El Amine Seddik, Igor Colin, Aladin Virmaux
HIGHLIGHT: This paper proposes a theoretical analysis of a Lasso-based classification algorithm.
- 939, **TITLE:** Extended Unconstrained Features Model for Exploring Deep Neural Collapse
<https://proceedings.mlr.press/v162/tirer22a.html>
AUTHORS: Tom Tirer, Joan Bruna
HIGHLIGHT: In this paper, we further analyze and extend the UFM.
- 940, **TITLE:** Object Permanence Emerges in a Random Walk along Memory
<https://proceedings.mlr.press/v162/tokmakov22a.html>
AUTHORS: Pavel Tokmakov, Allan Jabri, Jie Li, Adrien Gaidon
HIGHLIGHT: This paper proposes a self-supervised objective for learning representations that localize objects under occlusion - a property known as object permanence.
- 941, **TITLE:** Generic Coreset for Scalable Learning of Monotonic Kernels: Logistic Regression, Sigmoid and more
<https://proceedings.mlr.press/v162/tolochinsky22a.html>
AUTHORS: Elad Tolochinsky, Ibrahim Jubran, Dan Feldman
HIGHLIGHT: In this work we provide: (i) A lower bound which proves that there are sets with no coresets smaller than $\frac{n}{\epsilon} \log \frac{1}{\epsilon}$ for general monotonic loss functions.
- 942, **TITLE:** Failure and success of the spectral bias prediction for Laplace Kernel Ridge Regression: the case of low-dimensional data
<https://proceedings.mlr.press/v162/tomasini22a.html>
AUTHORS: Umberto M Tomasini, Antonio Sclocchi, Matthieu Wyart
HIGHLIGHT: To clarify when the spectral bias prediction holds, we first focus on a one-dimensional model where rigorous results are obtained and then use scaling arguments to generalize and test our findings in higher dimensions.
- 943, **TITLE:** Quantifying and Learning Linear Symmetry-Based Disentanglement
<https://proceedings.mlr.press/v162/tonnaer22a.html>
AUTHORS: Loek Tonnaer, Luis Armando Perez Rey, Vlado Menkovski, Mike Holenderski, Jim Portegies
HIGHLIGHT: We propose D_{LSBD} , a mathematically sound metric to quantify LSBD, and provide a practical implementation for $SO(2)$ groups.

- 944, TITLE: A Temporal-Difference Approach to Policy Gradient Estimation
<https://proceedings.mlr.press/v162/tosatto22a.html>
AUTHORS: Samuele Tosatto, Andrew Patterson, Martha White, Rupam Mahmood
HIGHLIGHT: In this paper, we propose a new approach of reconstructing the policy gradient from the start state without requiring a particular sampling strategy.
- 945, TITLE: Simple and near-optimal algorithms for hidden stratification and multi-group learning
<https://proceedings.mlr.press/v162/tosh22a.html>
AUTHORS: Christopher J Tosh, Daniel Hsu
HIGHLIGHT: This paper studies the structure of solutions to the multi-group learning problem, and provides simple and near-optimal algorithms for the learning problem.
- 946, TITLE: Design-Bench: Benchmarks for Data-Driven Offline Model-Based Optimization
<https://proceedings.mlr.press/v162/trabucco22a.html>
AUTHORS: Brandon Trabucco, Xinyang Geng, Aviral Kumar, Sergey Levine
HIGHLIGHT: However, the lack of standardized benchmarks in this emerging field is making progress difficult to track. To address this, we present Design-Bench, a benchmark for offline MBO with a unified evaluation protocol and reference implementations of recent methods.
- 947, TITLE: AnyMorph: Learning Transferable Policies By Inferring Agent Morphology
<https://proceedings.mlr.press/v162/trabucco22b.html>
AUTHORS: Brandon Trabucco, Mariano Phielipp, Glen Berseth
HIGHLIGHT: This is a challenging problem that required previous approaches to use hand-designed descriptions of the new agent’s morphology. Instead of hand-designing this description, we propose a data-driven method that learns a representation of morphology directly from the reinforcement learning objective.
- 948, TITLE: Detecting Adversarial Examples Is (Nearly) As Hard As Classifying Them
<https://proceedings.mlr.press/v162/tramer22a.html>
AUTHORS: Florian Tramer
HIGHLIGHT: We prove a hardness reduction between detection and classification of adversarial examples: given a robust detector for attacks at distance ϵ (in some metric), we show how to build a similarly robust (but inefficient) classifier for attacks at distance $\epsilon/2$.
- 949, TITLE: Nesterov Accelerated Shuffling Gradient Method for Convex Optimization
<https://proceedings.mlr.press/v162/tran22a.html>
AUTHORS: Trang H Tran, Katya Scheinberg, Lam M Nguyen
HIGHLIGHT: In this paper, we propose Nesterov Accelerated Shuffling Gradient (NASG), a new algorithm for the convex finite-sum minimization problems.
- 950, TITLE: A Completely Tuning-Free and Robust Approach to Sparse Precision Matrix Estimation
<https://proceedings.mlr.press/v162/tran22b.html>
AUTHORS: Chau Tran, Guo Yu
HIGHLIGHT: We propose a completely tuning-free approach for estimating sparse Gaussian graphical models.
- 951, TITLE: Tackling covariate shift with node-based Bayesian neural networks
<https://proceedings.mlr.press/v162/trinh22a.html>
AUTHORS: Trung Q Trinh, Markus Heinonen, Luigi Acerbi, Samuel Kaski
HIGHLIGHT: Node-based BNNs have recently been introduced as scalable alternatives, which induce epistemic uncertainty by multiplying each hidden node with latent random variables, while learning a point-estimate of the weights. In this paper, we interpret these latent noise variables as implicit representations of simple and domain-agnostic data perturbations during training, producing BNNs that perform well under covariate shift due to input corruptions.
- 952, TITLE: Fenrir: Physics-Enhanced Regression for Initial Value Problems
<https://proceedings.mlr.press/v162/tronarp22a.html>
AUTHORS: Filip Tronarp, Nathanael Bosch, Philipp Hennig
HIGHLIGHT: We show how probabilistic numerics can be used to convert an initial value problem into a Gauss–Markov process parametrised by the dynamics of the initial value problem.
- 953, TITLE: Interpretable Off-Policy Learning via Hyperbox Search
<https://proceedings.mlr.press/v162/tschermutter22a.html>
AUTHORS: Daniel Tschermutter, Tobias Hatt, Stefan Feuerriegel
HIGHLIGHT: In this paper, we propose an algorithm for interpretable off-policy learning via hyperbox search.
- 954, TITLE: FriendlyCore: Practical Differentially Private Aggregation
<https://proceedings.mlr.press/v162/tsfadia22a.html>
AUTHORS: Eliad Tsfadia, Edith Cohen, Haim Kaplan, Yishay Mansour, Uri Stemmer
HIGHLIGHT: We propose a simple and practical tool FriendlyCore that takes a set of points \mathcal{D} from an unrestricted (pseudo) metric space as input.
- 955, TITLE: Pairwise Conditional Gradients without Swap Steps and Sparser Kernel Herding

- <https://proceedings.mlr.press/v162/tsuji22a.html>
AUTHORS: Kazuma K Tsuji, Ken?Ichiro Tanaka, Sebastian Pokutta
HIGHLIGHT: We propose a new variant of PCG, the so-called Blended Pairwise Conditional Gradients (BPCG).
- 956, TITLE: Prototype Based Classification from Hierarchy to Fairness
<https://proceedings.mlr.press/v162/tucker22a.html>
AUTHORS: Mycal Tucker, Julie A. Shah
HIGHLIGHT: Our contribution in this work is a new neural network architecture, the concept subspace network (CSN), which generalizes existing specialized classifiers to produce a unified model capable of learning a spectrum of multi-concept relationships.
- 957, TITLE: Consensus Multiplicative Weights Update: Learning to Learn using Projector-based Game Signatures
<https://proceedings.mlr.press/v162/vadori22a.html>
AUTHORS: Nelson Vadori, Rahul Savani, Thomas Spooner, Sumitra Ganesh
HIGHLIGHT: In doing so, we introduce CMWU, a new algorithm that extends consensus optimization to the constrained case, has local convergence guarantees for zero-sum bimatrix games, and show that it enjoys competitive performance on both zero-sum games with constant coefficients and across a spectrum of games when its coefficients are learnt.
- 958, TITLE: Self-Supervised Models of Audio Effectively Explain Human Cortical Responses to Speech
<https://proceedings.mlr.press/v162/vaidya22a.html>
AUTHORS: Aditya R Vaidya, Shailee Jain, Alexander Huth
HIGHLIGHT: In this work, we capitalize on the progress of self-supervised speech representation learning (SSL) to create new state-of-the-art models of the human auditory system.
- 959, TITLE: Path-Gradient Estimators for Continuous Normalizing Flows
<https://proceedings.mlr.press/v162/vaitl22a.html>
AUTHORS: Lorenz Vaitl, Kim Andrea Nicoli, Shinichi Nakajima, Pan Kessel
HIGHLIGHT: In many applications, this regime can however not be reached by a simple Gaussian variational distribution. In this work, we overcome this crucial limitation by proposing a path-gradient estimator for the considerably more expressive variational family of continuous normalizing flows.
- 960, TITLE: Improved Convergence Rates for Sparse Approximation Methods in Kernel-Based Learning
<https://proceedings.mlr.press/v162/vakili22a.html>
AUTHORS: Sattar Vakili, Jonathan Scarlett, Da-Shan Shiu, Alberto Bernacchia
HIGHLIGHT: In this work, we provide novel confidence intervals for the Nyström method and the sparse variational Gaussian process approximation method, which we establish using novel interpretations of the approximate (surrogate) posterior variance of the models.
- 961, TITLE: EDEN: Communication-Efficient and Robust Distributed Mean Estimation for Federated Learning
<https://proceedings.mlr.press/v162/vargaftik22a.html>
AUTHORS: Shay Vargaftik, Ran Ben Basat, Amit Portnoy, Gal Mendelson, Yaniv Ben Itzhak, Michael Mitzenmacher
HIGHLIGHT: In this work, we propose a robust DME technique named EDEN that naturally handles heterogeneous communication budgets and packet losses.
- 962, TITLE: Towards Noise-adaptive, Problem-adaptive (Accelerated) Stochastic Gradient Descent
<https://proceedings.mlr.press/v162/vaswani22a.html>
AUTHORS: Sharan Vaswani, Benjamin Dubois-Taine, Reza Babanezhad
HIGHLIGHT: We aim to make stochastic gradient descent (SGD) adaptive to (i) the noise σ^2 in the stochastic gradients and (ii) problem-dependent constants.
- 963, TITLE: Correlation Clustering via Strong Triadic Closure Labeling: Fast Approximation Algorithms and Practical Lower Bounds
<https://proceedings.mlr.press/v162/veldt22a.html>
AUTHORS: Nate Veldt
HIGHLIGHT: Correlation clustering is a widely studied framework for clustering based on pairwise similarity and dissimilarity scores, but its best approximation algorithms rely on impractical linear programming relaxations. We present faster approximation algorithms that avoid these relaxations, for two well-studied special cases: cluster editing and cluster deletion.
- 964, TITLE: The CLRS Algorithmic Reasoning Benchmark
<https://proceedings.mlr.press/v162/velickovic22a.html>
AUTHORS: Petar Velickovic, Aditya Puigdomenech Badia, David Budden, Razvan Pascanu, Andrea Banino, Misha Dashevskiy, Raia Hadsell, Charles Blundell
HIGHLIGHT: To consolidate progress and work towards unified evaluation, we propose the CLRS Algorithmic Reasoning Benchmark, covering classical algorithms from the Introduction to Algorithms textbook.
- 965, TITLE: Bregman Power k-Means for Clustering Exponential Family Data
<https://proceedings.mlr.press/v162/vellal22a.html>
AUTHORS: Adithya Vellal, Saptarshi Chakraborty, Jason Q Xu
HIGHLIGHT: In this paper, we bridge these new algorithmic advances to classical work on hard clustering under Bregman divergences, which enjoy a bijection to exponential family distributions and are thus well-suited for clustering objects arising from a breadth of data generating mechanisms.

- 966, TITLE: Estimation in Rotationally Invariant Generalized Linear Models via Approximate Message Passing
<https://proceedings.mlr.press/v162/venkataramanan22a.html>
AUTHORS: Ramji Venkataramanan, Kevin K[?]gler, Marco Mondelli
HIGHLIGHT: We propose a novel family of approximate message passing (AMP) algorithms for signal estimation, and rigorously characterize their performance in the high-dimensional limit via a state evolution recursion.
- 967, TITLE: Bayesian Optimization under Stochastic Delayed Feedback
<https://proceedings.mlr.press/v162/verma22a.html>
AUTHORS: Arun Verma, Zhongxiang Dai, Bryan Kian Hsiang Low
HIGHLIGHT: In this paper, we consider the BO under stochastic delayed feedback problem.
- 968, TITLE: VarScene: A Deep Generative Model for Realistic Scene Graph Synthesis
<https://proceedings.mlr.press/v162/verma22b.html>
AUTHORS: Tathagat Verma, Abir De, Yateesh Agrawal, Vishwa Vinay, Soumen Chakrabarti
HIGHLIGHT: In contrast, scene graphs have much larger object and relation vocabularies, and their semantics are latent. To address this challenge, we propose a variational autoencoder for scene graphs, which is optimized for the maximum mean discrepancy (MMD) between the ground truth scene graph distribution and distribution of the generated scene graphs.
- 969, TITLE: Calibrated Learning to Defer with One-vs-All Classifiers
<https://proceedings.mlr.press/v162/verma22c.html>
AUTHORS: Rajeev Verma, Eric Nalisnick
HIGHLIGHT: We propose an L2D system based on one-vs-all classifiers that is able to produce calibrated probabilities of expert correctness.
- 970, TITLE: Regret Bounds for Stochastic Shortest Path Problems with Linear Function Approximation
<https://proceedings.mlr.press/v162/vial22a.html>
AUTHORS: Daniel Vial, Advait Parulekar, Sanjay Shakkottai, R Srikanth
HIGHLIGHT: We propose an algorithm that uses linear function approximation (LFA) for stochastic shortest path (SSP).
- 971, TITLE: On Implicit Bias in Overparameterized Bilevel Optimization
<https://proceedings.mlr.press/v162/vicol22a.html>
AUTHORS: Paul Vicol, Jonathan P Lorraine, Fabian Pedregosa, David Duvenaud, Roger B Grosse
HIGHLIGHT: Inspired by recent studies of the implicit bias induced by optimization algorithms in single-level optimization, we investigate the implicit bias of different gradient-based algorithms for jointly optimizing the inner and outer parameters.
- 972, TITLE: Multiclass learning with margin: exponential rates with no bias-variance trade-off
<https://proceedings.mlr.press/v162/vigogna22a.html>
AUTHORS: Stefano Vigogna, Giacomo Meanti, Ernesto De Vito, Lorenzo Rosasco
HIGHLIGHT: We study the behavior of error bounds for multiclass classification under suitable margin conditions.
- 973, TITLE: Addressing Optimism Bias in Sequence Modeling for Reinforcement Learning
<https://proceedings.mlr.press/v162/villaflor22a.html>
AUTHORS: Adam R Villaflor, Zhe Huang, Swapnil Pande, John M Dolan, Jeff Schneider
HIGHLIGHT: Thus, in adversarial or stochastic environments, these methods lead to overly optimistic behavior that can be dangerous in safety-critical systems like autonomous driving. In this work, we propose a method that addresses this optimism bias by explicitly disentangling the policy and world models, which allows us at test time to search for policies that are robust to multiple possible futures in the environment.
- 974, TITLE: Bayesian Nonparametrics for Offline Skill Discovery
<https://proceedings.mlr.press/v162/villecroze22a.html>
AUTHORS: Valentin Villecroze, Harry Braviner, Panteha Naderian, Chris Maddison, Gabriel Loaiza-Ganem
HIGHLIGHT: We first propose a method for offline learning of options (a particular skill framework) exploiting advances in variational inference and continuous relaxations. We then highlight an unexplored connection between Bayesian nonparametrics and offline skill discovery, and show how to obtain a nonparametric version of our model.
- 975, TITLE: Hermite Polynomial Features for Private Data Generation
<https://proceedings.mlr.press/v162/vinaroz22a.html>
AUTHORS: Margarita Vinaroz, Mohammad-Amin Charusaie, Frederik Harder, Kamil Adamczewski, Mi Jung Park
HIGHLIGHT: To improve the sensitivity, we propose to replace random features with Hermite polynomial features.
- 976, TITLE: What Can Linear Interpolation of Neural Network Loss Landscapes Tell Us?
<https://proceedings.mlr.press/v162/vlaar22a.html>
AUTHORS: Tiffany J Vlaar, Jonathan Frankle
HIGHLIGHT: In this paper, we put inferences of this kind to the test, systematically evaluating how linear interpolation and final performance vary when altering the data, choice of initialization, and other optimizer and architecture design choices.
- 977, TITLE: Multirate Training of Neural Networks
<https://proceedings.mlr.press/v162/vlaar22b.html>
AUTHORS: Tiffany J Vlaar, Benedict Leimkuhler

- HIGHLIGHT:** We propose multirate training of neural networks: partitioning neural network parameters into “fast” and “slow” parts which are trained on different time scales, where slow parts are updated less frequently.
- 978, **TITLE:** Provably Adversarially Robust Nearest Prototype Classifiers
<https://proceedings.mlr.press/v162/voracek22a.html>
AUTHORS: V?clav Vor?cek, Matthias Hein
HIGHLIGHT: In this paper we provide a complete discussion on the complexity when using ℓ_p -distances for decision and ℓ_q -threat models for certification for $p, q \in \{1, 2, \infty\}$.
- 979, **TITLE:** First-Order Regret in Reinforcement Learning with Linear Function Approximation: A Robust Estimation Approach
<https://proceedings.mlr.press/v162/wagenmaker22a.html>
AUTHORS: Andrew J Wagenmaker, Yifang Chen, Max Simchowitz, Simon Du, Kevin Jamieson
HIGHLIGHT: While such bounds exist in many settings, they have proven elusive in reinforcement learning with large state spaces. In this work we address this gap, and show that it is possible to obtain regret scaling as $\widetilde{O}(\sqrt{d^3 H^3 \cdot V_1^* \cdot K} + d^{3.5} H^3 \log K)$ in reinforcement learning with large state spaces, namely the linear MDP setting.
- 980, **TITLE:** Reward-Free RL is No Harder Than Reward-Aware RL in Linear Markov Decision Processes
<https://proceedings.mlr.press/v162/wagenmaker22b.html>
AUTHORS: Andrew J Wagenmaker, Yifang Chen, Max Simchowitz, Simon Du, Kevin Jamieson
HIGHLIGHT: To our knowledge, our approach is the first computationally efficient algorithm to achieve optimal d dependence in linear MDPs, even in the single-reward PAC setting.
- 981, **TITLE:** Training Characteristic Functions with Reinforcement Learning: XAI-methods play Connect Four
<https://proceedings.mlr.press/v162/waldchen22a.html>
AUTHORS: Stephan Waldchen, Sebastian Pokutta, Felix Huber
HIGHLIGHT: We propose a setup to directly train characteristic functions in the form of neural networks to play simple two-player games.
- 982, **TITLE:** Retroformer: Pushing the Limits of End-to-end Retrosynthesis Transformer
<https://proceedings.mlr.press/v162/wan22a.html>
AUTHORS: Yue Wan, Chang-Yu Hsieh, Ben Liao, Shengyu Zhang
HIGHLIGHT: In this paper, we propose Retroformer, a novel Transformer-based architecture for retrosynthesis prediction without relying on any cheminformatics tools for molecule editing.
- 983, **TITLE:** Safe Exploration for Efficient Policy Evaluation and Comparison
<https://proceedings.mlr.press/v162/wan22b.html>
AUTHORS: Runzhe Wan, Branislav Kveton, Rui Song
HIGHLIGHT: This paper initiates the study of efficient and safe data collection for bandit policy evaluation. We formulate the problem and investigate its several representative variants.
- 984, **TITLE:** Greedy based Value Representation for Optimal Coordination in Multi-agent Reinforcement Learning
<https://proceedings.mlr.press/v162/wan22c.html>
AUTHORS: Lipeng Wan, Zeyang Liu, Xingyu Chen, Xuguang Lan, Nanning Zheng
HIGHLIGHT: In this paper, we derive the expression of the joint Q value function of LVD and MVD.
- 985, **TITLE:** Towards Evaluating Adaptivity of Model-Based Reinforcement Learning Methods
<https://proceedings.mlr.press/v162/wan22d.html>
AUTHORS: Yi Wan, Ali Rahimi-Kalahroudi, Janarthanan Rajendran, Ida Momennejad, Sarath Chandar, Harm H Van Seijen
HIGHLIGHT: Combined with prior work that made a similar observation about the other popular model-based method, MuZero, a trend appears to emerge, suggesting that current deep model-based methods have serious limitations. We dive deeper into the causes of this poor performance, by identifying elements that hurt adaptive behavior and linking these to underlying techniques frequently used in deep model-based RL.
- 986, **TITLE:** Fast Lossless Neural Compression with Integer-Only Discrete Flows
<https://proceedings.mlr.press/v162/wang22a.html>
AUTHORS: Siyu Wang, Jianfei Chen, Chongxuan Li, Jun Zhu, Bo Zhang
HIGHLIGHT: In this work, we propose Integer-Only Discrete Flows (IODF) an efficient neural compressor with integer-only arithmetic.
- 987, **TITLE:** Accelerating Shapley Explanation via Contributive Cooperator Selection
<https://proceedings.mlr.press/v162/wang22b.html>
AUTHORS: Guanchu Wang, Yu-Neng Chuang, Mengnan Du, Fan Yang, Quan Zhou, Pushkar Tripathi, Xuanting Cai, Xia Hu
HIGHLIGHT: Even though Shapley value provides an effective explanation for a DNN model prediction, the computation relies on the enumeration of all possible input feature coalitions, which leads to the exponentially growing complexity. To address this problem, we propose a novel method SHEAR to significantly accelerate the Shapley explanation for DNN models, where only a few coalitions of input features are involved in the computation.

- 988, TITLE: Denoised MDPs: Learning World Models Better Than the World Itself
<https://proceedings.mlr.press/v162/wang22c.html>
AUTHORS: Tongzhou Wang, Simon Du, Antonio Torralba, Phillip Isola, Amy Zhang, Yuandong Tian
HIGHLIGHT: In this work, we categorize information out in the wild into four types based on controllability and relation with reward, and formulate useful information as that which is both controllable and reward-relevant.
- 989, TITLE: Neural Implicit Dictionary Learning via Mixture-of-Expert Training
<https://proceedings.mlr.press/v162/wang22d.html>
AUTHORS: Peihao Wang, Zhiwen Fan, Tianlong Chen, Zhangyang Wang
HIGHLIGHT: In this paper, we present a generic INR framework that achieves both data and training efficiency by learning a Neural Implicit Dictionary (NID) from a data collection and representing INR as a functional combination of wavelets sampled from the dictionary.
- 990, TITLE: Robust Models Are More Interpretable Because Attributions Look Normal
<https://proceedings.mlr.press/v162/wang22e.html>
AUTHORS: Zifan Wang, Matt Fredrikson, Anupam Datta
HIGHLIGHT: We show that smooth decision boundaries play an important role in this enhanced interpretability, as the model's input gradients around data points will more closely align with boundaries' normal vectors when they are smooth.
- 991, TITLE: Disentangling Disease-related Representation from Obscure for Disease Prediction
<https://proceedings.mlr.press/v162/wang22f.html>
AUTHORS: Chu-Ran Wang, Fei Gao, Fandong Zhang, Fangwei Zhong, Yizhou Yu, Yizhou Wang
HIGHLIGHT: In this paper, to learn the representations for identifying obscured lesions, we propose a disentanglement learning strategy under the guidance of alpha blending generation in an encoder-decoder framework (DAB-Net).
- 992, TITLE: Solving Stackelberg Prediction Game with Least Squares Loss via Spherically Constrained Least Squares Reformulation
<https://proceedings.mlr.press/v162/wang22g.html>
AUTHORS: Jiali Wang, Wen Huang, Rujun Jiang, Xudong Li, Alex L Wang
HIGHLIGHT: In this paper, we explore an alternative reformulation of the SPG-LS.
- 993, TITLE: VLMixer: Unpaired Vision-Language Pre-training via Cross-Modal CutMix
<https://proceedings.mlr.press/v162/wang22h.html>
AUTHORS: Teng Wang, Wenhao Jiang, Zhichao Lu, Feng Zheng, Ran Cheng, Chengguo Yin, Ping Luo
HIGHLIGHT: This paper proposes a data augmentation method, namely cross-modal CutMix (CMC), for implicit cross-modal alignment learning in unpaired VLP.
- 994, TITLE: DynaMixer: A Vision MLP Architecture with Dynamic Mixing
<https://proceedings.mlr.press/v162/wang22i.html>
AUTHORS: Ziyu Wang, Wenhao Jiang, Yiming M Zhu, Li Yuan, Yibing Song, Wei Liu
HIGHLIGHT: To this end, this paper presents an efficient MLP-like network architecture, dubbed DynaMixer, resorting to dynamic information fusion.
- 995, TITLE: Improving Screening Processes via Calibrated Subset Selection
<https://proceedings.mlr.press/v162/wang22j.html>
AUTHORS: Lequn Wang, Thorsten Joachims, Manuel Gomez Rodriguez
HIGHLIGHT: In this paper, we investigate what guarantees a screening classifier can provide, independently of whether it is constructed manually or trained.
- 996, TITLE: The Geometry of Robust Value Functions
<https://proceedings.mlr.press/v162/wang22k.html>
AUTHORS: Kaixin Wang, Navdeep Kumar, Kuangqi Zhou, Bryan Hooi, Jiashi Feng, Shie Mannor
HIGHLIGHT: In this paper, we study the geometry of the robust value space for the more general Robust MDPs (RMDPs) setting, where transition uncertainties are considered.
- 997, TITLE: What Dense Graph Do You Need for Self-Attention?
<https://proceedings.mlr.press/v162/wang22l.html>
AUTHORS: Yuxin Wang, Chu-Tak Lee, Qipeng Guo, Zhangyue Yin, Yunhua Zhou, Xuanjing Huang, Xipeng Qiu
HIGHLIGHT: In this paper, we propose Normalized Information Payload (NIP), a graph scoring function measuring information transfer on graph, which provides an analysis tool for trade-offs between performance and complexity.
- 998, TITLE: Improved Certified Defenses against Data Poisoning with (Deterministic) Finite Aggregation
<https://proceedings.mlr.press/v162/wang22m.html>
AUTHORS: Wenxiao Wang, Alexander J Levine, Soheil Feizi
HIGHLIGHT: In this work, we propose an improved certified defense against general poisoning attacks, namely Finite Aggregation.
- 999, TITLE: Understanding Gradual Domain Adaptation: Improved Analysis, Optimal Path and Beyond
<https://proceedings.mlr.press/v162/wang22n.html>

- AUTHORS: Haoxiang Wang, Bo Li, Han Zhao
HIGHLIGHT: In this work, we analyze gradual self-training under more general and relaxed assumptions, and prove a significantly improved generalization bound as $\tilde{O}(\epsilon_0 + T\Delta + T\sqrt{n} + 1/\sqrt{nT})$, where Δ is the average distributional distance between consecutive domains.
- 1000, TITLE: Communication-Efficient Adaptive Federated Learning
<https://proceedings.mlr.press/v162/wang22o.html>
AUTHORS: Yujia Wang, Lu Lin, Jinghui Chen
HIGHLIGHT: In this paper, we propose a novel communication-efficient adaptive federated learning method (FedCAMS) with theoretical convergence guarantees.
- 1001, TITLE: Provable Acceleration of Heavy Ball beyond Quadratics for a Class of Polyak-Lojasiewicz Functions when the Non-Convexity is Averaged-Out
<https://proceedings.mlr.press/v162/wang22p.html>
AUTHORS: Jun-Kun Wang, Chi-Heng Lin, Andre Wibisono, Bin Hu
HIGHLIGHT: In this work, we develop some new techniques that help show acceleration beyond quadratics, which is achieved by analyzing how the change of the Hessian at two consecutive time points affects the convergence speed.
- 1002, TITLE: Robustness Verification for Contrastive Learning
<https://proceedings.mlr.press/v162/wang22q.html>
AUTHORS: Zekai Wang, Weiwei Liu
HIGHLIGHT: However, the robustness metric used in these methods is linked to attack algorithms, image labels and downstream tasks, all of which may affect the consistency and reliability of robustness metric for CL. To address these problems, this paper proposes a novel Robustness Verification framework for Contrastive Learning (RVCL).
- 1003, TITLE: Convergence and Recovery Guarantees of the K-Subspaces Method for Subspace Clustering
<https://proceedings.mlr.press/v162/wang22r.html>
AUTHORS: Peng Wang, Huikang Liu, Anthony Man-Cho So, Laura Balzano
HIGHLIGHT: In this work, we present local convergence analysis and a recovery guarantee for KSS, assuming data are generated by the semi-random union of subspaces model, where N points are randomly sampled from $K \geq 2$ overlapping subspaces.
- 1004, TITLE: NP-Match: When Neural Processes meet Semi-Supervised Learning
<https://proceedings.mlr.press/v162/wang22s.html>
AUTHORS: Jianfeng Wang, Thomas Lukasiewicz, Daniela Massiceti, Xiaolin Hu, Vladimir Pavlovic, Alexandros Neophytou
HIGHLIGHT: In this work, we adjust neural processes (NPs) to the semi-supervised image classification task, resulting in a new method named NP-Match.
- 1005, TITLE: Iterative Double Sketching for Faster Least-Squares Optimization
<https://proceedings.mlr.press/v162/wang22t.html>
AUTHORS: Rui Wang, Yanyan Ouyang, Wangli Xu
HIGHLIGHT: We generalize the iterative Hessian sketching (IHS) algorithm and propose a new sketching framework named iterative double sketching (IDS) which uses approximations for both the gradient and the Hessian in each iteration.
- 1006, TITLE: What Language Model Architecture and Pretraining Objective Works Best for Zero-Shot Generalization?
<https://proceedings.mlr.press/v162/wang22u.html>
AUTHORS: Thomas Wang, Adam Roberts, Daniel Hesslow, Teven Le Scao, Hyung Won Chung, Iz Beltagy, Julien Launay, Colin Raffel
HIGHLIGHT: In this work, we present a large-scale evaluation of modeling choices and their impact on zero-shot generalization.
- 1007, TITLE: Improving Task-free Continual Learning by Distributionally Robust Memory Evolution
<https://proceedings.mlr.press/v162/wang22v.html>
AUTHORS: Zhenyi Wang, Li Shen, Le Fang, Qiuling Suo, Tiehang Duan, Mingchen Gao
HIGHLIGHT: To address these problems, for the first time, we propose a principled memory evolution framework to dynamically evolve the memory data distribution by making the memory buffer gradually harder to be memorized with distributionally robust optimization (DRO).
- 1008, TITLE: Risk-Averse No-Regret Learning in Online Convex Games
<https://proceedings.mlr.press/v162/wang22w.html>
AUTHORS: Zifan Wang, Yi Shen, Michael Zavlanos
HIGHLIGHT: To address this challenge, we propose a new online risk-averse learning algorithm that relies on one-point zeroth-order estimation of the CVaR gradients computed using CVaR values that are estimated by appropriately sampling the cost functions.
- 1009, TITLE: Provable Domain Generalization via Invariant-Feature Subspace Recovery
<https://proceedings.mlr.press/v162/wang22x.html>
AUTHORS: Haoxiang Wang, Haozhe Si, Bo Li, Han Zhao
HIGHLIGHT: In this paper, we propose to achieve domain generalization with Invariant-feature Subspace Recovery (ISR).

- 1010, TITLE: ProgFed: Effective, Communication, and Computation Efficient Federated Learning by Progressive Training
<https://proceedings.mlr.press/v162/wang22y.html>
AUTHORS: Hui-Po Wang, Sebastian Stich, Yang He, Mario Fritz
HIGHLIGHT: In contrast, we propose ProgFed, the first progressive training framework for efficient and effective federated learning.
- 1011, TITLE: Model-based Meta Reinforcement Learning using Graph Structured Surrogate Models and Amortized Policy Search
<https://proceedings.mlr.press/v162/wang22z.html>
AUTHORS: Qi Wang, Herke Van Hoof
HIGHLIGHT: In this paper, we take a closer look at this framework and propose a new posterior sampling based approach that consists of a new model to identify task dynamics together with an amortized policy optimization step.
- 1012, TITLE: Approximately Equivariant Networks for Imperfectly Symmetric Dynamics
<https://proceedings.mlr.press/v162/wang22aa.html>
AUTHORS: Rui Wang, Robin Walters, Rose Yu
HIGHLIGHT: We explore approximately equivariant networks which are biased towards preserving symmetry but are not strictly constrained to do so.
- 1013, TITLE: Three-stage Evolution and Fast Equilibrium for SGD with Non-degenerate Critical Points
<https://proceedings.mlr.press/v162/wang22ab.html>
AUTHORS: Yi Wang, Zhiren Wang
HIGHLIGHT: We justify the fast equilibrium conjecture on stochastic gradient descent from (Li et al. 2020) under the assumptions that critical points are non-degenerate and the stochastic noise is a standard Gaussian.
- 1014, TITLE: Understanding Instance-Level Impact of Fairness Constraints
<https://proceedings.mlr.press/v162/wang22ac.html>
AUTHORS: Jialu Wang, Xin Eric Wang, Yang Liu
HIGHLIGHT: Building on the concept of influence function, a measure that characterizes the impact of a training example on the target model and its predictive performance, this work studies the influence of training examples when fairness constraints are imposed.
- 1015, TITLE: Tractable Uncertainty for Structure Learning
<https://proceedings.mlr.press/v162/wang22ad.html>
AUTHORS: Benjie Wang, Matthew R Wicker, Marta Kwiatkowska
HIGHLIGHT: In this work, we present Tractable Uncertainty for Structure learning (TRUST), a framework for approximate posterior inference that relies on probabilistic circuits as a representation of our posterior belief.
- 1016, TITLE: Causal Dynamics Learning for Task-Independent State Abstraction
<https://proceedings.mlr.press/v162/wang22ae.html>
AUTHORS: Zizhao Wang, Xuesu Xiao, Zifan Xu, Yuke Zhu, Peter Stone
HIGHLIGHT: In this paper, we introduce Causal Dynamics Learning for Task-Independent State Abstraction (CDL), which first learns a theoretically proved causal dynamics model that removes unnecessary dependencies between state variables and the action, thus generalizing well to unseen states.
- 1017, TITLE: Multiple-Play Stochastic Bandits with Shareable Finite-Capacity Arms
<https://proceedings.mlr.press/v162/wang22af.html>
AUTHORS: Xuchuang Wang, Hong Xie, John C. S. Lui
HIGHLIGHT: We generalize the multiple-play multi-armed bandits (MP-MAB) problem with a shareable arms setting, in which several plays can share the same arm.
- 1018, TITLE: Generative Coarse-Graining of Molecular Conformations
<https://proceedings.mlr.press/v162/wang22ag.html>
AUTHORS: Wujie Wang, Minkai Xu, Chen Cai, Benjamin K Miller, Tess Smidt, Yusu Wang, Jian Tang, Rafael Gomez-Bombarelli
HIGHLIGHT: Inspired by the recent progress in generative models and equivariant networks, we propose a novel model that rigorously embeds the vital probabilistic nature and geometrical consistency requirements of the backmapping transformation.
- 1019, TITLE: Nonparametric Embeddings of Sparse High-Order Interaction Events
<https://proceedings.mlr.press/v162/wang22ah.html>
AUTHORS: Zheng Wang, Yiming Xu, Conor Tillinghast, Shibo Li, Akil Narayan, Shandian Zhe
HIGHLIGHT: In this paper, we propose Nonparametric Embeddings of Sparse High-order interaction events (NESH).
- 1020, TITLE: When Are Linear Stochastic Bandits Attackable?
<https://proceedings.mlr.press/v162/wang22ai.html>
AUTHORS: Huazheng Wang, Haifeng Xu, Hongning Wang
HIGHLIGHT: We study adversarial attacks on linear stochastic bandits: by manipulating the rewards, an adversary aims to control the behaviour of the bandit algorithm.

- 1021, TITLE: DRAGONN: Distributed Randomized Approximate Gradients of Neural Networks
<https://proceedings.mlr.press/v162/wang22aj.html>
AUTHORS: Zhuang Wang, Zhaozhuo Xu, Xinyu Wu, Anshumali Shrivastava, T. S. Eugene Ng
HIGHLIGHT: In this paper, we propose DRAGONN, a randomized hashing algorithm for GS in DDT.
- 1022, TITLE: Finite-Sum Coupled Compositional Stochastic Optimization: Theory and Applications
<https://proceedings.mlr.press/v162/wang22ak.html>
AUTHORS: Bokun Wang, Tianbao Yang
HIGHLIGHT: The contribution of this paper is to provide a comprehensive convergence analysis of a simple stochastic algorithm for both non-convex and convex objectives.
- 1023, TITLE: OFA: Unifying Architectures, Tasks, and Modalities Through a Simple Sequence-to-Sequence Learning Framework
<https://proceedings.mlr.press/v162/wang22al.html>
AUTHORS: Peng Wang, An Yang, Rui Men, Junyang Lin, Shuai Bai, Zhikang Li, Jianxin Ma, Chang Zhou, Jingren Zhou, Hongxia Yang
HIGHLIGHT: In this work, we pursue a unified paradigm for multimodal pretraining to break the shackles of complex task/modality-specific customization.
- 1024, TITLE: How Powerful are Spectral Graph Neural Networks
<https://proceedings.mlr.press/v162/wang22am.html>
AUTHORS: Xiyuan Wang, Muhan Zhang
HIGHLIGHT: Inspired by the analysis, we propose JacobiConv, which uses Jacobi basis due to its orthogonality and flexibility to adapt to a wide range of weight functions.
- 1025, TITLE: Thompson Sampling for Robust Transfer in Multi-Task Bandits
<https://proceedings.mlr.press/v162/wang22an.html>
AUTHORS: Zhi Wang, Chicheng Zhang, Kamalika Chaudhuri
HIGHLIGHT: In this work, we present a TS-type algorithm for a more general online multi-task learning protocol, which extends the concurrent setting.
- 1026, TITLE: Individual Reward Assisted Multi-Agent Reinforcement Learning
<https://proceedings.mlr.press/v162/wang22ao.html>
AUTHORS: Li Wang, Yupeng Zhang, Yujing Hu, Weixun Wang, Chongjie Zhang, Yang Gao, Jianye Hao, Tangjie Lv, Changjie Fan
HIGHLIGHT: In this paper, we propose Individual Reward Assisted Team Policy Learning (IRAT), which learns two policies for each agent from the dense individual reward and the sparse team reward with discrepancy constraints for updating the two policies mutually.
- 1027, TITLE: Removing Batch Normalization Boosts Adversarial Training
<https://proceedings.mlr.press/v162/wang22ap.html>
AUTHORS: Haotao Wang, Aston Zhang, Shuai Zheng, Xingjian Shi, Mu Li, Zhangyang Wang
HIGHLIGHT: Although the dominant approach is to extend BN to capture this mixture of distribution, we propose to completely eliminate this bottleneck by removing all BN layers in AT.
- 1028, TITLE: Partial and Asymmetric Contrastive Learning for Out-of-Distribution Detection in Long-Tailed Recognition
<https://proceedings.mlr.press/v162/wang22aq.html>
AUTHORS: Haotao Wang, Aston Zhang, Yi Zhu, Shuai Zheng, Mu Li, Alex J Smola, Zhangyang Wang
HIGHLIGHT: To solve this problem, we propose Partial and Asymmetric Supervised Contrastive Learning (PASCL), which explicitly encourages the model to distinguish between tail-class in-distribution samples and OOD samples.
- 1029, TITLE: Nonparametric Factor Trajectory Learning for Dynamic Tensor Decomposition
<https://proceedings.mlr.press/v162/wang22ar.html>
AUTHORS: Zheng Wang, Shandian Zhe
HIGHLIGHT: However, current methods always assume the factor representations of the entities in each tensor mode are static, and never consider their temporal evolution. To fill this gap, we propose NONparametric FACTor Trajectory learning for dynamic tensor decomposition (NONFAT).
- 1030, TITLE: Thompson Sampling for (Combinatorial) Pure Exploration
<https://proceedings.mlr.press/v162/wang22as.html>
AUTHORS: Siwei Wang, Jun Zhu
HIGHLIGHT: To deal with this challenge, we explore the idea of Thompson Sampling (TS) that uses independent random samples instead of the upper confidence bounds, and design the first TS-based algorithm TS-Explore for (combinatorial) pure exploration.
- 1031, TITLE: Policy Gradient Method For Robust Reinforcement Learning
<https://proceedings.mlr.press/v162/wang22at.html>
AUTHORS: Yue Wang, Shaofeng Zou
HIGHLIGHT: This paper develops the first policy gradient method with global optimality guarantee and complexity analysis for robust reinforcement learning under model mismatch.

- 1032, TITLE: Certifying Out-of-Domain Generalization for Blackbox Functions
<https://proceedings.mlr.press/v162/weber22a.html>
AUTHORS: Maurice G Weber, Linyi Li, Boxin Wang, Zhikuan Zhao, Bo Li, Ce Zhang
HIGHLIGHT: In this paper, we focus on the problem of certifying distributional robustness for blackbox models and bounded loss functions, and propose a novel certification framework based on the Hellinger distance.
- 1033, TITLE: More Than a Toy: Random Matrix Models Predict How Real-World Neural Representations Generalize
<https://proceedings.mlr.press/v162/wei22a.html>
AUTHORS: Alexander Wei, Wei Hu, Jacob Steinhardt
HIGHLIGHT: Our findings suggest that random matrix theory, rather than just being a toy model, may be central to understanding the properties of neural representations in practice.
- 1034, TITLE: To Smooth or Not? When Label Smoothing Meets Noisy Labels
<https://proceedings.mlr.press/v162/wei22b.html>
AUTHORS: Jiaheng Wei, Hangyu Liu, Tongliang Liu, Gang Niu, Masashi Sugiyama, Yang Liu
HIGHLIGHT: We provide understandings for the properties of LS and NLS when learning with noisy labels.
- 1035, TITLE: Open-Sampling: Exploring Out-of-Distribution data for Re-balancing Long-tailed datasets
<https://proceedings.mlr.press/v162/wei22c.html>
AUTHORS: Hongxin Wei, Lue Tao, Renchunzi Xie, Lei Feng, Bo An
HIGHLIGHT: In this work, we theoretically show that out-of-distribution data can still be leveraged to augment the minority classes from a Bayesian perspective.
- 1036, TITLE: Mitigating Neural Network Overconfidence with Logit Normalization
<https://proceedings.mlr.press/v162/wei22d.html>
AUTHORS: Hongxin Wei, Renchunzi Xie, Hao Cheng, Lei Feng, Bo An, Yixuan Li
HIGHLIGHT: However, neural networks are known to suffer from the overconfidence issue, where they produce abnormally high confidence for both in- and out-of-distribution inputs. In this work, we show that this issue can be mitigated through Logit Normalization (LogitNorm)—a simple fix to the cross-entropy loss—by enforcing a constant vector norm on the logits in training.
- 1037, TITLE: Koopman Q-learning: Offline Reinforcement Learning via Symmetries of Dynamics
<https://proceedings.mlr.press/v162/weissenbacher22a.html>
AUTHORS: Matthias Weissenbacher, Samarth Sinha, Animesh Garg, Kawahara Yoshinobu
HIGHLIGHT: Current algorithms over-fit to the training dataset and as a consequence perform poorly when deployed to out-of-distribution generalizations of the environment. We aim to address these limitations by learning a Koopman latent representation which allows us to infer symmetries of the system’s underlying dynamic.
- 1038, TITLE: Fishing for User Data in Large-Batch Federated Learning via Gradient Magnification
<https://proceedings.mlr.press/v162/wen22a.html>
AUTHORS: Yuxin Wen, Jonas A. Geiping, Liam Fowl, Micah Goldblum, Tom Goldstein
HIGHLIGHT: We introduce a new strategy that dramatically elevates existing attacks to operate on batches of arbitrarily large size, and without architectural modifications.
- 1039, TITLE: BabelTower: Learning to Auto-parallelized Program Translation
<https://proceedings.mlr.press/v162/wen22b.html>
AUTHORS: Yuanbo Wen, Qi Guo, Qiang Fu, Xiaqing Li, Jianxing Xu, Yanlin Tang, Yongwei Zhao, Xing Hu, Zidong Du, Ling Li, Chao Wang, Xuehai Zhou, Yunji Chen
HIGHLIGHT: In this paper, we propose a learning-based framework, i.e., BabelTower, to address this problem. We first create a large-scale dataset consisting of compute-intensive function-level monolingual corpora.
- 1040, TITLE: Random Forest Density Estimation
<https://proceedings.mlr.press/v162/wen22c.html>
AUTHORS: Hongwei Wen, Hanyuan Hang
HIGHLIGHT: We propose a density estimation algorithm called random forest density estimation (RFDE) based on random trees where the split of cell is along the midpoint of the randomly chosen dimension.
- 1041, TITLE: Fighting Fire with Fire: Avoiding DNN Shortcuts through Priming
<https://proceedings.mlr.press/v162/wen22d.html>
AUTHORS: Chuan Wen, Jianing Qian, Jierui Lin, Jiaye Teng, Dinesh Jayaraman, Yang Gao
HIGHLIGHT: In this paper, we show empirically that DNNs can be coaxed to avoid poor shortcuts by providing an additional “priming” feature computed from key input features, usually a coarse output estimate.
- 1042, TITLE: Preconditioning for Scalable Gaussian Process Hyperparameter Optimization
<https://proceedings.mlr.press/v162/wenger22a.html>
AUTHORS: Jonathan Wenger, Geoff Pleiss, Philipp Hennig, John Cunningham, Jacob Gardner
HIGHLIGHT: Iterative numerical techniques are becoming popular to scale to larger datasets, relying on the conjugate gradient method (CG) for the linear solves and stochastic trace estimation for the log-determinant. This work introduces new algorithmic and theoretical insights for preconditioning these computations.

- 1043, TITLE: Measure Estimation in the Barycentric Coding Model
<https://proceedings.mlr.press/v162/werenski22a.html>
AUTHORS: Matthew E Werenski, Ruijie Jiang, Abiy Tasissa, Shuchin Aeron, James M Murphy
HIGHLIGHT: This paper considers the problem of measure estimation under the barycentric coding model (BCM), in which an unknown measure is assumed to belong to the set of Wasserstein-2 barycenters of a finite set of known measures.
- 1044, TITLE: COLA: Consistent Learning with Opponent-Learning Awareness
<https://proceedings.mlr.press/v162/willi22a.html>
AUTHORS: Timon Willi, Alistair Hp Letcher, Johannes Treutlein, Jakob Foerster
HIGHLIGHT: First, we formalize consistency and show that higher-order LOLA (HOLA) solves LOLA's inconsistency problem if it converges. Second, we correct a claim made in the literature by Schfer and Anandkumar (2019), proving that Competitive Gradient Descent (CGD) does not recover HOLA as a series expansion (and fails to solve the consistency problem). Third, we propose a new method called Consistent LOLA (COLA), which learns update functions that are consistent under mutual opponent shaping.
- 1045, TITLE: Distributional Hamilton-Jacobi-Bellman Equations for Continuous-Time Reinforcement Learning
<https://proceedings.mlr.press/v162/wiltzer22a.html>
AUTHORS: Harley E Wiltzer, David Meger, Marc G. Bellemare
HIGHLIGHT: Our derivation highlights additional terms due to statistical diffusivity which arise from the proper handling of distributions in the continuous-time setting. Based on this, we propose a tractable algorithm for approximately solving the distributional HJB based on a JKO scheme, which can be implemented in an online, control algorithm.
- 1046, TITLE: Easy Variational Inference for Categorical Models via an Independent Binary Approximation
<https://proceedings.mlr.press/v162/wojnowicz22a.html>
AUTHORS: Michael T Wojnowicz, Shuchin Aeron, Eric L Miller, Michael Hughes
HIGHLIGHT: We pursue tractable Bayesian analysis of generalized linear models (GLMs) for categorical data.
- 1047, TITLE: Continual Learning with Guarantees via Weight Interval Constraints
<https://proceedings.mlr.press/v162/wolczyk22a.html>
AUTHORS: Maciej Wolczyk, Karol Piczak, Bartosz Wójcik, Lukasz Pustelnik, Pawel Morawiecki, Jacek Tabor, Tomasz Trzcinski, Przemyslaw Spurek
HIGHLIGHT: In this work, we show how to put bounds on forgetting by reformulating continual learning of a model as a continual contraction of its parameter space.
- 1048, TITLE: A Deep Learning Approach for the Segmentation of Electroencephalography Data in Eye Tracking Applications
<https://proceedings.mlr.press/v162/wolf22a.html>
AUTHORS: Lukas Wolf, Ard Kastrati, Martyna B Plomecka, Jie-Ming Li, Dustin Klebe, Alexander Veicht, Roger Wattenhofer, Nicolas Langer
HIGHLIGHT: Here, we introduce DETRtime, a novel framework for time-series segmentation that creates ocular event detectors that do not require additionally recorded eye-tracking modality and rely solely on EEG data.
- 1049, TITLE: Leverage Score Sampling for Tensor Product Matrices in Input Sparsity Time
<https://proceedings.mlr.press/v162/woodruff22a.html>
AUTHORS: David Woodruff, Amir Zandieh
HIGHLIGHT: We propose an input sparsity time sampling algorithm that can spectrally approximate the Gram matrix corresponding to the q-fold column-wise tensor product of q matrices using a nearly optimal number of samples, improving upon all previously known methods by poly(q) factors.
- 1050, TITLE: Model soups: averaging weights of multiple fine-tuned models improves accuracy without increasing inference time
<https://proceedings.mlr.press/v162/wortsman22a.html>
AUTHORS: Mitchell Wortsman, Gabriel Ilharco, Samir Ya Gadre, Rebecca Roelofs, Raphael Gontijo-Lopes, Ari S Morcos, Hongseok Namkoong, Ali Farhadi, Yair Carmon, Simon Kornblith, Ludwig Schmidt
HIGHLIGHT: The conventional recipe for maximizing model accuracy is to (1) train multiple models with various hyperparameters and (2) pick the individual model which performs best on a held-out validation set, discarding the remainder. In this paper, we revisit the second step of this procedure in the context of fine-tuning large pre-trained models, where fine-tuned models often appear to lie in a single low error basin.
- 1051, TITLE: Metric-Fair Classifier Derandomization
<https://proceedings.mlr.press/v162/wu22a.html>
AUTHORS: Jimmy Wu, Yatong Chen, Yang Liu
HIGHLIGHT: In this paper, we initiate a systematic study of classifier derandomization with metric fairness guarantees.
- 1052, TITLE: Structural Entropy Guided Graph Hierarchical Pooling
<https://proceedings.mlr.press/v162/wu22b.html>
AUTHORS: Junran Wu, Xueyuan Chen, Ke Xu, Shangzhe Li
HIGHLIGHT: In this work, inspired by structural entropy, we propose a hierarchical pooling approach, SEP, to tackle the two issues.

- 1053, TITLE: Self-supervised Models are Good Teaching Assistants for Vision Transformers
<https://proceedings.mlr.press/v162/wu22c.html>
AUTHORS: Haiyan Wu, Yuting Gao, Yinqi Zhang, Shaohui Lin, Yuan Xie, Xing Sun, Ke Li
HIGHLIGHT: Specifically, we propose a head-level knowledge distillation method that selects the most important head of the supervised teacher and self-supervised teaching assistant, and let the student mimic the attention distribution of these two heads, so as to make the student focus on the relationship between tokens deemed by the teacher and the teacher assistant.
- 1054, TITLE: Characterizing and Overcoming the Greedy Nature of Learning in Multi-modal Deep Neural Networks
<https://proceedings.mlr.press/v162/wu22d.html>
AUTHORS: Nan Wu, Stanislaw Jastrzebski, Kyunghyun Cho, Krzysztof J Geras
HIGHLIGHT: We propose an algorithm to balance the conditional learning speeds between modalities during training and demonstrate that it indeed addresses the issue of greedy learning.
- 1055, TITLE: Instrumental Variable Regression with Confounder Balancing
<https://proceedings.mlr.press/v162/wu22e.html>
AUTHORS: Anpeng Wu, Kun Kuang, Bo Li, Fei Wu
HIGHLIGHT: In this paper, we propose a Confounder Balanced IV Regression (CB-IV) algorithm to jointly remove the bias from the unmeasured confounders and balance the observed confounders.
- 1056, TITLE: MemSR: Training Memory-efficient Lightweight Model for Image Super-Resolution
<https://proceedings.mlr.press/v162/wu22f.html>
AUTHORS: Kailu Wu, Chung-Kuei Lee, Kaisheng Ma
HIGHLIGHT: This paper aims at calculating a winning initialization from a complex teacher network for a plain student network, which can provide performance comparable to complex models.
- 1057, TITLE: Delay-Adaptive Step-sizes for Asynchronous Learning
<https://proceedings.mlr.press/v162/wu22g.html>
AUTHORS: Xuyang Wu, Sindri Magnusson, Hamid Reza Feyzmahdavian, Mikael Johansson
HIGHLIGHT: In this paper, we show that it is possible to use learning rates that depend on the actual time-varying delays in the system.
- 1058, TITLE: Variational nearest neighbor Gaussian process
<https://proceedings.mlr.press/v162/wu22h.html>
AUTHORS: Luhuan Wu, Geoff Pleiss, John P Cunningham
HIGHLIGHT: In this work, we instead exploit a sparse approximation of the precision matrix.
- 1059, TITLE: Understanding Policy Gradient Algorithms: A Sensitivity-Based Approach
<https://proceedings.mlr.press/v162/wu22i.html>
AUTHORS: Shuang Wu, Ling Shi, Jun Wang, Guangjian Tian
HIGHLIGHT: In particular, we derive PG in a unified framework, precisely clarify the relation between PG implementation and theory, and echo back the findings by \citeauthor{nota2020policy}.
- 1060, TITLE: DAVINZ: Data Valuation using Deep Neural Networks at Initialization
<https://proceedings.mlr.press/v162/wu22j.html>
AUTHORS: Zhaoxuan Wu, Yao Shu, Bryan Kian Hsiang Low
HIGHLIGHT: To this end, we theoretically derive a domain-aware generalization bound to estimate the generalization performance of DNNs without model training.
- 1061, TITLE: Robust Deep Reinforcement Learning through Bootstrapped Opportunistic Curriculum
<https://proceedings.mlr.press/v162/wu22k.html>
AUTHORS: Junlin Wu, Yevgeniy Vorobeychik
HIGHLIGHT: We propose Bootstrapped Opportunistic Adversarial Curriculum Learning (BCL), a novel flexible adversarial curriculum learning framework for robust reinforcement learning.
- 1062, TITLE: Revisiting Consistency Regularization for Deep Partial Label Learning
<https://proceedings.mlr.press/v162/wu22l.html>
AUTHORS: Dong-Dong Wu, Deng-Bao Wang, Min-Ling Zhang
HIGHLIGHT: In this paper, we revisit a simple idea namely consistency regularization, which has been shown effective in traditional PLL literature, to guide the training of deep models.
- 1063, TITLE: Flowformer: Linearizing Transformers with Conservation Flows
<https://proceedings.mlr.press/v162/wu22m.html>
AUTHORS: Haixu Wu, Jialong Wu, Jiehui Xu, Jianmin Wang, Mingsheng Long
HIGHLIGHT: In this paper, we linearize Transformers free from specific inductive biases based on the flow network theory.
- 1064, TITLE: Nearly Optimal Policy Optimization with Stable at Any Time Guarantee
<https://proceedings.mlr.press/v162/wu22n.html>
AUTHORS: Tianhao Wu, Yunchang Yang, Han Zhong, Liwei Wang, Simon Du, Jiantao Jiao
HIGHLIGHT: To bridge such a gap, we propose a novel algorithm Reference-based Policy Optimization with Stable at Any Time guarantee (RPO-SAT), which features the property “Stable at Any Time”.

- 1065, TITLE: RetrievalGuard: Provably Robust 1-Nearest Neighbor Image Retrieval
<https://proceedings.mlr.press/v162/wu22o.html>
AUTHORS: Yihan Wu, Hongyang Zhang, Heng Huang
HIGHLIGHT: In this paper, we aim to design a provably robust image retrieval model which keeps the most important evaluation metric Recall@1 invariant to adversarial perturbation.
- 1066, TITLE: Last Iterate Risk Bounds of SGD with Decaying Stepsize for Overparameterized Linear Regression
<https://proceedings.mlr.press/v162/wu22p.html>
AUTHORS: Jingfeng Wu, Difan Zou, Vladimir Braverman, Quanquan Gu, Sham Kakade
HIGHLIGHT: In this paper, we provide a problem-dependent analysis on the last iterate risk bounds of SGD with decaying stepsize, for (overparameterized) linear regression problems.
- 1067, TITLE: Optimal Clustering with Noisy Queries via Multi-Armed Bandit
<https://proceedings.mlr.press/v162/xia22a.html>
AUTHORS: Jinghui Xia, Zengfeng Huang
HIGHLIGHT: In this work, we obtain the first matching upper and lower bounds for a wide range of parameters.
- 1068, TITLE: ProGCL: Rethinking Hard Negative Mining in Graph Contrastive Learning
<https://proceedings.mlr.press/v162/xia22b.html>
AUTHORS: Jun Xia, Lirong Wu, Ge Wang, Jintao Chen, Stan Z. Li
HIGHLIGHT: To remedy this deficiency, we propose an effective method, dubbed \textbf{ProGCL}, to estimate the probability of a negative being true one, which constitutes a more suitable measure for negatives' hardness together with similarity.
- 1069, TITLE: Synergy and Symmetry in Deep Learning: Interactions between the Data, Model, and Inference Algorithm
<https://proceedings.mlr.press/v162/xiao22a.html>
AUTHORS: Lechao Xiao, Jeffrey Pennington
HIGHLIGHT: While previous efforts have investigated this question by studying the data (D), model (M), and inference algorithm (I) as independent modules, in this paper we analyze the triplet (D,M,I) as an integrated system and identify important synergies that help mitigate the curse of dimensionality.
- 1070, TITLE: Identification of Linear Non-Gaussian Latent Hierarchical Structure
<https://proceedings.mlr.press/v162/xie22a.html>
AUTHORS: Feng Xie, Biwei Huang, Zhengming Chen, Yangbo He, Zhi Geng, Kun Zhang
HIGHLIGHT: Accordingly, this paper investigates the problem of discovering the hidden causal variables and estimating the causal structure, including both the causal relations among latent variables and those between latent and measured variables.
- 1071, TITLE: COAT: Measuring Object Compositionality in Emergent Representations
<https://proceedings.mlr.press/v162/xie22b.html>
AUTHORS: Sirui Xie, Ari S Morcos, Song-Chun Zhu, Ramakrishna Vedantam
HIGHLIGHT: We propose to directly measure compositionality in the representation space as a form of objections, making such evaluations tractable for a wider class of models.
- 1072, TITLE: Robust Policy Learning over Multiple Uncertainty Sets
<https://proceedings.mlr.press/v162/xie22c.html>
AUTHORS: Annie Xie, Shagun Sodhani, Chelsea Finn, Joelle Pineau, Amy Zhang
HIGHLIGHT: Towards a more general solution, we formulate the multi-set robustness problem to learn a policy robust to different perturbation sets.
- 1073, TITLE: Adaptive Inertia: Disentangling the Effects of Adaptive Learning Rate and Momentum
<https://proceedings.mlr.press/v162/xie22d.html>
AUTHORS: Zeke Xie, Xinrui Wang, Huishuai Zhang, Issei Sato, Masashi Sugiyama
HIGHLIGHT: However, it is empirically known that Adam often generalizes worse than Stochastic Gradient Descent (SGD). The purpose of this paper is to unveil the mystery of this behavior in the diffusion theoretical framework.
- 1074, TITLE: Self-Supervised Representation Learning via Latent Graph Prediction
<https://proceedings.mlr.press/v162/xie22e.html>
AUTHORS: Yaochen Xie, Zhao Xu, Shuiwang Ji
HIGHLIGHT: In this work, we propose the LaGraph, a theoretically grounded predictive SSL framework based on latent graph prediction.
- 1075, TITLE: Efficient Computation of Higher-Order Subgraph Attribution via Message Passing
<https://proceedings.mlr.press/v162/xiong22a.html>
AUTHORS: Ping Xiong, Thomas Schnake, Grégoire Montavon, Klaus-Robert Müller, Shinichi Nakajima
HIGHLIGHT: GNN-LRP gives a relevance attribution of walks between nodes at each layer, and the subgraph attribution is expressed as a sum over exponentially many such walks. In this work, we demonstrate that such exponential complexity can be avoided.
- 1076, TITLE: A Self-Play Posterior Sampling Algorithm for Zero-Sum Markov Games
<https://proceedings.mlr.press/v162/xiong22b.html>

- AUTHORS: Wei Xiong, Han Zhong, Chengshuai Shi, Cong Shen, Tong Zhang
HIGHLIGHT: This work focuses on a distinct approach of posterior sampling, which is celebrated in many bandits and reinforcement learning settings but remains under-explored for MGs.
- 1077, TITLE: Importance Weighted Kernel Bayes' Rule
<https://proceedings.mlr.press/v162/xu22a.html>
AUTHORS: Liyuan Xu, Yutian Chen, Arnaud Doucet, Arthur Gretton
HIGHLIGHT: We study a nonparametric approach to Bayesian computation via feature means, where the expectation of prior features is updated to yield expected posterior features, based on regression from kernel or neural net features of the observations.
- 1078, TITLE: Learning to Separate Voices by Spatial Regions
<https://proceedings.mlr.press/v162/xu22b.html>
AUTHORS: Alan Xu, Romit Roy Choudhury
HIGHLIGHT: We propose a two-stage self-supervised framework in which overheard voices from earphones are pre-processed to extract relatively clean personalized signals, which are then used to train a region-wise separation model.
- 1079, TITLE: Detached Error Feedback for Distributed SGD with Random Sparsification
<https://proceedings.mlr.press/v162/xu22c.html>
AUTHORS: An Xu, Heng Huang
HIGHLIGHT: In this work, we study distributed SGD with random block-wise sparsification as the gradient compressor, which is ring-allreduce compatible and highly computation-efficient but leads to inferior performance.
- 1080, TITLE: Accurate Quantization of Measures via Interacting Particle-based Optimization
<https://proceedings.mlr.press/v162/xu22d.html>
AUTHORS: Lantian Xu, Anna Korba, Dejan Slepcev
HIGHLIGHT: In particular, we prove general upper bounds on the quantization error of MMD and KSD at rates which significantly outperform quantization by i.i.d. samples.
- 1081, TITLE: Unified Fourier-based Kernel and Nonlinearity Design for Equivariant Networks on Homogeneous Spaces
<https://proceedings.mlr.press/v162/xu22e.html>
AUTHORS: Yinshuang Xu, Jiahui Lei, Edgar Dobriban, Kostas Daniilidis
HIGHLIGHT: We introduce a unified framework for group equivariant networks on homogeneous spaces derived from a Fourier perspective.
- 1082, TITLE: Inferring Cause and Effect in the Presence of Heteroscedastic Noise
<https://proceedings.mlr.press/v162/xu22f.html>
AUTHORS: Sascha Xu, Osman A Mian, Alexander Marx, Jilles Vreeken
HIGHLIGHT: We propose to partition the domain of the cause into multiple segments where the noise indeed is dependent.
- 1083, TITLE: Prompting Decision Transformer for Few-Shot Policy Generalization
<https://proceedings.mlr.press/v162/xu22g.html>
AUTHORS: Mengdi Xu, Yikang Shen, Shun Zhang, Yuchen Lu, Ding Zhao, Joshua Tenenbaum, Chuang Gan
HIGHLIGHT: We propose a Prompt-based Decision Transformer (Prompt-DT), which leverages the sequential modeling ability of the Transformer architecture and the prompt framework to achieve few-shot adaptation in offline RL.
- 1084, TITLE: Analyzing and Mitigating Interference in Neural Architecture Search
<https://proceedings.mlr.press/v162/xu22h.html>
AUTHORS: Jin Xu, Xu Tan, Kaitao Song, Renqian Luo, Yichong Leng, Tao Qin, Tie-Yan Liu, Jian Li
HIGHLIGHT: In this paper, we investigate the interference issue by sampling different child models and calculating the gradient similarity of shared operators, and observe that: 1) the interference on a shared operator between two child models is positively correlated with the number of different operators between them; 2) the interference is smaller when the inputs and outputs of the shared operator are more similar.
- 1085, TITLE: On the Statistical Benefits of Curriculum Learning
<https://proceedings.mlr.press/v162/xu22i.html>
AUTHORS: Ziping Xu, Ambuj Tewari
HIGHLIGHT: In this paper, we study the benefits of CL in the multitask linear regression problem under both structured and unstructured settings.
- 1086, TITLE: A Difference Standardization Method for Mutual Transfer Learning
<https://proceedings.mlr.press/v162/xu22j.html>
AUTHORS: Haoqing Xu, Meng Wang, Beilun Wang
HIGHLIGHT: However, previous studies about mutual transfer learning either suffer from high computational complexity or oversimplified hypothesis. To overcome these challenges, in this paper, we propose the Difference Standardization Method (DiffS) for mutual transfer learning.
- 1087, TITLE: SkexGen: Autoregressive Generation of CAD Construction Sequences with Disentangled Codebooks
<https://proceedings.mlr.press/v162/xu22k.html>
AUTHORS: Xiang Xu, Karl D.D. Willis, Joseph G Lambourne, Chin-Yi Cheng, Pradeep Kumar Jayaraman, Yasutaka Furukawa

HIGHLIGHT: We present SkexGen, a novel autoregressive generative model for computer-aided design (CAD) construction sequences containing sketch-and-extrude modeling operations.

1088, **TITLE:** Discriminator-Weighted Offline Imitation Learning from Suboptimal Demonstrations

<https://proceedings.mlr.press/v162/xu22l.html>

AUTHORS: Haoran Xu, Xianyuan Zhan, Honglei Yin, Huiling Qin

HIGHLIGHT: In this paper, we aim to address the problem without additional steps of reward learning and offline RL training for the case when demonstrations contain a large proportion of suboptimal data.

1089, **TITLE:** Adversarial Attack and Defense for Non-Parametric Two-Sample Tests

<https://proceedings.mlr.press/v162/xu22m.html>

AUTHORS: Xilie Xu, Jingfeng Zhang, Feng Liu, Masashi Sugiyama, Mohan Kankanhalli

HIGHLIGHT: To enable TST-agnostic attacks, we propose an ensemble attack (EA) framework that jointly minimizes the different types of test criteria.

1090, **TITLE:** Adversarially Robust Models may not Transfer Better: Sufficient Conditions for Domain Transferability from the View of Regularization

<https://proceedings.mlr.press/v162/xu22n.html>

AUTHORS: Xiaojun Xu, Jacky Y Zhang, Evelyn Ma, Hyun Ho Son, Sanmi Koyejo, Bo Li

HIGHLIGHT: In this paper, we explore the relationship between regularization and domain transferability considering different factors such as norm regularization and data augmentations (DA).

1091, **TITLE:** A Theoretical Analysis on Independence-driven Importance Weighting for Covariate-shift Generalization

<https://proceedings.mlr.press/v162/xu22o.html>

AUTHORS: Renzhe Xu, Xingxuan Zhang, Zheyang Shen, Tong Zhang, Peng Cui

HIGHLIGHT: Recently, independence-driven importance weighting algorithms in stable learning literature have shown empirical effectiveness to deal with covariate-shift generalization on several learning models, including regression algorithms and deep neural networks, while their theoretical analyses are missing. In this paper, we theoretically prove the effectiveness of such algorithms by explaining them as feature selection processes.

1092, **TITLE:** Langevin Monte Carlo for Contextual Bandits

<https://proceedings.mlr.press/v162/xu22p.html>

AUTHORS: Pan Xu, Hongkai Zheng, Eric V Mazumdar, Kamyar Azizzadenesheli, Animashree Anandkumar

HIGHLIGHT: We propose an efficient posterior sampling algorithm, viz., Langevin Monte Carlo Thompson Sampling (LMC-TS), that uses Markov Chain Monte Carlo (MCMC) methods to directly sample from the posterior distribution in contextual bandits.

1093, **TITLE:** Investigating Why Contrastive Learning Benefits Robustness against Label Noise

<https://proceedings.mlr.press/v162/xue22a.html>

AUTHORS: Yihao Xue, Kyle Whitecross, Baharan Mirzasoleiman

HIGHLIGHT: In this work, we rigorously prove that the representation matrix learned by contrastive learning boosts robustness, by having: (i) one prominent singular value corresponding to each sub-class in the data, and significantly smaller remaining singular values; and (ii) a large alignment between the prominent singular vectors and the clean labels of each sub-class.

1094, **TITLE:** Diversified Adversarial Attacks based on Conjugate Gradient Method

<https://proceedings.mlr.press/v162/yamamura22a.html>

AUTHORS: Keiichiro Yamamura, Haruki Sato, Nariaki Tateiwa, Nozomi Hata, Toru Mitsutake, Issa Oe, Hiroki Ishikura, Katsuki Fujisawa

HIGHLIGHT: Although existing methods based on the steepest descent have achieved high attack success rates, ill-conditioned problems occasionally reduce their performance. To address this limitation, we utilize the conjugate gradient (CG) method, which is effective for this type of problem, and propose a novel attack algorithm inspired by the CG method, named the Auto Conjugate Gradient (ACG) attack.

1095, **TITLE:** Cycle Representation Learning for Inductive Relation Prediction

<https://proceedings.mlr.press/v162/yan22a.html>

AUTHORS: Zuoyu Yan, Tengfei Ma, Liangcai Gao, Zhi Tang, Chao Chen

HIGHLIGHT: In this paper, based on the mathematics of algebraic topology, we propose a novel solution for inductive relation prediction, an important learning task for knowledge graph completion.

1096, **TITLE:** Optimally Controllable Perceptual Lossy Compression

<https://proceedings.mlr.press/v162/yan22b.html>

AUTHORS: Zeyu Yan, Fei Wen, Peilin Liu

HIGHLIGHT: In this paper, we present a nontrivial finding that only two decoders are sufficient for optimally achieving arbitrary (an infinite number of different) D-P tradeoff.

1097, **TITLE:** Active fairness auditing

<https://proceedings.mlr.press/v162/yan22c.html>

AUTHORS: Tom Yan, Chicheng Zhang

HIGHLIGHT: In this paper, we initiate the study of query-based auditing algorithms that can estimate the demographic parity of ML models in a query-efficient manner.

- 1098, TITLE: Self-Organized Polynomial-Time Coordination Graphs
<https://proceedings.mlr.press/v162/yang22a.html>
AUTHORS: Qianlan Yang, Weijun Dong, Zhizhou Ren, Jianhao Wang, Tonghan Wang, Chongjie Zhang
HIGHLIGHT: To bypass this systematic hardness, this paper proposes a novel method, named Self-Organized Polynomial-time Coordination Graphs (SOP-CG), which uses structured graph classes to guarantee the accuracy and the computational efficiency of collaborated action selection.
- 1099, TITLE: Regularizing a Model-based Policy Stationary Distribution to Stabilize Offline Reinforcement Learning
<https://proceedings.mlr.press/v162/yang22b.html>
AUTHORS: Shentao Yang, Yihao Feng, Shujian Zhang, Mingyuan Zhou
HIGHLIGHT: To avoid the detrimental impact of distribution mismatch, we regularize the undiscounted stationary distribution of the current policy towards the offline data during the policy optimization process.
- 1100, TITLE: A Psychological Theory of Explainability
<https://proceedings.mlr.press/v162/yang22c.html>
AUTHORS: Scott Cheng-Hsin Yang, Nils Erik Tomas Folke, Patrick Shafto
HIGHLIGHT: We propose a psychological theory of how humans draw conclusions from saliency maps, the most common form of XAI explanation, which for the first time allows for precise prediction of explainee inference conditioned on explanation.
- 1101, TITLE: Omni-Granular Ego-Semantic Propagation for Self-Supervised Graph Representation Learning
<https://proceedings.mlr.press/v162/yang22d.html>
AUTHORS: Ling Yang, Shenda Hong
HIGHLIGHT: In this paper, we propose Omni-Granular Ego-Semantic Propagation for Self-Supervised Graph Representation Learning (OEPG).
- 1102, TITLE: Unsupervised Time-Series Representation Learning with Iterative Bilinear Temporal-Spectral Fusion
<https://proceedings.mlr.press/v162/yang22e.html>
AUTHORS: Ling Yang, Shenda Hong
HIGHLIGHT: In this paper, we propose a unified framework, namely Bilinear Temporal-Spectral Fusion (BTSF).
- 1103, TITLE: Searching for BurgerFormer with Micro-Meso-Macro Space Design
<https://proceedings.mlr.press/v162/yang22f.html>
AUTHORS: Longxing Yang, Yu Hu, Shun Lu, Zihao Sun, Jilin Mei, Yinhe Han, Xiaowei Li
HIGHLIGHT: By revisiting typical search spaces, we design micro-meso-macro space to search for Transformer-like architectures, namely BurgerFormer.
- 1104, TITLE: Efficient Variance Reduction for Meta-learning
<https://proceedings.mlr.press/v162/yang22g.html>
AUTHORS: Hansi Yang, James Kwok
HIGHLIGHT: In this paper, we propose a novel approach that integrates variance reduction with first-order meta-learning algorithms such as Reptile.
- 1105, TITLE: Injecting Logical Constraints into Neural Networks via Straight-Through Estimators
<https://proceedings.mlr.press/v162/yang22h.html>
AUTHORS: Zhun Yang, Joohyung Lee, Chiyoun Park
HIGHLIGHT: More specifically, we design a systematic way to represent discrete logical constraints as a loss function; minimizing this loss using gradient descent via a straight-through-estimator updates the neural network's weights in the direction that the binarized outputs satisfy the logical constraints.
- 1106, TITLE: Locally Sparse Neural Networks for Tabular Biomedical Data
<https://proceedings.mlr.press/v162/yang22i.html>
AUTHORS: Junchen Yang, Ofir Lindenbaum, Yuval Kluger
HIGHLIGHT: Practitioners in this domain prefer linear or tree-based models over neural networks since the latter are harder to interpret and tend to overfit when applied to tabular datasets. To address these neural networks' shortcomings, we propose an intrinsically interpretable network for heterogeneous biomedical data.
- 1107, TITLE: Not All Poisons are Created Equal: Robust Training against Data Poisoning
<https://proceedings.mlr.press/v162/yang22j.html>
AUTHORS: Yu Yang, Tian Yu Liu, Baharan Mirzasoleiman
HIGHLIGHT: In this work, we propose an efficient defense mechanism that significantly reduces the success rate of various data poisoning attacks, and provides theoretical guarantees for the performance of the model.
- 1108, TITLE: Does the Data Induce Capacity Control in Deep Learning?
<https://proceedings.mlr.press/v162/yang22k.html>
AUTHORS: Rubing Yang, Jialin Mao, Pratik Chaudhari
HIGHLIGHT: We show that the input correlation matrix of typical classification datasets has an eigenspectrum where, after a sharp initial drop, a large number of small eigenvalues are distributed uniformly over an exponentially large range.
- 1109, TITLE: Informed Learning by Wide Neural Networks: Convergence, Generalization and Sampling Complexity
<https://proceedings.mlr.press/v162/yang22l.html>

- AUTHORS: Jianyi Yang, Shaolei Ren
HIGHLIGHT: In this paper, we consider an informed deep neural network (DNN) with over-parameterization and domain knowledge integrated into its training objective function, and study how and why domain knowledge benefits the performance.
- 1110, TITLE: Linear Bandit Algorithms with Sublinear Time Complexity
<https://proceedings.mlr.press/v162/yang22m.html>
AUTHORS: Shuo Yang, Tongzheng Ren, Sanjay Shakkottai, Eric Price, Inderjit S. Dhillon, Sujay Sanghavi
HIGHLIGHT: We propose two linear bandits algorithms with per-step complexity sublinear in the number of arms SK .
- 1111, TITLE: A New Perspective on the Effects of Spectrum in Graph Neural Networks
<https://proceedings.mlr.press/v162/yang22n.html>
AUTHORS: Mingqi Yang, Yanming Shen, Rui Li, Heng Qi, Qiang Zhang, Baocai Yin
HIGHLIGHT: Inspired by this, we propose the correlation-free architecture which naturally removes the correlation issue among different channels, making it possible to utilize more sophisticated filters within each channel.
- 1112, TITLE: Fourier Learning with Cyclical Data
<https://proceedings.mlr.press/v162/yang22o.html>
AUTHORS: Yingxiang Yang, Zhihan Xiong, Tianyi Liu, Taiqing Wang, Chong Wang
HIGHLIGHT: In this paper, we have designed a novel approach to overcome the aforementioned shortcomings.
- 1113, TITLE: Estimating Instance-dependent Bayes-label Transition Matrix using a Deep Neural Network
<https://proceedings.mlr.press/v162/yang22p.html>
AUTHORS: Shuo Yang, Erkun Yang, Bo Han, Yang Liu, Min Xu, Gang Niu, Tongliang Liu
HIGHLIGHT: Motivated by that classifiers mostly output Bayes optimal labels for prediction, in this paper, we study to directly model the transition from Bayes optimal labels to noisy labels (i.e., Bayes-label transition matrix (BLTM)) and learn a classifier to predict Bayes optimal labels.
- 1114, TITLE: A Study of Face Obfuscation in ImageNet
<https://proceedings.mlr.press/v162/yang22q.html>
AUTHORS: Kaiyu Yang, Jacqueline H. Yau, Li Fei-Fei, Jia Deng, Olga Russakovsky
HIGHLIGHT: In this paper, we explore the effects of face obfuscation on the popular ImageNet challenge visual recognition benchmark.
- 1115, TITLE: Anarchic Federated Learning
<https://proceedings.mlr.press/v162/yang22r.html>
AUTHORS: Haibo Yang, Xin Zhang, Prashant Khanduri, Jia Liu
HIGHLIGHT: Toward this end, we propose two Anarchic Federated Averaging (AFA) algorithms with two-sided learning rates for both cross-device and cross-silo settings, which are named AFA-CD and AFA-CS, respectively.
- 1116, TITLE: Identity-Disentangled Adversarial Augmentation for Self-supervised Learning
<https://proceedings.mlr.press/v162/yang22s.html>
AUTHORS: Kaiwen Yang, Tianyi Zhou, Xinmei Tian, Dacheng Tao
HIGHLIGHT: In this paper, we study a simple adversarial augmentation method that can modify training data to be hard positives/negatives without distorting the key information about their original identities.
- 1117, TITLE: Learning from a Learning User for Optimal Recommendations
<https://proceedings.mlr.press/v162/yao22a.html>
AUTHORS: Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang, Haifeng Xu
HIGHLIGHT: In this paper, we formalize a model to capture such "learning users" and design an efficient system-side learning solution, coined Noise-Robust Active Ellipsoid Search (RAES), to confront the challenges brought by the non-stationary feedback from such a learning user.
- 1118, TITLE: Improving Out-of-Distribution Robustness via Selective Augmentation
<https://proceedings.mlr.press/v162/yao22b.html>
AUTHORS: Huaxiu Yao, Yu Wang, Sai Li, Linjun Zhang, Weixin Liang, James Zou, Chelsea Finn
HIGHLIGHT: In this paper, we specifically consider the problems of subpopulation shifts (e.g., imbalanced data) and domain shifts.
- 1119, TITLE: NLP From Scratch Without Large-Scale Pretraining: A Simple and Efficient Framework
<https://proceedings.mlr.press/v162/yao22c.html>
AUTHORS: Xingcheng Yao, Yanan Zheng, Xiaocong Yang, Zhilin Yang
HIGHLIGHT: We propose a simple and efficient learning framework, TLM, that does not rely on large-scale pretraining.
- 1120, TITLE: Feature Space Particle Inference for Neural Network Ensembles
<https://proceedings.mlr.press/v162/yashima22a.html>
AUTHORS: Shingo Yashima, Teppei Suzuki, Kohta Ishikawa, Ikuro Sato, Rei Kawakami
HIGHLIGHT: In this study, we propose to optimize particles in the feature space where activations of a specific intermediate layer lie to alleviate the abovementioned difficulties.
- 1121, TITLE: Centroid Approximation for Bootstrap: Improving Particle Quality at Inference

<https://proceedings.mlr.press/v162/ye22a.html>

AUTHORS: Mao Ye, Qiang Liu

HIGHLIGHT: In this work, we propose an efficient method to explicitly optimize a small set of high quality “centroid” points to better approximate the ideal bootstrap distribution.

1122, TITLE: Be Like Water: Adaptive Floating Point for Machine Learning

<https://proceedings.mlr.press/v162/yeh22a.html>

AUTHORS: Thomas Yeh, Max Sterner, Zerlina Lai, Brandon Chuang, Alexander Ihler

HIGHLIGHT: In this paper, we propose a novel numerical representation, Adaptive Floating Point (AFP), that dynamically adjusts to the characteristics of deep learning data.

1123, TITLE: QSFL: A Two-Level Uplink Communication Optimization Framework for Federated Learning

<https://proceedings.mlr.press/v162/yi22a.html>

AUTHORS: Liping Yi, Wang Gang, Liu Xiaoguang

HIGHLIGHT: As a solution, we propose a novel FL framework named QSFL, towards optimizing FL uplink (client-to-server) communication at both client and model levels.

1124, TITLE: De novo mass spectrometry peptide sequencing with a transformer model

<https://proceedings.mlr.press/v162/yilmaz22a.html>

AUTHORS: Melih Yilmaz, William Fondrie, Wout Bittremieux, Sewoong Oh, William S Noble

HIGHLIGHT: We propose a simple yet powerful method for de novo peptide sequencing, Casanovo, that uses a transformer framework to map directly from a sequence of observed peaks (a mass spectrum) to a sequence of amino acids (a peptide).

1125, TITLE: Bayesian Nonparametric Learning for Point Processes with Spatial Homogeneity: A Spatial Analysis of NBA Shot Locations

<https://proceedings.mlr.press/v162/yin22a.html>

AUTHORS: Fan Yin, Jieying Jiao, Jun Yan, Guanyu Hu

HIGHLIGHT: We present a novel nonparametric Bayesian method for learning the underlying intensity surface built upon a combination of Dirichlet process and Markov random field.

1126, TITLE: Bitwidth Heterogeneous Federated Learning with Progressive Weight Dequantization

<https://proceedings.mlr.press/v162/yoon22a.html>

AUTHORS: Jaehong Yoon, Geon Park, Wonyong Jeong, Sung Ju Hwang

HIGHLIGHT: BHFL brings in a new challenge, that the aggregation of model parameters with different bitwidths could result in severe performance degeneration, especially for high-bitwidth models. To tackle this problem, we propose ProWD framework, which has a trainable weight dequantizer at the central server that progressively reconstructs the low-bitwidth weights into higher bitwidth weights, and finally into full-precision weights.

1127, TITLE: ShiftAddNAS: Hardware-Inspired Search for More Accurate and Efficient Neural Networks

<https://proceedings.mlr.press/v162/you22a.html>

AUTHORS: Haoran You, Baopu Li, Shi Huihong, Yonggan Fu, Yingyan Lin

HIGHLIGHT: To this end, this work advocates hybrid NNs that consist of both powerful yet costly multiplications and efficient yet less powerful operators for marrying the best of both worlds, and proposes ShiftAddNAS, which can automatically search for more accurate and more efficient NNs.

1128, TITLE: Molecular Representation Learning via Heterogeneous Motif Graph Neural Networks

<https://proceedings.mlr.press/v162/you22a.html>

AUTHORS: Zhaoning Yu, Hongyang Gao

HIGHLIGHT: However, most existing methods deal with molecular graphs individually while neglecting their connections, such as motif-level relationships. We propose a novel molecular graph representation learning method by constructing a heterogeneous motif graph to address this issue.

1129, TITLE: Understanding Robust Overfitting of Adversarial Training and Beyond

<https://proceedings.mlr.press/v162/you22b.html>

AUTHORS: Chaojian Yu, Bo Han, Li Shen, Jun Yu, Chen Gong, Mingming Gong, Tongliang Liu

HIGHLIGHT: Given these observations, we further designed data ablation adversarial training and identify that some small-loss data which are not worthy of the adversary strength cause robust overfitting in the strong adversary mode. To relieve this issue, we propose minimum loss constrained adversarial training (MLCAT): in a minibatch, we learn large-loss data as usual, and adopt additional measures to increase the loss of the small-loss data.

1130, TITLE: How to Leverage Unlabeled Data in Offline Reinforcement Learning

<https://proceedings.mlr.press/v162/you22c.html>

AUTHORS: Tianhe Yu, Aviral Kumar, Yevgen Chebotar, Karol Hausman, Chelsea Finn, Sergey Levine

HIGHLIGHT: In this paper, we find that, perhaps surprisingly, a much simpler method that simply applies zero rewards to unlabeled data leads to effective data sharing both in theory and in practice, without learning any reward model at all.

1131, TITLE: Reachability Constrained Reinforcement Learning

<https://proceedings.mlr.press/v162/you22d.html>

AUTHORS: Dongjie Yu, Haitong Ma, Shengbo Li, Jianyu Chen

HIGHLIGHT: To deal with this problem, this paper proposes the reachability CRL (RCRL) method by using reachability analysis to establish the novel self-consistency condition and characterize the feasible sets.

1132, **TITLE:** Topology-Aware Network Pruning using Multi-stage Graph Embedding and Reinforcement Learning
<https://proceedings.mlr.press/v162/yu22e.html>

AUTHORS: Sixing Yu, Arya Mazaheri, Ali Jannesari

HIGHLIGHT: In this paper, we propose a novel multi-stage graph embedding technique based on graph neural networks (GNNs) to identify DNN topologies and use reinforcement learning (RL) to find a suitable compression policy.

1133, **TITLE:** The Combinatorial Brain Surgeon: Pruning Weights That Cancel One Another in Neural Networks
<https://proceedings.mlr.press/v162/yu22f.html>

AUTHORS: Xin Yu, Thiago Serra, Srikumar Ramalingam, Shandian Zhe

HIGHLIGHT: We propose a tractable heuristic for solving the combinatorial extension of OBS, in which we select weights for simultaneous removal, and we combine it with a single-pass systematic update of unpruned weights.

1134, **TITLE:** GraphFM: Improving Large-Scale GNN Training via Feature Momentum
<https://proceedings.mlr.press/v162/yu22g.html>

AUTHORS: Haiyang Yu, Limei Wang, Bokun Wang, Meng Liu, Tianbao Yang, Shuiwang Ji

HIGHLIGHT: Here, we propose a new technique, named feature momentum (FM), that uses a momentum step to incorporate historical embeddings when updating feature representations.

1135, **TITLE:** Latent Diffusion Energy-Based Model for Interpretable Text Modelling
<https://proceedings.mlr.press/v162/yu22h.html>

AUTHORS: Peiyu Yu, Sirui Xie, Xiaojian Ma, Baoxiong Jia, Bo Pang, Ruiqi Gao, Yixin Zhu, Song-Chun Zhu, Ying Nian Wu

HIGHLIGHT: Inspired by the recent efforts that leverage diffusion recovery likelihood learning as a cure for the sampling issue, we introduce a novel symbiosis between the diffusion models and latent space EBMs in a variational learning framework, coined as the latent diffusion energy-based model.

1136, **TITLE:** Predicting Out-of-Distribution Error with the Projection Norm
<https://proceedings.mlr.press/v162/yu22i.html>

AUTHORS: Yaodong Yu, Zitong Yang, Alexander Wei, Yi Ma, Jacob Steinhardt

HIGHLIGHT: We propose a metric—Projection Norm—to predict a model’s performance on out-of-distribution (OOD) data without access to ground truth labels.

1137, **TITLE:** Robust Task Representations for Offline Meta-Reinforcement Learning via Contrastive Learning
<https://proceedings.mlr.press/v162/yuan22a.html>

AUTHORS: Haoqi Yuan, Zongqing Lu

HIGHLIGHT: Existing offline meta-reinforcement learning algorithms cannot distinguish these factors, making task representations unstable to the change of behavior policies. To address this problem, we propose a contrastive learning framework for task representations that are robust to the distribution mismatch of behavior policies in training and test.

1138, **TITLE:** Provable Stochastic Optimization for Global Contrastive Learning: Small Batch Does Not Harm Performance
<https://proceedings.mlr.press/v162/yuan22b.html>

AUTHORS: Zhuoning Yuan, Yuexin Wu, Zi-Hao Qiu, Xianzhi Du, Lijun Zhang, Denny Zhou, Tianbao Yang

HIGHLIGHT: In this paper, we study contrastive learning from an optimization perspective, aiming to analyze and address a fundamental issue of existing contrastive learning methods that either rely on a large batch size or a large dictionary of feature vectors.

1139, **TITLE:** Neural Tangent Kernel Empowered Federated Learning
<https://proceedings.mlr.press/v162/yue22a.html>

AUTHORS: Kai Yue, Richeng Jin, Ryan Pilgrim, Chau-Wai Wong, Dror Baron, Huaiyu Dai

HIGHLIGHT: In this paper, we propose a novel FL paradigm empowered by the NTK framework.

1140, **TITLE:** Time Is MattEr: Temporal Self-supervision for Video Transformers
<https://proceedings.mlr.press/v162/yun22a.html>

AUTHORS: Sukmin Yun, Jaehyung Kim, Dongyoon Han, Hwanjun Song, Jung-Woo Ha, Jinwoo Shin

HIGHLIGHT: Based on the observations, we design simple yet effective self-supervised tasks for video models to learn temporal dynamics better.

1141, **TITLE:** Pure Noise to the Rescue of Insufficient Data: Improving Imbalanced Classification by Training on Random Noise Images
<https://proceedings.mlr.press/v162/zada22a.html>

AUTHORS: Shiran Zada, Itay Benou, Michal Irani

HIGHLIGHT: Despite remarkable progress on visual recognition tasks, deep neural-nets still struggle to generalize well when training data is scarce or highly imbalanced, rendering them extremely vulnerable to real-world examples. In this paper, we present a surprisingly simple yet highly effective method to mitigate this limitation: using pure noise images as additional training data.

1142, **TITLE:** Adaptive Conformal Predictions for Time Series
<https://proceedings.mlr.press/v162/zaffran22a.html>

AUTHORS: Margaux Zaffran, Olivier Feron, Yannig Goode, Julie Josse, Aymeric Dieuleveut

- HIGHLIGHT:** We propose a parameter-free method, AgACI, that adaptively builds upon ACI based on online expert aggregation.
- 1143, **TITLE:** Actor-Critic based Improper Reinforcement Learning
<https://proceedings.mlr.press/v162/zaki22a.html>
AUTHORS: Mohammadi Zaki, Avi Mohan, Aditya Gopalan, Shie Mannor
HIGHLIGHT: Towards this, we propose two algorithms: (1) a Policy Gradient-based approach; and (2) an algorithm that can switch between a simple Actor-Critic (AC) based scheme and a Natural Actor-Critic (NAC) scheme depending on the available information.
- 1144, **TITLE:** Stabilizing Q-learning with Linear Architectures for Provable Efficient Learning
<https://proceedings.mlr.press/v162/zanette22a.html>
AUTHORS: Andrea Zanette, Martin Wainwright
HIGHLIGHT: This work proposes an exploration variant of the basic Q-learning protocol with linear function approximation.
- 1145, **TITLE:** Multi Resolution Analysis (MRA) for Approximate Self-Attention
<https://proceedings.mlr.press/v162/zeng22a.html>
AUTHORS: Zhanpeng Zeng, Sourav Pal, Jeffery Kline, Glenn M Fung, Vikas Singh
HIGHLIGHT: In this paper, we revisit classical Multiresolution Analysis (MRA) concepts such as Wavelets, whose potential value in this setting remains underexplored thus far.
- 1146, **TITLE:** Efficient PAC Learning from the Crowd with Pairwise Comparisons
<https://proceedings.mlr.press/v162/zeng22b.html>
AUTHORS: Shiwei Zeng, Jie Shen
HIGHLIGHT: In this paper, we show that by leveraging the more easily acquired pairwise comparison queries, it is possible to exponentially reduce the label complexity while retaining the overall query complexity and runtime.
- 1147, **TITLE:** Multi-Grained Vision Language Pre-Training: Aligning Texts with Visual Concepts
<https://proceedings.mlr.press/v162/zeng22c.html>
AUTHORS: Yan Zeng, Xinsong Zhang, Hang Li
HIGHLIGHT: To this end, we propose a new method called X-VLM to perform ‘multi-grained vision language pre-training.’
- 1148, **TITLE:** Position Prediction as an Effective Pretraining Strategy
<https://proceedings.mlr.press/v162/zhai22a.html>
AUTHORS: Shuangfei Zhai, Navdeep Jaitly, Jason Ramapuram, Dan Busbridge, Tatiana Likhomanenko, Joseph Y Cheng, Walter Talbott, Chen Huang, Hanlin Goh, Joshua M Suskind
HIGHLIGHT: In this paper, we propose a novel, but surprisingly simple alternative to content reconstruction – that of predicting locations from content, without providing positional information for it.
- 1149, **TITLE:** Anytime Information Cascade Popularity Prediction via Self-Exciting Processes
<https://proceedings.mlr.press/v162/zhang22a.html>
AUTHORS: Xi Zhang, Akshay Aravamudan, Georgios C Anagnostopoulos
HIGHLIGHT: In this paper, for general, marked Hawkes point processes, we present closed-form expressions for the mean and variance of future event counts, conditioned on observed events.
- 1150, **TITLE:** Understanding Clipping for Federated Learning: Convergence and Client-Level Differential Privacy
<https://proceedings.mlr.press/v162/zhang22b.html>
AUTHORS: Xinwei Zhang, Xiangyi Chen, Mingyi Hong, Steven Wu, Jinfeng Yi
HIGHLIGHT: In this paper, we first empirically demonstrate that the clipped FedAvg can perform surprisingly well even with substantial data heterogeneity when training neural networks, which is partly because the clients’ updates become similar for several popular deep architectures. Based on this key observation, we provide the convergence analysis of a differential private (DP) FedAvg algorithm and highlight the relationship between clipping bias and the distribution of the clients’ updates.
- 1151, **TITLE:** Collaboration of Experts: Achieving 80% Top-1 Accuracy on ImageNet with 100M FLOPs
<https://proceedings.mlr.press/v162/zhang22c.html>
AUTHORS: Yikang Zhang, Zhuo Chen, Zhao Zhong
HIGHLIGHT: In this paper, we propose a Collaboration of Experts (CoE) framework to assemble the expertise of multiple networks towards a common goal.
- 1152, **TITLE:** PDE-Based Optimal Strategy for Unconstrained Online Learning
<https://proceedings.mlr.press/v162/zhang22d.html>
AUTHORS: Zhiyu Zhang, Ashok Cutkosky, Ioannis Paschalidis
HIGHLIGHT: To streamline this workflow, we present a framework that generates new potential functions by solving a Partial Differential Equation (PDE).
- 1153, **TITLE:** Stochastic Continuous Submodular Maximization: Boosting via Non-oblivious Function
<https://proceedings.mlr.press/v162/zhang22e.html>
AUTHORS: Qixin Zhang, Zengde Deng, Zaiyi Chen, Haoyuan Hu, Yu Yang
HIGHLIGHT: In this paper, we revisit Stochastic Continuous Submodular Maximization in both offline and online settings, which can benefit wide applications in machine learning and operations research areas.

- 1154, TITLE: When and How Mixup Improves Calibration
<https://proceedings.mlr.press/v162/zhang22f.html>
AUTHORS: Linjun Zhang, Zhun Deng, Kenji Kawaguchi, James Zou
HIGHLIGHT: In this paper, we theoretically prove that Mixup improves calibration in high-dimensional settings by investigating natural statistical models.
- 1155, TITLE: UAST: Uncertainty-Aware Siamese Tracking
<https://proceedings.mlr.press/v162/zhang22g.html>
AUTHORS: Dawei Zhang, Yanwei Fu, Zhonglong Zheng
HIGHLIGHT: We argue that these approaches lack a clear probabilistic explanation, so it is desirable to model the uncertainty and ambiguity representation of target estimation. To address this issue, this paper presents an Uncertainty-Aware Siamese Tracker (UAST) by developing a novel distribution-based regression formulation with localization uncertainty.
- 1156, TITLE: Examining Scaling and Transfer of Language Model Architectures for Machine Translation
<https://proceedings.mlr.press/v162/zhang22h.html>
AUTHORS: Biao Zhang, Behrooz Ghorbani, Ankur Bapna, Yong Cheng, Xavier Garcia, Jonathan Shen, Orhan Firat
HIGHLIGHT: In this work, we thoroughly examine the role of several architectural design choices on the performance of LMs on bilingual, (massively) multilingual and zero-shot translation tasks, under systematic variations of data conditions and model sizes.
- 1157, TITLE: Revisiting End-to-End Speech-to-Text Translation From Scratch
<https://proceedings.mlr.press/v162/zhang22i.html>
AUTHORS: Biao Zhang, Barry Haddow, Rico Sennrich
HIGHLIGHT: However, transcripts are not always available, and how significant such pretraining is for E2E ST has rarely been studied in the literature. In this paper, we revisit this question and explore the extent to which the quality of E2E ST trained on speech-translation pairs alone can be improved.
- 1158, TITLE: A Stochastic Multi-Rate Control Framework For Modeling Distributed Optimization Algorithms
<https://proceedings.mlr.press/v162/zhang22j.html>
AUTHORS: Xinwei Zhang, Mingyi Hong, Sairaj Dhople, Nicola Elia
HIGHLIGHT: This work offers a fresh perspective to model, analyze, and design distributed optimization algorithms through the lens of stochastic multi-rate feedback control.
- 1159, TITLE: GALAXY: Graph-based Active Learning at the Extreme
<https://proceedings.mlr.press/v162/zhang22k.html>
AUTHORS: Jifan Zhang, Julian Katz-Samuels, Robert Nowak
HIGHLIGHT: We propose a new strategy for active learning called GALAXY (Graph-based Active Learning At the eXtreme), which blends ideas from graph-based active learning and deep learning.
- 1160, TITLE: Fairness Interventions as (Dis)Incentives for Strategic Manipulation
<https://proceedings.mlr.press/v162/zhang22l.html>
AUTHORS: Xueru Zhang, Mohammad Mahdi Khalili, Kun Jin, Parinaz Naghizadeh, Mingyan Liu
HIGHLIGHT: Existing works have largely examined these as two separate issues, e.g., by focusing on building ML algorithms robust to strategic manipulation, or on training a fair ML algorithm. In this study, we set out to understand the impact they each have on the other, and examine how to characterize fair policies in the presence of strategic behavior.
- 1161, TITLE: Role-based Multiplex Network Embedding
<https://proceedings.mlr.press/v162/zhang22m.html>
AUTHORS: Hegui Zhang, Gang Kou
HIGHLIGHT: However, existing multiplex network embedding methods neglect structural role information, which can be used to determine the structural similarity between nodes. To overcome this shortcoming, this work proposes a simple, effective, role-based embedding method for multiplex networks, called RMNE.
- 1162, TITLE: Dynamic Topic Models for Temporal Document Networks
<https://proceedings.mlr.press/v162/zhang22n.html>
AUTHORS: Delvin Ce Zhang, Hady Lauw
HIGHLIGHT: While existing topic models focus on the dynamics of individual documents, we propose two neural topic models aimed at learning unified topic distributions that incorporate both document dynamics and network structure.
- 1163, TITLE: Personalized Federated Learning via Variational Bayesian Inference
<https://proceedings.mlr.press/v162/zhang22o.html>
AUTHORS: Xu Zhang, Yinchuan Li, Wenpeng Li, Kaiyang Guo, Yunfeng Shao
HIGHLIGHT: Federated learning faces huge challenges from model overfitting due to the lack of data and statistical diversity among clients. To address these challenges, this paper proposes a novel personalized federated learning method via Bayesian variational inference named pFedBayes.
- 1164, TITLE: Federated Learning with Label Distribution Skew via Logits Calibration
<https://proceedings.mlr.press/v162/zhang22p.html>
AUTHORS: Jie Zhang, Zhiqi Li, Bo Li, Jianghe Xu, Shuang Wu, Shouhong Ding, Chao Wu

- HIGHLIGHT:** In this paper, we investigate the label distribution skew in FL, where the distribution of labels varies across clients.
- 1165, **TITLE:** Neural Network Weights Do Not Converge to Stationary Points: An Invariant Measure Perspective
<https://proceedings.mlr.press/v162/zhang22q.html>
AUTHORS: Jingzhao Zhang, Haochuan Li, Suvrit Sra, Ali Jadbabaie
HIGHLIGHT: Remarkably, however, we observe that even though the weights do not converge to stationary points, the progress in minimizing the loss function halts and training loss stabilizes. Inspired by this observation, we propose a new perspective based on ergodic theory of dynamical systems to explain it.
- 1166, **TITLE:** Beyond Worst-Case Analysis in Stochastic Approximation: Moment Estimation Improves Instance Complexity
<https://proceedings.mlr.press/v162/zhang22r.html>
AUTHORS: Jingzhao Zhang, Hongzhou Lin, Subhro Das, Suvrit Sra, Ali Jadbabaie
HIGHLIGHT: We study oracle complexity of gradient based methods for stochastic approximation problems.
- 1167, **TITLE:** Deep and Flexible Graph Neural Architecture Search
<https://proceedings.mlr.press/v162/zhang22s.html>
AUTHORS: Wentao Zhang, Zheyu Lin, Yu Shen, Yang Li, Zhi Yang, Bin Cui
HIGHLIGHT: This paper proposes DFG-NAS, a novel method that searches for deep and flexible GNN architectures.
- 1168, **TITLE:** A Langevin-like Sampler for Discrete Distributions
<https://proceedings.mlr.press/v162/zhang22t.html>
AUTHORS: Ruqi Zhang, Xingchao Liu, Qiang Liu
HIGHLIGHT: We propose discrete Langevin proposal (DLP), a simple and scalable gradient-based proposal for sampling complex high-dimensional discrete distributions.
- 1169, **TITLE:** Rich Feature Construction for the Optimization-Generalization Dilemma
<https://proceedings.mlr.press/v162/zhang22u.html>
AUTHORS: Jianyu Zhang, David Lopez-Paz, Leon Bottou
HIGHLIGHT: We propose to initialize the networks with a rich representation containing a palette of potentially useful features, ready to be used by even simple models.
- 1170, **TITLE:** Generative Flow Networks for Discrete Probabilistic Modeling
<https://proceedings.mlr.press/v162/zhang22v.html>
AUTHORS: Dinghui Zhang, Nikolay Malkin, Zhen Liu, Alexandra Volokhova, Aaron Courville, Yoshua Bengio
HIGHLIGHT: We present energy-based generative flow networks (EB-GFN), a novel probabilistic modeling algorithm for high-dimensional discrete data.
- 1171, **TITLE:** Neurotoxin: Durable Backdoors in Federated Learning
<https://proceedings.mlr.press/v162/zhang22w.html>
AUTHORS: Zhengming Zhang, Ashwinee Panda, Linyue Song, Yaoqing Yang, Michael Mahoney, Prateek Mittal, Ramchandran Kannan, Joseph Gonzalez
HIGHLIGHT: We propose Neurotoxin, a simple one-line backdoor attack that functions by attacking parameters that are changed less in magnitude during training.
- 1172, **TITLE:** Making Linear MDPs Practical via Contrastive Representation Learning
<https://proceedings.mlr.press/v162/zhang22x.html>
AUTHORS: Tianjun Zhang, Tongzheng Ren, Mengjiao Yang, Joseph Gonzalez, Dale Schuurmans, Bo Dai
HIGHLIGHT: Instead, we consider an alternative definition of linear MDPs that automatically ensures normalization while allowing efficient representation learning via contrastive estimation.
- 1173, **TITLE:** NAFS: A Simple yet Tough-to-beat Baseline for Graph Representation Learning
<https://proceedings.mlr.press/v162/zhang22y.html>
AUTHORS: Wentao Zhang, Zeang Sheng, Mingyu Yang, Yang Li, Yu Shen, Zhi Yang, Bin Cui
HIGHLIGHT: In this paper, we present node-adaptive feature smoothing (NAFS), a simple non-parametric method that constructs node representations without parameter learning.
- 1174, **TITLE:** Correct-N-Contrast: a Contrastive Approach for Improving Robustness to Spurious Correlations
<https://proceedings.mlr.press/v162/zhang22z.html>
AUTHORS: Michael Zhang, Nimit S Sohoni, Hongyang R Zhang, Chelsea Finn, Christopher Re
HIGHLIGHT: To improve worst-group performance on spuriously correlated data without training attribute labels, we propose Correct-N-Contrast (CNC), a contrastive approach to directly learn representations robust to spurious correlations.
- 1175, **TITLE:** Efficient Reinforcement Learning in Block MDPs: A Model-free Representation Learning approach
<https://proceedings.mlr.press/v162/zhang22aa.html>
AUTHORS: Xuezhou Zhang, Yuda Song, Masatoshi Uehara, Mengdi Wang, Alekh Agarwal, Wen Sun
HIGHLIGHT: We present BRIEE, an algorithm for efficient reinforcement learning in Markov Decision Processes with block-structured dynamics (i.e., Block MDPs), where rich observations are generated from a set of unknown latent states.
- 1176, **TITLE:** Partial Counterfactual Identification from Observational and Experimental Data

- <https://proceedings.mlr.press/v162/zhang22ab.html>
AUTHORS: Junzhe Zhang, Jin Tian, Elias Bareinboim
HIGHLIGHT: This paper investigates the problem of bounding counterfactual queries from an arbitrary collection of observational and experimental distributions and qualitative knowledge about the underlying data-generating model represented in the form of a causal diagram.
- 1177, TITLE: Set Norm and Equivariant Skip Connections: Putting the Deep in Deep Sets
<https://proceedings.mlr.press/v162/zhang22ac.html>
AUTHORS: Lily Zhang, Veronica Tozzo, John Higgins, Rajesh Ranganath
HIGHLIGHT: In this work, we mathematically and empirically analyze normalization layers and residual connections in the context of deep permutation invariant neural networks. Based on our analysis, we propose Deep Sets++ and Set Transformer++, deep models that reach comparable or better performance than their original counterparts on a diverse suite of tasks.
- 1178, TITLE: Learning to Estimate and Refine Fluid Motion with Physical Dynamics
<https://proceedings.mlr.press/v162/zhang22ad.html>
AUTHORS: Mingrui Zhang, Jianhong Wang, James B Thomole, Matthew Piggott
HIGHLIGHT: Here we propose an unsupervised learning based prediction-correction scheme for fluid flow estimation.
- 1179, TITLE: A Branch and Bound Framework for Stronger Adversarial Attacks of ReLU Networks
<https://proceedings.mlr.press/v162/zhang22ae.html>
AUTHORS: Huan Zhang, Shiqi Wang, Kaidi Xu, Yihan Wang, Suman Jana, Cho-Jui Hsieh, Zico Kolter
HIGHLIGHT: In this work, we systematically search adversarial examples in the activation space of ReLU networks to tackle hard instances where none of the existing adversarial attacks succeed.
- 1180, TITLE: A Simple yet Universal Strategy for Online Convex Optimization
<https://proceedings.mlr.press/v162/zhang22af.html>
AUTHORS: Lijun Zhang, Guanghui Wang, Jinfeng Yi, Tianbao Yang
HIGHLIGHT: However, they need to design and optimize one surrogate loss for each type of functions, making it difficult to exploit the structure of the problem and utilize existing algorithms. In this paper, we propose a simple strategy for universal online convex optimization, which avoids these limitations.
- 1181, TITLE: Low-Precision Stochastic Gradient Langevin Dynamics
<https://proceedings.mlr.press/v162/zhang22ag.html>
AUTHORS: Ruqi Zhang, Andrew Gordon Wilson, Christopher De Sa
HIGHLIGHT: In this paper, we provide the first study of low-precision Stochastic Gradient Langevin Dynamics (SGLD), showing that its costs can be significantly reduced without sacrificing performance, due to its intrinsic ability to handle system noise.
- 1182, TITLE: Expression might be enough: representing pressure and demand for reinforcement learning based traffic signal control
<https://proceedings.mlr.press/v162/zhang22ah.html>
AUTHORS: Liang Zhang, Qiang Wu, Jun Shen, Linyuan L?, Bo Du, Jianqing Wu
HIGHLIGHT: In this paper, we (1) present a novel, flexible and efficient method, namely advanced max pressure (Advanced-MP), taking both running and queuing vehicles into consideration to decide whether to change current signal phase; (2) inventively design the traffic movement representation with the efficient pressure and effective running vehicles from Advanced-MP, namely advanced traffic state (ATS); and (3) develop a reinforcement learning (RL) based algorithm template, called Advanced-XLight, by combining ATS with the latest RL approaches, and generate two RL algorithms, namely "Advanced-MPLight" and "Advanced-CoLight"; from Advanced-XLight.
- 1183, TITLE: Uncertainty Modeling in Generative Compressed Sensing
<https://proceedings.mlr.press/v162/zhang22ai.html>
AUTHORS: Yilang Zhang, Mengchu Xu, Xiaojun Mao, Jian Wang
HIGHLIGHT: However, the reconstruction capability of the generative model is fundamentally limited by the range of its generator, typically a small subset of the signal space of interest. To break this bottleneck and thus reconstruct those out-of-range signals, this paper presents a novel method called CS-BGM that can effectively expand the range of generator.
- 1184, TITLE: Building Robust Ensembles via Margin Boosting
<https://proceedings.mlr.press/v162/zhang22aj.html>
AUTHORS: Dinghuai Zhang, Hongyang Zhang, Aaron Courville, Yoshua Bengio, Pradeep Ravikumar, Arun Sai Suggala
HIGHLIGHT: In this work, we take a principled approach towards building robust ensembles.
- 1185, TITLE: Revisiting and Advancing Fast Adversarial Training Through The Lens of Bi-Level Optimization
<https://proceedings.mlr.press/v162/zhang22ak.html>
AUTHORS: Yihua Zhang, Guanhua Zhang, Prashant Khanduri, Mingyi Hong, Shiyu Chang, Sijia Liu
HIGHLIGHT: In this paper, we advance Fast-AT from the fresh perspective of bi-level optimization (BLO).
- 1186, TITLE: Off-Policy Fitted Q-Evaluation with Differentiable Function Approximators: Z-Estimation and Inference Theory
<https://proceedings.mlr.press/v162/zhang22al.html>
AUTHORS: Ruiqi Zhang, Xuezhou Zhang, Chengzhuo Ni, Mengdi Wang

HIGHLIGHT: We focus on FQE with general differentiable function approximators, making our theory applicable to neural function approximations.

1187, **TITLE:** ROCK: Causal Inference Principles for Reasoning about Commonsense Causality
<https://proceedings.mlr.press/v162/zhang22am.html>

AUTHORS: Jiayao Zhang, Hongming Zhang, Weijie Su, Dan Roth

HIGHLIGHT: We propose a novel framework, ROCK, to Reason O(A)bout Commonsense K(C)ausality, which utilizes temporal signals as incidental supervision, and balances confounding effects using temporal propensities that are analogous to propensity scores.

1188, **TITLE:** No-Regret Learning in Time-Varying Zero-Sum Games
<https://proceedings.mlr.press/v162/zhang22an.html>

AUTHORS: Mengxiao Zhang, Peng Zhao, Haipeng Luo, Zhi-Hua Zhou

HIGHLIGHT: Learning from repeated play in a fixed two-player zero-sum game is a classic problem in game theory and online learning. We consider a variant of this problem where the game payoff matrix changes over time, possibly in an adversarial manner.

1189, **TITLE:** PLATON: Pruning Large Transformer Models with Upper Confidence Bound of Weight Importance
<https://proceedings.mlr.press/v162/zhang22ao.html>

AUTHORS: Qingru Zhang, Simiao Zuo, Chen Liang, Alexander Bukharin, Pengcheng He, Weizhu Chen, Tuo Zhao

HIGHLIGHT: To resolve this issue, we propose PLATON, which captures the uncertainty of importance scores by upper confidence bound of importance estimation.

1190, **TITLE:** NysADMM: faster composite convex optimization via low-rank approximation
<https://proceedings.mlr.press/v162/zhao22a.html>

AUTHORS: Shipu Zhao, Zachary Frangella, Madeleine Udell

HIGHLIGHT: This paper develops a scalable new algorithm, called NysADMM, to minimize a smooth convex loss function with a convex regularizer.

1191, **TITLE:** Toward Compositional Generalization in Object-Oriented World Modeling
<https://proceedings.mlr.press/v162/zhao22b.html>

AUTHORS: Linfeng Zhao, Lingzhi Kong, Robin Walters, Lawson L.S. Wong

HIGHLIGHT: We introduce a conceptual environment, Object Library, and two instances, and deploy a principled pipeline to measure the generalization ability.

1192, **TITLE:** Dynamic Regret of Online Markov Decision Processes
<https://proceedings.mlr.press/v162/zhao22c.html>

AUTHORS: Peng Zhao, Long-Fei Li, Zhi-Hua Zhou

HIGHLIGHT: We investigate online Markov Decision Processes (MDPs) with adversarially changing loss functions and known transitions.

1193, **TITLE:** Learning to Solve PDE-constrained Inverse Problems with Graph Networks
<https://proceedings.mlr.press/v162/zhao22d.html>

AUTHORS: Qingqing Zhao, David B Lindell, Gordon Wetzstein

HIGHLIGHT: In many application domains across science and engineering, however, we are not only interested in a forward simulation but also in solving inverse problems with constraints defined by a partial differential equation (PDE). Here we explore GNNs to solve such PDE-constrained inverse problems.

1194, **TITLE:** Learning from Counterfactual Links for Link Prediction
<https://proceedings.mlr.press/v162/zhao22e.html>

AUTHORS: Tong Zhao, Gang Liu, Daheng Wang, Wenhao Yu, Meng Jiang

HIGHLIGHT: However, the causal relationship between the two variables was largely ignored for learning to predict links on a graph. In this work, we visit this factor by asking a counterfactual question: “would the link still exist if the graph structure became different from observation?”

1195, **TITLE:** Global Optimization Networks
<https://proceedings.mlr.press/v162/zhao22f.html>

AUTHORS: Sen Zhao, Erez Louidor, Maya Gupta

HIGHLIGHT: We propose to fit a new type of function called a global optimization network (GON), defined as any composition of an invertible function and a unimodal function, whose unique global maximizer can be inferred in $\mathcal{O}(D)$ time, and used as the estimate.

1196, **TITLE:** Certified Robustness Against Natural Language Attacks by Causal Intervention
<https://proceedings.mlr.press/v162/zhao22g.html>

AUTHORS: Haiteng Zhao, Chang Ma, Xinshuai Dong, Anh Tuan Luu, Zhi-Hong Deng, Hanwang Zhang

HIGHLIGHT: This paper follows a causal perspective to look into the adversarial vulnerability and proposes Causal Intervention by Semantic Smoothing (CISS), a novel framework towards robustness against natural language attacks.

1197, **TITLE:** Efficient Learning for AlphaZero via Path Consistency
<https://proceedings.mlr.press/v162/zhao22h.html>

- AUTHORS: Dengwei Zhao, Shikui Tu, Lei Xu
HIGHLIGHT: This paper aims at building powerful models under a limited amount of self-plays which can be utilized by a human throughout the lifetime.
- 1198, TITLE: Penalizing Gradient Norm for Efficiently Improving Generalization in Deep Learning
<https://proceedings.mlr.press/v162/zhao22i.html>
AUTHORS: Yang Zhao, Hao Zhang, Xiuyuan Hu
HIGHLIGHT: In this paper, we propose an effective method to improve the model generalization by additionally penalizing the gradient norm of loss function during optimization.
- 1199, TITLE: Ripple Attention for Visual Perception with Sub-quadratic Complexity
<https://proceedings.mlr.press/v162/zheng22a.html>
AUTHORS: Lin Zheng, Huijie Pan, Lingpeng Kong
HIGHLIGHT: To bridge the gap, we propose ripple attention, a sub-quadratic attention mechanism for vision transformers.
- 1200, TITLE: Linear Complexity Randomized Self-attention Mechanism
<https://proceedings.mlr.press/v162/zheng22b.html>
AUTHORS: Lin Zheng, Chong Wang, Lingpeng Kong
HIGHLIGHT: By combining the expressiveness in RA and the efficiency in RFA, we develop a novel linear complexity self-attention mechanism called linear randomized attention (LARA).
- 1201, TITLE: Online Decision Transformer
<https://proceedings.mlr.press/v162/zheng22c.html>
AUTHORS: Qinqing Zheng, Amy Zhang, Aditya Grover
HIGHLIGHT: We propose Online Decision Transformers (ODT), an RL algorithm based on sequence modeling that blends offline pretraining with online finetuning in a unified framework.
- 1202, TITLE: Learning Efficient and Robust Ordinary Differential Equations via Invertible Neural Networks
<https://proceedings.mlr.press/v162/zhi22a.html>
AUTHORS: Weiming Zhi, Tin Lai, Lionel Ott, Edwin V. Bonilla, Fabio Ramos
HIGHLIGHT: We propose to learn an ODE of interest from data by viewing its dynamics as a vector field related to another base vector field via a diffeomorphism (i.e., a differentiable bijection), represented by an invertible neural network (INN).
- 1203, TITLE: HyperTransformer: Model Generation for Supervised and Semi-Supervised Few-Shot Learning
<https://proceedings.mlr.press/v162/zhmoginov22a.html>
AUTHORS: Andrey Zhmoginov, Mark Sandler, Maksym Vladymyrov
HIGHLIGHT: In this work we propose a HyperTransformer, a Transformer-based model for supervised and semi-supervised few-shot learning that generates weights of a convolutional neural network (CNN) directly from support samples.
- 1204, TITLE: Describing Differences between Text Distributions with Natural Language
<https://proceedings.mlr.press/v162/zhong22a.html>
AUTHORS: Ruiqi Zhong, Charlie Snell, Dan Klein, Jacob Steinhardt
HIGHLIGHT: We propose to automatically summarize the differences by “learning a natural language hypothesis” given two distributions $\mathcal{D}_{\{0\}}$ and $\mathcal{D}_{\{1\}}$, we search for a description that is more often true for $\mathcal{D}_{\{1\}}$, e.g., “is military-related.”
- 1205, TITLE: Pessimistic Minimax Value Iteration: Provably Efficient Equilibrium Learning from Offline Datasets
<https://proceedings.mlr.press/v162/zhong22b.html>
AUTHORS: Han Zhong, Wei Xiong, Jiyuan Tan, Liwei Wang, Tong Zhang, Zhaoran Wang, Zhuoran Yang
HIGHLIGHT: We propose a pessimism-based algorithm, dubbed as pessimistic minimax value iteration (PMVI), which overcomes the distributional shift by constructing pessimistic estimates of the value functions for both players and outputs a policy pair by solving a correlated coarse equilibrium based on the two value functions.
- 1206, TITLE: Dimension-free Complexity Bounds for High-order Nonconvex Finite-sum Optimization
<https://proceedings.mlr.press/v162/zhou22a.html>
AUTHORS: Dongruo Zhou, Quanquan Gu
HIGHLIGHT: In this paper, we show that the polylogarithmic dimension dependence gap is not essential and can be closed.
- 1207, TITLE: A Hierarchical Bayesian Approach to Inverse Reinforcement Learning with Symbolic Reward Machines
<https://proceedings.mlr.press/v162/zhou22b.html>
AUTHORS: Weichao Zhou, Wenchao Li
HIGHLIGHT: We propose a hierarchical Bayesian approach for inferring the most likely assignments such that the concretized reward machine can discriminate expert demonstrated trajectories from other trajectories with high accuracy.
- 1208, TITLE: On the Optimization Landscape of Neural Collapse under MSE Loss: Global Optimality with Unconstrained Features
<https://proceedings.mlr.press/v162/zhou22c.html>
AUTHORS: Jinxin Zhou, Xiao Li, Tianyu Ding, Chong You, Qing Qu, Zhihui Zhu
HIGHLIGHT: In this work, we justify NC under the mean squared error (MSE) loss, where recent empirical evidence shows that it performs comparably or even better than the de-facto cross-entropy loss.

- 1209, TITLE: Model Agnostic Sample Reweighting for Out-of-Distribution Learning
<https://proceedings.mlr.press/v162/zhou22d.html>
AUTHORS: Xiao Zhou, Yong Lin, Renjie Pi, Weizhong Zhang, Renzhe Xu, Peng Cui, Tong Zhang
HIGHLIGHT: This work proposes a principled method, Model Agnostic samPLe rEweighting (MAPLE), to effectively address OOD problem, especially in overparameterized scenarios.
- 1210, TITLE: Sparse Invariant Risk Minimization
<https://proceedings.mlr.press/v162/zhou22e.html>
AUTHORS: Xiao Zhou, Yong Lin, Weizhong Zhang, Tong Zhang
HIGHLIGHT: In this paper, we propose a simple yet effective paradigm named Sparse Invariant Risk Minimization (SparseIRM) to address this contradiction.
- 1211, TITLE: Prototype-Anchored Learning for Learning with Imperfect Annotations
<https://proceedings.mlr.press/v162/zhou22f.html>
AUTHORS: Xiong Zhou, Xianming Liu, Deming Zhai, Junjun Jiang, Xin Gao, Xiangyang Ji
HIGHLIGHT: In this work, we thoroughly investigate the popular softmax loss and margin-based loss, and offer a feasible approach to tighten the generalization error bound by maximizing the minimal sample margin.
- 1212, TITLE: FEDformer: Frequency Enhanced Decomposed Transformer for Long-term Series Forecasting
<https://proceedings.mlr.press/v162/zhou22g.html>
AUTHORS: Tian Zhou, Ziqing Ma, Qingsong Wen, Xue Wang, Liang Sun, Rong Jin
HIGHLIGHT: To address these problems, we propose to combine Transformer with the seasonal-trend decomposition method, in which the decomposition method captures the global profile of time series while Transformers capture more detailed structures.
- 1213, TITLE: Probabilistic Bilevel Coreset Selection
<https://proceedings.mlr.press/v162/zhou22h.html>
AUTHORS: Xiao Zhou, Renjie Pi, Weizhong Zhang, Yong Lin, Zonghao Chen, Tong Zhang
HIGHLIGHT: In this work, for the first time we propose a continuous probabilistic bilevel formulation of coreset selection by learning a probabilistic weight for each training sample.
- 1214, TITLE: Approximate Frank-Wolfe Algorithms over Graph-structured Support Sets
<https://proceedings.mlr.press/v162/zhou22i.html>
AUTHORS: Baojian Zhou, Yifan Sun
HIGHLIGHT: In this paper, we consider approximate Frank-Wolfe (FW) algorithms to solve convex optimization problems over graph-structured support sets where the linear minimization oracle (LMO) cannot be efficiently obtained in general.
- 1215, TITLE: Improving Adversarial Robustness via Mutual Information Estimation
<https://proceedings.mlr.press/v162/zhou22j.html>
AUTHORS: Dawei Zhou, Nannan Wang, Xinbo Gao, Bo Han, Xiaoyu Wang, Yibing Zhan, Tongliang Liu
HIGHLIGHT: They are typically misled by adversarial samples to make wrong predictions. To alleviate this negative effect, in this paper, we investigate the dependence between outputs of the target model and input adversarial samples from the perspective of information theory, and propose an adversarial defense method.
- 1216, TITLE: Modeling Adversarial Noise for Adversarial Training
<https://proceedings.mlr.press/v162/zhou22k.html>
AUTHORS: Dawei Zhou, Nannan Wang, Bo Han, Tongliang Liu
HIGHLIGHT: Motivated by the fact that adversarial noise contains well-generalizing features and that the relationship between adversarial data and natural data can help infer natural data and make reliable predictions, in this paper, we study to model adversarial noise by learning the transition relationship between adversarial labels (i.e. the flipped labels used to generate adversarial data) and natural labels (i.e. the ground truth labels of the natural data).
- 1217, TITLE: Contrastive Learning with Boosted Memorization
<https://proceedings.mlr.press/v162/zhou22l.html>
AUTHORS: Zhihan Zhou, Jiangchao Yao, Yan-Feng Wang, Bo Han, Ya Zhang
HIGHLIGHT: Different from previous works, we explore this direction from an alternative perspective, i.e., the data perspective, and propose a novel Boosted Contrastive Learning (BCL) method.
- 1218, TITLE: Understanding The Robustness in Vision Transformers
<https://proceedings.mlr.press/v162/zhou22m.html>
AUTHORS: Daquan Zhou, Zhiding Yu, Enze Xie, Chaowei Xiao, Animashree Anandkumar, Jiashi Feng, Jose M. Alvarez
HIGHLIGHT: In this paper, we examine the role of self-attention in learning robust representations.
- 1219, TITLE: VLUE: A Multi-Task Multi-Dimension Benchmark for Evaluating Vision-Language Pre-training
<https://proceedings.mlr.press/v162/zhou22n.html>
AUTHORS: Wangchunshu Zhou, Yan Zeng, Shizhe Diao, Xinsong Zhang
HIGHLIGHT: To this end, we introduce the Vision-Language Understanding Evaluation (VLUE) benchmark, a multi-task multi-dimension benchmark for evaluating the generalization capabilities and the efficiency-performance trade-off ("Pareto SOTA") of VLP models.

- 1220, TITLE: Detecting Corrupted Labels Without Training a Model to Predict
<https://proceedings.mlr.press/v162/zhu22a.html>
AUTHORS: Zhaowei Zhu, Zihao Dong, Yang Liu
HIGHLIGHT: In this paper, from a more data-centric perspective, we propose a training-free solution to detect corrupted labels.
- 1221, TITLE: Contextual Bandits with Large Action Spaces: Made Practical
<https://proceedings.mlr.press/v162/zhu22b.html>
AUTHORS: Yinglun Zhu, Dylan J Foster, John Langford, Paul Mineiro
HIGHLIGHT: We present the first efficient, general-purpose algorithm for contextual bandits with continuous, linearly structured action spaces.
- 1222, TITLE: Neural-Symbolic Models for Logical Queries on Knowledge Graphs
<https://proceedings.mlr.press/v162/zhu22c.html>
AUTHORS: Zhaocheng Zhu, Mikhail Galkin, Zuobai Zhang, Jian Tang
HIGHLIGHT: In this paper, we propose Graph Neural Network Query Executor (GNN-QE), a neural-symbolic model that enjoys the advantages of both worlds.
- 1223, TITLE: Topology-aware Generalization of Decentralized SGD
<https://proceedings.mlr.press/v162/zhu22d.html>
AUTHORS: Tongtian Zhu, Fengxiang He, Lan Zhang, Zhengyang Niu, Mingli Song, Dacheng Tao
HIGHLIGHT: This paper studies the algorithmic stability and generalizability of decentralized stochastic gradient descent (D-SGD).
- 1224, TITLE: Resilient and Communication Efficient Learning for Heterogeneous Federated Systems
<https://proceedings.mlr.press/v162/zhu22e.html>
AUTHORS: Zhuangdi Zhu, Junyuan Hong, Steve Drew, Jiayu Zhou
HIGHLIGHT: In this work, we propose an FL scheme to address both challenges simultaneously.
- 1225, TITLE: On Numerical Integration in Neural Ordinary Differential Equations
<https://proceedings.mlr.press/v162/zhu22f.html>
AUTHORS: Aiqing Zhu, Pengzhan Jin, Beibei Zhu, Yifa Tang
HIGHLIGHT: In this paper, we propose the inverse modified differential equations (IMDE) to clarify the influence of numerical integration on training Neural ODE models.
- 1226, TITLE: When AUC meets DRO: Optimizing Partial AUC for Deep Learning with Non-Convex Convergence Guarantee
<https://proceedings.mlr.press/v162/zhu22g.html>
AUTHORS: Dixian Zhu, Gang Li, Bokun Wang, Xiaodong Wu, Tianbao Yang
HIGHLIGHT: In this paper, we propose systematic and efficient gradient-based methods for both one-way and two-way partial AUC (pAUC) maximization that are applicable to deep learning.
- 1227, TITLE: Contextual Bandits with Smooth Regret: Efficient Learning in Continuous Action Spaces
<https://proceedings.mlr.press/v162/zhu22h.html>
AUTHORS: Yinglun Zhu, Paul Mineiro
HIGHLIGHT: We propose a smooth regret notion for contextual bandits, which dominates previously proposed alternatives.
- 1228, TITLE: Residual-Based Sampling for Online Outlier-Robust PCA
<https://proceedings.mlr.press/v162/zhu22i.html>
AUTHORS: Tianhao Zhu, Jie Shen
HIGHLIGHT: In this paper, we study online ORPCA, an important variant that addresses the practical challenge that the data points arrive in a sequential manner and the goal is to recover the underlying subspace of the clean data with one pass of the data.
- 1229, TITLE: Region-Based Semantic Factorization in GANs
<https://proceedings.mlr.press/v162/zhu22j.html>
AUTHORS: Jiapeng Zhu, Yujun Shen, Yinghao Xu, Deli Zhao, Qifeng Chen
HIGHLIGHT: In this work, we present a highly efficient algorithm to factorize the latent semantics learned by GANs concerning an arbitrary image region.
- 1230, TITLE: Beyond Images: Label Noise Transition Matrix Estimation for Tasks with Lower-Quality Features
<https://proceedings.mlr.press/v162/zhu22k.html>
AUTHORS: Zhaowei Zhu, Jialu Wang, Yang Liu
HIGHLIGHT: We observe that tasks with lower-quality features fail to meet the anchor-point or clusterability condition, due to the coexistence of both uninformative and informative representations. To handle this issue, we propose a generic and practical information-theoretic approach to down-weight the less informative parts of the lower-quality features.
- 1231, TITLE: Towards Uniformly Superhuman Autonomy via Subdominance Minimization
<https://proceedings.mlr.press/v162/ziebart22a.html>
AUTHORS: Brian Ziebart, Sanjiban Choudhury, Xinyan Yan, Paul Vernaza
HIGHLIGHT: We instead assume demonstrations are of varying quality and seek to induce behavior that is unambiguously better (i.e., Pareto dominant or minimally subdominant) than all human demonstrations.

1232, TITLE: Inductive Matrix Completion: No Bad Local Minima and a Fast Algorithm

<https://proceedings.mlr.press/v162/zilber22a.html>

AUTHORS: Pini Zilber, Boaz Nadler

HIGHLIGHT: In this work, we make three contributions to the IMC problem: (i) we prove that under suitable conditions, the IMC optimization landscape has no bad local minima; (ii) we derive a simple scheme with theoretical guarantees to estimate the rank of the unknown matrix; and (iii) we propose GNIMC, a simple Gauss-Newton based method to solve the IMC problem, analyze its runtime and derive for it strong recovery guarantees.

1233, TITLE: Counterfactual Prediction for Outcome-Oriented Treatments

<https://proceedings.mlr.press/v162/zou22a.html>

AUTHORS: Hao Zou, Bo Li, Jiangang Han, Shuiping Chen, Xuetao Ding, Peng Cui

HIGHLIGHT: To overcome it, we establish a new objective of optimizing counterfactual prediction on outcome-oriented treatments, propose a novel Outcome-Oriented Sample Re-weighting (OOSR) method to make the predictive model concentrate more on outcome-oriented treatments, and theoretically analyze that our method can improve treatment selection towards the optimal one.

1234, TITLE: SpaceMAP: Visualizing High-Dimensional Data by Space Expansion

<https://proceedings.mlr.press/v162/zu22a.html>

AUTHORS: Xinrui Zu, Qian Tao

HIGHLIGHT: However, there exist intriguing, non-intuitive discrepancies between the geometry of high- and low-dimensional space. We look into such discrepancies and propose a novel visualization method called Space-based Manifold Approximation and Projection (SpaceMAP).