

# **International Conference on Machine Learning (ICML 2022)**

Proceedings of Machine Learning Research Volume 162

Baltimore, Maryland, USA  
17-23 July 2022

Part 1 of 33

## **Editors:**

**Kamalika Chaudhuri  
Stefanie Jegelka  
Le Song**

**Csaba Szepesvari  
Gang Niu  
Sivan Sabato**

ISBN: 978-1-7138-7101-9

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2022) by the Authors  
All rights reserved.

Printed with permission by Curran Associates, Inc. (2023)

For permission requests, please contact International Machine Learning Society (IMLS)  
at the address below.

International Machine Learning Society (IMLS)  
c/o Priscilla Rasmussen  
209 North Eighth Street  
Stroudsburg PA 18360  
USA

Phone: (570) 476-8006  
Fax: (570) 476-0860

rasmussen@ptd.net

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2633  
Email: curran@proceedings.com  
Web: www.proceedings.com

# TABLE OF CONTENTS

## PART 1

PAC-Bayesian Bounds on Rate-Efficient Classifiers .....	1
<i>Alhabib Abbas, Yiannis Andreopoulos</i>	
Sharp-MAML: Sharpness-Aware Model-Agnostic Meta Learning.....	10
<i>Momin Abbas, Quan Xiao, Lisha Chen, Pin-Yu Chen, Tianyi Chen</i>	
An Initial Alignment Between Neural Network and Target is Needed for Gradient Descent to Learn .....	33
<i>Emmanuel Abbe, Elisabetta Cornacchia, Jan Hazla, Christopher Marquis</i>	
Active Sampling for Min-Max Fairness .....	53
<i>Jacob D Abernethy, Pranjal Awasthi, Matthäus Kleindessner, Jamie Morgenstern, Chris Russell, Jie Zhang</i>	
Meaningfully Debugging Model Mistakes Using Conceptual Counterfactual Explanations .....	66
<i>Abubakar Abid, Mert Yuksekgonul, James Zou</i>	
Batched Dueling Bandits .....	89
<i>Arpit Agarwal, Rohan Ghuge, Viswanath Nagarajan</i>	
Hierarchical Shrinkage: Improving the Accuracy and Interpretability of Tree-Based Models.....	111
<i>Abhineet Agarwal, Yan Shuo Tan, Omer Ronen, Chandan Singh, Bin Yu</i>	
Deep Equilibrium Networks Are Sensitive to Initialization Statistics .....	136
<i>Atish Agarwala, Samuel S Schoenholz</i>	
Learning of Cluster-Based Feature Importance for Electronic Health Record Time-Series.....	161
<i>Henrique Aguiar, Mauro Santos, Peter Watkinson, Tingting Zhu</i>	
On the Convergence of the Shapley Value in Parametric Bayesian Learning Games .....	180
<i>Lucas Agussurja, Xinyi Xu, Bryan Kian Hsiang Low</i>	
Individual Preference Stability for Clustering .....	197
<i>Saba Ahmadi, Pranjal Awasthi, Samir Khuller, Matthäus Kleindessner, Jamie Morgenstern, Pattara Sukprasert, Ali Vakilian</i>	
Understanding the Unstable Convergence of Gradient Descent.....	247
<i>Kwangjun Ahn, Jingzhao Zhang, Suvrit Sra</i>	
Minimum Cost Intervention Design for Causal Effect Identification.....	258
<i>Sina Akbari, Jalal Etesami, Negar Kiyavash</i>	
How Faithful is Your Synthetic Data? Sample-Level Metrics for Evaluating and Auditing Generative Models.....	290
<i>Ahmed Alaa, Boris Van Breugel, Evgeny S. Saveliev, Mihaela Van Der Schaar</i>	
A Natural Actor-Critic Framework for Zero-Sum Markov Games.....	307
<i>Ahmet Alacaoglu, Luca Viano, Niao He, Volkan Cevher</i>	
Deploying Convolutional Networks on Untrusted Platforms Using 2D Holographic Reduced Representations .....	367
<i>Mohammad Mahmudul Alam, Edward Raff, Tim Oates, James Holt</i>	
Optimistic Linear Support and Successor Features as a Basis for Optimal Policy Transfer .....	394
<i>Lucas Nunes Alegre, Ana Bazzan, Bruno C. Da Silva</i>	
Structured Stochastic Gradient MCMC.....	414
<i>Antonios Alexos, Alex J Boyd, Stephan Mandt</i>	
XAI for Transformers: Better Explanations Through Conservative Propagation.....	435
<i>Ameen Ali, Thomas Schnake, Oliver Eberle, Grégoire Montavon, Klaus-Robert Müller, Lior Wolf</i>	
RUMs from Head-To-Head Contests .....	452
<i>Matteo Almanza, Flavio Chierichetti, Ravi Kumar, Alessandro Panconesi, Andrew Tomkins</i>	

Neuro-Symbolic Language Modeling with Automaton-Augmented Retrieval .....	468
<i>Uri Alon, Frank Xu, Junxian He, Sudipta Sengupta, Dan Roth, Graham Neubig</i>	
Minimax Classification Under Concept Drift with Multidimensional Adaptation and Performance Guarantees .....	486
<i>Verónica Álvarez, Santiago Mazuelas, Jose A Lozano</i>	
Scalable First-Order Bayesian Optimization Via Structured Automatic Differentiation .....	500
<i>Sebastian E Ament, Carla P Gomes</i>	
Public Data-Assisted Mirror Descent for Private Model Training .....	517
<i>Ehsan Amid, Arun Ganesh, Rajiv Mathews, Swaroop Ramaswamy, Shuang Song, Thomas Steinke, Thomas Steinke, Vinith M Suriyakumar, Om Thakkar, Abhradeep Thakurta</i>	
On Last-Iterate Convergence Beyond Zero-Sum Games .....	536
<i>Ioannis Anagnostides, Ioannis Panageas, Gabriele Farina, Tuomas Sandholm</i>	
Online Algorithms with Multiple Predictions.....	582
<i>Keerti Anand, Rong Ge, Amit Kumar, Debmalya Panigrahi</i>	
Learning to Hash Robustly, Guaranteed.....	599
<i>Alexandr Andoni, Daniel Beaglehole</i>	
Set Based Stochastic Subsampling.....	619
<i>Bruno Andreis, Seanie Lee, A. Tuan Nguyen, Juho Lee, Eunho Yang, Sung Ju Hwang</i>	
Towards Understanding Sharpness-Aware Minimization.....	639
<i>Maksym Andriushchenko, Nicolas Flammarion</i>	
Fair and Fast k-Center Clustering for Data Summarization .....	669
<i>Haris Angelidakis, Adam Kurpisz, Leon Sering, Rico Zenklusen</i>	
Interactive Correlation Clustering with Existential Cluster Constraints.....	703
<i>Rico Angell, Nicholas Monath, Nishant Yadav, Andrew McCallum</i>	
Image-To-Image Regression with Distribution-Free Uncertainty Quantification and Applications in Imaging.....	717
<i>Anastasios N Angelopoulos, Amit Pal Kohli, Stephen Bates, Michael Jordan, Jitendra Malik, Thayer Alshaabi, Srigokul Upadhyayula, Yaniv Romano</i>	
AdaGrad Avoids Saddle Points.....	731
<i>Kimon Antonakopoulos, Panayotis Mertikopoulos, Georgios Piliouras, Xiao Wang</i>	

## PART 2

UnderGrad: A Universal Black-Box Optimization Method with Almost Dimension-Free Convergence Rate Guarantees .....	772
<i>Kimon Antonakopoulos, Dong Quan Vu, Volkan Cevher, Kfir Levy, Panayotis Mertikopoulos</i>	
Adapting the Linearised Laplace Model Evidence for Modern Deep Learning .....	796
<i>Javier Antoran, David Janz, James U Allingham, Erik Daxberger, Riccardo Rb Barbano, Eric Nalisnick, Jose Miguel Hernandez-Lobato</i>	
EAT-C: Environment-Adversarial sub-Task Curriculum for Efficient Reinforcement Learning .....	822
<i>Shuang Ao, Tianyi Zhou, Jing Jiang, Guodong Long, Xuan Song, Chengqi Zhang</i>	
Online Balanced Experimental Design.....	844
<i>David Arbour, Drew Dimmery, Tung Mai, Anup Rao</i>	
VariGrow: Variational Architecture Growing for Task-Agnostic Continual Learning Based on Bayesian Novelty .....	865
<i>Randy Ardywibowo, Zepeng Huo, Zhangyang Wang, Bobak J Mortazavi, Shuai Huang, Xiaoning Qian</i>	
Thresholded Lasso Bandit .....	878
<i>Kaito Ariu, Kenshi Abe, Alexandre Proutiere</i>	
Gradient Based Clustering.....	929
<i>Aleksandar Armacki, Dragana Bajovic, Dusan Jakovetic, Soumya Kar</i>	

Understanding Gradient Descent on the Edge of Stability in Deep Learning .....	948
<i>Sanjeev Arora, Zhiyuan Li, Abhishek Panigrahi</i>	
Private Optimization in the Interpolation Regime: Faster Rates and Hardness Results .....	1025
<i>Hilal Asi, Karan Chadha, Gary Cheng, John Duchi</i>	
Optimal Algorithms for Mean Estimation Under Local Differential Privacy.....	1046
<i>Hilal Asi, Vitaly Feldman, Kunal Talwar</i>	
Asymptotically-Optimal Gaussian Bandits with Side Observations .....	1057
<i>Alexia Atsidakou, Orestis Papadigenopoulos, Constantine Caramanis, Sujay Sanghavi, Sanjay Shakkottai</i>	
Congested Bandits: Optimal Routing Via Short-Term Resets .....	1078
<i>Pranjal Awasthi, Kush Bhatia, Sreenivas Gollapudi, Kostas Kollias</i>	
Do More Negative Samples Necessarily Hurt in Contrastive Learning? .....	1101
<i>Pranjal Awasthi, Nishanth Dikkala, Pritish Kamath</i>	
H-Consistency Bounds for Surrogate Loss Minimizers .....	1117
<i>Pranjal Awasthi, Anqi Mao, Mehryar Mohri, Yutao Zhong</i>	
Iterative Hard Thresholding with Adaptive Regularization: Sparser Solutions Without Sacrificing Runtime .....	1175
<i>Kyriakos Axiotis, Maxim Sviridenko</i>	
Proving Theorems Using Incremental Learning and Hindsight Experience Replay .....	1198
<i>Eser Aygün, Ankit Anand, Laurent Orseau, Xavier Glorot, Stephen M McAleer, Vlad Firoiu, Lei M Zhang, Doina Precup, Shihab Mourad</i>	
Near-Optimal Rate of Consistency for Linear Models with Missing Values.....	1211
<i>Alexis Ayme, Claire Boyer, Aymeric Dieuleveut, Erwan Scornet</i>	
How Tempering Fixes Data Augmentation in Bayesian Neural Networks.....	1244
<i>Gregor Bachmann, Lorenzo Noci, Thomas Hofmann</i>	
ASAP.SGD: Instance-Based Adaptiveness to Staleness in Asynchronous SGD .....	1261
<i>Karl Bäckström, Marina Papatriantafilou, Philippas Tsigas</i>	
From Noisy Prediction to True Label: Noisy Prediction Calibration Via Generative Model .....	1277
<i>Heesun Bae, Seungjae Shin, Byeonghu Na, Joonho Jang, Kyungwoo Song, Il-Chul Moon</i>	
Data2vec: A General Framework for Self-Supervised Learning in Speech, Vision and Language .....	1298
<i>Alexei Baevski, Wei-Ning Hsu, Qiantong Xu, Arun Babu, Jiatao Gu, Michael Auli</i>	
End-To-End Balancing for Causal Continuous Treatment-Effect Estimation .....	1313
<i>Taha Bahadori, Eric Tchetgen Tchetgen, David Heckerman</i>	
A Hierarchical Transitive-Aligned Graph Kernel for Un-Attributed Graphs .....	1327
<i>Lu Bai, Lixin Cui, Hancock Edwin</i>	
Near-Optimal Learning of Extensive-Form Games with Imperfect Information .....	1337
<i>Yu Bai, Chi Jin, Song Mei, Tiancheng Yu</i>	
Gaussian Mixture Variational Autoencoder with Contrastive Learning for Multi-Label Classification .....	1383
<i>Junwen Bai, Shufeng Kong, Carla P Gomes</i>	
A3T: Alignment-Aware Acoustic and Text Pretraining for Speech Synthesis and Editing.....	1399
<i>He Bai, Renjie Zheng, Junkun Chen, Mingbo Ma, Xintong Li, Liang Huang</i>	
Stability Based Generalization Bounds for Exponential Family Langevin Dynamics .....	1412
<i>Arindam Banerjee, Tiancong Chen, Xinyan Li, Yingxue Zhou</i>	
Certified Neural Network Watermarks with Randomized Smoothing.....	1450
<i>Arpit Bansal, Ping-Yeh Chiang, Michael J Curry, Rajiv Jain, Curtis Wigington, Varun Manjunatha, John P Dickerson, Tom Goldstein</i>	
Data Scaling Laws in NMT: The Effect of Noise and Architecture .....	1466
<i>Yamini Bansal, Behrooz Ghorbani, Ankush Garg, Biao Zhang, Colin Cherry, Behnam Neyshabur, Orhan Firat</i>	
Learning Stable Classifiers by Transferring Unstable Features.....	1483
<i>Yujia Bao, Shiyu Chang, Dr. Regina Barzilay</i>	

Fast Composite Optimization and Statistical Recovery in Federated Learning .....	1508
<i>Yajie Bao, Michael Crawshaw, Shan Luo, Mingrui Liu</i>	
Generative Modeling for Multi-Task Visual Learning .....	1537
<i>Zhipeng Bao, Martial Hebert, Yu-Xiong Wang</i>	
Estimating the Optimal Covariance with Imperfect Mean in Diffusion Probabilistic Models .....	1555
<i>Fan Bao, Chongxuan Li, Jiacheng Sun, Jun Zhu, Bo Zhang</i>	

### PART 3

On the Surrogate Gap Between Contrastive and Supervised Losses.....	1585
<i>Han Bao, Yoshihiro Nagano, Kento Nozawa</i>	
Representation Topology Divergence: A Method for Comparing Neural Network Representations. ....	1607
<i>Serguei Barannikov, Ilya Trofimov, Nikita Balabin, Evgeny Burnaev</i>	
Sparse Mixed Linear Regression with Guarantees: Taming an Intractable Problem with Inexact Relaxation.....	1627
<i>Adarsh Barik, Jean Honorio</i>	
Neural Fisher Discriminant Analysis: Optimal Neural Network Embeddings in Polynomial Time.....	1647
<i>Burak Bartan, Mert Pilanci</i>	
Fictitious Play and Best-Response Dynamics in Identical Interest and Zero-Sum Stochastic Games .....	1664
<i>Lucas Baudin, Rida Laraki</i>	
Information Discrepancy in Strategic Learning.....	1691
<i>Yahav Bechavod, Chara Podimata, Steven Wu, Juba Ziani</i>	
On the Hidden Biases of Policy Mirror Ascent in Continuous Action Spaces .....	1716
<i>Amrit Singh Bedi, Souradip Chakraborty, Anjaly Parayil, Brian M Sadler, Pratap Tokekar, Alec Koppel</i>	
Imitation Learning by Estimating Expertise of Demonstrators .....	1732
<i>Mark Beliaev, Andy Shih, Stefano Ermon, Dorsa Sadigh, Ramtin Pedarsani</i>	
Matching Normalizing Flows and Probability Paths on Manifolds.....	1749
<i>Heli Ben-Hamu, Samuel Cohen, Joey Bose, Brandon Amos, Maximillian Nickel, Aditya Grover, Ricky T. Q. Chen, Yaron Lipman</i>	
Stochastic Contextual Dueling Bandits Under Linear Stochastic Transitivity Models .....	1764
<i>Viktor Bengs, Aadirupa Saha, Eyke Hüllermeier</i>	
Neural Inverse Kinematic.....	1787
<i>Raphael Bensch, Shir Gur, Nitsan Blau, Lior Wolf</i>	
Volatility Based Kernels and Moving Average Means for Accurate Forecasting with Gaussian Processes .....	1798
<i>Gregory Benton, Wesley Maddox, Andrew Gordon Wilson</i>	
Gradient Descent on Neurons and Its Link to Approximate Second-Order Optimization.....	1817
<i>Frederik Benzing</i>	
Safe Learning in Tree-Form Sequential Decision Making: Handling Hard and Soft Constraints .....	1854
<i>Martino Bernasconi, Federico Cacciamani, Matteo Castiglioni, Alberto Marchesi, Nicola Gatti, Francesco Trovò</i>	
Skin Deep Unlearning: Artefact and Instrument Debiasing in the Context of Melanoma Classification.....	1874
<i>Peter Bevan, Amir Atapour-Abarghouei</i>	
Approximate Bayesian Computation with Domain Expert in the Loop.....	1893
<i>Ayush Bharti, Louis Filstroff, Samuel Kaski</i>	
Minimax M-Estimation Under Adversarial Contamination .....	1906
<i>Sujay Bhatt, Guanhua Fang, Ping Li, Gennady Samorodnitsky</i>	
Nearly Optimal Catoni's M-Estimator for Infinite Variance .....	1925
<i>Sujay Bhatt, Guanhua Fang, Ping Li, Gennady Samorodnitsky</i>	

Personalization Improves Privacy-Accuracy Tradeoffs in Federated Learning.....	1945
<i>Alberto Bietti, Chen-Yu Wei, Miroslav Dudik, John Langford, Steven Wu</i>	
Non-Vacuous Generalisation Bounds for Shallow Neural Networks .....	1963
<i>Felix Biggs, Benjamin Guedj</i>	
Structure-Preserving GANs.....	1982
<i>Jeremiah Birrell, Markos Katsoulakis, Luc Rey-Bellet, Wei Zhu</i>	
Scalable Spike-And-Slab.....	2021
<i>Niloy Biswas, Lester Mackey, Xiao-Li Meng</i>	
Breaking Down Out-Of-Distribution Detection: Many Methods Based on OOD Training Data Estimate a Combination of the Same Core Quantities .....	2041
<i>Julian Bitterwolf, Alexander Meinke, Maximilian Augustin, Matthias Hein</i>	
A Query-Optimal Algorithm for Finding Counterfactuals.....	2075
<i>Guy Blanc, Caleb Koch, Jane Lange, Li-Yang Tan</i>	
Popular Decision Tree Algorithms Are Provably Noise Tolerant .....	2091
<i>Guy Blanc, Jane Lange, Ali Malik, Li-Yang Tan</i>	
Optimizing Sequential Experimental Design with Deep Reinforcement Learning .....	2107
<i>Tom Blau, Edwin V. Bonilla, Iadine Chades, Amir Dezfouli</i>	
Lagrangian Method for Q-Function Learning (with Applications to Machine Translation).....	2129
<i>Huang Bojun</i>	
Generalized Results for the Existence and Consistency of the MLE in the Bradley-Terry-Luce Model .....	2160
<i>Heejong Bong, Alessandro Rinaldo</i>	
How to Train Your Wide Neural Network Without Backprop: An Input-Weight Alignment Perspective.....	2178
<i>Akhilan Boopathy, Ila Fiete</i>	
Improving Language Models by Retrieving from Trillions of Tokens .....	2206
<i>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</i>	
Lie Point Symmetry Data Augmentation for Neural PDE Solvers .....	2241
<i>Johannes Brandstetter, Max Welling, Daniel E Worrall</i>	
An Iterative Clustering Algorithm for the Contextual Stochastic Block Model with Optimality Guarantees .....	2257
<i>Guillaume Braun, Hemant Tyagi, Christophe Biernacki</i>	
Tractable Dendritic RNNs for Reconstructing Nonlinear Dynamical Systems .....	2292
<i>Manuel Brenner, Florian Hess, Jonas M Mikhaeil, Leonard F Bereska, Zahra Monfared, Po- Chen Kuo, Daniel Durstewitz</i>	
Learning to Predict Graphs with Fused Gromov-Wasserstein Barycenters.....	2321
<i>Luc Brogat-Motte, Rémi Flamary, Celine Brouard, Juho Rousu, Florence D'Alché-Buc</i>	
Efficient Learning of CNNs Using Patch Based Features .....	2336
<i>Alon Brutzkus, Amir Globerson, Eran Malach, Alon Regev Netser, Shai Shalev-Schwartz</i>	
Causal Structure-Based Root Cause Analysis of Outliers .....	2357
<i>Kailash Budhathoki, Lenon Minorics, Patrick Bloebaum, Dominik Janzing</i>	
IGLUE: A Benchmark for Transfer Learning Across Modalities, Tasks, and Languages .....	2370
<i>Emanuele Bugliarello, Fangyu Liu, Jonas Pfeiffer, Siva Reddy, Desmond Elliott, Edoardo Maria Ponti, Ivan Vulic</i>	
Interactive Inverse Reinforcement Learning for Cooperative Games .....	2393
<i>Thomas Kleine Büning, Anne-Marie George, Christos Dimitrakakis</i>	

Convolutional and Residual Networks Provably Contain Lottery Tickets .....	2414
<i>Rebekka Burkholz</i>	

#### PART 4

Near-Optimal Algorithms for Autonomous Exploration and Multi-Goal Stochastic Shortest Path .....	2434
<i>Haoyuan Cai, Tengyu Ma, Simon Du</i>	
Convergence of Invariant Graph Networks .....	2457
<i>Chen Cai, Yusu Wang</i>	
Reinforcement Learning from Partial Observation: Linear Function Approximation with Provable Sample Efficiency .....	2485
<i>Qi Cai, Zhuoran Yang, Zhaoran Wang</i>	
Scaling Gaussian Process Optimization by Evaluating a Few Unique Candidates Multiple Times .....	2523
<i>Daniele Calandriello, Luigi Carratino, Alessandro Lazaric, Michal Valko, Lorenzo Rosasco</i>	
Adaptive Gaussian Process Change Point Detection.....	2542
<i>Edoardo Caldarelli, Philippe Wenk, Stefan Bauer, Andreas Krause</i>	
Measuring Dissimilarity with Diffeomorphism Invariance .....	2572
<i>Théophile Cantelobre, Carlo Ciliberto, Benjamin Guedj, Alessandro Rudi</i>	
A Model-Agnostic Randomized Learning Framework Based on Random Hypothesis Subspace Sampling.....	2597
<i>Yiting Cao, Chao Lan</i>	
Gaussian Process Uniform Error Bounds with Unknown Hyperparameters for Safety-Critical Applications.....	2609
<i>Alexandre Capone, Armin Lederer, Sandra Hirche</i>	
Burst-Dependent Plasticity and Dendritic Amplification Support Target-Based Learning and Hierarchical Imitation Learning .....	2625
<i>Cristiano Capone, Cosimo Lupo, Paolo Muratore, Pier Stanislao Paolucci</i>	
A Marriage Between Adversarial Team Games and 2-Player Games: Enabling Abstractions, No-Regret Learning, and Subgame Solving .....	2638
<i>Luca Carminati, Federico Cacciamani, Marco Ciccone, Nicola Gatti</i>	
RECAPP: Crafting a More Efficient Catalyst for Convex Optimization .....	2658
<i>Yair Carmon, Arun Jambulapati, Yujia Jin, Aaron Sidford</i>	
Estimating and Penalizing Induced Preference Shifts in Recommender Systems.....	2686
<i>Micah D Carroll, Anca Dragan, Stuart Russell, Dylan Hadfield-Menell</i>	
YourTTS: Towards Zero-Shot Multi-Speaker TTS and Zero-Shot Voice Conversion for Everyone.....	2709
<i>Edresson Casanova, Julian Weber, Christopher D Shulby, Arnaldo Candido Junior, Eren Gölge, Moacir A Ponti</i>	
The Infinite Contextual Graph Markov Model.....	2721
<i>Daniele Castellana, Federico Errica, Davide Bacciu, Alessio Micheli</i>	
Compressed-VFL: Communication-Efficient Learning with Vertically Partitioned Data .....	2738
<i>Timothy J Castiglia, Anirban Das, Shiqiang Wang, Stacy Patterson</i>	
Online Learning with Knapsacks: The Best of Both Worlds .....	2767
<i>Matteo Castiglioni, Andrea Celli, Christian Kroer</i>	
Stabilizing Off-Policy Deep Reinforcement Learning from Pixels .....	2784
<i>Edoardo Cetin, Philip J Ball, Stephen Roberts, Oya Celikutan</i>	
Accelerated, Optimal and Parallel: Some Results on Model-Based Stochastic Optimization.....	2811
<i>Karan Chadha, Gary Cheng, John Duchi</i>	
Robust Imitation Learning Against Variations in Environment Dynamics.....	2828
<i>Jongseong Chae, Seungyul Han, Whiyoung Jung, Myungsik Cho, Sungho Choi, Youngchul Sung</i>	
Fairness with Adaptive Weights .....	2853
<i>Junyi Chai, Xiaoqian Wang</i>	



UNIREX: A Unified Learning Framework for Language Model Rationale Extraction .....	2867
<i>Aaron Chan, Maziar Sanjabi, Lambert Mathias, Liang Tan, Shaoliang Nie, Xiaochang Peng, Xiang Ren, Hamed Firooz</i>	
Revisiting Label Smoothing and Knowledge Distillation Compatibility: What Was Missing? .....	2890
<i>Keshigeyan Chandrasegaran, Ngoc-Trung Tran, Yunqing Zhao, Ngai-Man Cheung</i>	
Style Equalization: Unsupervised Learning of Controllable Generative Sequence Models.....	2917
<i>Jen-Hao Rick Chang, Ashish Shrivastava, Hema Koppula, Xiaoshuai Zhang, Oncel Tuzel</i>	
Learning Bellman Complete Representations for Offline Policy Evaluation .....	2938
<i>Jonathan Chang, Kaiwen Wang, Nathan Kallus, Wen Sun</i>	
Sample Efficient Learning of Predictors that Complement Humans.....	2972
<i>Mohammad-Amin Charusaie, Hussein Mozannar, David Sontag, Samira Samadi</i>	
Nyström Kernel Mean Embeddings .....	3006
<i>Antoine Chatalic, Nicolas Schreuder, Lorenzo Rosasco, Alessandro Rudi</i>	
Coarsening the Granularity: Towards Structurally Sparse Lottery Tickets.....	3025
<i>Tianlong Chen, Xuxi Chen, Xiaolong Ma, Yanzhi Wang, Zhangyang Wang</i>	
Learning Domain Adaptive Object Detection with Probabilistic Teacher .....	3040
<i>Meilin Chen, Weijie Chen, Shicai Yang, Jie Song, Xinchao Wang, Lei Zhang, Yunfeng Yan, Donglian Qi, Yueting Zhuang, Di Xie, Shiliang Pu</i>	
The Fundamental Price of Secure Aggregation in Differentially Private Federated Learning .....	3056
<i>Wei-Ning Chen, Christopher A Choquette Choo, Peter Kairouz, Ananda Theertha Suresh</i>	
Perfectly Balanced: Improving Transfer and Robustness of Supervised Contrastive Learning .....	3090
<i>Mayee Chen, Daniel Y Fu, Avanika Narayan, Michael Zhang, Zhao Song, Kayvon Fatahalian, Christopher Re</i>	
Strategies for Safe Multi-Armed Bandits with Logarithmic Regret and Risk .....	3123
<i>Tianrui Chen, Aditya Gangrade, Venkatesh Saligrama</i>	
On the Sample Complexity of Learning Infinite-Horizon Discounted Linear Kernel MDPs.....	3149
<i>Yuanzhou Chen, Jiafan He, Quanquan Gu</i>	
Streaming Algorithms for Support-Aware Histograms.....	3184
<i>Justin Chen, Piotr Indyk, Tal Wagner</i>	
Improved No-Regret Algorithms for Stochastic Shortest Path with Linear MDP.....	3204
<i>Liyu Chen, Rahul Jain, Haipeng Luo</i>	
Learning Infinite-Horizon Average-Reward Markov Decision Process with Constraints .....	3246
<i>Liyu Chen, Rahul Jain, Haipeng Luo</i>	

## PART 5

Active Multi-Task Representation Learning.....	3271
<i>Yifang Chen, Kevin Jamieson, Simon Du</i>	
On Collective Robustness of Bagging Against Data Poisoning .....	3299
<i>Ruoxin Chen, Zenan Li, Jie Li, Junchi Yan, Chentao Wu</i>	
Online Active Regression .....	3320
<i>Cheng Chen, Yi Li, Yiming Sun</i>	
Selling Data to a Machine Learner: Pricing Via Costly Signaling .....	3336
<i>Junjie Chen, Minming Li, Haifeng Xu</i>	
ME-GAN: Learning Panoptic Electrocardio Representations for Multi-View ECG Synthesis Conditioned on Heart Diseases.....	3360
<i>Jintai Chen, Kuanlun Liao, Kun Wei, Haochao Ying, Danny Z Chen, Jian Wu</i>	
Weisfeiler-Lehman Meets Gromov-Wasserstein .....	3371
<i>Samantha Chen, Sunhyuk Lim, Facundo Memoli, Zhengchao Wan, Yusu Wang</i>	
On Non-Local Convergence Analysis of Deep Linear Networks.....	3417
<i>Kun Chen, Dachao Lin, Zhihua Zhang</i>	

Flow-Based Recurrent Belief State Learning for POMDPs .....	3444
<i>Xiaoyu Chen, Yao Mark Mu, Ping Luo, Shengbo Li, Jianyu Chen</i>	
Structure-Aware Transformer for Graph Representation Learning .....	3469
<i>Dexiong Chen, Leslie O'Bray, Karsten Borgwardt</i>	
The Poisson Binomial Mechanism for Unbiased Federated Learning with Secure Aggregation .....	3490
<i>Wei-Ning Chen, Ayfer Ozgur, Peter Kairouz</i>	
Learning Mixtures of Linear Dynamical Systems .....	3507
<i>Yanxi Chen, H. Vincent Poor</i>	
On Well-Posedness and Minimax Optimal Rates of Nonparametric Q-Function Estimation in Off-Policy Evaluation .....	3558
<i>Xiaohong Chen, Zhengling Qi</i>	
Faster Fundamental Graph Algorithms Via Learned Predictions .....	3583
<i>Justin Chen, Sandeep Silwal, Ali Vakilian, Fred Zhang</i>	
Improve Single-Point Zeroth-Order Optimization Using High-Pass and Low-Pass Filters .....	3603
<i>Xin Chen, Yujie Tang, Na Li</i>	
Deep Variational Graph Convolutional Recurrent Network for Multivariate Time Series Anomaly Detection .....	3621
<i>Wenchao Chen, Long Tian, Bo Chen, Liang Dai, Zhibin Duan, Mingyuan Zhou</i>	
Auxiliary Learning with Joint Task and Data Scheduling .....	3634
<i>Hong Chen, Xin Wang, Chaoyu Guan, Yue Liu, Wenwu Zhu</i>	
Optimization-Induced Graph Implicit Nonlinear Diffusion .....	3648
<i>Qi Chen, Yifei Wang, Yisen Wang, Jiansheng Yang, Zhouchen Lin</i>	
Robust Meta-Learning with Sampling Noise and Label Noise Via Eigen-Reptile .....	3662
<i>Dong Chen, Lingfei Wu, Siliang Tang, Xiao Yun, Bo Long, Yueting Zhuang</i>	
Adaptive Model Design for Markov Decision Process .....	3679
<i>Siyu Chen, Donglin Yang, Jiayang Li, Senmiao Wang, Zhuoran Yang, Zhaoran Wang</i>	
State Transition of Dendritic Spines Improves Learning of Sparse Spiking Neural Networks .....	3701
<i>Yanqi Chen, Zhaofei Yu, Wei Fang, Zhengyu Ma, Tiejun Huang, Yonghong Tian</i>	
Efficient Online ML API Selection for Multi-Label Classification Tasks .....	3716
<i>Lingjiao Chen, Matei Zaharia, James Zou</i>	
Data-Efficient Double-Win Lottery Tickets from Robust Pre-Training .....	3747
<i>Tianlong Chen, Zhenyu Zhang, Sijia Liu, Yang Zhang, Shiyu Chang, Zhangyang Wang</i>	
Linearity Grafting: Relaxed Neuron Pruning Helps Certifiable Robustness .....	3760
<i>Tianlong Chen, Huan Zhang, Zhenyu Zhang, Shiyu Chang, Sijia Liu, Pin-Yu Chen, Zhangyang Wang</i>	
Human-In-The-Loop: Provably Efficient Preference-Based Reinforcement Learning with General Function Approximation .....	3773
<i>Xiaoyu Chen, Han Zhong, Zhuoran Yang, Zhaoran Wang, Liwei Wang</i>	
Sample and Communication-Efficient Decentralized Actor-Critic Algorithms with Finite-Time Analysis .....	3794
<i>Ziyi Chen, Yi Zhou, Rong-Rong Chen, Shaofeng Zou</i>	
Task-Aware Privacy Preservation for Multi-Dimensional Data .....	3835
<i>Jiangnan Cheng, Ao Tang, Sandeep Chinchali</i>	
Adversarially Trained Actor Critic for Offline Reinforcement Learning .....	3852
<i>Ching-An Cheng, Tengyang Xie, Nan Jiang, Alekh Agarwal</i>	
Quantum-Inspired Algorithms from Randomized Numerical Linear Algebra .....	3879
<i>Nadiia Chepurko, Kenneth Clarkson, Lior Horesh, Honghao Lin, David Woodruff</i>	
RieszNet and ForestRiesz: Automatic Debiased Machine Learning with Neural Nets and Random Forests .....	3901
<i>Victor Chernozhukov, Whitney Newey, Victor M Quintas-Martinez, Vasilis Syrgkanis</i>	
Self-Supervised Learning with Random-Projection Quantizer for Speech Recognition .....	3915
<i>Chung-Cheng Chiu, James Qin, Yu Zhang, Jiahui Yu, Yonghui Wu</i>	

Discrete Probabilistic Inverse Optimal Transport.....	3925
<i>Wei-Ting Chiu, Pei Wang, Patrick Shafto</i>	
Selective Network Linearization for Efficient Private Inference.....	3947
<i>Minsu Cho, Ameya Joshi, Brandon Reagen, Siddharth Garg, Chinmay Hegde</i>	
From block-Toeplitz Matrices to Differential Equations on Graphs: Towards a General Theory for Scalable Masked Transformers.....	3962
<i>Krzysztof Choromanski, Han Lin, Haoxian Chen, Tianyi Zhang, Arijit Sehanobish, Valerii Likhoshesterov, Jack Parker-Holder, Tamas Sarlos, Adrian Weller, Thomas Weingarten</i>	
Shuffle Private Linear Contextual Bandits .....	3984
<i>Sayak Ray Chowdhury, Xingyu Zhou</i>	
DNA: Domain Generalization with Diversified Neural Averaging .....	4010
<i>Xu Chu, Yujie Jin, Wenwu Zhu, Yasha Wang, Xin Wang, Shanghang Zhang, Hong Mei</i>	
TPC: Transformation-Specific Smoothing for Point Cloud Models.....	4035
<i>Wenda Chu, Linyi Li, Bo Li</i>	
Unified Scaling Laws for Routed Language Models.....	4057
<i>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</i>	
Context-Aware Drift Detection.....	4087
<i>Oliver Cobb, Arnaud Van Looveren</i>	

## PART 6

On the Robustness of CountSketch to Adaptive Inputs .....	4112
<i>Edith Cohen, Xin Lyu, Jelani Nelson, Tamas Sarlos, Moshe Shechner, Uri Stemmer</i>	
Diffusion Bridges Vector Quantized Variational Autoencoders.....	4141
<i>Max Cohen, Guillaume Quispe, Sylvain Le Corff, Charles Ollion, Eric Moulines</i>	
Online and Consistent Correlation Clustering .....	4157
<i>Vincent Cohen-Addad, Silvio Lattanzi, Andreas Maggiori, Nikos Parotsidis</i>	
Massively Parallel k-Means Clustering for Perturbation Resilient Instances.....	4180
<i>Vincent Cohen-Addad, Vahab Mirrokni, Peilin Zhong</i>	
One-Pass Diversified Sampling with Application to Terabyte-Scale Genomic Sequence Streams .....	4202
<i>Benjamin Coleman, Benito Geordie, Li Chou, R. A. Leo Elworth, Todd J. Treangen, Anshumali Shrivastava</i>	
Transfer and Marginalize: Explaining Away Label Noise with Privileged Information.....	4219
<i>Mark Collier, Rodolphe Jenatton, Effrosyni Kokiopoulou, Jesse Berent</i>	
MAML and ANIL Provably Learn Representations .....	4238
<i>Liam Collins, Aryan Mokhtari, Sewoong Oh, Sanjay Shakkottai</i>	
Entropic Causal Inference: Graph Identifiability.....	4311
<i>Spencer Compton, Kristjan Greenewald, Dmitriy A Katz, Murat Kocaoglu</i>	
Mitigating Gender Bias in Face Recognition Using the Von Mises-Fisher Mixture Model.....	4344
<i>Jean-Rémy Conti, Nathan Noiry, Stephan Clemenccon, Vincent Despiegel, Stéphane Gentric</i>	
Counterfactual Transportability: A Formal Approach .....	4370
<i>Juan D Correa, Sanghack Lee, Elias Bareinboim</i>	
Label-Free Explainability for Unsupervised Models.....	4391
<i>Jonathan Crabbé, Mihaela Van Der Schaar</i>	
Evaluating the Adversarial Robustness of Adaptive Test-Time Defenses .....	4421
<i>Francesco Croce, Sven Gowal, Thomas Brunner, Evan Shelhamer, Matthias Hein, Taylan Cemgil</i>	

Adversarial Robustness Against Multiple and Single lp-Threat Models Via Quick Fine-Tuning of Robust Classifiers.....	4436
<i>Francesco Croce, Matthias Hein</i>	
Self-Conditioning Pre-Trained Language Models.....	4455
<i>Xavier Suau, Luca Zappella, Nicholas Apostoloff</i>	
Only Tails Matter: Average-Case Universality and Robustness in the Convex Regime.....	4474
<i>Leonardo Cunha, Gauthier Gidel, Fabian Pedregosa, Damien Scieur, Courtney Paquette</i>	
Principal Component Flows .....	4492
<i>Edmond Cunningham, Adam Cobb, Susmit Jha</i>	
Deep Symbolic Regression for Recurrence Prediction.....	4520
<i>St'ephane D'Ascoli, Pierre-Alexandre Kamienny, Guillaume Lample, Francois Charton</i>	
Continuous Control with Action Quantization from Demonstrations .....	4537
<i>Robert Dadashi, Léonard Hussenot, Damien Vincent, Sertan Girgin, Anton Raichuk, Matthieu Geist, Olivier Pietquin</i>	
Dialog Inpainting: Turning Documents into Dialogs .....	4558
<i>Zhuyun Dai, Arun Tejasvi Chaganty, Vincent Y Zhao, Aida Amini, Qazi Mamunur Rashid, Mike Green, Kelvin Guu</i>	
DisPFL: Towards Communication-Efficient Personalized Federated Learning Via Decentralized Sparse Training.....	4587
<i>Rong Dai, Li Shen, Fengxiang He, Xinmei Tian, Dacheng Tao</i>	
Marginal Distribution Adaptation for Discrete Sets Via Module-Oriented Divergence Minimization.....	4605
<i>Hanjun Dai, Mengjiao Yang, Yuan Xue, Dale Schuurmans, Bo Dai</i>	
Balancing Sample Efficiency and Suboptimality in Inverse Reinforcement Learning.....	4618
<i>Angelo Damiani, Giorgio Manganini, Alberto Maria Metelli, Marcello Restelli</i>	
Understanding Robust Generalization in Learning Regular Languages.....	4630
<i>Soham Dan, Osbert Bastani, Dan Roth</i>	
Unsupervised Image Representation Learning with Deep Latent Particles.....	4644
<i>Tal Daniel, Aviv Tamar</i>	
Guarantees for Epsilon-Greedy Reinforcement Learning with Function Approximation .....	4666
<i>Chris Dann, Yishay Mansour, Mehryar Mohri, Ayush Sekhari, Karthik Sridharan</i>	
Monarch: Expressive Structured Matrices for Efficient and Accurate Training.....	4690
<i>Tri Dao, Beidi Chen, Nimit S Sohoni, Arjun Desai, Michael Poli, Jessica Grogan, Alexander Liu, Aniruddh Rao, Atri Rudra, Christopher Re</i>	
Score-Guided Intermediate Level Optimization: Fast Langevin Mixing for Inverse Problems .....	4722
<i>Giannis Daras, Yuval Dagan, Alex Dimakis, Constantinos Daskalakis</i>	
Test-Time Training Can Close the Natural Distribution Shift Performance Gap in Deep Learning Based Compressed Sensing.....	4754
<i>Mohammad Zalbagi Darestani, Jiayu Liu, Reinhard Heckel</i>	
Knowledge Base Question Answering by Case-Based Reasoning Over Subgraphs .....	4777
<i>Rajarshi Das, Ameya Godbole, Ankita Naik, Elliot Tower, Manzil Zaheer, Hannaneh Hajishirzi, Robin Jia, Andrew McCallum</i>	
Framework for Evaluating Faithfulness of Local Explanations .....	4794
<i>Sanjoy Dasgupta, Nave Frost, Michal Moshkovitz</i>	
Distinguishing Rule and Exemplar-Based Generalization in Learning Systems.....	4816
<i>Ishita Dasgupta, Erin Grant, Tom Griffiths</i>	
Robust Multi-Objective Bayesian Optimization Under Input Noise .....	4831
<i>Samuel Daulton, Sait Cakmak, Maximilian Balandat, Michael A. Osborne, Enlu Zhou, Eytan Bakshy</i>	
Attentional Meta-Learners for Few-Shot Polythetic Classification.....	4867
<i>Ben J Day, Ramon Viñas Torné, Nikola Simidjievski, Pietro Lió</i>	
Adversarial Vulnerability of Randomized Ensembles .....	4890
<i>Hassan Dbouk, Naresh Shanbhag</i>	

Born-Infeld (BI) for AI: Energy-Conserving Descent (ECD) for Optimization.....	4918
<i>Giuseppe Bruno De Luca, Eva Silverstein</i>	
Error-Driven Input Modulation: Solving the Credit Assignment Problem Without a Backward Pass .....	4937
<i>Giorgia Dellaferrera, Gabriel Kreiman</i>	

## PART 7

DreamerPro: Reconstruction-Free Model-Based Reinforcement Learning with Prototypical Representations .....	4956
<i>Fei Deng, Ingoon Jang, Sungjin Ahn</i>	
NeuralEF: Deconstructing Kernels by Deep Neural Networks .....	4976
<i>Zhijie Deng, Jiaxin Shi, Jun Zhu</i>	
Deep Causal Metric Learning.....	4993
<i>Xiang Deng, Zhongfei Zhang</i>	
On the Convergence of Inexact Predictor-Corrector Methods for Linear Programming.....	5007
<i>Gregory Dexter, Agniva Chowdhury, Haim Avron, Petros Drineas</i>	
Analysis of Stochastic Processes Through Replay Buffers .....	5039
<i>Shirli Di-Castro, Shie Mannor, Dotan Di Castro</i>	
Streaming Algorithms for High-Dimensional Robust Statistics .....	5061
<i>Ilias Diakonikolas, Daniel M. Kane, Ankit Pensia, Thanasis Pittas</i>	
Learning General Halfspaces with Adversarial Label Noise Via Online Gradient Descent .....	5118
<i>Ilias Diakonikolas, Vasilis Kontonis, Christos Tzamos, Nikos Zarifis</i>	
Variational Feature Pyramid Networks.....	5142
<i>Panagiotis Dimitrakopoulos, Giorgos Sfikas, Christophoros Nikou</i>	
Understanding Doubly Stochastic Clustering.....	5153
<i>Tianjiao Ding, Derek Lim, Rene Vidal, Benjamin D Haeffele</i>	
Independent Policy Gradient for Large-Scale Markov Potential Games: Sharper Rates, Function Approximation, and Game-Agnostic Convergence.....	5166
<i>Dongsheng Ding, Chen-Yu Wei, Kaiqing Zhang, Mihailo Jovanovic</i>	
Generalization and Robustness Implications in Object-Centric Learning.....	5221
<i>Andrea Dittadi, Samuele S Papa, Michele De Vita, Bernhard Schölkopf, Ole Winther, Francesco Locatello</i>	
Fair Generalized Linear Models with a Convex Penalty .....	5286
<i>Hyungrak Do, Preston Putzel, Axel S Martin, Padhraic Smyth, Judy Zhong</i>	
Bayesian Learning with Information Gain Provably Bounds Risk for a Robust Adversarial Defense.....	5309
<i>Bao Gia Doan, Ehsan M. Abbasnejad, Javen Qinfeng Shi, Damith C. Ranasinghe</i>	
On the Adversarial Robustness of Causal Algorithmic Recourse.....	5324
<i>Ricardo Dominguez-Olmedo, Amir H Karimi, Bernhard Schölkopf</i>	
Finding the Task-Optimal Low-Bit Sub-Distribution in Deep Neural Networks .....	5343
<i>Runpei Dong, Zhanhong Tan, Mengdi Wu, Linfeng Zhang, Kaisheng Ma</i>	
PACE: A Parallelizable Computation Encoder for Directed Acyclic Graphs .....	5360
<i>Zehao Dong, Muhan Zhang, Fuhai Li, Yixin Chen</i>	
Privacy for Free: How Does Dataset Condensation Help Privacy?.....	5378
<i>Tian Dong, Bo Zhao, Lingjuan Lyu</i>	
Fast Rates for Noisy Interpolation Require Rethinking the Effect of Inductive Bias.....	5397
<i>Konstantin Donhauser, Nicolò Ruggeri, Stefan Stojanovic, Fanny Yang</i>	
Adapting to Mixing Time in Stochastic Optimization with Markovian Data.....	5429
<i>Ron Dorfman, Kfir Yehuda Levy</i>	
TACTiS: Transformer-Attentional Copulas for Time Series .....	5447
<i>Alexandre Drouin, Étienne Marcotte, Nicolas Chapados</i>	
Branching Reinforcement Learning .....	5494
<i>Yihan Du, Wei Chen</i>	

Bayesian Imitation Learning for End-To-End Mobile Manipulation .....	5531
<i>Yuqing Du, Daniel Ho, Alex Alemi, Eric Jang, Mohi Khansari</i>	
GLaM: Efficient Scaling of Language Models with Mixture-Of-Experts .....	5547
<i>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</i>	
Learning Iterative Reasoning Through Energy Minimization .....	5570
<i>Yilun Du, Shuang Li, Joshua Tenenbaum, Igor Mordatch</i>	
SE(3) Equivariant Graph Neural Networks with Complete Local Frames .....	5583
<i>Weitao Du, He Zhang, Yuanqi Du, Qi Meng, Wei Chen, Nanning Zheng, Bin Shao, Tie-Yan Liu</i>	
A Context-Integrated Transformer-Based Neural Network for Auction Design.....	5609
<i>Zhijian Duan, Jingwu Tang, Yutong Yin, Zhe Feng, Xiang Yan, Manzil Zaheer, Xiaotie Deng</i>	
Augment with Care: Contrastive Learning for Combinatorial Problems .....	5627
<i>Haonan Duan, Pashootan Vaezipoor, Max B Paulus, Yangjun Ruan, Chris Maddison</i>	
Parametric Visual Program Induction with Function Modularization .....	5643
<i>Xuguang Duan, Xin Wang, Ziwei Zhang, Wenwu Zhu</i>	
Bayesian Deep Embedding Topic Meta-Learner .....	5659
<i>Zhibin Duan, Yishi Xu, Jianqiao Sun, Bo Chen, Wenchao Chen, Chaojie Wang, Mingyuan Zhou</i>	
Deletion Robust Submodular Maximization Over Matroids .....	5671
<i>Paul Duetting, Federico Fusco, Silvio Lattanzi, Ashkan Norouzi-Fard, Morteza Zadimoghaddam</i>	
From Data to Funct: Your Data Point is a Function and You Can Treat it Like One .....	5694
<i>Emilien Dupont, Hyunjik Kim, S. M. Ali Eslami, Danilo Jimenez Rezende, Dan Rosenbaum</i>	
Efficient Low Rank Convex Bounds for Pairwise Discrete Graphical Models .....	5726
<i>Valentin Durante, George Katsirelos, Thomas Schiex</i>	
Robust Counterfactual Explanations for Tree-Based Ensembles .....	5742
<i>Sanghamitra Dutta, Jason Long, Saumitra Mishra, Cecilia Tilli, Daniele Magazzeni</i>	
On the Difficulty of Defending Self-Supervised Learning Against Model Extraction.....	5757
<i>Adam Dziedziec, Nikita Dhawan, Muhammad Ahmad Kaleem, Jonas Guan, Nicolas Papernot</i>	
LIMO: Latent Inceptionism for Targeted Molecule Generation.....	5777
<i>Peter Eckmann, Kunyang Sun, Bo Zhao, Mudong Feng, Michael Gilson, Rose Yu</i>	

## PART 8

Inductive Biases and Variable Creation in Self-Attention Mechanisms .....	5793
<i>Benjamin L Edelman, Surbhi Goel, Sham Kakade, Cyril Zhang</i>	
Provable Reinforcement Learning with a Short-Term Memory .....	5832
<i>Yonathan Efroni, Chi Jin, Akshay Krishnamurthy, Sobhan Miryoosefi</i>	
Sparsity in Partially Controllable Linear Systems .....	5851
<i>Yonathan Efroni, Sham Kakade, Akshay Krishnamurthy, Cyril Zhang</i>	
FedNew: A Communication-Efficient and Privacy-Preserving Newton-Type Method for Federated Learning .....	5861
<i>Anis Elgabli, Chaouki Ben Issaid, Amrit Singh Bedi, Ketan Rajawat, Mehdi Bennis, Vaneet Aggarwal</i>	
PathGCN: Learning General Graph Spatial Operators from Paths.....	5878
<i>Moshe Eliasof, Eldad Haber, Eran Treister</i>	
Discrete Tree Flows Via Tree-Structured Permutations.....	5892
<i>Mai Elkady, Hyung Zin Lim, David I Inouye</i>	

For Learning in Symmetric Teams, Local Optima Are Global Nash Equilibria.....	5924
<i>Scott Emmons, Caspar Oesterheld, Andrew Critch, Vincent Conitzer, Stuart Russell</i>	
Streaming Algorithm for Monotone k-Submodular Maximization with Cardinality Constraints .....	5944
<i>Alina Ene, Huy Nguyen</i>	
Towards Scaling Difference Target Propagation by Learning Backprop Targets .....	5968
<i>Maxence M Ernoult, Fabrice Normandin, Abhinav Moudgil, Sean Spinney, Eugene Belilovsky, Irina Rish, Blake Richards, Yoshua Bengio</i>	
Understanding Dataset Difficulty with V-Usable Information .....	5988
<i>Kawin Ethayarajh, Yejin Choi, Swabha Swayamdipta</i>	
Head2Toe: Utilizing Intermediate Representations for Better Transfer Learning .....	6009
<i>Utku Evci, Vincent Dumoulin, Hugo Larochelle, Michael C Mozer</i>	
Variational Sparse Coding with Learned Thresholding .....	6034
<i>Kion Fallah, Christopher J Rozell</i>	
Training Discrete Deep Generative Models Via Gapped Straight-Through Estimator.....	6059
<i>Ting-Han Fan, Ta-Chung Chi, Alexander I. Rudnicky, Peter J Ramadge</i>	
DRIBO: Robust Deep Reinforcement Learning Via Multi-View Information Bottleneck.....	6074
<i>Jiameng Fan, Wenchao Li</i>	
Generalized Data Distribution Iteration.....	6103
<i>Jiajun Fan, Changnan Xiao</i>	
Variational Wasserstein Gradient Flow.....	6185
<i>Jiaojiao Fan, Qinsheng Zhang, Amirhossein Taghvaei, Yongxin Chen</i>	
Data Determines Distributional Robustness in Contrastive Language Image Pre-Training (CLIP).....	6216
<i>Alex Fang, Gabriel Ilharco, Mitchell Wortsman, Yuhao Wan, Vaishaal Shankar, Achal Dave, Ludwig Schmidt</i>	
Bayesian Continuous-Time Tucker Decomposition .....	6235
<i>Shikai Fang, Akil Narayan, Robert Kirby, Shandian Zhe</i>	
Byzantine Machine Learning Made Easy by Resilient Averaging of Momentums .....	6246
<i>Sadegh Farhadkhani, Rachid Guerraoui, Nirupam Gupta, Rafael Pinot, John Stephan</i>	
An Equivalence Between Data Poisoning and Byzantine Gradient Attacks.....	6284
<i>Sadegh Farhadkhani, Rachid Guerraoui, Lê Nguyễn Hoàng, Oscar Villemaud</i>	
Investigating Generalization by Controlling Normalized Margin .....	6324
<i>Alexander R Farhang, Jeremy D Bernstein, Kushal Tirumala, Yang Liu, Yisong Yue</i>	
Kernelized Multiplicative Weights for 0/1-Polyhedral Games: Bridging the Gap Between Learning in Extensive-Form and Normal-Form Games .....	6337
<i>Gabriele Farina, Chung-Wei Lee, Haipeng Luo, Christian Kroer</i>	
Local Linear Convergence of Douglas-Rachford for Linear Programming: A Probabilistic Analysis.....	6358
<i>Oisin Faust, Hamza Fawzi</i>	
Matching Structure for Dual Learning .....	6373
<i>Hao Fei, Shengqiong Wu, Yafeng Ren, Meishan Zhang</i>	
Cascaded Gaps: Towards Logarithmic Regret for Risk-Sensitive Reinforcement Learning .....	6392
<i>Yingjie Fei, Ruitu Xu</i>	
Private Frequency Estimation Via Projective Geometry .....	6418
<i>Vitaly Feldman, Jelani Nelson, Huy Nguyen, Kunal Talwar</i>	
An Intriguing Property of Geophysics Inversion.....	6434
<i>Yinan Feng, Yinpeng Chen, Shihang Feng, Peng Jin, Zicheng Liu, Youzuo Lin</i>	
Principled Knowledge Extrapolation with GANs .....	6447
<i>Ruili Feng, Jie Xiao, Kecheng Zheng, Deli Zhao, Jingren Zhou, Qibin Sun, Zheng-Jun Zha</i>	
A Resilient Distributed Boosting Algorithm.....	6465
<i>Yuval Filmus, Idan Mehalel, Shay Moran</i>	
Model-Value Inconsistency as a Signal for Epistemic Uncertainty.....	6474
<i>Angelos Filos, Eszter Vértés, Zita Marinho, Gregory Farquhar, Diana Borsa, Abram Friesen, Feryal Behbahani, Tom Schaul, Andre Barreto, Simon Osindero</i>	

Coordinated Double Machine Learning .....	6499
<i>Nitai Fingerhut, Matteo Sesia, Yaniv Romano</i>	
Conformal Prediction Sets with Limited False Positives .....	6514
<i>Adam Fisch, Tal Schuster, Tommi Jaakkola, Dr. Regina Barzilay</i>	
Fast Population-Based Reinforcement Learning on a Single Machine .....	6533
<i>Arthur Flajolet, Claire Bizon Monroc, Karim Beguir, Thomas Pierrot</i>	
Fast Relative Entropy Coding with A* Coding .....	6548
<i>Gergely Flamich, Stratis Markou, Jose Miguel Hernandez-Lobato</i>	
Contrastive Mixture of Posteriors for Counterfactual Inference, Data Integration and Fairness.....	6578
<i>Adam Foster, Arpi Vezer, Craig A. Glastonbury, Paidi Creed, Samer Abujudeh, Aaron Sim</i>	
Label Ranking Through Nonparametric Regression .....	6622
<i>Dimitris Fotakis, Alkis Kalavasis, Eleni Psaroudaki</i>	

## PART 9

A Neural Tangent Kernel Perspective of GANs .....	6660
<i>Jean-Yves Franceschi, Emmanuel De Bézenac, Ibrahim Ayed, Mickael Chen, Sylvain Lamprier, Patrick Gallinari</i>	
Extracting Latent State Representations with Linear Dynamics from Rich Observations .....	6705
<i>Abraham Frandsen, Rong Ge, Holden Lee</i>	
SPDY: Accurate Pruning with Speedup Guarantees.....	6726
<i>Elias Frantar, Dan Alistarh</i>	
Revisiting the Effects of Stochasticity for Hamiltonian Samplers .....	6744
<i>Giulio Franzese, Dimitrios Milios, Maurizio Filippone, Pietro Michiardi</i>	
Bregman Neural Networks .....	6779
<i>Jordan Frecon, Gilles Gasso, Massimiliano Pontil, Saverio Salzo</i>	
(Non-)Convergence Results for Predictive Coding Networks.....	6793
<i>Simon Frieder, Thomas Lukasiewicz</i>	
Scaling Structured Inference with Randomization .....	6811
<i>Yao Fu, John Cunningham, Mirella Lapata</i>	
Greedy When Sure and Conservative When Uncertain About the Opponents .....	6829
<i>Haobo Fu, Ye Tian, Hongxiang Yu, Weiming Liu, Shuang Wu, Jiechao Xiong, Ying Wen, Kai Li, Junliang Xing, Qiang Fu, Wei Yang</i>	
DepthShrinker: A New Compression Paradigm Towards Boosting Real-Hardware Efficiency of Compact Neural Networks .....	6849
<i>Yonggan Fu, Haichuan Yang, Jiayi Yuan, Meng Li, Cheng Wan, Raghuraman Krishnamoorthi, Vikas Chandra, Yingyan Lin</i>	
Revisiting Some Common Practices in Cooperative Multi-Agent Reinforcement Learning .....	6863
<i>Wei Fu, Chao Yu, Zelai Xu, Jiaqi Yang, Yi Wu</i>	
P-Laplacian Based Graph Neural Networks.....	6878
<i>Guoji Fu, Peilin Zhao, Yatao Bian</i>	
Why Should I Trust You, Bellman? the Bellman Error is a Poor Replacement for Value Error.....	6918
<i>Scott Fujimoto, David Meger, Doina Precup, Ofir Nachum, Shixiang Shane Gu</i>	
Robin Hood and Matthew Effects: Differential Privacy Has Disparate Impact on Synthetic Data .....	6944
<i>Georgi Ganev, Bristena Oprisanu, Emiliano De Cristofaro</i>	
The Complexity of k-Means Clustering When Little is Known .....	6960
<i>Robert Ganian, Thekla Hamm, Viktoriia Korchemna, Karolina Okrasa, Kirill Simonov</i>	
IDYNO: Learning Nonparametric DAGs from Interventional Dynamic Data.....	6988
<i>Tian Gao, Debarun Bhattacharjya, Elliot Nelson, Miao Liu, Yue Yu</i>	
Loss Function Learning for Domain Generalization by Implicit Gradient.....	7002
<i>Boyan Gao, Henry Gouk, Yongxin Yang, Timothy Hospedales</i>	



On the Convergence of Local Stochastic Compositional Gradient Descent with Momentum .....	7017
<i>Hongchang Gao, Junyi Li, Heng Huang</i>	
Deep Reference Priors: What is the Best Way to Pretrain a Model? .....	7036
<i>Yansong Gao, Rahul Ramesh, Pratik Chaudhari</i>	
On the Equivalence Between Temporal and Static Equivariant Graph Representations .....	7052
<i>Jianfei Gao, Bruno Ribeiro</i>	
Generalizing Gaussian Smoothing for Random Search .....	7077
<i>Katelyn Gao, Ozan Sener</i>	
Rethinking Image-Scaling Attacks: The Interplay Between Vulnerabilities in Machine Learning Systems.....	7102
<i>Yue Gao, Ilya Shumailov, Kassem Fawaz</i>	
Lazy Estimation of Variable Importance for Large Neural Networks .....	7122
<i>Yue Gao, Abby Stevens, Garvesh Raskutti, Rebecca Willett</i>	
Fast and Reliable Evaluation of Adversarial Robustness with Minimum-Margin Attack .....	7144
<i>Ruize Gao, Jiongxiao Wang, Kaiwen Zhou, Feng Liu, Binghui Xie, Gang Niu, Bo Han, James Cheng</i>	
Value Function Based Difference-Of-Convex Algorithm for Bilevel Hyperparameter Selection Problems.....	7164
<i>Lucy L Gao, Jane Ye, Haian Yin, Shangzhi Zeng, Jin Zhang</i>	
Learning to Incorporate Texture Saliency Adaptive Attention to Image Cartoonization.....	7183
<i>Xiang Gao, Yuqi Zhang, Yingjie Tian</i>	
Stochastic Smoothing of the top-K Calibrated Hinge Loss for Deep Imbalanced Classification.....	7208
<i>Camille Garcin, Maximilien Servajean, Alexis Joly, Joseph Salmon</i>	
PAGE-PG: A Simple and Loopless Variance-Reduced Policy Gradient Method with Probabilistic Gradient Estimation.....	7223
<i>Matilde Gargiani, Andrea Zanelli, Andrea Martinelli, Tyler Summers, John Lygeros</i>	
The Power of First-Order Smooth Optimization for Black-Box Non-Smooth Problems.....	7241
<i>Alexander Gasnikov, Anton Novitskii, Vasili Novitskii, Farshed Abdukhakimov, Dmitry Kamzolov, Aleksandr Beznosikov, Martin Takac, Pavel Dvurechensky, Bin Gu</i>	
A Functional Information Perspective on Model Interpretation .....	7266
<i>Itai Gat, Nitay Calderon, Roi Reichart, Tamir Hazan</i>	
UniRank: Unimodal Bandit Algorithms for Online Ranking .....	7279
<i>Camille-Sovanneary Gauthier, Romaric Gaudel, Elisa Fromont</i>	
Variational Inference with Locally Enhanced Bounds for Hierarchical Models .....	7310
<i>Tomas Geffner, Justin Domke</i>	
Inducing Causal Structure for Interpretable Neural Networks .....	7324
<i>Atticus Geiger, Zhengxuan Wu, Hanson Lu, Josh Rozner, Elisa Kreiss, Thomas Icard, Noah Goodman, Christopher Potts</i>	
Achieving Minimax Rates in Pool-Based Batch Active Learning.....	7339
<i>Claudio Gentile, Zhilei Wang, Tong Zhang</i>	
Near-Exact Recovery for Tomographic Inverse Problems Via Deep Learning .....	7368
<i>Martin Genzel, Ingo Gühring, Jan Macdonald, Maximilian März</i>	
Online Learning for Min Sum Set Cover and Pandora's Box .....	7382
<i>Evangelia Gergatsouli, Christos Tzamos</i>	
Equivariance Versus Augmentation for Spherical Images.....	7404
<i>Jan Gerken, Oscar Carlsson, Hampus Linander, Fredrik Ohlsson, Christoffer Petersson, Daniel Persson</i>	
A Regret Minimization Approach to Multi-Agent Control.....	7422
<i>Udaya Ghai, Udari Madhushani, Naomi Leonard, Elad Hazan</i>	
Blocks Assemble! Learning to Assemble with Large-Scale Structured Reinforcement Learning.....	7435
<i>Seyed Kamyar Seyed Ghasemipour, Satoshi Kataoka, Byron David, Daniel Freeman, Shixiang Shane Gu, Igor Mordatch</i>	

Faster Privacy Accounting Via Evolving Discretization .....	7470
<i>Badih Ghazi, Pritish Kamath, Ravi Kumar, Pasin Manurangsi</i>	

## PART 10

Plug-In Inversion: Model-Agnostic Inversion for Vision with Data Augmentations .....	7484
<i>Amin Ghiasi, Hamid Kazemi, Steven Reich, Chen Zhu, Micah Goldblum, Tom Goldstein</i>	
Offline RL Policies Should Be Trained to Be Adaptive .....	7513
<i>Dibya Ghosh, Anurag Ajay, Pulkit Agrawal, Sergey Levine</i>	
Breaking the $\sqrt{T}$ Barrier: Instance-Independent Logarithmic Regret in Stochastic Contextual Linear Bandits .....	7531
<i>Avishek Ghosh, Abishek Sankararaman</i>	
SCHA-VAE: Hierarchical Context Aggregation for Few-Shot Generation.....	7550
<i>Giorgio Giannone, Ole Winther</i>	
A Joint Exponential Mechanism for Differentially Private Top-K .....	7570
<i>Jennifer Gillenwater, Matthew Joseph, Andres Munoz, Monica Ribero Diaz</i>	
Neuro-Symbolic Hierarchical Rule Induction .....	7583
<i>Claire Glanois, Zhaohui Jiang, Xuening Feng, Paul Weng, Matthieu Zimmer, Dong Li, Wulong Liu, Jianye Hao</i>	
It's Raw! Audio Generation with State-Space Models .....	7616
<i>Karan Goel, Albert Gu, Chris Donahue, Christopher Re</i>	
RankSim: Ranking Similarity Regularization for Deep Imbalanced Regression .....	7634
<i>Yu Gong, Greg Mori, Fred Tung</i>	
How to Fill the Optimum Set? Population Gradient Descent with Harmless Diversity .....	7650
<i>Chengyue Gong, Lemeng Wu, Qiang Liu</i>	
Partial Label Learning Via Label Influence Function.....	7665
<i>Xiuwen Gong, Dong Yuan, Wei Bao</i>	
Secure Distributed Training at Scale .....	7679
<i>Eduard Gorbunov, Alexander Borzunov, Michael Diskin, Max Ryabinin</i>	
Retrieval-Augmented Reinforcement Learning.....	7740
<i>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</i>	
The State of Sparse Training in Deep Reinforcement Learning .....	7766
<i>Laura Graesser, Utku Evci, Erich Elsen, Pablo Samuel Castro</i>	
Causal Inference Through the Structural Causal Marginal Problem .....	7793
<i>Luigi Gresele, Julius Von Kügelgen, Jonas Kübler, Elke Kirschbaum, Bernhard Schölkopf, Dominik Janzing</i>	
Mirror Learning: A Unifying Framework of Policy Optimisation.....	7825
<i>Jakub Grudzien, Christian A Schroeder De Witt, Jakob Foerster</i>	
Adapting K-Means Algorithms for Outliers .....	7845
<i>Christoph Grunau, Václav Rozhon</i>	
Variational Mixtures of ODEs for Inferring Cellular Gene Expression Dynamics.....	7887
<i>Yichen Gu, David T Blaauw, Joshua Welch</i>	
Learning Pseudometric-Based Action Representations for Offline Reinforcement Learning .....	7902
<i>Pengjie Gu, Mengchen Zhao, Chen Chen, Dong Li, Jianye Hao, Bo An</i>	
NeuroFluid: Fluid Dynamics Grounding with Particle-Driven Neural Radiance Fields .....	7919
<i>Shanyan Guan, Huayu Deng, Yunbo Wang, Xiaokang Yang</i>	
Fast-Rate PAC-Bayesian Generalization Bounds for Meta-Learning .....	7930
<i>Jiechao Guan, Zhiwu Lu</i>	
Leveraging Approximate Symbolic Models for Reinforcement Learning Via Skill Diversity.....	7949
<i>Lin Guan, Sarath Sreedharan, Subbarao Kambhampati</i>	

Large-Scale Graph Neural Architecture Search.....	7968
<i>Chaoyu Guan, Xin Wang, Hong Chen, Ziwei Zhang, Wenwu Zhu</i>	
Identifiability Conditions for Domain Adaptation.....	7982
<i>Ishaan Gulrajani, Tatsunori Hashimoto</i>	
A Parametric Class of Approximate Gradient Updates for Policy Optimization.....	7998
<i>Ramki Gummadi, Saurabh Kumar, Junfeng Wen, Dale Schuurmans</i>	
Provably Efficient Offline Reinforcement Learning for Partially Observable Markov Decision Processes .....	8016
<i>Hongyi Guo, Qi Cai, Yufeng Zhang, Zhuoran Yang, Zhaoran Wang</i>	
No-Regret Learning in Partially-Informed Auctions .....	8039
<i>Wenshuo Guo, Michael Jordan, Ellen Vitercik</i>	
Bounding Training Data Reconstruction in Private (Deep) Learning.....	8056
<i>Chuan Guo, Brian Karrer, Kamalika Chaudhuri, Laurens Van Der Maaten</i>	
Adversarially Trained Neural Representations Are Already as Robust as Biological Neural Representations .....	8072
<i>Chong Guo, Michael Lee, Guillaume Leclerc, Joel Dapello, Yug Rao, Aleksander Madry, James Dicarlo</i>	
Class-Imbalanced Semi-Supervised Learning with Adaptive Thresholding.....	8082
<i>Lan-Zhe Guo, Yu-Feng Li</i>	
Deep Squared Euclidean Approximation to the Levenshtein Distance for DNA Storage .....	8095
<i>Alan J. X. Guo, Cong Liang, Qing-Hu Hou</i>	
Online Continual Learning Through Mutual Information Maximization.....	8109
<i>Yiduo Guo, Bing Liu, Dongyan Zhao</i>	
Fast Provably Robust Decision Trees and Boosting.....	8127
<i>Jun-Qi Guo, Ming-Zhuo Teng, Wei Gao, Zhi-Hua Zhou</i>	
Understanding and Improving Knowledge Graph Embedding for Entity Alignment .....	8145
<i>Lingbing Guo, Qiang Zhang, Zequn Sun, Mingyang Chen, Wei Hu, Huajun Chen</i>	
NISPA: Neuro-Inspired Stability-Plasticity Adaptation for Continual Learning in Sparse Networks.....	8157
<i>Mustafa B Gurbuz, Constantine Dovrolis</i>	
Active Learning on a Budget: Opposite Strategies Suit High and Low Budgets .....	8175
<i>Guy Hacoen, Avihu Dekel, Daphna Weinshall</i>	
You Only Cut Once: Boosting Data Augmentation with a Single Cut .....	8196
<i>Junlin Han, Pengfei Fang, Weihao Li, Jie Hong, Mohammad Ali Armin, Ian Reid, Lars Petersson, Hongdong Li</i>	
Scalable MCMC Sampling for Nonsymmetric Determinantal Point Processes .....	8213
<i>Insu Han, Mike Gartrell, Elvis Dohmatob, Amin Karbasi</i>	
G-Mixup: Graph Data Augmentation for Graph Classification.....	8230
<i>Xiaotian Han, Zhimeng Jiang, Ninghao Liu, Xia Hu</i>	
Private Streaming SCO in $\ell_p$ Geometry with Applications in High Dimensional Online Decision Making .....	8249
<i>Yuxuan Han, Zhicong Liang, Zhipeng Liang, Yang Wang, Yuan Yao, Jiheng Zhang</i>	
Off-Policy Reinforcement Learning with Delayed Rewards.....	8280
<i>Beining Han, Zhizhou Ren, Zuofan Wu, Yuan Zhou, Jian Peng</i>	
Adversarial Attacks on Gaussian Process Bandits.....	8304
<i>Eric Han, Jonathan Scarlett</i>	

## PART 11

Random Gegenbauer Features for Scalable Kernel Methods .....	8330
<i>Insu Han, Amir Zandieh, Haim Avron</i>	
Stochastic Reweighted Gradient Descent.....	8359
<i>Ayoub El Hanchi, David Stephens, Chris Maddison</i>	

Dual Perspective of Label-Specific Feature Learning for Multi-Label Classification .....	8375
<i>Jun-Yi Hang, Min-Ling Zhang</i>	
Temporal Difference Learning for Model Predictive Control .....	8387
<i>Nicklas A Hansen, Hao Su, Xiaolong Wang</i>	
Bisimulation Makes Analogies in Goal-Conditioned Reinforcement Learning .....	8407
<i>Philippe Hansen-Estruch, Amy Zhang, Ashvin Nair, Patrick Yin, Sergey Levine</i>	
TURF: Two-Factor, Universal, Robust, Fast Distribution Learning Algorithm .....	8427
<i>Yi Hao, Ayush Jain, Alon Orlitsky, Vaishakh Ravindrakumar</i>	
Contextual Information-Directed Sampling .....	8446
<i>Botao Hao, Tor Lattimore, Chao Qin</i>	
GSmooth: Certified Robustness Against Semantic Transformations Via Generalized Randomized Smoothing .....	8465
<i>Zhongkai Hao, Chengyang Ying, Yinpeng Dong, Hang Su, Jian Song, Jun Zhu</i>	
Implicit Regularization with Polynomial Growth in Deep Tensor Factorization .....	8484
<i>Kais Hariz, Hachem Kadri, Stephane Ayache, Maher Moakher, Thierry Artieres</i>	
Strategic Instrumental Variable Regression: Recovering Causal Relationships from Strategic Responses .....	8502
<i>Keegan Harris, Dung Daniel T Ngo, Logan Stapleton, Hoda Heidari, Steven Wu</i>	
C*-Algebra Net: A New Approach Generalizing Neural Network Parameters to C*-Algebra .....	8523
<i>Yuka Hashimoto, Zhao Wang, Tomoko Matsui</i>	
General-Purpose, Long-Context Autoregressive Modeling with Perceiver AR .....	8535
<i>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</i>	
On Distribution Shift in Learning-Based Bug Detectors .....	8559
<i>Jingxuan He, Luca Beurer-Kellner, Martin Vechev</i>	
GNNRank: Learning Global Rankings from Pairwise Comparisons Via Directed Graph Neural Networks .....	8581
<i>Yixuan He, Quan Gan, David Wipf, Gesine D Reinert, Junchi Yan, Mihai Cucuringu</i>	
Exploring the Gap Between Collapsed & Whitened Features in Self-Supervised Learning .....	8613
<i>Bobby He, Mete Ozay</i>	
Sparse Double Descent: Where Network Pruning Aggravates Overfitting .....	8635
<i>Zheng He, Zeke Xie, Quanzhi Zhu, Zengchang Qin</i>	
A Reduction from Linear Contextual Bandits Lower Bounds to Estimations Lower Bounds.....	8660
<i>Jiahao He, Jiheng Zhang, Rachel Q. Zhang</i>	
HyperPrompt: Prompt-Based Task-Conditioning of Transformers .....	8678
<i>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</i>	
Label-Descriptive Patterns and Their Application to Characterizing Classification Errors .....	8691
<i>Michael A. Hedderich, Jonas Fischer, Dietrich Klakow, Jilles Vreeken</i>	
NOMU: Neural Optimization-Based Model Uncertainty .....	8708
<i>Jakob M Heiss, Jakob Weissteiner, Hanna S Wutte, Sven Seuken, Josef Teichmann</i>	
Scaling Out-Of-Distribution Detection for Real-World Settings.....	8759
<i>Dan Hendrycks, Steven Basart, Mantas Mazeika, Andy Zou, Joseph Kwon, Mohammadreza Mostajabi, Jacob Steinhardt, Dawn Song</i>	
Generalization Bounds Using Lower Tail Exponents in Stochastic Optimizers.....	8774
<i>Liam Hodgkinson, Umut Simsekli, Rajiv Khanna, Michael Mahoney</i>	
Unsupervised Detection of Contextualized Embedding Bias with Application to Ideology .....	8796
<i>Valentin Hofmann, Janet Pierrehumbert, Hinrich Schütze</i>	
Neural Laplace: Learning Diverse Classes of Differential Equations in the Laplace Domain .....	8811
<i>Samuel I Holt, Zhaozhi Qian, Mihaela Van Der Schaar</i>	

Deep Hierarchy in Bandits .....	8833
<i>Joey Hong, Branislav Kveton, Sumeet Katariya, Manzil Zaheer, Mohammad Ghavamzadeh</i>	
DAdaQuant: Doubly-Adaptive Quantization for Communication-Efficient Federated Learning .....	8852
<i>Robert Hönig, Yiren Zhao, Robert Mullins</i>	
Equivariant Diffusion for Molecule Generation in 3D .....	8867
<i>Emiel Hoogeboom, Vi'ctor Garcia Satorras, Clément Vignac, Max Welling</i>	
Conditional GANs with Auxiliary Discriminative Classifier .....	8888
<i>Liang Hou, Qi Cao, Huawei Shen, Siyuan Pan, Xiaoshuang Li, Xueqi Cheng</i>	
AdaAUC: End-To-End Adversarial AUC Optimization Against Long-Tail Problems.....	8903
<i>Wenzheng Hou, Qianqian Xu, Zhiyong Yang, Shilong Bao, Yuan He, Qingming Huang</i>	
Wide Bayesian Neural Networks Have a Simple Weight Posterior: Theory and Accelerated Sampling.....	8926
<i>Jiri Hron, Roman Novak, Jeffrey Pennington, Jascha Sohl-Dickstein</i>	
Learning Inverse Folding from Millions of Predicted Structures .....	8946
<i>Chloe Hsu, Robert Verkuil, Jason Liu, Zeming Lin, Brian Hie, Tom Sercu, Adam Lerer, Alexander Rives</i>	
Nearly Minimax Optimal Reinforcement Learning with Linear Function Approximation .....	8971
<i>Pihe Hu, Yu Chen, Longbo Huang</i>	
Neuron Dependency Graphs: A Causal Abstraction of Neural Networks.....	9020
<i>Yaojie Hu, Jin Tian</i>	
Policy Diagnosis Via Measuring Role Diversity in Cooperative Multi-Agent RL.....	9041
<i>Siyi Hu, Chuanlong Xie, Xiaodan Liang, Xiaojun Chang</i>	
On the Role of Discount Factor in Offline Reinforcement Learning.....	9072
<i>Hao Hu, Yiqin Yang, Qianchuan Zhao, Chongjie Zhang</i>	
Transformer Quality in Linear Time.....	9099
<i>Weizhe Hua, Zihang Dai, Hanxiao Liu, Quoc Le</i>	
Language Models as Zero-Shot Planners: Extracting Actionable Knowledge for Embodied Agents .....	9118
<i>Wenlong Huang, Pieter Abbeel, Deepak Pathak, Igor Mordatch</i>	
Forward Operator Estimation in Generative Models with Kernel Transfer Operators .....	9148
<i>Zhichun Huang, Rudrasis Chakraborty, Vikas Singh</i>	

## PART 12

Adaptive Best-Of-Both-Worlds Algorithm for Heavy-Tailed Multi-Armed Bandits .....	9173
<i>Jiatai Huang, Yan Dai, Longbo Huang</i>	
Frustratingly Easy Transferability Estimation .....	9201
<i>Long-Kai Huang, Junzhou Huang, Yu Rong, Qiang Yang, Ying Wei</i>	
Modality Competition: What Makes Joint Training of Multi-Modal Network Fail in Deep Learning? (Provably).....	9226
<i>Yu Huang, Junyang Lin, Chang Zhou, Hongxia Yang, Longbo Huang</i>	
Action-Sufficient State Representation Learning for Control with Structural Constraints.....	9260
<i>Biwei Huang, Chaochao Lu, Liu Leqi, Jose Miguel Hernandez-Lobato, Clark Glymour, Bernhard Schölkopf, Kun Zhang</i>	
3DLinker: An E(3) Equivariant Variational Autoencoder for Molecular Linker Design.....	9280
<i>Yinan Huang, Xingang Peng, Jianzhu Ma, Muhan Zhang</i>	
SDQ: Stochastic Differentiable Quantization with Mixed Precision.....	9295
<i>Xijie Huang, Zhiqiang Shen, Shichao Li, Zechun Liu, Hu Xianghong, Jeffrey Wicaksana, Eric Xing, Kwang-Ting Cheng</i>	
Tackling Data Heterogeneity: A New Unified Framework for Decentralized SGD with Sample-Induced Topology .....	9310
<i>Yan Huang, Ying Sun, Zehan Zhu, Changzhi Yan, Jinming Xu</i>	

Efficient Representation Learning Via Adaptive Context Pooling .....	9346
<i>Chen Huang, Walter Talbott, Navdeep Jaitly, Joshua M Susskind</i>	
On the Learning of Non-Autoregressive Transformers .....	9356
<i>Fei Huang, Tianhua Tao, Hao Zhou, Lei Li, Minlie Huang</i>	
Going Deeper into Permutation-Sensitive Graph Neural Networks .....	9377
<i>Zhongyu Huang, Yingheng Wang, Chaozhuo Li, Huiguang He</i>	
Directed Acyclic Transformer for Non-Autoregressive Machine Translation .....	9410
<i>Fei Huang, Hao Zhou, Yang Liu, Hang Li, Minlie Huang</i>	
Unsupervised Ground Metric Learning Using Wasserstein Singular Vectors .....	9429
<i>Geert-Jan Huizing, Laura Cantini, Gabriel Peyré</i>	
Robust Kernel Density Estimation with Median-Of-Means Principle .....	9444
<i>Pierre Humbert, Batiste Le Bars, Ludovic Minvielle</i>	
A Data-Driven Approach for Learning to Control Computers .....	9466
<i>Peter C Humphreys, David Raposo, Tobias Pohlen, Gregory Thornton, Rachita Chhaparia, Alistair Muldal, Josh Abramson, Petko Georgiev, Adam Santoro, Timothy Lillicrap</i>	
Proximal Denoiser for Convergent Plug-And-Play Optimization with Nonconvex Regularization .....	9483
<i>Samuel Hurault, Arthur Leclaire, Nicolas Papadakis</i>	
Inverse Contextual Bandits: Learning How Behavior Evolves Over Time .....	9506
<i>Alihan Hüyük, Daniel Jarrett, Mihaela Van Der Schaar</i>	
Datamodels: Understanding Predictions with Data and Data with Predictions .....	9525
<i>Andrew Ilyas, Sung Min Park, Logan Engstrom, Guillaume Leclerc, Aleksander Madry</i>	
Parsimonious Learning-Augmented Caching .....	9588
<i>Sungjin Im, Ravi Kumar, Aditya Petety, Manish Purohit</i>	
Bayesian Optimization for Distributionally Robust Chance-Constrained Problem.....	9602
<i>Yu Inatsu, Shion Takeno, Masayuki Karasuyama, Ichiro Takeuchi</i>	
LeNSE: Learning to Navigate Subgraph Embeddings for Large-Scale Combinatorial Optimisation .....	9622
<i>David Ireland, Giovanni Montana</i>	
The Dual Form of Neural Networks Revisited: Connecting Test Time Predictions to Training Patterns Via Spotlights of Attention .....	9639
<i>Kazuki Irie, Róbert Csordás, Jürgen Schmidhuber</i>	
A Modern Self-Referential Weight Matrix that Learns to Modify Itself .....	9660
<i>Kazuki Irie, Imanol Schlag, Róbert Csordás, Jürgen Schmidhuber</i>	
Revisiting Online Submodular Minimization: Gap-Dependent Regret Bounds, Best of Both Worlds and Adversarial Robustness.....	9678
<i>Shinji Ito</i>	
Modeling Strong and Human-Like Gameplay with KL-Regularized Search.....	9695
<i>Athul Paul Jacob, David J Wu, Gabriele Farina, Adam Lerer, Hengyuan Hu, Anton Bakhtin, Jacob Andreas, Noam Brown</i>	
A Deep Convolutional Neural Network that is Invariant to Time Rescaling.....	9729
<i>Brandon G Jacques, Zoran Tiganj, Aakash Sarkar, Marc Howard, Per Sederberg</i>	
Input Dependent Sparse Gaussian Processes.....	9739
<i>Bahram Jafrasteh, Carlos Villacampa-Calvo, Daniel Hernandez-Lobato</i>	
Regret Minimization with Performative Feedback.....	9760
<i>Meena Jagadeesan, Tijana Zrnic, Celestine Mender-Dünner</i>	
Biological Sequence Design with GFlowNets.....	9786
<i>Moksh Jain, Emmanuel Bengio, Alex Hernandez-Garcia, Jarrid Rector-Brooks, Bonaventure F. P. Dossou, Chanakya Ajit Ekbote, Jie Fu, Tianyu Zhang, Michael Kilgour, Dinghuai Zhang, Lena Simine, Payel Das, Yoshua Bengio</i>	
Combining Diverse Feature Priors .....	9802
<i>Saachi Jain, Dimitris Tsipras, Aleksander Madry</i>	
Training Your Sparse Neural Network Better with Any Mask .....	9833
<i>Ajay Kumar Jaiswal, Haoyu Ma, Tianlong Chen, Ying Ding, Zhangyang Wang</i>	

Sequential Covariate Shift Detection Using Classifier Two-Sample Tests.....	9845
<i>Sooyong Jang, Sangdon Park, Insup Lee, Osbert Bastani</i>	
Surrogate Likelihoods for Variational Annealed Importance Sampling .....	9881
<i>Martin Jankowiak, Du Phan</i>	
Planning with Diffusion for Flexible Behavior Synthesis .....	9902
<i>Michael Janner, Yilun Du, Joshua Tenenbaum, Sergey Levine</i>	
HyperImpute: Generalized Iterative Imputation with Automatic Model Selection .....	9916
<i>Daniel Jarrett, Bogdan C Cebere, Tennison Liu, Alicia Curth, Mihaela Van Der Schaar</i>	
Mitigating Modality Collapse in Multimodal VAEs Via Impartial Optimization.....	9938
<i>Adrian Javaloy, Maryam Meghdadi, Isabel Valera</i>	
Towards Understanding How Momentum Improves Generalization in Deep Learning.....	9965
<i>Samy Jelassi, Yuanzhi Li</i>	

### PART 13

MASER: Multi-Agent Reinforcement Learning with Subgoals Generated from Experience Replay Buffer.....	10041
<i>Jeewon Jeon, Woojun Kim, Whiyoung Jung, Youngchul Sung</i>	
An Exact Symbolic Reduction of Linear Smart Predict+Optimize to Mixed Integer Linear Programming .....	10053
<i>Jihwan Jeong, Parth Jaggi, Andrew Butler, Scott Sanner</i>	
Agnostic Learnability of Halfspaces Via Logistic Loss .....	10068
<i>Ziwei Ji, Kwangjun Ahn, Pranjal Awasthi, Satyen Kale, Stefani Karp</i>	
Improving Policy Optimization with Generalist-Specialist Learning.....	10104
<i>Zhiwei Jia, Xuanlin Li, Zhan Ling, Shuang Liu, Yiran Wu, Hao Su</i>	
Translatotron 2: High-Quality Direct Speech-To-Speech Translation with Voice Preservation.....	10120
<i>Ye Jia, Michelle Tadmor Ramanovich, Tal Remez, Roi Pomerantz</i>	
Online Learning and Pricing with Reusable Resources: Linear Bandits with Sub-Exponential Rewards.....	10135
<i>Huiwen Jia, Cong Shi, Siqian Shen</i>	
The Role of Deconfounding in Meta-Learning .....	10161
<i>Yinjie Jiang, Zhengyu Chen, Kun Kuang, Luotian Yuan, Xinhai Ye, Zhihua Wang, Fei Wu, Ying Wei</i>	
Subspace Learning for Effective Meta-Learning.....	10177
<i>Weisen Jiang, James Kwok, Yu Zhang</i>	
Optimal Algorithms for Stochastic Multi-Level Compositional Optimization.....	10195
<i>Wei Jiang, Bokun Wang, Yibo Wang, Lijun Zhang, Tianbao Yang</i>	
Antibody-Antigen Docking and Design Via Hierarchical Structure Refinement .....	10217
<i>Wengong Jin, Dr. Regina Barzilay, Tommi Jaakkola</i>	
Sharpened Quasi-Newton Methods: Faster Superlinear Rate and Larger Local Convergence Neighborhood.....	10228
<i>Qiujiang Jin, Alec Koppel, Ketan Rajawat, Aryan Mokhtari</i>	
The Power of Exploiter: Provable Multi-Agent RL in Large State Spaces .....	10251
<i>Chi Jin, Qinghua Liu, Tiancheng Yu</i>	
Domain Adaptation for Time Series Forecasting Via Attention Sharing .....	10280
<i>Xiaoyong Jin, Youngsuk Park, Danielle Maddix, Hao Wang, Yuyang Wang</i>	
Accelerated Federated Learning with Decoupled Adaptive Optimization .....	10298
<i>Jiayin Jin, Jiaxiang Ren, Yang Zhou, Lingjuan Lyu, Ji Liu, Dejing Dou</i>	
Supervised Off-Policy Ranking.....	10323
<i>Yue Jin, Yue Zhang, Tao Qin, Xudong Zhang, Jian Yuan, Houqiang Li, Tie-Yan Liu</i>	
Input-Agnostic Certified Group Fairness Via Gaussian Parameter Smoothing.....	10340
<i>Jiayin Jin, Zeru Zhang, Yang Zhou, Lingfei Wu</i>	

Score-Based Generative Modeling of Graphs Via the System of Stochastic Differential Equations .....	10362
<i>Jaehyeong Jo, Seul Lee, Sung Ju Hwang</i>	
Choosing Answers in Epsilon-Best-Answer Identification for Linear Bandits .....	10384
<i>Marc Jourdan, Rémy Degenne</i>	
Robust Fine-Tuning of Deep Neural Networks with Hessian-Based Generalization Guarantees .....	10431
<i>Haotian Ju, Dongyue Li, Hongyang R Zhang</i>	
Robust Alignment of Cross-Session Recordings of Neural Population Activity by Behaviour Via Unsupervised Domain Adaptation.....	10462
<i>Justin Jude, Matthew Perich, Lee Miller, Matthias Hennig</i>	
On Measuring Causal Contributions Via Do-Interventions.....	10476
<i>Yonghan Jung, Shiva Kasiviswanathan, Jin Tian, Dominik Janzing, Patrick Bloebaum, Elias Bareinboim</i>	
Efficient Approximate Inference for Stationary Kernel on Frequency Domain .....	10502
<i>Yohan Jung, Kyungwoo Song, Jinkyoo Park</i>	
Sketching Algorithms and Lower Bounds for Ridge Regression .....	10539
<i>Praneeth Kacham, David Woodruff</i>	
Flashlight: Enabling Innovation in Tools for Machine Learning.....	10557
<i>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</i>	
Learning-Based Optimisation of Particle Accelerators Under Partial Observability Without Real- World Training.....	10575
<i>Jan Kaiser, Oliver Stein, Annika Eichler</i>	
Stochastic Deep Networks with Linear Competing Units for Model-Agnostic Meta-Learning.....	10586
<i>Konstantinos Kalais, Sotirios Chatzis</i>	
Doubly Robust Distributionally Robust Off-Policy Evaluation and Learning .....	10598
<i>Nathan Kallus, Xiaojie Mao, Kaiwen Wang, Zhengyuan Zhou</i>	
Improved Rates for Differentially Private Stochastic Convex Optimization with Heavy-Tailed Data.....	10633
<i>Gautam Kamath, Xingtu Liu, Huanyu Zhang</i>	
Comprehensive Analysis of Negative Sampling in Knowledge Graph Representation Learning .....	10661
<i>Hidetaka Kamigaito, Katsuhiko Hayashi</i>	
Matching Learned Causal Effects of Neural Networks with Domain Priors.....	10676
<i>Sai Srinivas Kancheti, Abbavaram Gowtham Reddy, Vineeth N Balasubramanian, Amit Sharma</i>	
Deduplicating Training Data Mitigates Privacy Risks in Language Models.....	10697
<i>Nikhil Kandpal, Eric Wallace, Colin Raffel</i>	
Lyapunov Density Models: Constraining Distribution Shift in Learning-Based Control.....	10708
<i>Katie Kang, Paula Gradu, Jason J Choi, Michael Janner, Claire Tomlin, Sergey Levine</i>	
Forget-Free Continual Learning with Winning Subnetworks.....	10734
<i>Haeyong Kang, Rusty John Lloyd Mina, Sultan Rizky Hikmawan Madjid, Jaehong Yoon, Mark Hasegawa-Johnson, Sung Ju Hwang, Chang D. Yoo</i>	
Differentially Private Approximate Quantiles .....	10751
<i>Haim Kaplan, Shachar Schnapp, Uri Stemmer</i>	
Simultaneous Graph Signal Clustering and Graph Learning.....	10762
<i>Abdullah Karaaslanli, Selin Aviyente</i>	
Composing Partial Differential Equations with Physics-Aware Neural Networks .....	10773
<i>Matthias Karlbauer, Timothy Praditia, Sebastian Otte, Sergey Oladyshkin, Wolfgang Nowak, Martin V. Butz</i>	
Meta-Learning Hypothesis Spaces for Sequential Decision-Making .....	10802
<i>Parnian Kassarai, Jonas Rothfuss, Andreas Krause</i>	
FOCUS: Familiar Objects in Common and Uncommon Settings .....	10825
<i>Priyatham Kattakinda, Soheil Feizi</i>	



Training OOD Detectors in Their Natural Habitats.....	10848
<i>Julian Katz-Samuels, Julia B Nakhleh, Robert Nowak, Yixuan Li</i>	

## PART 14

Robustness Implies Generalization Via Data-Dependent Generalization Bounds.....	10866
<i>Kenji Kawaguchi, Zhun Deng, Kyle Luh, Jiaoyang Huang</i>	
Generating Distributional Adversarial Examples to Evade Statistical Detectors .....	10896
<i>Yigitcan Kaya, Muhammad Bilal Zafar, Sergul Aydore, Nathalie Rauschmayr, Krishnaram Kenthapadi</i>	
Secure Quantized Training for Deep Learning.....	10913
<i>Marcel Keller, Ke Sun</i>	
A Convergent and Dimension-Independent Min-Max Optimization Algorithm .....	10940
<i>Vijay Keswani, Oren Mangoubi, Sushant Sachdeva, Nisheeth K. Vishnoi</i>	
Neural Network Poisson Models for Behavioural and Neural Spike Train Data.....	10975
<i>Moein Khajehnejad, Forough Habibollahi, Richard Nock, Ehsan Arabzadeh, Peter Dayan, Amir Dezfouli</i>	
Federated Reinforcement Learning: Linear Speedup Under Markovian Sampling.....	10998
<i>Sajad Khodadadian, Pranay Sharma, Gauri Joshi, Siva Theja Maguluri</i>	
Multi-Level Branched Regularization for Federated Learning .....	11059
<i>Jinkyu Kim, Geeho Kim, Bohyung Han</i>	
Learning Fair Representation with a Parametric Integral Probability Metric.....	11075
<i>Dongha Kim, Kunwoong Kim, Insung Kong, Ilsang Ohn, Yongdai Kim</i>	
Dataset Condensation Via Efficient Synthetic-Data Parameterization .....	11103
<i>Jang-Hyun Kim, Jinuk Kim, Seong Joon Oh, Sangdoon Yun, Hwanjun Song, Joonhyun Jeong, Jung-Woo Ha, Hyun Oh Song</i>	
Guided-TTS: A Diffusion Model for Text-To-Speech Via Classifier Guidance .....	11120
<i>Heeseung Kim, Sungwon Kim, Sungroh Yoon</i>	
Variational On-The-Fly Personalization .....	11135
<i>Jangho Kim, Jun-Tae Lee, Simyung Chang, Nojun Kwak</i>	
Fisher SAM: Information Geometry and Sharpness Aware Minimisation .....	11149
<i>Minyoung Kim, Da Li, Shell X Hu, Timothy Hospedales</i>	
ViT-NeT: Interpretable Vision Transformers with Neural Tree Decoder.....	11163
<i>Sangwon Kim, Jaeyeal Nam, Byoung Chul Ko</i>	
Sanity Simulations for Saliency Methods.....	11174
<i>Joon Sik Kim, Gregory Plumb, Ameet Talwalkar</i>	
Soft Truncation: A Universal Training Technique of Score-Based Diffusion Model for High Precision Score Estimation.....	11202
<i>Dongjun Kim, Seungjae Shin, Kyungwoo Song, Wanmo Kang, Il-Chul Moon</i>	
Rotting Infinitely Many-Armed Bandits .....	11230
<i>Jung-Hun Kim, Milan Vojnovic, Se-Young Yun</i>	
Accelerated Gradient Methods for Geodesically Convex Optimization: Tractable Algorithms and Convergence Analysis .....	11256
<i>Jungbin Kim, Insoon Yang</i>	
Generalizing to New Physical Systems Via Context-Informed Dynamics Model.....	11284
<i>Matthieu Kirchmeyer, Yuan Yin, Jeremie Dona, Nicolas Baskiotis, Alain Rakotomamonjy, Patrick Gallinari</i>	
SoQal: Selective Oracle Questioning for Consistency Based Active Learning of Cardiac Signals.....	11303
<i>Dani Kiyasseh, Tingting Zhu, David A Clifton</i>	
Curriculum Reinforcement Learning Via Constrained Optimal Transport.....	11342
<i>Pascal Klink, Haoyi Yang, Carlo D'Eramo, Jan Peters, Joni Pajarinen</i>	

Exploiting Redundancy: Separable Group Convolutional Networks on Lie Groups .....	11360
<i>David M. Knigge, David W Romero, Erik J Bekkers</i>	
Revisiting Contrastive Learning Through the Lens of Neighborhood Component Analysis: An Integrated Framework .....	11388
<i>Ching-Yun Ko, Jeet Mohapatra, Sijia Liu, Pin-Yu Chen, Luca Daniel, Lily Weng</i>	
Transfer Learning in Differential Privacy’s Hybrid-Model.....	11414
<i>Refael Kohen, Or Sheffet</i>	
Markov Chain Monte Carlo for Continuous-Time Switching Dynamical Systems .....	11431
<i>Lukas Köhs, Bastian Alt, Heinz Koepl</i>	
Partial Disentanglement for Domain Adaptation.....	11456
<i>Lingjing Kong, Shaoran Xie, Weiran Yao, Yujia Zheng, Guangyi Chen, Petar Stojanov, Victor Akinwande, Kun Zhang</i>	
Simultaneously Learning Stochastic and Adversarial Bandits with General Graph Feedback.....	11474
<i>Fang Kong, Yichi Zhou, Shuai Li</i>	
Adaptive Data Analysis with Correlated Observations .....	11484
<i>Aryeh Kontorovich, Menachem Sadigurschi, Uri Stemmer</i>	
Controlling Conditional Language Models Without Catastrophic Forgetting .....	11500
<i>Tomasz Korbak, Hady Elsahar, German Kruszewski, Marc Dymetman</i>	
Batch Greenhorn Algorithm for Entropic-Regularized Multimarginal Optimal Transport: Linear Rate of Convergence and Iteration Complexity .....	11530
<i>Vladimir R. Kostic, Saverio Salzo, Massimiliano Pontil</i>	
Certified Adversarial Robustness Under the Bounded Support Set.....	11560
<i>Yiwen Kou, Qinyuan Zheng, Yisen Wang</i>	
Exact Learning of Preference Structure: Single-Peaked Preferences and Beyond .....	11599
<i>Sonja Kraiczy, Edith Elkind</i>	
Reconstructing Nonlinear Dynamical Systems from Multi-Modal Time Series .....	11614
<i>Daniel Kramer, Philine L Bommer, Carlo Tombolini, Georgia Koppe, Daniel Durstewitz</i>	
Probabilistic ODE Solutions in Millions of Dimensions.....	11635
<i>Nicholas Krämer, Nathanael Bosch, Jonathan Schmidt, Philipp Hennig</i>	
Active Nearest Neighbor Regression Through Delaunay Refinement .....	11651
<i>Alexander Kravberg, Giovanni Luca Marchetti, Vladislav Polianskii, Anastasiia Varava, Florian T. Pokorny, Danica Kragic</i>	
Functional Generalized Empirical Likelihood Estimation for Conditional Moment Restrictions.....	11666
<i>Heiner Kremer, Jia-Jie Zhu, Krikamol Muandet, Bernhard Schölkopf</i>	
Calibrated and Sharp Uncertainties in Deep Learning Via Density Estimation .....	11684
<i>Volodymyr Kuleshov, Shachi Deshpande</i>	

## PART 15

ActiveHedge: Hedge Meets Active Learning .....	11695
<i>Bhuvesh Kumar, Jacob D Abernethy, Venkatesh Saligrama</i>	
Balancing Discriminability and Transferability for Source-Free Domain Adaptation .....	11711
<i>Jogendra Nath Kundu, Akshay R Kulkarni, Suvaansh Bhambri, Deepesh Mehta, Shreyas Anand Kulkarni, Varun Jampani, Venkatesh Babu Radhakrishnan</i>	
Showing Your Offline Reinforcement Learning Work: Online Evaluation Budget Matters.....	11730
<i>Vladislav Kurenkov, Sergey Kolesnikov</i>	
Equivariant Priors for Compressed Sensing with Unknown Orientation .....	11754
<i>Anna Kuzina, Kumar Pratik, Fabio Valerio Massoli, Arash Behboodi</i>	
Coordinated Attacks Against Contextual Bandits: Fundamental Limits and Defense Mechanisms.....	11773
<i>Jeongyeol Kwon, Yonathan Efroni, Constantine Caramanis, Shie Mannor</i>	
Large Batch Experience Replay .....	11791
<i>Thibault Lahire, Matthieu Geist, Emmanuel Rachelson</i>	

FedScale: Benchmarking Model and System Performance of Federated Learning at Scale .....	11815
<i>Fan Lai, Yinwei Dai, Sanjay Singapuram, Jiachen Liu, Xiangfeng Zhu, Harsha Madhyastha, Mosharaf Chowdhury</i>	
Smoothed Adaptive Weighting for Imbalanced Semi-Supervised Learning: Improve Reliability Against Unknown Distribution Data .....	11829
<i>Zhengfeng Lai, Chao Wang, Henry Gunawan, Sen-Ching S Cheung, Chen-Nee Chuah</i>	
Functional Output Regression with Infimal Convolution: Exploring the Huber and $\epsilon$ -Insensitive Losses .....	11845
<i>Alex Lambert, Dimitri Bouche, Zoltan Szabo, Florence D'Alché-Buc</i>	
Tell Me Why! Explanations Support Learning Relational and Causal Structure.....	11869
<i>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</i>	
Generative Cooperative Networks for Natural Language Generation .....	11892
<i>Sylvain Lamprier, Thomas Scialom, Antoine Chaffin, Vincent Claveau, Ewa Kijak, Jacopo Staiano, Benjamin Piwowarski</i>	
DSTAGNN: Dynamic Spatial-Temporal Aware Graph Neural Network for Traffic Flow Forecasting .....	11907
<i>Shiyong Lan, Yitong Ma, Weikang Huang, Wenwu Wang, Hongyu Yang, Pyang Li</i>	
Cooperative Online Learning in Stochastic and Adversarial MDPs.....	11919
<i>Tal Lancewicki, Aviv Rosenberg, Yishay Mansour</i>	
PINs: Progressive Implicit Networks for Multi-Scale Neural Representations.....	11970
<i>Zoe Landgraf, Alexander Sorkine Hornung, Ricardo S Cabral</i>	
Co-Training Improves Prompt-Based Learning for Large Language Models .....	11986
<i>Hunter Lang, Monica N Agrawal, Yoon Kim, David Sontag</i>	
Goal Misgeneralization in Deep Reinforcement Learning .....	12005
<i>Lauro Langosco Di Langosco, Jack Koch, Lee D Sharkey, Jacob Pfau, David Krueger</i>	
Marginal Tail-Adaptive Normalizing Flows.....	12021
<i>Mike Laszkiewicz, Johannes Lederer, Asja Fischer</i>	
Bregman Proximal Langevin Monte Carlo Via Bregman-Moreau Envelopes .....	12050
<i>Tim Tsz-Kit Lau, Han Liu</i>	
Scalable Deep Reinforcement Learning Algorithms for Mean Field Games .....	12079
<i>Mathieu Lauriere, Sarah Perrin, Sertan Girgin, Paul Muller, Ayush Jain, Theophile Cabannes, Georgios Piliouras, Julien Perolat, Romuald Elie, Olivier Pietquin, Matthieu Geist</i>	
Implicit Bias of Linear Equivariant Networks.....	12097
<i>Hannah Lawrence, Kristian Georgiev, Andrew Dienes, Bobak T. Kiani</i>	
Differentially Private Maximal Information Coefficients .....	12127
<i>John Lazarsfeld, Aaron Johnson, Emmanuel Adeniran</i>	
Entropic Gromov-Wasserstein Between Gaussian Distributions.....	12165
<i>Khang Le, Dung Q Le, Huy Nguyen, Dat Do, Tung Pham, Nhat Ho</i>	
Neurocoder: General-Purpose Computation Using Stored Neural Programs.....	12205
<i>Hung Le, Svetha Venkatesh</i>	
Convergence of Policy Gradient for Entropy Regularized MDPs with Neural Network Approximation in the Mean-Field Regime.....	12223
<i>James-Michael Leahy, Bekzhan Kerimkulov, David Siska, Lukasz Szpruch</i>	
A Random Matrix Analysis of Data Stream Clustering: Coping with Limited Memory Resources .....	12254
<i>Hugo Lebeau, Romain Couillet, Florent Chatelain</i>	
Neural Tangent Kernel Analysis of Deep Narrow Neural Networks .....	12283
<i>Jongmin Lee, Joo Young Choi, Ernest K Ryu, Albert No</i>	
Dataset Condensation with Contrastive Signals .....	12353
<i>Saehyung Lee, Sanghyuk Chun, Sangwon Jung, Sangdoon Yun, Sungroh Yoon</i>	
Confidence Score for Source-Free Unsupervised Domain Adaptation .....	12366
<i>Jonghyun Lee, Dahuin Jung, Junho Yim, Sungroh Yoon</i>	

A Statistical Manifold Framework for Point Cloud Data .....	12379
<i>Yonghyeon Lee, Seungyeon Kim, Jinwon Choi, Frank Park</i>	
Low-Complexity Deep Convolutional Neural Networks on Fully Homomorphic Encryption Using Multiplexed Parallel Convolutions .....	12404
<i>Eunsang Lee, Joon-Woo Lee, Junghyun Lee, Young-Sik Kim, Yongjune Kim, Jong-Seon No, Woosuk Choi</i>	
Statistical Inference with Implicit SGD: Proximal Robbins-Monro Vs. Polyak-Ruppert .....	12424
<i>Yoonhyung Lee, Sungdong Lee, Joong-Ho Won</i>	
Maslow’s Hammer in Catastrophic Forgetting: Node Re-Use Vs. Node Activation.....	12456
<i>Sebastian Lee, Stefano Sarao Mannelli, Claudia Clopath, Sebastian Goldt, Andrew Saxe</i>	
Query-Efficient and Scalable Black-Box Adversarial Attacks on Discrete Sequential Data Via Bayesian Optimization .....	12479
<i>Deokjae Lee, Seungyong Moon, Junhyeok Lee, Hyun Oh Song</i>	
Least Squares Estimation Using Sketched Data with Heteroskedastic Errors.....	12499
<i>Sokbae Lee, Serena Ng</i>	
Why the Rich Get Richer? on the Balancedness of Random Partition Models .....	12522
<i>Changwoo J Lee, Huiyan Sang</i>	

## PART 16

Model Selection in Batch Policy Optimization .....	12543
<i>Jonathan Lee, George Tucker, Ofir Nachum, Bo Dai</i>	
Supervised Learning with General Risk Functionals .....	12571
<i>Liu Leqi, Audrey Huang, Zachary Lipton, Kamyar Azizzadenesheli</i>	
Generalized Strategic Classification and the Case of Aligned Incentives .....	12594
<i>Sagi Levanon, Nir Rosenfeld</i>	
A Simple Unified Framework for High Dimensional Bandit Problems .....	12620
<i>Wenjie Li, Adarsh Barik, Jean Honorio</i>	
Robust Training of Neural Networks Using Scale Invariant Architectures .....	12657
<i>Zhiyuan Li, Srinadh Bhojanapalli, Manzil Zaheer, Sashank Reddi, Sanjiv Kumar</i>	
Spatial-Channel Token Distillation for Vision MLPs .....	12686
<i>Yanxi Li, Xinghao Chen, Minjing Dong, Yehui Tang, Yunhe Wang, Chang Xu</i>	
An Analytical Update Rule for General Policy Optimization .....	12697
<i>Hepeng Li, Nicholas Clavette, Haibo He</i>	
On Convergence of Gradient Descent Ascent: A Tight Local Analysis.....	12718
<i>Haochuan Li, Farzan Farnia, Subhro Das, Ali Jadbabaie</i>	
On the Finite-Time Performance of the Knowledge Gradient Algorithm .....	12742
<i>Yanwen Li, Siyang Gao</i>	
Phasic Self-Imitative Reduction for Sparse-Reward Goal-Conditioned Reinforcement Learning.....	12766
<i>Yunfei Li, Tian Gao, Jiaqi Yang, Huazhe Xu, Yi Wu</i>	
G2CN: Graph Gaussian Convolution Networks with Concentrated Graph Filters .....	12783
<i>Mingjie Li, Xiaojun Guo, Yifei Wang, Yisen Wang, Zhouchen Lin</i>	
Decomposing Temporal High-Order Interactions Via Latent ODEs .....	12798
<i>Shibo Li, Robert Kirby, Shandian Zhe</i>	
Neural Inverse Transform Sampler.....	12814
<i>Henry Li, Yuval Kluger</i>	
PLATINUM: Semi-Supervised Model Agnostic Meta-Learning Using Submodular Mutual Information.....	12827
<i>Changbin Li, Suraj Kothawade, Feng Chen, Rishabh Iyer</i>	
Deconfounded Value Decomposition for Multi-Agent Reinforcement Learning.....	12844
<i>Jiahui Li, Kun Kuang, Baoxiang Wang, Furu Li, Long Chen, Changjie Fan, Fei Wu, Jun Xiao</i>	

C-MinHash: Improving Minwise Hashing with Circulant Permutation.....	12858
<i>Xiaoyun Li, Ping Li</i>	
BLIP: Bootstrapping Language-Image Pre-Training for Unified Vision-Language Understanding and Generation .....	12889
<i>Junnan Li, Dongxu Li, Caiming Xiong, Steven Hoi</i>	
Restarted Nonconvex Accelerated Gradient Descent: No More Polylogarithmic Factor in the $O(\epsilon^{-7/4})$ Complexity.....	12902
<i>Huan Li, Zhouchen Lin</i>	
Achieving Fairness at No Utility Cost Via Data Reweighting with Influence.....	12918
<i>Peizhao Li, Hongfu Liu</i>	
High Probability Guarantees for Nonconvex Stochastic Gradient Descent with Heavy Tails.....	12932
<i>Shaojie Li, Yong Liu</i>	
MetAug: Contrastive Learning Via Meta Feature Augmentation .....	12965
<i>Jiangmeng Li, Wenwen Qiang, Changwen Zheng, Bing Su, Hui Xiong</i>	
PMIC: Improving Multi-Agent Reinforcement Learning with Progressive Mutual Information Collaboration.....	12980
<i>Pengyi Li, Hongyao Tang, Tianpei Yang, Xiaotian Hao, Tong Sang, Yan Zheng, Jianye Hao, Matthew E. Taylor, Wenyuan Tao, Zhen Wang</i>	
CerDEQ: Certifiable Deep Equilibrium Model.....	12999
<i>Mingjie Li, Yisen Wang, Zhouchen Lin</i>	
Generalization Guarantee of Training Graph Convolutional Networks with Graph Topology Sampling.....	13015
<i>Hongkang Li, Meng Wang, Sijia Liu, Pin-Yu Chen, Jinjun Xiong</i>	
Let Invariant Rationale Discovery Inspire Graph Contrastive Learning .....	13053
<i>Sihang Li, Xiang Wang, An Zhang, Yingxin Wu, Xiangnan He, Tat-Seng Chua</i>	
Difference Advantage Estimation for Multi-Agent Policy Gradients.....	13067
<i>Yueheng Li, Guangming Xie, Zongqing Lu</i>	
Private Adaptive Optimization with Side Information .....	13087
<i>Tian Li, Manzil Zaheer, Sashank Reddi, Virginia Smith</i>	
Permutation Search of Tensor Network Structures Via Local Sampling .....	13107
<i>Chao Li, Junhua Zeng, Zerui Tao, Qibin Zhao</i>	
Hessian-Free High-Resolution Nesterov Acceleration for Sampling .....	13126
<i>Ruilin Li, Hongyuan Zha, Molei Tao</i>	
Double Sampling Randomized Smoothing .....	13164
<i>Linyi Li, Jiawei Zhang, Tao Xie, Bo Li</i>	
HousE: Knowledge Graph Embedding with Householder Parameterization .....	13210
<i>Rui Li, Jianan Zhao, Chaozhuo Li, Di He, Yiqi Wang, Yuming Liu, Hao Sun, Senzhang Wang, Weiwei Deng, Yanming Shen, Xing Xie, Qi Zhang</i>	
Learning Multiscale Transformer Models for Sequence Generation.....	13226
<i>Bei Li, Tong Zheng, Yi Jing, Chengbo Jiao, Tong Xiao, Jingbo Zhu</i>	
Finding Global Homophily in Graph Neural Networks When Meeting Heterophily .....	13243
<i>Xiang Li, Renyu Zhu, Yao Cheng, Caihua Shan, Siqiang Luo, Dongsheng Li, Weining Qian</i>	
Fat-Tailed Variational Inference with Anisotropic Tail Adaptive Flows .....	13258
<i>Feynman Liang, Michael Mahoney, Liam Hodgkinson</i>	
Exploring and Exploiting Hubness Priors for High-Quality GAN Latent Sampling.....	13272
<i>Yuanbang Liang, Jing Wu, Yu-Kun Lai, Yipeng Qin</i>	
Reducing Variance in Temporal-Difference Value Estimation Via Ensemble of Deep Networks .....	13286
<i>Litian Liang, Yaosheng Xu, Stephen McAleer, Dailin Hu, Alexander Ihler, Pieter Abbeel, Roy Fox</i>	
TSPipe: Learn from Teacher Faster with Pipelines .....	13303
<i>Hwijoon Lim, Yechan Kim, Sukmin Yun, Jinwoo Shin, Dongsu Han</i>	

Order Constraints in Optimal Transport .....	13314
<i>Yu Chin Fabian Lim, Laura Wynter, Shiau Hong Lim</i>	
Flow-Guided Sparse Transformer for Video Deblurring .....	13335
<i>Jing Lin, Yuanhao Cai, Xiaowan Hu, Haoqian Wang, Youliang Yan, Xueyi Zou, Henghui Ding, Yulun Zhang, Radu Timofte, Luc Van Gool</i>	
Federated Learning with Positive and Unlabeled Data.....	13345
<i>Xinyang Lin, Hanting Chen, Yixing Xu, Chao Xu, Xiaolin Gui, Yiping Deng, Yunhe Wang</i>	
Decentralized Online Convex Optimization in Networked Systems.....	13357
<i>Yiheng Lin, Judy Gan, Guannan Qu, Yash Kanoria, Adam Wierman</i>	

## PART 17

Unsupervised Flow-Aligned Sequence-To-Sequence Learning for Video Restoration .....	13395
<i>Jing Lin, Xiaowan Hu, Yuanhao Cai, Haoqian Wang, Youliang Yan, Xueyi Zou, Yulun Zhang, Luc Van Gool</i>	
Constrained Gradient Descent: A Powerful and Principled Evasion Attack Against Neural Networks .....	13406
<i>Weiran Lin, Keane Lucas, Lujo Bauer, Michael K. Reiter, Mahmood Sharif</i>	
Learning Augmented Binary Search Trees .....	13432
<i>Honghao Lin, Tian Luo, David Woodruff</i>	
Online Nonsubmodular Minimization with Delayed Costs: From Full Information to Bandit Feedback.....	13442
<i>Tianyi Lin, Aldo Pacchiano, Yaodong Yu, Michael Jordan</i>	
Measuring the Effect of Training Data on Deep Learning Predictions Via Randomized Experiments .....	13469
<i>Jinkun Lin, Anqi Zhang, Mathias Lécuyer, Jinyang Li, Aurojit Panda, Siddhartha Sen</i>	
Interactively Learning Preference Constraints in Linear Bandits .....	13506
<i>David Lindner, Sebastian Tschiatschek, Katja Hofmann, Andreas Krause</i>	
Delayed Reinforcement Learning by Imitation .....	13529
<i>Pierre Liotet, Davide Maran, Lorenzo Bisi, Marcello Restelli</i>	
CITRIS: Causal Identifiability from Temporal Intervened Sequences .....	13558
<i>Phillip Lippe, Sara Magliacane, Sindy Löwe, Yuki M Asano, Taco Cohen, Stratis Gavves</i>	
StreamingQA: A Benchmark for Adaptation to New Knowledge Over Time in Question Answering Models.....	13605
<i>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</i>	
Distributionally Robust Q-Learning .....	13624
<i>Zijian Liu, Qinxun Bai, Jose Blanchet, Perry Dong, Wei Xu, Zhengqing Zhou, Zhengyuan Zhou</i>	
Constrained Variational Policy Optimization for Safe Reinforcement Learning.....	13645
<i>Zuxin Liu, Zhepeng Cen, Vladislav Isenbaev, Wei Liu, Steven Wu, Bo Li, Ding Zhao</i>	
Benefits of Overparameterized Convolutional Residual Networks: Function Approximation Under Smoothness Constraint .....	13670
<i>Hao Liu, Minshuo Chen, Siawpeng Er, Wenjing Liao, Tong Zhang, Tuo Zhao</i>	
Boosting Graph Structure Learning with Dummy Nodes.....	13705
<i>Xin Liu, Jiayang Cheng, Yangqiu Song, Xin Jiang</i>	
Equivalence Analysis Between Counterfactual Regret Minimization and Online Mirror Descent .....	13718
<i>Weiming Liu, Huacong Jiang, Bin Li, Houqiang Li</i>	
Deep Probability Estimation.....	13747
<i>Sheng Liu, Aakash Kaku, Weicheng Zhu, Matan Leibovich, Sreyas Mohan, Boyang Yu, Haoxiang Huang, Laure Zanna, Narges Razavian, Jonathan Niles-Weed, Carlos Fernandez-Granda</i>	
Gating Dropout: Communication-Efficient Regularization for Sparsely Activated Transformers .....	13783
<i>Rui Liu, Young Jin Kim, Alexandre Muzio, Hany Hassan</i>	

Simplex Neural Population Learning: Any-Mixture Bayes-Optimality in Symmetric Zero-Sum Games.....	13794
<i>Siqi Liu, Marc Lanctot, Luke Marris, Nicolas Heess</i>	
Rethinking Attention-Model Explainability Through Faithfulness Violation Test .....	13808
<i>Yibing Liu, Haoliang Li, Yangyang Guo, Chenqi Kong, Jing Li, Shiqi Wang</i>	
Optimization-Derived Learning with Essential Convergence Analysis of Training and Hyper-Training .....	13826
<i>Risheng Liu, Xuan Liu, Shangzhi Zeng, Jin Zhang, Yixuan Zhang</i>	
Deep Neural Network Fusion Via Graph Matching with Applications to Model Ensemble and Federated Learning.....	13858
<i>Chang Liu, Chenfei Lou, Runzhong Wang, Alan Yuhan Xi, Li Shen, Junchi Yan</i>	
Welfare Maximization in Competitive Equilibrium: Reinforcement Learning for Markov Exchange Economy.....	13871
<i>Zhihan Liu, Miao Lu, Zhaoran Wang, Michael Jordan, Zhuoran Yang</i>	
Generating 3D Molecules for Target Protein Binding .....	13913
<i>Meng Liu, Youzhi Luo, Kanji Uchino, Koji Maruhashi, Shuiwang Ji</i>	
Communication-Efficient Distributed Learning for Large Batch Optimization .....	13926
<i>Rui Liu, Barzan Mozafari</i>	
Adaptive Accelerated (Extra-)Gradient Methods with Variance Reduction .....	13948
<i>Zijian Liu, Ta Duy Nguyen, Alina Ene, Huy Nguyen</i>	
REvolveR: Continuous Evolutionary Models for Robot-To-Robot Policy Transfer .....	13996
<i>Xingyu Liu, Deepak Pathak, Kris Kitani</i>	
Kill a Bird with Two Stones: Closing the Convergence Gaps in Non-Strongly Convex Optimization by Directly Accelerated SVRG with Double Compensation and Snapshots .....	14009
<i>Yuan Yuan Liu, Fanhua Shang, Weixin An, Hongying Liu, Zhouchen Lin</i>	
Learning Markov Games with Adversarial Opponents: Efficient Algorithms and Fundamental Limits .....	14037
<i>Qinghua Liu, Yuanhao Wang, Chi Jin</i>	
Local Augmentation for Graph Neural Networks.....	14055
<i>Songtao Liu, Rex Ying, Hanze Dong, Lanqing Li, Tingyang Xu, Yu Rong, Peilin Zhao, Junzhou Huang, Dinghao Wu</i>	
Asking for Knowledge (AFK): Training RL Agents to Query External Knowledge Using Language.....	14074
<i>Iou-Jen Liu, Xingdi Yuan, Marc-Alexandre Côté, Pierre-Yves Oudeyer, Alexander Schwing</i>	
Learning from Demonstration: Provably Efficient Adversarial Policy Imitation with Linear Function Approximation.....	14095
<i>Zhihan Liu, Yufeng Zhang, Zuyue Fu, Zhuoran Yang, Zhaoran Wang</i>	
GACT: Activation Compressed Training for Generic Network Architectures .....	14140
<i>Xiaoxuan Liu, Lianmin Zheng, Dequan Wang, Yukuo Cen, Weize Chen, Xu Han, Jianfei Chen, Zhiyuan Liu, Jie Tang, Joey Gonzalez, Michael Mahoney, Alvin Cheung</i>	
Robust Training Under Label Noise by Over-Parameterization.....	14154
<i>Sheng Liu, Zhihui Zhu, Qing Qu, Chong You</i>	
Plan Your Target and Learn Your Skills: Transferable State-Only Imitation Learning Via Decoupled Policy Optimization.....	14174
<i>Minghuan Liu, Zhengbang Zhu, Yuzheng Zhuang, Weinan Zhang, Jianye Hao, Yong Yu, Jun Wang</i>	
On the Impossibility of Learning to Cooperate with Adaptive Partner Strategies in Repeated Games .....	14198
<i>Robert Loftin, Frans A Oliehoek</i>	
AutoIP: A United Framework to Integrate Physics into Gaussian Processes .....	14211
<i>Da Long, Zheng Wang, Aditi Krishnapriyan, Robert Kirby, Shandian Zhe, Michael Mahoney</i>	

## PART 18

Bayesian Model Selection, the Marginal Likelihood, and Generalization .....	14224
<i>Sanae Lotfi, Pavel Izmailov, Gregory Benton, Micah Goldblum, Andrew Gordon Wilson</i>	
Feature Learning and Signal Propagation in Deep Neural Networks .....	14249
<i>Yizhang Lou, Chris E Mingard, Soufiane Hayou</i>	
Fluctuations, Bias, Variance & Ensemble of Learners: Exact Asymptotics for Convex Losses in High-Dimension .....	14284
<i>Bruno Loureiro, Cedric Gerbelot, Maria Refinetti, Gabriele Sicuro, Florent Krzakala</i>	
A Single-Loop Gradient Descent and Perturbed Ascent Algorithm for Nonconvex Functional Constrained Optimization.....	14316
<i>Songtao Lu</i>	
Additive Gaussian Processes Revisited.....	14359
<i>Xiaoyu Lu, Alexis Boukouvalas, James Hensman</i>	
ModLaNets: Learning Generalisable Dynamics Via Modularity and Physical Inductive Bias .....	14385
<i>Yupu Lu, Shijie Lin, Guanqi Chen, Jia Pan</i>	
Model-Free Opponent Shaping .....	14399
<i>Christopher Lu, Timon Willi, Christian A Schroeder De Witt, Jakob Foerster</i>	
Multi-Slots Online Matching with High Entropy .....	14413
<i>Xingyu Lu, Qintong Wu, Wenliang Zhong</i>	
Maximum Likelihood Training for Score-Based Diffusion ODEs by High Order Denoising Score Matching.....	14430
<i>Cheng Lu, Kaiwen Zheng, Fan Bao, Jianfei Chen, Chongxuan Li, Jun Zhu</i>	
Orchestra: Unsupervised Federated Learning Via Globally Consistent Clustering .....	14462
<i>Ekdeep Lubana, Chi Ian Tang, Fahim Kawsar, Robert Dick, Akhil Mathur</i>	
A Rigorous Study of Integrated Gradients Method and Extensions to Internal Neuron Attributions .....	14486
<i>Daniel D Lundstrom, Tianjian Huang, Meisam Razaviyayn</i>	
BAMDT: Bayesian Additive Semi-Multivariate Decision Trees for Nonparametric Regression.....	14510
<i>Zhao Tang Luo, Huiyan Sang, Bani Mallick</i>	
Disentangled Federated Learning for Tackling Attributes Skew Via Invariant Aggregation and Diversity Transferring .....	14528
<i>Zhengquan Luo, Yunlong Wang, Zilei Wang, Zhenan Sun, Tieniu Tan</i>	
Channel Importance Matters in Few-Shot Image Classification .....	14543
<i>Xu Luo, Jing Xu, Zenglin Xu</i>	
Learning Dynamics and Generalization in Deep Reinforcement Learning .....	14561
<i>Clare Lyle, Mark Rowland, Will Dabney, Marta Kwiatkowska, Yarin Gal</i>	
On Finite-Sample Identifiability of Contrastive Learning-Based Nonlinear Independent Component Analysis.....	14583
<i>Qi Lyu, Xiao Fu</i>	
Pessimism Meets VCG: Learning Dynamic Mechanism Design Via Offline Reinforcement Learning .....	14602
<i>Boxiang Lyu, Zhaoran Wang, Mladen Kolar, Zhuoran Yang</i>	
Versatile Offline Imitation from Observations and Examples Via Regularized State-Occupancy Matching.....	14640
<i>Yecheng Ma, Andrew Shen, Dinesh Jayaraman, Osbert Bastani</i>	
Quantification and Analysis of Layer-Wise and Pixel-Wise Information Discarding .....	14665
<i>Haotian Ma, Hao Zhang, Fan Zhou, Yinqing Zhang, Quanshi Zhang</i>	
Interpretable Neural Networks with Frank-Wolfe: Sparse Relevance Maps and Relevance Orderings.....	14700
<i>Jan Macdonald, Mathieu E. Besançon, Sebastian Pokutta</i>	



A Tighter Analysis of Spectral Clustering, and Beyond .....	14718
<i>Peter Macgregor, He Sun</i>	
Zero-Shot Reward Specification Via Grounded Natural Language.....	14744
<i>Parsa Mahmoudieh, Deepak Pathak, Trevor Darrell</i>	
Feature Selection Using E-Values .....	14754
<i>Subhabrata Majumdar, Snigdhasu Chatterjee</i>	
Knowledge-Grounded Self-Rationalization Via Extractive and Natural Language Explanations.....	14775
<i>Bodhisattwa Prasad Majumder, Oana Camburu, Thomas Lukasiewicz, Julian McAuley</i>	
Nonparametric Involutive Markov Chain Monte Carlo.....	14791
<i>Carol Mak, Fabian Zaiser, Luke Ong</i>	
Architecture Agnostic Federated Learning for Neural Networks .....	14849
<i>Disha Makhija, Xing Han, Nhat Ho, Joydeep Ghosh</i>	
Robustness in Multi-Objective Submodular Optimization: A Quantile Approach .....	14860
<i>Cedric Malherbe, Kevin Scaman</i>	
More Efficient Sampling for Tensor Decomposition with Worst-Case Guarantees .....	14876
<i>Osman Asif Malik</i>	
Unaligned Supervision for Automatic Music Transcription in the Wild .....	14907
<i>Ben Maman, Amit H Bermano</i>	
Decision-Focused Learning: Through the Lens of Learning to Rank.....	14924
<i>Jayanta Mandi, Vi'ctor Bucarey, Maxime Mulamba Ke Tchomba, Tias Guns</i>	
Differentially Private Coordinate Descent for Composite Empirical Risk Minimization .....	14937
<i>Paul Mangold, Aurélien Bellet, Joseph Salmon, Marc Tommasi</i>	
Refined Convergence Rates for Maximum Likelihood Estimation Under Finite Mixture Models .....	14968
<i>Tudor Manole, Nhat Ho</i>	
On Improving Model-Free Algorithms for Decentralized Multi-Agent Reinforcement Learning .....	14996
<i>Weichao Mao, Lin Yang, Kaiqing Zhang, Tamer Basar</i>	
On the Effects of Artificial Data Modification .....	15039
<i>Antonia Marcu, Adam Prugel-Bennett</i>	
Personalized Federated Learning Through Local Memorization .....	15059
<i>Othmane Marfoq, Giovanni Neglia, Richard Vidal, Laetitia Kamani</i>	

## PART 19

Nested Bandits.....	15082
<i>Matthieu Martin, Panayotis Mertikopoulos, Thibaud Rahier, Houssam Zenati</i>	
Closed-Form Diffeomorphic Transformations for Time Series Alignment .....	15111
<i>Iñigo Martinez, Elisabeth Viles, Igor G. Olaizola</i>	
SPECTRE: Spectral Conditioning Helps to Overcome the Expressivity Limits of One-Shot Graph Generators .....	15148
<i>Karolis Martinkus, Andreas Loukas, Nathanaël Perraudin, Roger Wattenhofer</i>	
Modular Conformal Calibration .....	15169
<i>Charles Marx, Shengjia Zhao, Willie Neiswanger, Stefano Ermon</i>	
Continual Repeated Annealed Flow Transport Monte Carlo .....	15185
<i>Alex Matthews, Michael Arbel, Danilo Jimenez Rezende, Arnaud Doucet</i>	
How to Stay Curious While Avoiding Noisy TVs Using Aleatoric Uncertainty Estimation .....	15209
<i>Augustine Mavor-Parker, Kimberly Young, Caswell Barry, Lewis Griffin</i>	
How to Steer Your Adversary: Targeted and Efficient Model Stealing Defenses with Gradient Redirection .....	15230
<i>Mantas Mazeika, Bo Li, David Forsyth</i>	
Quant-BnB: A Scalable Branch-And-Bound Method for Optimal Decision Trees with Continuous Features .....	15244
<i>Rahul Mazumder, Xiang Meng, Haoyue Wang</i>	

Optimizing Tensor Network Contraction Using Reinforcement Learning .....	15267
<i>Eli Meir, Haggai Maron, Shie Mannor, Gal Chechik</i>	
Causal Transformer for Estimating Counterfactual Outcomes .....	15282
<i>Valentyn Melnychuk, Dennis Frauen, Stefan Feuerriegel</i>	
Steerable 3D Spherical Neurons .....	15319
<i>Pavlo Melnyk, Michael Felsberg, Mårten Wadenbäck</i>	
Transformers Are Meta-Reinforcement Learners .....	15329
<i>Luckeciano C Melo</i>	
ButterflyFlow: Building Invertible Layers with Butterfly Matrices .....	15349
<i>Chenlin Meng, Linqi Zhou, Kristy Choi, Tri Dao, Stefano Ermon</i>	
In Defense of Dual-Encoders for Neural Ranking .....	15365
<i>Aditya Menon, Sadeep Jayasumana, Ankit Singh Rawat, Seungyeon Kim, Sashank Reddi, Sanjiv Kumar</i>	
Equivariant Quantum Graph Circuits .....	15390
<i>Peter Mernyei, Konstantinos Meichanetzidis, Ismail Ilkan Ceylan</i>	
Stochastic Rising Bandits .....	15410
<i>Alberto Maria Metelli, Francesco Trovò, Matteo Pirola, Marcello Restelli</i>	
Minimizing Control for Credit Assignment with Strong Feedback .....	15447
<i>Alexander Meulemans, Matilde Tristany Farinha, Maria R. Cervera, João Sacramento, Benjamin F. Grewe</i>	
A Dynamical System Perspective for Lipschitz Neural Networks .....	15473
<i>Laurent Meunier, Blaise J Delattre, Alexandre Araujo, Alexandre Allauzen</i>	
Distribution Regression with Sliced Wasserstein Kernels .....	15490
<i>Dimitri Meunier, Massimiliano Pontil, Carlo Ciliberto</i>	
Interpretable and Generalizable Graph Learning Via Stochastic Attention Mechanism .....	15513
<i>Siqi Miao, Mia Liu, Pan Li</i>	
Modeling Structure with Undirected Neural Networks .....	15533
<i>Tsvetomila Mihaylova, Vlad Niculae, Andre Martins</i>	
Universal Hopfield Networks: A General Framework for Single-Shot Associative Memory Models .....	15550
<i>Beren Millidge, Tommaso Salvatori, Yuhang Song, Thomas Lukasiewicz, Rafal Bogacz</i>	
Learning Stochastic Shortest Path with Linear Function Approximation .....	15573
<i>Yifei Min, Jiafan He, Tianhao Wang, Quanquan Gu</i>	
Prioritized Training on Points that Are Learnable, Worth Learning, and Not Yet Learnt .....	15619
<i>Sören Mindermann, Jan M Brauner, Muhammed T Razzak, Mrinank Sharma, Andreas Kirsch, Winnie Xu, Benedikt Hölting, Aidan N Gomez, Adrien Morisot, Sebastian Farquhar, Yarin Gal</i>	
POEM: Out-Of-Distribution Detection with Posterior Sampling .....	15639
<i>Yifei Ming, Ying Fan, Yixuan Li</i>	
A Simple Reward-Free Approach to Constrained Reinforcement Learning .....	15655
<i>Sobhan Miryoosefi, Chi Jin</i>	
Wide Neural Networks Forget Less Catastrophically .....	15688
<i>Seyed Iman Mirzadeh, Arslan Chaudhry, Dong Yin, Huiyi Hu, Razvan Pascanu, Dilan Gorur, Mehrdad Farajtabar</i>	
Proximal and Federated Random Reshuffling .....	15707
<i>Konstantin Mishchenko, Ahmed Khaled, Peter Richtarik</i>	
ProxSkip: Yes! Local Gradient Steps Provably Lead to Communication Acceleration! Finally! .....	15739
<i>Konstantin Mishchenko, Grigory Malinovsky, Sebastian Stich, Peter Richtarik</i>	
Fast Convex Optimization for Two-Layer ReLU Networks: Equivalent Model Classes and Cone Decompositions .....	15759
<i>Aaron Mishkin, Arda Sahiner, Mert Pilanci</i>	
Memory-Based Model Editing at Scale .....	15806
<i>Eric Mitchell, Charles Lin, Antoine Bosselut, Christopher D Manning, Chelsea Finn</i>	

Invariant Ancestry Search.....	15821
<i>Phillip B Mogensen, Nikolaj Thams, Jonas Peters</i>	
Differentially Private Community Detection for Stochastic Block Models .....	15847
<i>Mohamed S Mohamed, Dung Nguyen, Anil Vullikanti, Ravi Tandon</i>	
A Multi-Objective / Multi-Task Learning Framework Induced by Pareto Stationarity .....	15884
<i>Michinari Momma, Chaosheng Dong, Jia Liu</i>	
EqR: Equivariant Representations for Data-Efficient Reinforcement Learning.....	15897
<i>Arnab Kumar Mondal, Vineet Jain, Kaleem Siddiqi, Siamak Ravanbakhsh</i>	

## PART 20

Feature and Parameter Selection in Stochastic Linear Bandits.....	15916
<i>Ahmadreza Moradipari, Berkay Turan, Yasin Abbasi-Yadkori, Mahnoosh Alizadeh, Mohammad Ghavamzadeh</i>	
Power-Law Escape Rate of SGD.....	15948
<i>Takashi Mori, Liu Ziyin, Kangqiao Liu, Masahito Ueda</i>	
Rethinking Fano’s Inequality in Ensemble Learning .....	15965
<i>Terufumi Morishita, Gaku Morio, Shota Horiguchi, Hiroaki Ozaki, Nobuo Nukaga</i>	
SpeqNets: Sparsity-Aware Permutation-Equivariant Graph Networks .....	16006
<i>Christopher Morris, Gaurav Rattan, Sandra Kiefer, Siamak Ravanbakhsh</i>	
CtrlFormer: Learning Transferable State Representation for Visual Control Via Transformer .....	16032
<i>Yao Mark Mu, Shoufa Chen, Mingyu Ding, Jianyu Chen, Runjian Chen, Ping Luo</i>	
Generalized Beliefs for Cooperative AI .....	16051
<i>Darius Muglich, Luisa M Zintgraf, Christian A Schroeder De Witt, Shimon Whiteson, Jakob Foerster</i>	
Bounding the Width of Neural Networks Via Coupled Initialization a Worst Case Analysis.....	16072
<i>Alexander Munteanu, Simon Omlor, Zhao Song, David Woodruff</i>	
Constants Matter: The Performance Gains of Active Learning.....	16112
<i>Stephen O Mussmann, Sanjoy Dasgupta</i>	
On the Generalization Analysis of Adversarial Learning .....	16163
<i>Waleed Mustafa, Yunwen Lei, Marius Kloft</i>	
Universal and Data-Adaptive Algorithms for Model Selection in Linear Contextual Bandits.....	16186
<i>Vidya K Muthukumar, Akshay Krishnamurthy</i>	
The Importance of Non-Markovianity in Maximum State Entropy Exploration .....	16212
<i>Mirco Mutti, Riccardo De Santi, Marcello Restelli</i>	
PAC-Net: A Model Pruning Approach to Inductive Transfer Learning.....	16229
<i>Sanghoon Myung, In Huh, Wonik Jang, Jae Myung Choe, Jisu Ryu, Daesin Kim, Kee-Eung Kim, Changwook Jeong</i>	
AutoSNN: Towards Energy-Efficient Spiking Neural Networks .....	16242
<i>Byunggook Na, Jisoo Mok, Seongsik Park, Dongjin Lee, Hyeokjun Choe, Sungroh Yoon</i>	
Implicit Bias of the Step Size in Linear Diagonal Neural Networks .....	16259
<i>Mor Shpigel Nacson, Kavya Ravichandran, Nathan Srebro, Daniel Soudry</i>	
DNNR: Differential Nearest Neighbors Regression.....	16285
<i>Youssef Nader, Leon Sixt, Tim Landgraf</i>	
Overcoming Oscillations in Quantization-Aware Training .....	16307
<i>Markus Nagel, Marios Fournarakis, Yelysei Bondarenko, Tijmen Blankevoort</i>	
Strategic Representation .....	16320
<i>Vineet Nair, Ganesh Ghalme, Inbal Talgam-Cohen, Nir Rosenfeld</i>	
Improving Ensemble Distillation with Weight Averaging and Diversifying Perturbation.....	16342
<i>Giung Nam, Hyungi Lee, Byeongho Heo, Juho Lee</i>	

Measuring Representational Robustness of Neural Networks Through Shared Invariances.....	16357
<i>Vedant Nanda, Till Speicher, Camila Kolling, John P Dickerson, Krishna Gummadi, Adrian Weller</i>	
Tight and Robust Private Mean Estimation with Few Users .....	16372
<i>Shyam Narayanan, Vahab Mirrokni, Hossein Esfandiari</i>	
Fast Aquatic Swimmer Optimization with Differentiable Projective Dynamics and Neural Network Hydrodynamic Models .....	16402
<i>Elvis Nava, John Z Zhang, Mike Yan Michelis, Tao Du, Pingchuan Ma, Benjamin F. Grewe, Wojciech Matusik, Robert Kevin Katzschmann</i>	
Multi-Task Learning as a Bargaining Game .....	16417
<i>Aviv Navon, Aviv Shamsian, Idan Achituve, Haggai Maron, Kenji Kawaguchi, Gal Chechik, Ethan Fetaya</i>	
Variational Inference for Infinitely Deep Neural Networks .....	16436
<i>Achille Nazaret, David Blei</i>	
Stable Conformal Prediction Sets.....	16451
<i>Eugene Ndiaye</i>	
Discovering Generalizable Spatial Goal Representations Via Graph-Based Active Reward Learning .....	16469
<i>Aviv Netanyahu, Tianmin Shu, Joshua Tenenbaum, Pulkit Agrawal</i>	
Sublinear-Time Clustering Oracle for Signed Graphs .....	16485
<i>Stefan Neumann, Pan Peng</i>	
Improved Regret for Differentially Private Exploration in Linear MDP.....	16518
<i>Dung Daniel T Ngo, Giuseppe Vietri, Steven Wu</i>	
A Framework for Learning to Request Rich and Contextually Useful Information from Humans .....	16542
<i>Khanh X Nguyen, Yonatan Bisk, Hal Daumé III</i>	
Transformer Neural Processes: Uncertainty-Aware Meta Learning Via Sequence Modeling .....	16558
<i>Tung Nguyen, Aditya Grover</i>	
Improving Transformers with Probabilistic Attention Keys .....	16584
<i>Tam Minh Nguyen, Tan Minh Nguyen, Dung D. D. Le, Duy Khuong Nguyen, Viet-Anh Tran, Richard Baraniuk, Nhat Ho, Stanley Osher</i>	
On Transportation of Mini-Batches: A Hierarchical Approach .....	16611
<i>Khai Nguyen, Dang Nguyen, Quoc Dinh Nguyen, Tung Pham, Hung Bui, Dinh Phung, Trung Le, Nhat Ho</i>	
Improving Mini-Batch Optimal Transport Via Partial Transportation.....	16645
<i>Khai Nguyen, Dang Nguyen, The-Anh Vu-Le, Tung Pham, Nhat Ho</i>	
Recurrent Model-Free RL Can Be a Strong Baseline for Many POMDPs.....	16680
<i>Tianwei Ni, Benjamin Eysenbach, Ruslan Salakhutdinov</i>	
Optimal Estimation of Policy Gradient Via Double Fitted Iteration.....	16713
<i>Chengzhuo Ni, Ruiqi Zhang, Xiang Ji, Xuezhou Zhang, Mengdi Wang</i>	

## PART 21

GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models.....	16773
<i>Alexander Quinn Nichol, Prafulla Dhariwal, Aditya Ramesh, Pranav Shyam, Pamela Mishkin, Bob McGrew, Ilya Sutskever, Mark Chen</i>	
Diffusion Models for Adversarial Purification .....	16794
<i>Weili Nie, Brandon Guo, Yujia Huang, Chaowei Xiao, Arash Vahdat, Animashree Anandkumar</i>	
The Primacy Bias in Deep Reinforcement Learning .....	16817
<i>Evgenii Nikishin, Max Schwarzer, Pierluca D'Oro, Pierre-Luc Bacon, Aaron Courville</i>	
Causal Conceptions of Fairness and Their Consequences.....	16837
<i>Hamed Nilforoshan, Johann D Gaebler, Ravi Shroff, Sharad Goel</i>	

Efficient Test-Time Model Adaptation Without Forgetting .....	16877
<i>Shuaicheng Niu, Jiaxiang Wu, Yifan Zhang, Yaofu Chen, Shijian Zheng, Peilin Zhao, Mingkui Tan</i>	
Generative Trees: Adversarial and Copycat .....	16895
<i>Richard Nock, Mathieu Guillame-Bert</i>	
Path-Aware and Structure-Preserving Generation of Synthetically Accessible Molecules .....	16941
<i>Juhwan Noh, Dae-Woong Jeong, Kiyoun Kim, Sehui Han, Moontae Lee, Honglak Lee, Yousung Jung</i>	
Utilizing Expert Features for Contrastive Learning of Time-Series Representations.....	16958
<i>Manuel T Nonnenmacher, Lukas Oldenburg, Ingo Steinwart, David Reeb</i>	
Tranception: Protein Fitness Prediction with Autoregressive Transformers and Inference-Time Retrieval .....	16979
<i>Pascal Notin, Mafalda Dias, Jonathan Frazer, Javier Marchena Hurtado, Aidan N Gomez, Debora Marks, Yarin Gal</i>	
Fast Finite Width Neural Tangent Kernel .....	17007
<i>Roman Novak, Jascha Sohl-Dickstein, Samuel S Schoenholz</i>	
Multicoated Supermasks Enhance Hidden Networks.....	17034
<i>Yasuyuki Okoshi, Ángel López Garcí'a-Arias, Kazutoshi Hirose, Kota Ando, Kazushi Kawamura, Thiem Van Chu, Masato Motomura, Jaehoon Yu</i>	
Generalized Leverage Scores: Geometric Interpretation and Applications .....	17045
<i>Bruno Ordozgoiti, Antonis Matakos, Aristides Gionis</i>	
Practical Almost-Linear-Time Approximation Algorithms for Hybrid and Overlapping Graph Clustering .....	17060
<i>Lorenzo Orecchia, Konstantinos Ameranis, Charalampos Tsourakakis, Kunal Talwar</i>	
Anticorrelated Noise Injection for Improved Generalization .....	17083
<i>Antonio Orvieto, Hans Kersting, Frank Proske, Francis Bach, Aurelien Lucchi</i>	
Scalable Deep Gaussian Markov Random Fields for General Graphs .....	17106
<i>Joel Oskarsson, Per Sidén, Fredrik Lindsten</i>	
Zero-Shot AutoML with Pretrained Models .....	17127
<i>Ekrem Öztürk, Fabio Ferreira, Hadi Jomaa, Lars Schmidt-Thieme, Josif Grabocka, Frank Hutter</i>	
History Compression Via Language Models in Reinforcement Learning .....	17145
<i>Fabian Paischer, Thomas Adler, Vihang Patil, Angela Bitto-Nemling, Markus Holzleitner, Sebastian Lehner, Hamid Eghbal-Zadeh, Sepp Hochreiter</i>	
A Study on the Ramanujan Graph Property of Winning Lottery Tickets .....	17175
<i>Bithika Pal, Arindam Biswas, Sudeshna Kolay, Pabitra Mitra, Biswajit Basu</i>	
On Learning Mixture of Linear Regressions in the Non-Realizable Setting.....	17191
<i>Soumyabrata Pal, Arya Mazumdar, Rajat Sen, Avishek Ghosh</i>	
Plan Better Amid Conservatism: Offline Multi-Agent Reinforcement Learning with Actor Rectification .....	17210
<i>Ling Pan, Longbo Huang, Tengyu Ma, Huazhe Xu</i>	
A Unified Weight Initialization Paradigm for Tensorial Convolutional Neural Networks .....	17227
<i>Yu Pan, Zeyong Su, Ao Liu, Wang Jingquan, Nannan Li, Zenglin Xu</i>	
Robustness and Accuracy Could Be Reconcilable by (Proper) Definition.....	17247
<i>Tianyu Pang, Min Lin, Xiao Yang, Jun Zhu, Shuicheng Yan</i>	
Towards Coherent and Consistent Use of Entities in Narrative Generation .....	17267
<i>Pinelopi Papalampidi, Kris Cao, Tomas Kocisky</i>	
Constrained Discrete Black-Box Optimization Using Mixed-Integer Programming .....	17284
<i>Theodore P Papalexopoulos, Christian Tjandraatmadja, Ross Anderson, Juan Pablo Vielma, David Belanger</i>	
A Theoretical Comparison of Graph Neural Network Extensions .....	17312
<i>Pál András Papp, Roger Wattenhofer</i>	

Validating Causal Inference Methods .....	17335
<i>Harsh Parikh, Carlos Varjao, Louise Xu, Eric Tchetgen Tchetgen</i>	
The Unsurprising Effectiveness of Pre-Trained Vision Models for Control.....	17348
<i>Simone Parisi, Aravind Rajeswaran, Senthil Purushwalkam, Abhinav Gupta</i>	
Learning Symmetric Embeddings for Equivariant World Models .....	17361
<i>Jung Yeon Park, Ondrej Biza, Linfeng Zhao, Jan-Willem Van De Meent, Robin Walters</i>	
Blurs Behave Like Ensembles: Spatial Smoothings to Improve Accuracy, Uncertainty, and Robustness.....	17379
<i>Namuk Park, Songkuk Kim</i>	
Exact Optimal Accelerated Complexity for Fixed-Point Iterations.....	17409
<i>Jisun Park, Ernest K Ryu</i>	
Kernel Methods for Radial Transformed Compositional Data with Many Zeros.....	17447
<i>Junyoung Park, Changwon Yoon, Cheolwoo Park, Jeongyoun Ahn</i>	
Evolving Curricula with Regret-Based Environment Design.....	17462
<i>Jack Parker-Holder, Minqi Jiang, Michael Dennis, Mikayel Samvelyan, Jakob Foerster, Edward Grefenstette, Tim Rocktäschel</i>	
Neural Language Models Are Not Born Equal to Fit Brain Data, but Training Helps .....	17488
<i>Alexandre Pasquiou, Yair Lakretz, John T Hale, Bertrand Thirion, Christophe Pallier</i>	
A New Similarity Measure for Covariate Shift with Applications to Nonparametric Regression.....	17506
<i>Reese Pathak, Cong Ma, Martin Wainwright</i>	
Align-RUDDER: Learning from Few Demonstrations by Reward Redistribution .....	17520
<i>Vihang Patil, Markus Hofmarcher, Marius-Constantin Dinu, Matthias Dorfer, Patrick M Blies, Johannes Brandstetter, José Arjona-Medina, Sepp Hochreiter</i>	
POET: Training Neural Networks on Tiny Devices with Integrated Rematerialization and Paging .....	17562
<i>Shishir G. Patil, Paras Jain, Prabal Dutta, Ion Stoica, Joseph Gonzalez</i>	
Learning to Cut by Looking Ahead: Cutting Plane Selection Via Imitation Learning .....	17573
<i>Max B Paulus, Giulia Zarpellon, Andreas Krause, Laurent Charlin, Chris Maddison</i>	
Neural Network Pruning Denoises the Features and Makes Local Connectivity Emerge in Visual Tasks.....	17590
<i>Franco Pellegrini, Giulio Biroli</i>	

## PART 22

Branchformer: Parallel MLP-Attention Architectures to Capture Local and Global Context for Speech Recognition and Understanding.....	17616
<i>Yifan Peng, Siddharth Dalmia, Ian Lane, Shinji Watanabe</i>	
Pocket2Mol: Efficient Molecular Sampling Based on 3D Protein Pockets.....	17633
<i>Xingang Peng, Shitong Luo, Jiaqi Guan, Qi Xie, Jian Peng, Jianzhu Ma</i>	
Differentiable Top-K Classification Learning .....	17645
<i>Felix Petersen, Hilde Kuehne, Christian Borgelt, Oliver Deussen</i>	
Multi-Scale Feature Learning Dynamics: Insights for Double Descent .....	17658
<i>Mohammad Pezeshki, Amartya Mitra, Yoshua Bengio, Guillaume Lajoie</i>	
A Differential Entropy Estimator for Training Neural Networks .....	17680
<i>Georg Pichler, Pierre Jean A. Colombo, Malik Boudiaf, Günther Koliander, Pablo Piantanida</i>	
Federated Learning with Partial Model Personalization.....	17705
<i>Krishna Pillutla, Kshitiz Malik, Abdel-Rahman Mohamed, Mike Rabbat, Maziar Sanjabi, Lin Xiao</i>	
Deep Networks on Toroids: Removing Symmetries Reveals the Structure of Flat Regions in the Landscape Geometry .....	17748
<i>Fabrizio Pittorino, Antonio Ferraro, Gabriele Perugini, Christoph Feinauer, Carlo Baldassi, Riccardo Zecchina</i>	

Geometric Multimodal Contrastive Representation Learning .....	17771
<i>Petra Poklukar, Miguel Vasco, Hang Yin, Francisco S. Melo, Ana Paiva, Danica Kragic</i>	
Constrained Offline Policy Optimization .....	17790
<i>Nicholas Polosky, Bruno C. Da Silva, Madalina Fiterau, Jithin Jagannath</i>	
Offline Meta-Reinforcement Learning with Online Self-Supervision.....	17800
<i>Vitchyr H Pong, Ashvin V Nair, Laura M Smith, Catherine Huang, Sergey Levine</i>	
Debiasser Beware: Pitfalls of Centering Regularized Transport Maps .....	17819
<i>Aram-Alexandre Pooladian, Marco Cuturi, Jonathan Niles-Weed</i>	
Adaptive Second Order Coresets for Data-Efficient Machine Learning .....	17837
<i>Omead Pooladzandi, David Davini, Baharan Mirzasoleiman</i>	
On the Practicality of Deterministic Epistemic Uncertainty.....	17859
<i>Janis Postels, Mattia Segù, Tao Sun, Luca Daniel Sieber, Luc Van Gool, Fisher Yu, Federico Tombari</i>	
A Simple Guard for Learned Optimizers.....	17899
<i>Isabeau Prémont-Schwarz, Jaroslav Vít, Jan Feyereisl</i>	
Hardness and Algorithms for Robust and Sparse Optimization .....	17915
<i>Eric Price, Sandeep Silwal, Samson Zhou</i>	
Nonlinear Feature Diffusion on Hypergraphs.....	17934
<i>Konstantin Prokopych, Austin R Benson, Francesco Tudisco</i>	
Universal Joint Approximation of Manifolds and Densities by Simple Injective Flows.....	17948
<i>Michael Puthawala, Matti Lassas, Ivan Dokmanic, Maarten De Hoop</i>	
The Teaching Dimension of Regularized Kernel Learners.....	17973
<i>Hong Qian, Xu-Hui Liu, Chen-Xi Su, Aimin Zhou, Yang Yu</i>	
ContentVec: An Improved Self-Supervised Speech Representation by Disentangling Speakers .....	17992
<i>Kaizhi Qian, Yang Zhang, Heting Gao, Junrui Ni, Cheng-I Lai, David Cox, Mark Hasegawa-Johnson, Shiyu Chang</i>	
Interventional Contrastive Learning with Meta Semantic Regularizer .....	18007
<i>Wenwen Qiang, Jiangmeng Li, Changwen Zheng, Bing Su, Hui Xiong</i>	
Sample-Efficient Reinforcement Learning with $\log\log(T)$ Switching Cost .....	18020
<i>Dan Qiao, Ming Yin, Ming Min, Yu-Xiang Wang</i>	
Generalizing to Evolving Domains with Latent Structure-Aware Sequential Autoencoder .....	18051
<i>Tiexin Qin, Shiqi Wang, Haoliang Li</i>	
Graph Neural Architecture Search Under Distribution Shifts .....	18072
<i>Yijian Qin, Xin Wang, Ziwei Zhang, Pengtao Xie, Wenwu Zhu</i>	
Spectral Representation of Robustness Measures for Optimization Under Input Uncertainty .....	18085
<i>Jixiang Qing, Tom Dhaene, Ivo Couckuyt</i>	
Large-Scale Stochastic Optimization of NDCG Surrogates for Deep Learning with Provable Convergence.....	18111
<i>Zi-Hao Qiu, Quanqi Hu, Yongjian Zhong, Lijun Zhang, Tianbao Yang</i>	
Latent Outlier Exposure for Anomaly Detection with Contaminated Data .....	18142
<i>Chen Qiu, Aodong Li, Marius Kloft, Maja Rudolph, Stephan Mandt</i>	
Contrastive UCB: Provably Efficient Contrastive Self-Supervised Learning in Online Reinforcement Learning.....	18157
<i>Shuang Qiu, Lingxiao Wang, Chenjia Bai, Zhuoran Yang, Zhaoran Wang</i>	
Fast and Provable Nonconvex Tensor RPCA .....	18200
<i>Haiquan Qiu, Yao Wang, Shaojie Tang, Deyu Meng, Quanming Yao</i>	
Generalized Federated Learning Via Sharpness Aware Minimization.....	18239
<i>Zhe Qu, Xingyu Li, Rui Duan, Yao Liu, Bo Tang, Zhuo Lu</i>	
Particle Transformer for Jet Tagging .....	18270
<i>Huilin Qu, Congqiao Li, Sitian Qian</i>	

Winning the Lottery Ahead of Time: Efficient Early Network Pruning .....	18282
<i>John Rachwan, Daniel Zügner, Bertrand Charpentier, Simon Geisler, Morgane Ayle, Stephan Günnemann</i>	
Convergence of Uncertainty Sampling for Active Learning .....	18299
<i>Anant Raj, Francis Bach</i>	
DeepSpeed-MoE: Advancing Mixture-Of-Experts Inference and Training to Power Next- Generation AI Scale.....	18321
<i>Samyam Rajbhandari, Conglong Li, Zhewei Yao, Minjia Zhang, Reza Yazdani Aminabadi, Ammar Ahmad Awan, Jeff Rasley, Yuxiong He</i>	
Fishr: Invariant Gradient Variances for Out-Of-Distribution Generalization .....	18336
<i>Alexandre Rame, Corentin Dancette, Matthieu Cord</i>	
A Closer Look at Smoothness in Domain Adversarial Training.....	18367
<i>Harsh Rangwani, Sumukh K Aithal, Mayank Mishra, Arihant Jain, Venkatesh Babu Radhakrishnan</i>	
Linear Adversarial Concept Erasure.....	18389
<i>Shauli Ravfogel, Michael Twiton, Yoav Goldberg, Ryan D Cotterell</i>	
Implicit Regularization in Hierarchical Tensor Factorization and Deep Convolutional Neural Networks .....	18411
<i>Noam Razin, Asaf Maman, Nadav Cohen</i>	

## PART 23

One-Pass Algorithms for MAP Inference of Nonsymmetric Determinantal Point Processes.....	18452
<i>Aravind Reddy, Ryan A. Rossi, Zhao Song, Anup Rao, Tung Mai, Nedim Lipka, Gang Wu, Eunyee Koh, Nesreen Ahmed</i>	
Universality of Winning Tickets: A Renormalization Group Perspective .....	18472
<i>William T Redman, Tianlong Chen, Zhangyang Wang, Akshunna S. Dogra</i>	
The Dynamics of Representation Learning in Shallow, Non-Linear Autoencoders .....	18488
<i>Maria Refinetti, Sebastian Goldt</i>	
Proximal Exploration for Model-Guided Protein Sequence Design.....	18509
<i>Zhizhou Ren, Jiahan Li, Fan Ding, Yuan Zhou, Jianzhu Ma, Jian Peng</i>	
Towards Theoretical Analysis of Transformation Complexity of ReLU DNNs.....	18526
<i>Jie Ren, Mingjie Li, Meng Zhou, Shih-Han Chan, Quanshi Zhang</i>	
Benchmarking and Analyzing Point Cloud Classification Under Corruptions.....	18548
<i>Jiawei Ren, Liang Pan, Ziwei Liu</i>	
A Unified View on PAC-Bayes Bounds for Meta-Learning.....	18565
<i>Arezou Rezazadeh</i>	
3PC: Three Point Compressors for Communication-Efficient Distributed Training and a Better Theory for Lazy Aggregation .....	18585
<i>Peter Richtarik, Igor Sokolov, Elnur Gasanov, Ilyas Fatkhullin, Zhize Li, Eduard Gorbunov</i>	
Robust SDE-Based Variational Formulations for Solving Linear PDEs Via Deep Learning .....	18638
<i>Lorenz Richter, Julius Berner</i>	
Probabilistically Robust Learning: Balancing Average and Worst-Case Performance.....	18656
<i>Alexander Robey, Luiz Chamon, George J. Pappas, Hamed Hassani</i>	
LyaNet: A Lyapunov Framework for Training Neural ODEs.....	18676
<i>Ivan Dario Jimenez Rodriguez, Aaron Ames, Yisong Yue</i>	
Short-Term Plasticity Neurons Learning to Learn and Forget.....	18693
<i>Hector Garcia Rodriguez, Qinghai Guo, Timoleon Moraitis</i>	
Function-Space Inference with Sparse Implicit Processes .....	18712
<i>Simon Rodri'guez-Santana, Bryan Zaldivar, Daniel Hernandez-Lobato</i>	



Score Matching Enables Causal Discovery of Nonlinear Additive Noise Models .....	18730
<i>Paul Rolland, Volkan Cevher, Matthäus Kleindessner, Chris Russell, Dominik Janzing, Bernhard Schölkopf, Francesco Locatello</i>	
Dual Decomposition of Convex Optimization Layers for Consistent Attention in Medical Images.....	18743
<i>Tom Ron, Tamir Hazan</i>	
A Consistent and Efficient Evaluation Strategy for Attribution Methods.....	18759
<i>Yao Rong, Tobias Leemann, Vadim Borisov, Gjergji Kasneci, Enkelejda Kasneci</i>	
Efficiently Learning the Topology and Behavior of a Networked Dynamical System Via Active Queries .....	18785
<i>Daniel J Rosenkrantz, Abhijin Adiga, Madhav Marathe, Zirou Qiu, S S Ravi, Richard Stearns, Anil Vullikanti</i>	
Learning to Infer Structures of Network Games .....	18798
<i>Emanuele Rossi, Federico Monti, Yan Leng, Michael Bronstein, Xiaowen Dong</i>	
Direct Behavior Specification Via Constrained Reinforcement Learning .....	18817
<i>Julien Roy, Roger Girgis, Joshua Romoff, Pierre-Luc Bacon, Chris J Pal</i>	
Constraint-Based Graph Network Simulator.....	18833
<i>Yulia Rubanova, Alvaro Sanchez-Gonzalez, Tobias Pfaff, Peter Battaglia</i>	
Continual Learning Via Sequential Function-Space Variational Inference .....	18860
<i>Tim G. J. Rudner, Freddie Bickford Smith, Qixuan Feng, Yee Whye Teh, Yarin Gal</i>	
Graph-Coupled Oscillator Networks .....	18877
<i>T. Konstantin Rusch, Ben Chamberlain, James Rowbottom, Siddhartha Mishra, Michael Bronstein</i>	
Hindering Adversarial Attacks with Implicit Neural Representations.....	18899
<i>Andrei A Rusu, Dan Andrei Calian, Sven Gowal, Raia Hadsell</i>	
Exploiting Independent Instruments: Identification and Distribution Generalization .....	18924
<i>Sorawit Saengkyongam, Leonard Henckel, Niklas Pfister, Jonas Peters</i>	
FedNL: Making Newton-Type Methods Applicable to Federated Learning .....	18948
<i>Mher Safaryan, Rustem Islamov, Xun Qian, Peter Richtarik</i>	
Versatile Dueling Bandits: Best-Of-Both World Analyses for Learning from Relative Preferences.....	19000
<i>Aadirupa Saha, Pierre Gaillard</i>	
Optimal and Efficient Dynamic Regret Algorithms for Non-Stationary Dueling Bandits .....	19016
<i>Aadirupa Saha, Shubham Gupta</i>	
Unraveling Attention Via Convex Duality: Analysis and Interpretations of Vision Transformers .....	19039
<i>Arda Sahiner, Tolga Ergen, Batu Ozturkler, John Pauly, Morteza Mardani, Mert Pilanci</i>	
Off-Policy Evaluation for Large Action Spaces Via Embeddings .....	19078
<i>Yuta Saito, Thorsten Joachims</i>	
Optimal Clipping and Magnitude-Aware Differentiation for Improved Quantization-Aware Training .....	19112
<i>Charbel Sakr, Steve Dai, Rangha Venkatesan, Brian Zimmer, William Dally, Brucek Khailany</i>	
A Convergence Theory for SVGD in the Population Limit Under Talagrand’s Inequality T1 .....	19128
<i>Adil Salim, Lukang Sun, Peter Richtarik</i>	
FITNESS: (Fine Tune on New and Similar Samples) to Detect Anomalies in Streams with Drift and Outliers .....	19142
<i>Abishek Sankararaman, Balakrishnan Narayanaswamy, Vikramank Y Singh, Zhao Song</i>	
The Algebraic Path Problem for Graph Metrics .....	19167
<i>Enrique Fita Sanmarti’n, Sebastian Damrich, Fred Hamprecht</i>	
LSB: Local Self-Balancing MCMC in Discrete Spaces.....	19194
<i>Emanuele Sansone</i>	
PoF: Post-Training of Feature Extractor for Improving Generalization.....	19210
<i>Ikuro Sato, Yamada Ryota, Masayuki Tanaka, Nakamasa Inoue, Rei Kawakami</i>	
Re-Evaluating Word Mover’s Distance .....	19220
<i>Ryoma Sato, Makoto Yamada, Hisashi Kashima</i>	

Understanding Contrastive Learning Requires Incorporating Inductive Biases.....	19239
<i>Nikunj Saunshi, Jordan Ash, Surbhi Goel, Dipendra Misra, Cyril Zhang, Sanjeev Arora, Sham Kakade, Akshay Krishnamurthy</i>	
The Neural Race Reduction: Dynamics of Abstraction in Gated Networks.....	19276
<i>Andrew Saxe, Shagun Sodhani, Sam Jay Lewallen</i>	

## PART 24

Convergence Rates of Non-Convex Stochastic Gradient Descent Under a Generic Lojasiewicz Condition and Local Smoothness .....	19299
<i>Kevin Scaman, Cedric Malherbe, Ludovic Dos Santos</i>	
An Asymptotic Test for Conditional Independence Using Analytic Kernel Embeddings .....	19317
<i>Meyer Scetbon, Laurent Meunier, Yaniv Romano</i>	
Linear-Time Gromov Wasserstein Distances Using Low Rank Couplings and Costs.....	19336
<i>Meyer Scetbon, Gabriel Peyré, Marco Cuturi</i>	
Streaming Inference for Infinite Feature Models .....	19355
<i>Rylan Schaeffer, Yilun Du, Gabrielle K Liu, Ila Fiete</i>	
Modeling Irregular Time Series with Continuous Recurrent Units.....	19377
<i>Mona Schirmer, Mazin Eltayeb, Stefan Lessmann, Maja Rudolph</i>	
Structure Preserving Neural Networks: A Case Study in the Entropy Closure of the Boltzmann Equation .....	19395
<i>Steffen Schothöfer, Tianbai Xiao, Martin Frank, Cory Hauck</i>	
Improving Robustness Against Real-World and Worst-Case Distribution Shifts Through Decision Region Quantification .....	19423
<i>Leo Schwinn, Leon Bungert, An Nguyen, René Raab, Falk Pulsmeier, Doina Precup, Bjoern Eskofier, Dario Zanca</i>	
Symmetric Machine Theory of Mind .....	19439
<i>Melanie Sclar, Graham Neubig, Yonatan Bisk</i>	
Data-SUITE: Data-Centric Identification of In-Distribution Incongruous Examples .....	19456
<i>Nabeel Seedat, Jonathan Crabbé, Mihaela Van Der Schaar</i>	
Continuous-Time Modeling of Counterfactual Outcomes Using Neural Controlled Differential Equations.....	19486
<i>Nabeel Seedat, Fergus Imrie, Alexis Bellot, Zhaozhi Qian, Mihaela Van Der Schaar</i>	
Neural Tangent Kernel Beyond the Infinite-Width Limit: Effects of Depth and Initialization.....	19511
<i>Mariia Seleznova, Gitta Kutyniok</i>	
Reinforcement Learning with Action-Free Pre-Training from Videos .....	19550
<i>Younggyo Seo, Kimin Lee, Stephen L James, Pieter Abbeel</i>	
Efficient Model-Based Multi-Agent Reinforcement Learning Via Optimistic Equilibrium Computation .....	19569
<i>Pier Giuseppe Sessa, Maryam Kamgarpour, Andreas Krause</i>	
Selective Regression Under Fairness Criteria .....	19587
<i>Abhin Shah, Yuheng Bu, Joshua K Lee, Subhro Das, Rameswar Panda, Prasanna Sattigeri, Gregory W Wornell</i>	
Utility Theory for Sequential Decision Making .....	19605
<i>Mehran Shakerinava, Siamak Ravanbakhsh</i>	
Translating Robot Skills: Learning Unsupervised Skill Correspondences Across Robots .....	19615
<i>Tanmay Shankar, Yixin Lin, Aravind Rajeswaran, Vikash Kumar, Stuart Anderson, Jean Oh</i>	
A State-Distribution Matching Approach to Non-Episodic Reinforcement Learning .....	19634
<i>Archit Sharma, Rehaan Ahmad, Chelsea Finn</i>	
Content Addressable Memory Without Catastrophic Forgetting by Heteroassociation with a Fixed Scaffold .....	19647
<i>Sugandha Sharma, Sarthak Chandra, Ila Fiete</i>	

Federated Minimax Optimization: Improved Convergence Analyses and Algorithms .....	19672
<i>Pranay Sharma, Rohan Panda, Gauri Joshi, Pramod Varshney</i>	
DNS: Determinantal Point Process Based Neural Network Sampler for Ensemble Reinforcement Learning .....	19720
<i>Hassam Sheikh, Kizza Frisbee, Mariano Phielipp</i>	
Instance Dependent Regret Analysis of Kernelized Bandits .....	19736
<i>Shubhanshu Shekhar, Tara Javidi</i>	
Data Augmentation as Feature Manipulation .....	19762
<i>Ruoqi Shen, Sebastien Bubeck, Suriya Gunasekar</i>	
Metric-Fair Active Learning .....	19798
<i>Jie Shen, Nan Cui, Jing Wang</i>	
PDO-S3DCNNs: Partial Differential Operator Based Steerable 3D CNNs .....	19816
<i>Zhengyang Shen, Tao Hong, Qi She, Jinwen Ma, Zhouchen Lin</i>	
Connect, Not Collapse: Explaining Contrastive Learning for Unsupervised Domain Adaptation .....	19836
<i>Kendrick Shen, Robbie M Jones, Ananya Kumar, Sang Michael Xie, Jeff Z. Haochen, Tengyu Ma, Percy Liang</i>	
Constrained Optimization with Dynamic Bound-Scaling for Effective NLP Backdoor Defense .....	19868
<i>Guangyu Shen, Yingqi Liu, Guan hong Tao, Qiuling Xu, Zhuo Zhang, Shengwei An, Shiqing Ma, Xiangyu Zhang</i>	
Staged Training for Transformer Language Models .....	19882
<i>Sheng Shen, Pete Walsh, Kurt Keutzer, Jesse Dodge, Matthew Peters, Iz Beltagy</i>	
Deep Network Approximation in Terms of Intrinsic Parameters .....	19898
<i>Zuowei Shen, Haizhao Yang, Shijun Zhang</i>	
Gradient-Free Method for Heavily Constrained Nonconvex Optimization .....	19924
<i>Wanli Shi, Hongchang Gao, Bin Gu</i>	
Global Optimization of K-Center Clustering .....	19945
<i>Mingfei Shi, Kaixun Hua, Jiayang Ren, Yankai Cao</i>	
Pessimistic Q-Learning for Offline Reinforcement Learning: Towards Optimal Sample Complexity .....	19956
<i>Laixi Shi, Gen Li, Yuting Wei, Yuxin Chen, Yuejie Chi</i>	
Adversarial Masking for Self-Supervised Learning .....	20015
<i>Yuge Shi, N Siddharth, Philip Torr, Adam R Kosiorek</i>	
Visual Attention Emerges from Recurrent Sparse Reconstruction .....	20030
<i>Baifeng Shi, Yale Song, Neel Joshi, Trevor Darrell, Xin Wang</i>	
A Minimax Learning Approach to Off-Policy Evaluation in Confounded Partially Observable Markov Decision Processes .....	20046
<i>Chengchun Shi, Masatoshi Uehara, Jiawei Huang, Nan Jiang</i>	
Robust Group Synchronization Via Quadratic Programming .....	20084
<i>Yunpeng Shi, Cole M Wyeth, Gilad Lerman</i>	
Log-Euclidean Signatures for Intrinsic Distances Between Unaligned Datasets .....	20095
<i>Tal Shnitzer, Mikhail Yurochkin, Kristjan Greenewald, Justin M Solomon</i>	
Scalable Computation of Causal Bounds .....	20114
<i>Madhumitha Shridharan, Garud Iyengar</i>	

## PART 25

Bit Prioritization in Variational Autoencoders Via Progressive Coding .....	20130
<i>Rui Shu, Stefano Ermon</i>	
Fair Representation Learning Through Implicit Path Alignment .....	20145
<i>Changjian Shui, Qi Chen, Jiaqi Li, Boyu Wang, Christian Gagné</i>	
Faster Algorithms for Learning Convex Functions .....	20165
<i>Ali Siahkamari, Durmus Alp Emre Acar, Christopher Liao, Kelly L Geyer, Venkatesh Saligrama, Brian Kulis</i>	

Coin Flipping Neural Networks .....	20184
<i>Yuval Sieradzki, Nitzan Hodos, Gal Yehuda, Assaf Schuster</i>	
Reverse Engineering the Neural Tangent Kernel.....	20204
<i>James Benjamin Simon, Sajant Anand, Mike Deweese</i>	
Demystifying the Adversarial Robustness of Random Transformation Defenses .....	20221
<i>Chawin Sitawarin, Zachary J Golan-Strieb, David Wagner</i>	
Smoothed Adversarial Linear Contextual Bandits with Knapsacks .....	20242
<i>Vidyashankar Sivakumar, Shiliang Zuo, Arindam Banerjee</i>	
GenLabel: Mixup Relabeling Using Generative Models.....	20267
<i>Jy-Yong Sohn, Liang Shang, Hongxu Chen, Jaekyun Moon, Dimitris Papailiopoulos, Kangwook Lee</i>	
Communicating Via Markov Decision Processes.....	20303
<i>Samuel Sokota, Christian A Schroeder De Witt, Maximilian Igl, Luisa M Zintgraf, Philip Torr, Martin Strohmeier, Zico Kolter, Shimon Whiteson, Jakob Foerster</i>	
The Multivariate Community Hawkes Model for Dependent Relational Events in Continuous-Time Networks .....	20318
<i>Hadeel Soliman, Lingfei Zhao, Zhipeng Huang, Subhadeep Paul, Kevin S Xu</i>	
Disentangling Sources of Risk for Distributional Multi-Agent Reinforcement Learning .....	20336
<i>Kyunghwan Son, Junsu Kim, Sungsoo Ahn, Roben D Delos Reyes, Yung Yi, Jinwoo Shin</i>	
TAM: Topology-Aware Margin Loss for Class-Imbalanced Node Classification .....	20358
<i>Jaeyun Song, Joonhyung Park, Eunho Yang</i>	
A General Recipe for Likelihood-Free Bayesian Optimization.....	20373
<i>Jiaming Song, Lantao Yu, Willie Neiswanger, Stefano Ermon</i>	
Fully-Connected Network on Noncompact Symmetric Space and Ridgelet Transform Based on Helgason-Fourier Analysis .....	20394
<i>Sho Sonoda, Isao Ishikawa, Masahiro Ikeda</i>	
Saute RL: Almost Surely Safe Reinforcement Learning Using State Augmentation .....	20412
<i>Aivar Sootla, Alexander I Cowen-Rivers, Taher Jafferjee, Ziyang Wang, David H Mguni, Jun Wang, Haitham Ammar</i>	
Lightweight Projective Derivative Codes for Compressed Asynchronous Gradient Descent .....	20433
<i>Pedro J Soto, Ilija Ilmer, Haibin Guan, Jun Li</i>	
Accelerating Bayesian Optimization for Biological Sequence Design with Denoising Autoencoders .....	20448
<i>Samuel Stanton, Wesley Maddox, Nate Gruver, Phillip Maffettone, Emily Delaney, Peyton Greenside, Andrew Gordon Wilson</i>	
3D Infomax Improves GNNs for Molecular Property Prediction.....	20468
<i>Hannes Stärk, Dominique Beaini, Gabriele Corso, Prudencio Tossou, Christian Dallago, Stephan Günnemann, Pietro Lió</i>	
EquiBind: Geometric Deep Learning for Drug Binding Structure Prediction.....	20492
<i>Hannes Stärk, Octavian Ganea, Lagnajit Pattanaik, Dr. Regina Barzilay, Tommi Jaakkola</i>	
Plug & Play Attacks: Towards Robust and Flexible Model Inversion Attacks.....	20511
<i>Lukas Struppek, Dominik Hintersdorf, Antonio De Almeida Correia, Antonia Adler, Kristian Kersting</i>	
Scaling-Up Diverse Orthogonal Convolutional Networks by a Paraunitary Framework .....	20535
<i>Jiahao Su, Wonmin Byeon, Furong Huang</i>	
Divergence-Regularized Multi-Agent Actor-Critic .....	20569
<i>Kefan Su, Zongqing Lu</i>	
Influence-Augmented Local Simulators: A Scalable Solution for Fast Deep RL in Large Networked Systems.....	20593
<i>Miguel Suau, Jinke He, Matthijs T. J. Spaan, Frans Oliehoek</i>	
Improved StyleGAN-V2 Based Inversion for Out-Of-Distribution Images.....	20614
<i>Rakshith Subramanyam, Vivek Narayanaswamy, Mark Naufel, Andreas Spanias, Jayaraman J. Thiagarajan</i>	

Continuous-Time Analysis of Accelerated Gradient Methods Via Conservation Laws in Dilated Coordinate Systems .....	20629
<i>Jaewook J Suh, Gyumin Roh, Ernest K Ryu</i>	
Do Differentiable Simulators Give Better Policy Gradients? .....	20657
<i>Hyung Ju Suh, Max Simchowitz, Kaiqing Zhang, Russ Tedrake</i>	
Intriguing Properties of Input-Dependent Randomized Smoothing .....	20686
<i>Peter Sükeni 'k, Aleksei Kuvshinov, Stephan Günnemann</i>	
Cliff Diving: Exploring Reward Surfaces in Reinforcement Learning Environments .....	20733
<i>Ryan Sullivan, Jordan K Terry, Benjamin Black, John P Dickerson</i>	
AGNAS: Attention-Guided Micro and Macro-Architecture Search.....	20766
<i>Zihao Sun, Yu Hu, Shun Lu, Longxing Yang, Jilin Mei, Yinhe Han, Xiaowei Li</i>	
Adaptive Random Walk Gradient Descent for Decentralized Optimization .....	20779
<i>Tao Sun, Dongsheng Li, Bao Wang</i>	
MAE-DET: Revisiting Maximum Entropy Principle in Zero-Shot NAS for Efficient Object Detection .....	20799
<i>Zhenhong Sun, Ming Lin, Xiuyu Sun, Zhiyu Tan, Hao Li, Rong Jin</i>	
Out-Of-Distribution Detection with Deep Nearest Neighbors .....	20816
<i>Yiyu Sun, Yifei Ming, Xiaojin Zhu, Yixuan Li</i>	
Black-Box Tuning for Language-Model-As-A-Service .....	20830
<i>Tianxiang Sun, Yunfan Shao, Hong Qian, Xuanjing Huang, Xipeng Qiu</i>	
Correlated Quantization for Distributed Mean Estimation and Optimization .....	20845
<i>Ananda Theertha Suresh, Ziteng Sun, Jae Ro, Felix Yu</i>	
Causal Imitation Learning Under Temporally Correlated Noise .....	20866
<i>Gokul Swamy, Sanjiban Choudhury, Drew Bagnell, Steven Wu</i>	
Being Properly Improper .....	20880
<i>Tyler Sypherd, Richard Nock, Lalitha Sankar</i>	
Distributionally-Aware Kernelized Bandit Problems for Risk Aversion .....	20922
<i>Sho Takemori</i>	
Sequential and Parallel Constrained Max-Value Entropy Search Via Information Lower Bound.....	20949
<i>Shion Takeno, Tomoyuki Tamura, Kazuki Shitara, Masayuki Karasuyama</i>	

## PART 26

SQ-VAE: Variational Bayes on Discrete Representation with Self-Annealed Stochastic Quantization .....	20976
<i>Yuhta Takida, Takashi Shibuya, Weihsiang Liao, Chieh-Hsin Lai, Junki Ohmura, Toshimitsu Uesaka, Naoki Murata, Shusuke Takahashi, Toshiyuki Kumakura, Yuki Mitsufuji</i>	
A Tree-Based Model Averaging Approach for Personalized Treatment Effect Estimation from Heterogeneous Data Sources .....	21002
<i>Xiaoqing Tan, Chung-Chou H. Chang, Ling Zhou, Lu Tang</i>	
N-Penetrate: Active Learning of Neural Collision Handler for Complex 3D Mesh Deformations .....	21026
<i>Qingyang Tan, Zherong Pan, Breannan Smith, Takaaki Shiratori, Dinesh Manocha</i>	
Biased Gradient Estimate with Drastic Variance Reduction for Meta Reinforcement Learning .....	21039
<i>Yunhao Tang</i>	
Rethinking Graph Neural Networks for Anomaly Detection .....	21065
<i>Jianheng Tang, Jiajin Li, Ziqi Gao, Jia Li</i>	
Deep Safe Incomplete Multi-View Clustering: Theorem and Algorithm .....	21079
<i>Huayi Tang, Yong Liu</i>	
Virtual Homogeneity Learning: Defending Against Data Heterogeneity in Federated Learning .....	21100
<i>Zhenheng Tang, Yonggang Zhang, Shaohuai Shi, Xin He, Bo Han, Xiaowen Chu</i>	
Cross-Space Active Learning on Graph Convolutional Networks .....	21122
<i>Yufei Tao, Hao Wu, Shiyuan Deng</i>	

FedNest: Federated Bilevel, Minimax, and Compositional Optimization .....	21135
<i>Davoud Ataee Tarzanagh, Mingchen Li, Christos Thrampoulidis, Samet Oymak</i>	
Efficient Distributionally Robust Bayesian Optimization with Worst-Case Sensitivity .....	21169
<i>Sebastian Shenghong Tay, Chuan Sheng Foo, Urano Daisuke, Richalynn Leong, Bryan Kian Hsiang Low</i>	
LIDL: Local Intrinsic Dimension Estimation Using Approximate Likelihood .....	21194
<i>Piotr Tempczyk, Rafal Michaluk, Lukasz Garncarek, Przemyslaw Spurek, Jacek Tabor, Adam Golinski</i>	
LCANets: Lateral Competition Improves Robustness Against Corruption and Attack .....	21221
<i>Michael Teti, Garrett Kenyon, Ben Migliori, Juston Moore</i>	
Reverse Engineering $\ell_p$ Attacks: A Block-Sparse Optimization Approach with Recovery Guarantees .....	21242
<i>Darshan Thaker, Paris Giampouras, Rene Vidal</i>	
Generalised Policy Improvement with Geometric Policy Composition .....	21261
<i>Shantanu Thakoor, Mark Rowland, Diana Borsa, Will Dabney, Remi Munos, Andre Barreto</i>	
Algorithms for the Communication of Samples .....	21297
<i>Lucas Theis, Noureldin Y Ahmed</i>	
Consistent Polyhedral Surrogates for Top-K Classification and Variants .....	21318
<i>Anish Thilagar, Rafael Frongillo, Jessica J Finocchiaro, Emma Goodwill</i>	
On the Finite-Time Complexity and Practical Computation of Approximate Stationarity Concepts of Lipschitz Functions .....	21349
<i>Lai Tian, Kaiwen Zhou, Anthony Man-Cho So</i>	
From Dirichlet to Rubin: Optimistic Exploration in RL Without Bonuses .....	21369
<i>Daniil Tiapkin, Denis Belomestny, Eric Moulines, Alexey Naumov, Sergey Samsonov, Yunhao Tang, Michal Valko, Pierre Menard</i>	
Nonparametric Sparse Tensor Factorization with Hierarchical Gamma Processes .....	21421
<i>Conor Tillinghast, Zheng Wang, Shandian Zhe</i>	
Deciphering Lasso-Based Classification Through a Large Dimensional Analysis of the Iterative Soft-Thresholding Algorithm .....	21438
<i>Malik Tiomoko, Ekkehard Schnoor, Mohamed El Amine Seddik, Igor Colin, Aladin Virmaux</i>	
Extended Unconstrained Features Model for Exploring Deep Neural Collapse .....	21467
<i>Tom Tirer, Joan Bruna</i>	
Object Permanence Emerges in a Random Walk Along Memory .....	21495
<i>Pavel Tokmakov, Allan Jabri, Jie Li, Adrien Gaidon</i>	
Generic Coreset for Scalable Learning of Monotonic Kernels: Logistic Regression, Sigmoid and More .....	21509
<i>Elad Tolochinsky, Ibrahim Jubran, Dan Feldman</i>	
Failure and Success of the Spectral Bias Prediction for Laplace Kernel Ridge Regression: The Case of Low-Dimensional Data .....	21537
<i>Umberto M Tomasini, Antonio Sclocchi, Matthieu Wyart</i>	
Quantifying and Learning Linear Symmetry-Based Disentanglement .....	21573
<i>Loek Tonnaer, Luis Armando Perez Rey, Vlado Menkovski, Mike Holenderski, Jim Portegies</i>	
A Temporal-Difference Approach to Policy Gradient Estimation .....	21598
<i>Samuele Tosatto, Andrew Patterson, Martha White, Rupam Mahmood</i>	
Simple and Near-Optimal Algorithms for Hidden Stratification and Multi-Group Learning .....	21622
<i>Christopher J Tosh, Daniel Hsu</i>	
Design-Bench: Benchmarks for Data-Driven Offline Model-Based Optimization .....	21647
<i>Brandon Trabucco, Xinyang Geng, Aviral Kumar, Sergey Levine</i>	
AnyMorph: Learning Transferable Policies by Inferring Agent Morphology .....	21666
<i>Brandon Trabucco, Mariano Phielipp, Glen Berseth</i>	
Detecting Adversarial Examples is (Nearly) as Hard as Classifying Them .....	21681
<i>Florian Tramer</i>	

Nesterov Accelerated Shuffling Gradient Method for Convex Optimization.....	21692
<i>Trang H Tran, Katya Scheinberg, Lam M Nguyen</i>	
A Completely Tuning-Free and Robust Approach to Sparse Precision Matrix Estimation .....	21722
<i>Chau Tran, Guo Yu</i>	
Tackling Covariate Shift with Node-Based Bayesian Neural Networks .....	21740
<i>Trung Q Trinh, Markus Heinonen, Luigi Acerbi, Samuel Kaski</i>	
Fenrir: Physics-Enhanced Regression for Initial Value Problems .....	21765
<i>Filip Tronarp, Nathanael Bosch, Philipp Hennig</i>	
Interpretable Off-Policy Learning Via Hyperbox Search .....	21784
<i>Daniel Tschernutter, Tobias Hatt, Stefan Feuerriegel</i>	

## PART 27

FriendlyCore: Practical Differentially Private Aggregation .....	21817
<i>Eliad Tsfadia, Edith Cohen, Haim Kaplan, Yishay Mansour, Uri Stemmer</i>	
Pairwise Conditional Gradients Without Swap Steps and Sparser Kernel Herding .....	21853
<i>Kazuma K Tsuji, Ken 'Ichiro Tanaka, Sebastian Pokutta</i>	
Prototype Based Classification from Hierarchy to Fairness .....	21873
<i>Mycal Tucker, Julie A. Shah</i>	
Consensus Multiplicative Weights Update: Learning to Learn Using Projector-Based Game Signatures .....	21890
<i>Nelson Vadori, Rahul Savani, Thomas Spooner, Sumitra Ganesh</i>	
Self-Supervised Models of Audio Effectively Explain Human Cortical Responses to Speech .....	21916
<i>Aditya R Vaidya, Shailee Jain, Alexander Huth</i>	
Path-Gradient Estimators for Continuous Normalizing Flows .....	21934
<i>Lorenz Vaitl, Kim Andrea Nicoli, Shinichi Nakajima, Pan Kessel</i>	
Improved Convergence Rates for Sparse Approximation Methods in Kernel-Based Learning .....	21949
<i>Sattar Vakili, Jonathan Scarlett, Da-Shan Shiu, Alberto Bernacchia</i>	
EDEN: Communication-Efficient and Robust Distributed Mean Estimation for Federated Learning .....	21973
<i>Shay Vargafik, Ran Ben Basat, Amit Portnoy, Gal Mendelson, Yaniv Ben Itzhak, Michael Mitzenmacher</i>	
Towards Noise-Adaptive, Problem-Adaptive (Accelerated) Stochastic Gradient Descent .....	22004
<i>Sharan Vaswani, Benjamin Dubois-Taine, Reza Babanezhad</i>	
Correlation Clustering Via Strong Triadic Closure Labeling: Fast Approximation Algorithms and Practical Lower Bounds .....	22049
<i>Nate Veldt</i>	
The CLRS Algorithmic Reasoning Benchmark.....	22073
<i>Petar Velickovic, Adrià Puigdomènech Badia, David Budden, Razvan Pascanu, Andrea Banino, Misha Dashevskiy, Raia Hadsell, Charles Blundell</i>	
Bregman Power k-Means for Clustering Exponential Family Data .....	22092
<i>Adithya Vellal, Saptarshi Chakraborty, Jason Q Xu</i>	
Estimation in Rotationally Invariant Generalized Linear Models Via Approximate Message Passing .....	22109
<i>Ramji Venkataramanan, Kevin Kögler, Marco Mondelli</i>	
Bayesian Optimization Under Stochastic Delayed Feedback.....	22134
<i>Arun Verma, Zhongxiang Dai, Bryan Kian Hsiang Low</i>	
VarScene: A Deep Generative Model for Realistic Scene Graph Synthesis .....	22157
<i>Tathagat Verma, Abir De, Yateesh Agrawal, Vishwa Vinay, Soumen Chakrabarti</i>	
Calibrated Learning to Defer with One-Vs-All Classifiers .....	22173
<i>Rajeev Verma, Eric Nalisnick</i>	
Regret Bounds for Stochastic Shortest Path Problems with Linear Function Approximation.....	22192
<i>Daniel Vial, Advait Parulekar, Sanjay Shakkottai, R Srikant</i>	

On Implicit Bias in Overparameterized Bilevel Optimization .....	22223
<i>Paul Vicol, Jonathan P Lorraine, Fabian Pedregosa, David Duvenaud, Roger B Grosse</i>	
Multiclass Learning with Margin: Exponential Rates with No Bias-Variance Trade-Off .....	22249
<i>Stefano Vigogna, Giacomo Meanti, Ernesto De Vito, Lorenzo Rosasco</i>	
Addressing Optimism Bias in Sequence Modeling for Reinforcement Learning.....	22259
<i>Adam R Villaflor, Zhe Huang, Swapnil Pande, John M Dolan, Jeff Schneider</i>	
Bayesian Nonparametrics for Offline Skill Discovery .....	22273
<i>Valentin Vilecroze, Harry Braviner, Panteha Naderian, Chris Maddison, Gabriel Loaiza-Ganem</i>	
Hermite Polynomial Features for Private Data Generation .....	22289
<i>Margarita Vinaroz, Mohammad-Amin Charusaie, Frederik Harder, Kamil Adamczewski, Mi Jung Park</i>	
What Can Linear Interpolation of Neural Network Loss Landscapes Tell Us?.....	22314
<i>Tiffany J Vlaar, Jonathan Frankle</i>	
Multirate Training of Neural Networks .....	22331
<i>Tiffany J Vlaar, Benedict Leimkuhler</i>	
Provably Adversarially Robust Nearest Prototype Classifiers.....	22350
<i>Václav Voráček, Matthias Hein</i>	
First-Order Regret in Reinforcement Learning with Linear Function Approximation: A Robust Estimation Approach .....	22373
<i>Andrew J Wagenmaker, Yifang Chen, Max Simchowitz, Simon Du, Kevin Jamieson</i>	
Reward-Free RL is No Harder than Reward-Aware RL in Linear Markov Decision Processes.....	22419
<i>Andrew J Wagenmaker, Yifang Chen, Max Simchowitz, Simon Du, Kevin Jamieson</i>	
Training Characteristic Functions with Reinforcement Learning: XAI-Methods Play Connect Four.....	22446
<i>Stephan Wäldchen, Sebastian Pokutta, Felix Huber</i>	
Retroformer: Pushing the Limits of End-To-End Retrosynthesis Transformer .....	22464
<i>Yue Wan, Chang-Yu Hsieh, Ben Liao, Shengyu Zhang</i>	
Safe Exploration for Efficient Policy Evaluation and Comparison .....	22480
<i>Runzhe Wan, Branislav Kveton, Rui Song</i>	
Greedy Based Value Representation for Optimal Coordination in Multi-Agent Reinforcement Learning .....	22501
<i>Lipeng Wan, Zeyang Liu, Xingyu Chen, Xuguang Lan, Nanning Zheng</i>	
Towards Evaluating Adaptivity of Model-Based Reinforcement Learning Methods.....	22525
<i>Yi Wan, Ali Rahimi-Kalahroudi, Janarthanan Rajendran, Ida Momennejad, Sarath Chandar, Harm H Van Seijen</i>	
Fast Lossless Neural Compression with Integer-Only Discrete Flows.....	22551
<i>Siyu Wang, Jianfei Chen, Chongxuan Li, Jun Zhu, Bo Zhang</i>	
Accelerating Shapley Explanation Via Contributive Cooperator Selection.....	22565
<i>Guanchu Wang, Yu-Neng Chuang, Mengnan Du, Fan Yang, Quan Zhou, Pushkar Tripathi, Xuanting Cai, Xia Hu</i>	
Denoised MDPs: Learning World Models Better than the World Itself .....	22580
<i>Tongzhou Wang, Simon Du, Antonio Torralba, Phillip Isola, Amy Zhang, Yuandong Tian</i>	
Neural Implicit Dictionary Learning Via Mixture-Of-Expert Training .....	22602
<i>Peihao Wang, Zhiwen Fan, Tianlong Chen, Zhangyang Wang</i>	
Robust Models Are More Interpretable Because Attributions Look Normal .....	22614
<i>Zifan Wang, Matt Fredrikson, Anupam Datta</i>	
Disentangling Disease-Related Representation from Obscure for Disease Prediction.....	22641
<i>Chu-Ran Wang, Fei Gao, Fandong Zhang, Fangwei Zhong, Yizhou Yu, Yizhou Wang</i>	



## PART 28

Solving Stackelberg Prediction Game with Least Squares Loss Via Spherically Constrained Least Squares Reformulation .....	22654
<i>Jiali Wang, Wen Huang, Rujun Jiang, Xudong Li, Alex L Wang</i>	
VLMixer: Unpaired Vision-Language Pre-Training Via Cross-Modal CutMix .....	22669
<i>Teng Wang, Wenhao Jiang, Zhichao Lu, Feng Zheng, Ran Cheng, Chengguo Yin, Ping Luo</i>	
DynaMixer: A Vision MLP Architecture with Dynamic Mixing .....	22680
<i>Ziyu Wang, Wenhao Jiang, Yiming M Zhu, Li Yuan, Yibing Song, Wei Liu</i>	
Improving Screening Processes Via Calibrated Subset Selection.....	22691
<i>Lequn Wang, Thorsten Joachims, Manuel Gomez Rodriguez</i>	
The Geometry of Robust Value Functions.....	22716
<i>Kaixin Wang, Navdeep Kumar, Kuangqi Zhou, Bryan Hooi, Jiashi Feng, Shie Mannor</i>	
What Dense Graph Do You Need for Self-Attention? .....	22741
<i>Yuxin Wang, Chu-Tak Lee, Qipeng Guo, Zhangyue Yin, Yunhua Zhou, Xuanjing Huang, Xipeng Qiu</i>	
Improved Certified Defenses Against Data Poisoning with (Deterministic) Finite Aggregation .....	22758
<i>Wenxiao Wang, Alexander J Levine, Soheil Feizi</i>	
Understanding Gradual Domain Adaptation: Improved Analysis, Optimal Path and Beyond .....	22773
<i>Haoxiang Wang, Bo Li, Han Zhao</i>	
Communication-Efficient Adaptive Federated Learning .....	22791
<i>Yujia Wang, Lu Lin, Jinghui Chen</i>	
Provable Acceleration of Heavy Ball Beyond Quadratics for a Class of Polyak-Lojasiewicz Functions When the Non-Convexity is Averaged-Out .....	22828
<i>Jun-Kun Wang, Chi-Heng Lin, Andre Wibisono, Bin Hu</i>	
Robustness Verification for Contrastive Learning .....	22854
<i>Zekai Wang, Weiwei Liu</i>	
Convergence and Recovery Guarantees of the K-Subspaces Method for Subspace Clustering .....	22873
<i>Peng Wang, Huikang Liu, Anthony Man-Cho So, Laura Balzano</i>	
NP-Match: When Neural Processes Meet Semi-Supervised Learning .....	22908
<i>Jianfeng Wang, Thomas Lukasiewicz, Daniela Massiceti, Xiaolin Hu, Vladimir Pavlovic, Alexandros Neophytou</i>	
Iterative Double Sketching for Faster Least-Squares Optimization .....	22924
<i>Rui Wang, Yanyan Ouyang, Wangli Xu</i>	
What Language Model Architecture and Pretraining Objective Works Best for Zero-Shot Generalization?.....	22953
<i>Thomas Wang, Adam Roberts, Daniel Hesslow, Teven Le Scao, Hyung Won Chung, Iz Beltagy, Julien Launay, Colin Raffel</i>	
Improving Task-Free Continual Learning by Distributionally Robust Memory Evolution.....	22974
<i>Zhenyi Wang, Li Shen, Le Fang, Qiuling Suo, Tiehang Duan, Mingchen Gao</i>	
Risk-Averse No-Regret Learning in Online Convex Games .....	22988
<i>Zifan Wang, Yi Shen, Michael Zavlanos</i>	
Provable Domain Generalization Via Invariant-Feature Subspace Recovery .....	23007
<i>Haoxiang Wang, Haozhe Si, Bo Li, Han Zhao</i>	
ProgFed: Effective, Communication, and Computation Efficient Federated Learning by Progressive Training .....	23023
<i>Hui-Po Wang, Sebastian Stich, Yang He, Mario Fritz</i>	
Model-Based Meta Reinforcement Learning Using Graph Structured Surrogate Models and Amortized Policy Search .....	23044
<i>Qi Wang, Herke Van Hoof</i>	

Approximately Equivariant Networks for Imperfectly Symmetric Dynamics .....	23067
<i>Rui Wang, Robin Walters, Rose Yu</i>	
Three-Stage Evolution and Fast Equilibrium for SGD with Non-Degenerate Critical Points .....	23081
<i>Yi Wang, Zhiren Wang</i>	
Understanding Instance-Level Impact of Fairness Constraints .....	23103
<i>Jialu Wang, Xin Eric Wang, Yang Liu</i>	
Tractable Uncertainty for Structure Learning .....	23120
<i>Benjie Wang, Matthew R Wicker, Marta Kwiatkowska</i>	
Causal Dynamics Learning for Task-Independent State Abstraction .....	23140
<i>Zizhao Wang, Xuesu Xiao, Zifan Xu, Yuke Zhu, Peter Stone</i>	
Multiple-Play Stochastic Bandits with Shareable Finite-Capacity Arms .....	23170
<i>Xuchuang Wang, Hong Xie, John C. S. Lui</i>	
Generative Coarse-Graining of Molecular Conformations.....	23202
<i>Wujie Wang, Minkai Xu, Chen Cai, Benjamin K Miller, Tess Smidt, Yusu Wang, Jian Tang, Rafael Gomez-Bombarelli</i>	
Nonparametric Embeddings of Sparse High-Order Interaction Events.....	23226
<i>Zheng Wang, Yiming Xu, Conor Tillinghast, Shibo Li, Akil Narayan, Shandian Zhe</i>	
When Are Linear Stochastic Bandits Attackable? .....	23243
<i>Huazheng Wang, Haifeng Xu, Hongning Wang</i>	
DRAGONN: Distributed Randomized Approximate Gradients of Neural Networks .....	23263
<i>Zhuang Wang, Zhaozhao Xu, Xinyu Wu, Anshumali Shrivastava, T. S. Eugene Ng</i>	
Finite-Sum Coupled Compositional Stochastic Optimization: Theory and Applications.....	23281
<i>Bokun Wang, Tianbao Yang</i>	
OFA: Unifying Architectures, Tasks, and Modalities Through a Simple Sequence-To-Sequence Learning Framework .....	23307
<i>Peng Wang, An Yang, Rui Men, Junyang Lin, Shuai Bai, Zhikang Li, Jianxin Ma, Chang Zhou, Jingren Zhou, Hongxia Yang</i>	
How Powerful Are Spectral Graph Neural Networks.....	23330
<i>Xiyuan Wang, Muhan Zhang</i>	
Thompson Sampling for Robust Transfer in Multi-Task Bandits.....	23352
<i>Zhi Wang, Chicheng Zhang, Kamalika Chaudhuri</i>	
Individual Reward Assisted Multi-Agent Reinforcement Learning .....	23406
<i>Li Wang, Yupeng Zhang, Yujing Hu, Weixun Wang, Chongjie Zhang, Yang Gao, Jianye Hao, Tangjie Lv, Changjie Fan</i>	
Removing Batch Normalization Boosts Adversarial Training.....	23422
<i>Haotao Wang, Aston Zhang, Shuai Zheng, Xingjian Shi, Mu Li, Zhangyang Wang</i>	
Partial and Asymmetric Contrastive Learning for Out-Of-Distribution Detection in Long-Tailed Recognition .....	23435
<i>Haotao Wang, Aston Zhang, Yi Zhu, Shuai Zheng, Mu Li, Alex J Smola, Zhangyang Wang</i>	
Nonparametric Factor Trajectory Learning for Dynamic Tensor Decomposition .....	23448
<i>Zheng Wang, Shandian Zhe</i>	
Thompson Sampling for (Combinatorial) Pure Exploration .....	23459
<i>Siwei Wang, Jun Zhu</i>	
Policy Gradient Method for Robust Reinforcement Learning.....	23473
<i>Yue Wang, Shaofeng Zou</i>	

## PART 29

Certifying Out-Of-Domain Generalization for Blackbox Functions .....	23516
<i>Maurice G Weber, Linyi Li, Boxin Wang, Zhikuan Zhao, Bo Li, Ce Zhang</i>	

More than a Toy: Random Matrix Models Predict How Real-World Neural Representations Generalize.....	23538
<i>Alexander Wei, Wei Hu, Jacob Steinhardt</i>	
To Smooth Or Not? When Label Smoothing Meets Noisy Labels.....	23578
<i>Jiaheng Wei, Hangyu Liu, Tongliang Liu, Gang Niu, Masashi Sugiyama, Yang Liu</i>	
Open-Sampling: Exploring Out-Of-Distribution Data for Re-Balancing Long-Tailed Datasets.....	23604
<i>Hongxin Wei, Lue Tao, Renchunzi Xie, Lei Feng, Bo An</i>	
Mitigating Neural Network Overconfidence with Logit Normalization.....	23620
<i>Hongxin Wei, Renchunzi Xie, Hao Cheng, Lei Feng, Bo An, Yixuan Li</i>	
Koopman Q-Learning: Offline Reinforcement Learning Via Symmetries of Dynamics.....	23634
<i>Matthias Weissenbacher, Samarth Sinha, Animesh Garg, Kawahara Yoshinobu</i>	
Fishing for User Data in Large-Batch Federated Learning Via Gradient Magnification.....	23657
<i>Yuxin Wen, Jonas A. Geiping, Liam Fowl, Micah Goldblum, Tom Goldstein</i>	
BabelTower: Learning to Auto-Parallelized Program Translation.....	23674
<i>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</i>	
Random Forest Density Estimation.....	23690
<i>Hongwei Wen, Hanyuan Hang</i>	
Fighting Fire with Fire: Avoiding DNN Shortcuts Through Priming.....	23712
<i>Chuan Wen, Jianing Qian, Jierui Lin, Jiaye Teng, Dinesh Jayaraman, Yang Gao</i>	
Preconditioning for Scalable Gaussian Process Hyperparameter Optimization.....	23740
<i>Jonathan Wenger, Geoff Pleiss, Philipp Hennig, John Cunningham, Jacob Gardner</i>	
Measure Estimation in the Barycentric Coding Model.....	23770
<i>Matthew Werenski, Ruijie Jiang, Abiy Tasissa, Shuchin Aeron, James M Murphy</i>	
COLA: Consistent Learning with Opponent-Learning Awareness.....	23793
<i>Timon Willi, Alistair Hp Letcher, Johannes Treutlein, Jakob Foerster</i>	
Distributional Hamilton-Jacobi-Bellman Equations for Continuous-Time Reinforcement Learning.....	23821
<i>Harley E Wiltzer, David Meger, Marc G. Bellemare</i>	
Easy Variational Inference for Categorical Models Via an Independent Binary Approximation.....	23846
<i>Michael T Wojnowicz, Shuchin Aeron, Eric L Miller, Michael Hughes</i>	
Continual Learning with Guarantees Via Weight Interval Constraints.....	23886
<i>Maciej Wolczyk, Karol Piczak, Bartosz Wójcik, Lukasz Pustelnik, Pawel Morawiecki, Jacek Tabor, Tomasz Trzcinski, Przemyslaw Spurek</i>	
A Deep Learning Approach for the Segmentation of Electroencephalography Data in Eye Tracking Applications.....	23901
<i>Lukas Wolf, Ard Kastrati, Martyna B Plomecka, Jie-Ming Li, Dustin Klebe, Alexander Veicht, Roger Wattenhofer, Nicolas Langer</i>	
Leverage Score Sampling for Tensor Product Matrices in Input Sparsity Time.....	23922
<i>David Woodruff, Amir Zandieh</i>	
Model Soups: Averaging Weights of Multiple Fine-Tuned Models Improves Accuracy Without Increasing Inference Time.....	23954
<i>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</i>	
Metric-Fair Classifier Derandomization.....	23988
<i>Jimmy Wu, Yatong Chen, Yang Liu</i>	
Structural Entropy Guided Graph Hierarchical Pooling.....	24006
<i>Junran Wu, Xueyuan Chen, Ke Xu, Shangzhe Li</i>	
Self-Supervised Models Are Good Teaching Assistants for Vision Transformers.....	24020
<i>Haiyan Wu, Yuting Gao, Yinqi Zhang, Shaohui Lin, Yuan Xie, Xing Sun, Ke Li</i>	

Characterizing and Overcoming the Greedy Nature of Learning in Multi-Modal Deep Neural Networks .....	24032
<i>Nan Wu, Stanislaw Jastrzebski, Kyunghyun Cho, Krzysztof J Geras</i>	
Instrumental Variable Regression with Confounder Balancing .....	24045
<i>Anpeng Wu, Kun Kuang, Bo Li, Fei Wu</i>	
MemSR: Training Memory-Efficient Lightweight Model for Image Super-Resolution .....	24065
<i>Kailu Wu, Chung-Kuei Lee, Kaisheng Ma</i>	
Delay-Adaptive Step-Sizes for Asynchronous Learning .....	24082
<i>Xuyang Wu, Sindri Magnusson, Hamid Reza Feyzmahdavian, Mikael Johansson</i>	
Variational Nearest Neighbor Gaussian Process.....	24103
<i>Luhuan Wu, Geoff Pleiss, John P Cunningham</i>	
Understanding Policy Gradient Algorithms: A Sensitivity-Based Approach.....	24120
<i>Shuang Wu, Ling Shi, Jun Wang, Guangjian Tian</i>	
DAVINZ: Data Valuation Using Deep Neural Networks at Initialization .....	24139
<i>Zhaoxuan Wu, Yao Shu, Bryan Kian Hsiang Low</i>	
Robust Deep Reinforcement Learning Through Bootstrapped Opportunistic Curriculum .....	24166
<i>Junlin Wu, Yevgeniy Vorobeychik</i>	
Revisiting Consistency Regularization for Deep Partial Label Learning .....	24201
<i>Dong-Dong Wu, Deng-Bao Wang, Min-Ling Zhang</i>	
Flowformer: Linearizing Transformers with Conservation Flows .....	24215
<i>Haixu Wu, Jialong Wu, Jiehui Xu, Jianmin Wang, Mingsheng Long</i>	
Nearly Optimal Policy Optimization with Stable at Any Time Guarantee .....	24232
<i>Tianhao Wu, Yunchang Yang, Han Zhong, Liwei Wang, Simon Du, Jiantao Jiao</i>	
RetrievalGuard: Provably Robust 1-Nearest Neighbor Image Retrieval.....	24255
<i>Yihan Wu, Hongyang Zhang, Heng Huang</i>	
Last Iterate Risk Bounds of SGD with Decaying Stepsize for Overparameterized Linear Regression.....	24269
<i>Jingfeng Wu, Difan Zou, Vladimir Braverman, Quanquan Gu, Sham Kakade</i>	
Optimal Clustering with Noisy Queries Via Multi-Armed Bandit .....	24304
<i>Jinghui Xia, Zengfeng Huang</i>	
ProGCL: Rethinking Hard Negative Mining in Graph Contrastive Learning .....	24321
<i>Jun Xia, Lirong Wu, Ge Wang, Jintao Chen, Stan Z. Li</i>	
Synergy and Symmetry in Deep Learning: Interactions Between the Data, Model, and Inference Algorithm .....	24336
<i>Lechao Xiao, Jeffrey Pennington</i>	

### PART 30

Identification of Linear Non-Gaussian Latent Hierarchical Structure .....	24359
<i>Feng Xie, Biwei Huang, Zhengming Chen, Yangbo He, Zhi Geng, Kun Zhang</i>	
COAT: Measuring Object Compositionality in Emergent Representations.....	24377
<i>Sirui Xie, Ari S Morcos, Song-Chun Zhu, Ramakrishna Vedantam</i>	
Robust Policy Learning Over Multiple Uncertainty Sets .....	24403
<i>Annie Xie, Shagun Sodhani, Chelsea Finn, Joelle Pineau, Amy Zhang</i>	
Adaptive Inertia: Disentangling the Effects of Adaptive Learning Rate and Momentum .....	24419
<i>Zeke Xie, Xinrui Wang, Huishuai Zhang, Issei Sato, Masashi Sugiyama</i>	
Self-Supervised Representation Learning Via Latent Graph Prediction.....	24449
<i>Yaochen Xie, Zhao Xu, Shuiwang Ji</i>	
Efficient Computation of Higher-Order Subgraph Attribution Via Message Passing .....	24467
<i>Ping Xiong, Thomas Schnake, Grégoire Montavon, Klaus-Robert Müller, Shinichi Nakajima</i>	
A Self-Play Posterior Sampling Algorithm for Zero-Sum Markov Games .....	24485
<i>Wei Xiong, Han Zhong, Chengshuai Shi, Cong Shen, Tong Zhang</i>	

Importance Weighted Kernel Bayes' Rule.....	24513
<i>Liyuan Xu, Yutian Chen, Arnaud Doucet, Arthur Gretton</i>	
Learning to Separate Voices by Spatial Regions .....	24528
<i>Alan Xu, Romit Roy Choudhury</i>	
Detached Error Feedback for Distributed SGD with Random Sparsification .....	24539
<i>An Xu, Heng Huang</i>	
Accurate Quantization of Measures Via Interacting Particle-Based Optimization.....	24565
<i>Lantian Xu, Anna Korba, Dejan Slepcev</i>	
Unified Fourier-Based Kernel and Nonlinearity Design for Equivariant Networks on Homogeneous Spaces.....	24585
<i>Yinshuang Xu, Jiahui Lei, Edgar Dobriban, Kostas Daniilidis</i>	
Inferring Cause and Effect in the Presence of Heteroscedastic Noise .....	24604
<i>Sascha Xu, Osman A Mian, Alexander Marx, Jilles Vreeken</i>	
Prompting Decision Transformer for Few-Shot Policy Generalization.....	24620
<i>Mengdi Xu, Yikang Shen, Shun Zhang, Yuchen Lu, Ding Zhao, Joshua Tenenbaum, Chuang Gan</i>	
Analyzing and Mitigating Interference in Neural Architecture Search .....	24635
<i>Jin Xu, Xu Tan, Kaitao Song, Renqian Luo, Yichong Leng, Tao Qin, Tie-Yan Liu, Jian Li</i>	
On the Statistical Benefits of Curriculum Learning .....	24652
<i>Ziping Xu, Ambuj Tewari</i>	
A Difference Standardization Method for Mutual Transfer Learning .....	24672
<i>Haoqing Xu, Meng Wang, Beilun Wang</i>	
SkexGen: Autoregressive Generation of CAD Construction Sequences with Disentangled Codebooks.....	24687
<i>Xiang Xu, Karl D. D. Willis, Joseph G Lambourne, Chin-Yi Cheng, Pradeep Kumar Jayaraman, Yasutaka Furukawa</i>	
Discriminator-Weighted Offline Imitation Learning from Suboptimal Demonstrations .....	24714
<i>Haoran Xu, Xianyuan Zhan, Honglei Yin, Huling Qin</i>	
Adversarial Attack and Defense for Non-Parametric Two-Sample Tests .....	24732
<i>Xilie Xu, Jingfeng Zhang, Feng Liu, Masashi Sugiyama, Mohan Kankanhalli</i>	
Adversarially Robust Models May Not Transfer Better: Sufficient Conditions for Domain Transferability from the View of Regularization.....	24759
<i>Xiaojun Xu, Jacky Y Zhang, Evelyn Ma, Hyun Ho Son, Sanmi Koyejo, Bo Li</i>	
A Theoretical Analysis on Independence-Driven Importance Weighting for Covariate-Shift Generalization .....	24792
<i>Renzhe Xu, Xingxuan Zhang, Zheyang Shen, Tong Zhang, Peng Cui</i>	
Langevin Monte Carlo for Contextual Bandits.....	24819
<i>Pan Xu, Hongkai Zheng, Eric V Mazumdar, Kamyar Azizzadenesheli, Animashree Anandkumar</i>	
Investigating Why Contrastive Learning Benefits Robustness Against Label Noise .....	24840
<i>Yihao Xue, Kyle Whitecross, Baharan Mirzasoleiman</i>	
Diversified Adversarial Attacks Based on Conjugate Gradient Method.....	24861
<i>Keiichiro Yamamura, Haruki Sato, Nariaki Tateiwa, Nozomi Hata, Toru Mitsutake, Issa Oe, Hiroki Ishikura, Katsuki Fujisawa</i>	
Cycle Representation Learning for Inductive Relation Prediction .....	24884
<i>Zuoyu Yan, Tengfei Ma, Liangcai Gao, Zhi Tang, Chao Chen</i>	
Optimally Controllable Perceptual Lossy Compression.....	24900
<i>Zeyu Yan, Fei Wen, Peilin Liu</i>	
Active Fairness Auditing .....	24918
<i>Tom Yan, Chicheng Zhang</i>	
Self-Organized Polynomial-Time Coordination Graphs .....	24952
<i>Qianlan Yang, Weijun Dong, Zhizhou Ren, Jianhao Wang, Tonghan Wang, Chongjie Zhang</i>	

Regularizing a Model-Based Policy Stationary Distribution to Stabilize Offline Reinforcement Learning .....	24969
<i>Shentao Yang, Yihao Feng, Shujian Zhang, Mingyuan Zhou</i>	
A Psychological Theory of Explainability .....	24996
<i>Scott Cheng-Hsin Yang, Nils Erik Tomas Folke, Patrick Shafto</i>	
Omni-Granular Ego-Semantic Propagation for Self-Supervised Graph Representation Learning .....	25011
<i>Ling Yang, Shenda Hong</i>	
Unsupervised Time-Series Representation Learning with Iterative Bilinear Temporal-Spectral Fusion .....	25027
<i>Ling Yang, Shenda Hong</i>	
Searching for BurgerFormer with Micro-Meso-Macro Space Design .....	25044
<i>Longxing Yang, Yu Hu, Shun Lu, Zihao Sun, Jilin Mei, Yinhe Han, Xiaowei Li</i>	
Efficient Variance Reduction for Meta-Learning .....	25059
<i>Hansi Yang, James Kwok</i>	
Injecting Logical Constraints into Neural Networks Via Straight-Through Estimators .....	25085
<i>Zhun Yang, Joohyung Lee, Chiyoun Park</i>	
Locally Sparse Neural Networks for Tabular Biomedical Data .....	25112
<i>Junchen Yang, Ofir Lindenbaum, Yuval Kluger</i>	
Not All Poisons Are Created Equal: Robust Training Against Data Poisoning .....	25143
<i>Yu Yang, Tian Yu Liu, Baharan Mirzasoleiman</i>	
Does the Data Induce Capacity Control in Deep Learning? .....	25155
<i>Rubing Yang, Jialin Mao, Pratik Chaudhari</i>	

### PART 31

Informed Learning by Wide Neural Networks: Convergence, Generalization and Sampling Complexity .....	25187
<i>Jianyi Yang, Shaolei Ren</i>	
Linear Bandit Algorithms with Sublinear Time Complexity .....	25230
<i>Shuo Yang, Tongzheng Ren, Sanjay Shakkottai, Eric Price, Inderjit S. Dhillon, Sujay Sanghavi</i>	
A New Perspective on the Effects of Spectrum in Graph Neural Networks .....	25250
<i>Mingqi Yang, Yanming Shen, Rui Li, Heng Qi, Qiang Zhang, Baocai Yin</i>	
Fourier Learning with Cyclical Data .....	25269
<i>Yingxiang Yang, Zhihan Xiong, Tianyi Liu, Taiqing Wang, Chong Wang</i>	
Estimating Instance-Dependent Bayes-Label Transition Matrix Using a Deep Neural Network .....	25291
<i>Shuo Yang, Erkun Yang, Bo Han, Yang Liu, Min Xu, Gang Niu, Tongliang Liu</i>	
A Study of Face Obfuscation in ImageNet .....	25302
<i>Kaiyu Yang, Jacqueline H. Yau, Li Fei-Fei, Jia Deng, Olga Russakovsky</i>	
Anarchic Federated Learning .....	25320
<i>Haibo Yang, Xin Zhang, Prashant Khanduri, Jia Liu</i>	
Identity-Disentangled Adversarial Augmentation for Self-Supervised Learning .....	25353
<i>Kaiwen Yang, Tianyi Zhou, Xinmei Tian, Dacheng Tao</i>	
Learning from a Learning User for Optimal Recommendations .....	25371
<i>Fan Yao, Chuanhao Li, Denis Nekipelov, Hongning Wang, Haifeng Xu</i>	
Improving Out-Of-Distribution Robustness Via Selective Augmentation .....	25396
<i>Huaxiu Yao, Yu Wang, Sai Li, Linjun Zhang, Weixin Liang, James Zou, Chelsea Finn</i>	
NLP from Scratch Without Large-Scale Pretraining: A Simple and Efficient Framework .....	25427
<i>Xingcheng Yao, Yanan Zheng, Xiaocong Yang, Zhilin Yang</i>	
Feature Space Particle Inference for Neural Network Ensembles .....	25441
<i>Shingo Yashima, Tepei Suzuki, Kohta Ishikawa, Ikuro Sato, Rei Kawakami</i>	
Centroid Approximation for Bootstrap: Improving Particle Quality at Inference .....	25458
<i>Mao Ye, Qiang Liu</i>	

Be Like Water: Adaptive Floating Point for Machine Learning .....	25479
<i>Thomas Yeh, Max Sterner, Zerlina Lai, Brandon Chuang, Alexander Ihler</i>	
QSFL: A Two-Level Uplink Communication Optimization Framework for Federated Learning .....	25490
<i>Liping Yi, Wang Gang, Liu Xiaoguang</i>	
De Novo Mass Spectrometry Peptide Sequencing with a Transformer Model .....	25503
<i>Melih Yilmaz, William Fondrie, Wout Bittremieux, Sewoong Oh, William S Noble</i>	
Bayesian Nonparametric Learning for Point Processes with Spatial Homogeneity: A Spatial Analysis of NBA Shot Locations.....	25512
<i>Fan Yin, Jieying Jiao, Jun Yan, Guanyu Hu</i>	
Bitwidth Heterogeneous Federated Learning with Progressive Weight Dequantization .....	25541
<i>Jaehong Yoon, Geon Park, Wonyong Jeong, Sung Ju Hwang</i>	
ShiftAddNAS: Hardware-Inspired Search for More Accurate and Efficient Neural Networks .....	25555
<i>Haoran You, Baopu Li, Shi Huihong, Yonggan Fu, Yingyan Lin</i>	
Molecular Representation Learning Via Heterogeneous Motif Graph Neural Networks .....	25570
<i>Zhaoning Yu, Hongyang Gao</i>	
Understanding Robust Overfitting of Adversarial Training and Beyond .....	25584
<i>Chaojian Yu, Bo Han, Li Shen, Jun Yu, Chen Gong, Mingming Gong, Tongliang Liu</i>	
How to Leverage Unlabeled Data in Offline Reinforcement Learning .....	25600
<i>Tianhe Yu, Aviral Kumar, Yevgen Chebotar, Karol Hausman, Chelsea Finn, Sergey Levine</i>	
Reachability Constrained Reinforcement Learning.....	25625
<i>Dongjie Yu, Haitong Ma, Shengbo Li, Jianyu Chen</i>	
Topology-Aware Network Pruning Using Multi-Stage Graph Embedding and Reinforcement Learning .....	25645
<i>Sixing Yu, Arya Mazaheri, Ali Jannesari</i>	
The Combinatorial Brain Surgeon: Pruning Weights that Cancel One Another in Neural Networks.....	25657
<i>Xin Yu, Thiago Serra, Srikumar Ramalingam, Shandian Zhe</i>	
GraphFM: Improving Large-Scale GNN Training Via Feature Momentum .....	25673
<i>Haiyang Yu, Limei Wang, Bokun Wang, Meng Liu, Tianbao Yang, Shuiwang Ji</i>	
Latent Diffusion Energy-Based Model for Interpretable Text Modelling .....	25691
<i>Peiyu Yu, Sirui Xie, Xiaojian Ma, Baoxiong Jia, Bo Pang, Ruiqi Gao, Yixin Zhu, Song-Chun Zhu, Ying Nian Wu</i>	
Predicting Out-Of-Distribution Error with the Projection Norm.....	25710
<i>Yaodong Yu, Zitong Yang, Alexander Wei, Yi Ma, Jacob Steinhardt</i>	
Robust Task Representations for Offline Meta-Reinforcement Learning Via Contrastive Learning .....	25736
<i>Haoqi Yuan, Zongqing Lu</i>	
Provable Stochastic Optimization for Global Contrastive Learning: Small Batch Does Not Harm Performance.....	25749
<i>Zhuoning Yuan, Yuexin Wu, Zi-Hao Qiu, Xianzhi Du, Lijun Zhang, Denny Zhou, Tianbao Yang</i>	
Neural Tangent Kernel Empowered Federated Learning .....	25772
<i>Kai Yue, Richeng Jin, Ryan Pilgrim, Chau-Wai Wong, Dror Baron, Huaiyu Dai</i>	
Time is MattEr: Temporal Self-Supervision for Video Transformers.....	25793
<i>Sukmin Yun, Jaehyung Kim, Dongyoon Han, Hwanjun Song, Jung-Woo Ha, Jinwoo Shin</i>	
Pure Noise to the Rescue of Insufficient Data: Improving Imbalanced Classification by Training on Random Noise Images.....	25806
<i>Shiran Zada, Itay Benou, Michal Irani</i>	
Adaptive Conformal Predictions for Time Series.....	25823
<i>Margaux Zaffran, Olivier Feron, Yannig Goude, Julie Josse, Aymeric Dieuleveut</i>	
Actor-Critic Based Improper Reinforcement Learning .....	25856
<i>Mohammadi Zaki, Avi Mohan, Aditya Gopalan, Shie Mannor</i>	
Stabilizing Q-Learning with Linear Architectures for Provable Efficient Learning.....	25909
<i>Andrea Zanette, Martin Wainwright</i>	

Multi Resolution Analysis (MRA) for Approximate Self-Attention .....	25944
<i>Zhanpeng Zeng, Sourav Pal, Jeffery Kline, Glenn M Fung, Vikas Singh</i>	
Efficient PAC Learning from the Crowd with Pairwise Comparisons .....	25962
<i>Shiwei Zeng, Jie Shen</i>	
Multi-Grained Vision Language Pre-Training: Aligning Texts with Visual Concepts.....	25983
<i>Yan Zeng, Xinsong Zhang, Hang Li</i>	
Position Prediction as an Effective Pretraining Strategy .....	25999
<i>Shuangfei Zhai, Navdeep Jaitly, Jason Ramapuram, Dan Busbridge, Tatiana Likhomanenko, Joseph Y Cheng, Walter Talbott, Chen Huang, Hanlin Goh, Joshua M Susskind</i>	
Anytime Information Cascade Popularity Prediction Via Self-Exciting Processes.....	26017
<i>Xi Zhang, Akshay Aravamudan, Georgios C Anagnostopoulos</i>	

## PART 32

Understanding Clipping for Federated Learning: Convergence and Client-Level Differential Privacy.....	26037
<i>Xinwei Zhang, Xiangyi Chen, Mingyi Hong, Steven Wu, Jinfeng Yi</i>	
Collaboration of Experts: Achieving 80% Top-1 Accuracy on ImageNet with 100M FLOPs .....	26057
<i>Yikang Zhang, Zhuo Chen, Zhao Zhong</i>	
PDE-Based Optimal Strategy for Unconstrained Online Learning .....	26074
<i>Zhiyu Zhang, Ashok Cutkosky, Ioannis Paschalidis</i>	
Stochastic Continuous Submodular Maximization: Boosting Via Non-Oblivious Function.....	26105
<i>Qixin Zhang, Zengde Deng, Zaiyi Chen, Haoyuan Hu, Yu Yang</i>	
When and How Mixup Improves Calibration .....	26124
<i>Linjun Zhang, Zhun Deng, Kenji Kawaguchi, James Zou</i>	
UAST: Uncertainty-Aware Siamese Tracking.....	26150
<i>Dawei Zhang, Yanwei Fu, Zhonglong Zheng</i>	
Examining Scaling and Transfer of Language Model Architectures for Machine Translation.....	26165
<i>Biao Zhang, Behrooz Ghorbani, Ankur Bapna, Yong Cheng, Xavier Garcia, Jonathan Shen, Orhan Firat</i>	
Revisiting End-To-End Speech-To-Text Translation from Scratch .....	26182
<i>Biao Zhang, Barry Haddow, Rico Sennrich</i>	
A Stochastic Multi-Rate Control Framework for Modeling Distributed Optimization Algorithms .....	26195
<i>Xinwei Zhang, Mingyi Hong, Sairaj Dhople, Nicola Elia</i>	
GALAXY: Graph-Based Active Learning at the Extreme.....	26212
<i>Jifan Zhang, Julian Katz-Samuels, Robert Nowak</i>	
Fairness Interventions as (Dis)Incentives for Strategic Manipulation.....	26228
<i>Xueru Zhang, Mohammad Mahdi Khalili, Kun Jin, Parinaz Naghizadeh, Mingyan Liu</i>	
Role-Based Multiplex Network Embedding .....	26254
<i>Hegui Zhang, Gang Kou</i>	
Dynamic Topic Models for Temporal Document Networks.....	26270
<i>Delvin Ce Zhang, Hady Lauw</i>	
Personalized Federated Learning Via Variational Bayesian Inference .....	26282
<i>Xu Zhang, Yinchuan Li, Wenpeng Li, Kaiyang Guo, Yunfeng Shao</i>	
Federated Learning with Label Distribution Skew Via Logits Calibration .....	26300
<i>Jie Zhang, Zhiqi Li, Bo Li, Jianghe Xu, Shuang Wu, Shouhong Ding, Chao Wu</i>	
Neural Network Weights Do Not Converge to Stationary Points: An Invariant Measure Perspective.....	26319
<i>Jingzhao Zhang, Haochuan Li, Suvrit Sra, Ali Jadbabaie</i>	
Beyond Worst-Case Analysis in Stochastic Approximation: Moment Estimation Improves Instance Complexity .....	26336
<i>Jingzhao Zhang, Hongzhou Lin, Subhro Das, Suvrit Sra, Ali Jadbabaie</i>	



Deep and Flexible Graph Neural Architecture Search.....	26351
<i>Wentao Zhang, Zheyu Lin, Yu Shen, Yang Li, Zhi Yang, Bin Cui</i>	
A Langevin-Like Sampler for Discrete Distributions.....	26364
<i>Ruqi Zhang, Xingchao Liu, Qiang Liu</i>	
Rich Feature Construction for the Optimization-Generalization Dilemma.....	26386
<i>Jianguo Zhang, David Lopez-Paz, Leon Bottou</i>	
Generative Flow Networks for Discrete Probabilistic Modeling.....	26401
<i>Dinghuai Zhang, Nikolay Malkin, Zhen Liu, Alexandra Volokhova, Aaron Courville, Yoshua Bengio</i>	
Neurotoxin: Durable Backdoors in Federated Learning.....	26418
<i>Zhengming Zhang, Ashwinee Panda, Linyue Song, Yaoqing Yang, Michael Mahoney, Prateek Mittal, Ramchandran Kannan, Joseph Gonzalez</i>	
Making Linear MDPs Practical Via Contrastive Representation Learning.....	26436
<i>Tianjun Zhang, Tongzheng Ren, Mengjiao Yang, Joseph Gonzalez, Dale Schuurmans, Bo Dai</i>	
NAFS: A Simple Yet Tough-To-Beat Baseline for Graph Representation Learning.....	26456
<i>Wentao Zhang, Zeang Sheng, Mingyu Yang, Yang Li, Yu Shen, Zhi Yang, Bin Cui</i>	
Correct-N-Contrast: A Contrastive Approach for Improving Robustness to Spurious Correlations.....	26473
<i>Michael Zhang, Nimit S Sohoni, Hongyang R Zhang, Chelsea Finn, Christopher Re</i>	
Efficient Reinforcement Learning in Block MDPs: A Model-Free Representation Learning Approach.....	26506
<i>Xuezhou Zhang, Yuda Song, Masatoshi Uehara, Mengdi Wang, Alekh Agarwal, Wen Sun</i>	
Partial Counterfactual Identification from Observational and Experimental Data.....	26537
<i>Junzhe Zhang, Jin Tian, Elias Bareinboim</i>	
Set Norm and Equivariant Skip Connections: Putting the Deep in Deep Sets.....	26548
<i>Lily Zhang, Veronica Tozzo, John Higgins, Rajesh Ranganath</i>	
Learning to Estimate and Refine Fluid Motion with Physical Dynamics.....	26564
<i>Mingrui Zhang, Jianhong Wang, James B Thomole, Matthew Piggott</i>	
A Branch and Bound Framework for Stronger Adversarial Attacks of ReLU Networks.....	26580
<i>Huan Zhang, Shiqi Wang, Kaidi Xu, Yihan Wang, Suman Jana, Cho-Jui Hsieh, Zico Kolter</i>	
A Simple Yet Universal Strategy for Online Convex Optimization.....	26594
<i>Lijun Zhang, Guanghui Wang, Jinfeng Yi, Tianbao Yang</i>	
Low-Precision Stochastic Gradient Langevin Dynamics.....	26613
<i>Ruqi Zhang, Andrew Gordon Wilson, Christopher De Sa</i>	
Expression Might Be Enough: Representing Pressure and Demand for Reinforcement Learning Based Traffic Signal Control.....	26634
<i>Liang Zhang, Qiang Wu, Jun Shen, Linyuan Lü, Bo Du, Jianqing Wu</i>	
Uncertainty Modeling in Generative Compressed Sensing.....	26644
<i>Yilang Zhang, Mengchu Xu, Xiaojun Mao, Jian Wang</i>	
Building Robust Ensembles Via Margin Boosting.....	26658
<i>Dinghuai Zhang, Hongyang Zhang, Aaron Courville, Yoshua Bengio, Pradeep Ravikumar, Arun Sai Suggala</i>	
Revisiting and Advancing Fast Adversarial Training Through the Lens of Bi-Level Optimization.....	26682
<i>Yihua Zhang, Guanhua Zhang, Prashant Khanduri, Mingyi Hong, Shiyu Chang, Sijia Liu</i>	
Off-Policy Fitted Q-Evaluation with Differentiable Function Approximators: Z-Estimation and Inference Theory.....	26702
<i>Ruiqi Zhang, Xuezhou Zhang, Chengzhuo Ni, Mengdi Wang</i>	
ROCK: Causal Inference Principles for Reasoning About Commonsense Causality.....	26739
<i>Jiayao Zhang, Hongming Zhang, Weijie Su, Dan Roth</i>	
No-Regret Learning in Time-Varying Zero-Sum Games.....	26761
<i>Mengxiao Zhang, Peng Zhao, Haipeng Luo, Zhi-Hua Zhou</i>	

PLATON: Pruning Large Transformer Models with Upper Confidence Bound of Weight Importance.....	26798
<i>Qingru Zhang, Simiao Zuo, Chen Liang, Alexander Bukharin, Pengcheng He, Weizhu Chen, Tuo Zhao</i>	
NysADMM: Faster Composite Convex Optimization Via Low-Rank Approximation.....	26813
<i>Shipu Zhao, Zachary Frangella, Madeleine Udell</i>	
Toward Compositional Generalization in Object-Oriented World Modeling.....	26830
<i>Linfeng Zhao, Lingzhi Kong, Robin Walters, Lawson L. S. Wong</i>	
Dynamic Regret of Online Markov Decision Processes.....	26854
<i>Peng Zhao, Long-Fei Li, Zhi-Hua Zhou</i>	

### PART 33

Learning to Solve PDE-Constrained Inverse Problems with Graph Networks.....	26884
<i>Qingqing Zhao, David B Lindell, Gordon Wetzstein</i>	
Learning from Counterfactual Links for Link Prediction.....	26900
<i>Tong Zhao, Gang Liu, Daheng Wang, Wenhao Yu, Meng Jiang</i>	
Global Optimization Networks.....	26916
<i>Sen Zhao, Erez Louidor, Maya Gupta</i>	
Certified Robustness Against Natural Language Attacks by Causal Intervention.....	26947
<i>Haiteng Zhao, Chang Ma, Xinshuai Dong, Anh Tuan Luu, Zhi-Hong Deng, Hanwang Zhang</i>	
Efficient Learning for AlphaZero Via Path Consistency.....	26960
<i>Dengwei Zhao, Shikui Tu, Lei Xu</i>	
Penalizing Gradient Norm for Efficiently Improving Generalization in Deep Learning.....	26971
<i>Yang Zhao, Hao Zhang, Xiuyuan Hu</i>	
Ripple Attention for Visual Perception with Sub-Quadratic Complexity.....	26982
<i>Lin Zheng, Huijie Pan, Lingpeng Kong</i>	
Linear Complexity Randomized Self-Attention Mechanism.....	27000
<i>Lin Zheng, Chong Wang, Lingpeng Kong</i>	
Online Decision Transformer.....	27031
<i>Qinqing Zheng, Amy Zhang, Aditya Grover</i>	
Learning Efficient and Robust Ordinary Differential Equations Via Invertible Neural Networks.....	27049
<i>Weiming Zhi, Tin Lai, Lionel Ott, Edwin V. Bonilla, Fabio Ramos</i>	
HyperTransformer: Model Generation for Supervised and Semi-Supervised Few-Shot Learning.....	27064
<i>Andrey Zhmoginov, Mark Sandler, Maksym Vladymyrov</i>	
Describing Differences Between Text Distributions with Natural Language.....	27088
<i>Ruiqi Zhong, Charlie Snell, Dan Klein, Jacob Steinhardt</i>	
Pessimistic Minimax Value Iteration: Provably Efficient Equilibrium Learning from Offline Datasets.....	27106
<i>Han Zhong, Wei Xiong, Jiyuan Tan, Liwei Wang, Tong Zhang, Zhaoran Wang, Zhuoran Yang</i>	
Dimension-Free Complexity Bounds for High-Order Nonconvex Finite-Sum Optimization.....	27132
<i>Dongruo Zhou, Quanquan Gu</i>	
A Hierarchical Bayesian Approach to Inverse Reinforcement Learning with Symbolic Reward Machines.....	27148
<i>Weichao Zhou, Wenchao Li</i>	
On the Optimization Landscape of Neural Collapse Under MSE Loss: Global Optimality with Unconstrained Features.....	27168
<i>Jinxin Zhou, Xiao Li, Tianyu Ding, Chong You, Qing Qu, Zhihui Zhu</i>	
Model Agnostic Sample Reweighting for Out-Of-Distribution Learning.....	27192
<i>Xiao Zhou, Yong Lin, Renjie Pi, Weizhong Zhang, Renzhe Xu, Peng Cui, Tong Zhang</i>	
Sparse Invariant Risk Minimization.....	27211
<i>Xiao Zhou, Yong Lin, Weizhong Zhang, Tong Zhang</i>	

Prototype-Anchored Learning for Learning with Imperfect Annotations.....	27234
<i>Xiong Zhou, Xianming Liu, Deming Zhai, Junjun Jiang, Xin Gao, Xiangyang Ji</i>	
FEDformer: Frequency Enhanced Decomposed Transformer for Long-Term Series Forecasting.....	27257
<i>Tian Zhou, Ziqing Ma, Qingsong Wen, Xue Wang, Liang Sun, Rong Jin</i>	
Probabilistic Bilevel Coreset Selection .....	27276
<i>Xiao Zhou, Renjie Pi, Weizhong Zhang, Yong Lin, Zonghao Chen, Tong Zhang</i>	
Approximate Frank-Wolfe Algorithms Over Graph-Structured Support Sets .....	27292
<i>Baojian Zhou, Yifan Sun</i>	
Improving Adversarial Robustness Via Mutual Information Estimation.....	27327
<i>Dawei Zhou, Nannan Wang, Xinbo Gao, Bo Han, Xiaoyu Wang, Yibing Zhan, Tongliang Liu</i>	
Modeling Adversarial Noise for Adversarial Training .....	27342
<i>Dawei Zhou, Nannan Wang, Bo Han, Tongliang Liu</i>	
Contrastive Learning with Boosted Memorization.....	27356
<i>Zhihan Zhou, Jiangchao Yao, Yan-Feng Wang, Bo Han, Ya Zhang</i>	
Understanding the Robustness in Vision Transformers .....	27367
<i>Daquan Zhou, Zhiding Yu, Enze Xie, Chaowei Xiao, Animashree Anandkumar, Jiashi Feng, Jose M. Alvarez</i>	
VLUE: A Multi-Task Multi-Dimension Benchmark for Evaluating Vision-Language Pre-Training .....	27384
<i>Wangchunshu Zhou, Yan Zeng, Shizhe Diao, Xinsong Zhang</i>	
Detecting Corrupted Labels Without Training a Model to Predict .....	27401
<i>Zhaowei Zhu, Zihao Dong, Yang Liu</i>	
Contextual Bandits with Large Action Spaces: Made Practical .....	27417
<i>Yinglun Zhu, Dylan J Foster, John Langford, Paul Mineiro</i>	
Neural-Symbolic Models for Logical Queries on Knowledge Graphs.....	27443
<i>Zhaocheng Zhu, Mikhail Galkin, Zuobai Zhang, Jian Tang</i>	
Topology-Aware Generalization of Decentralized SGD.....	27468
<i>Tongtian Zhu, Fengxiang He, Lan Zhang, Zhengyang Niu, Mingli Song, Dacheng Tao</i>	
Resilient and Communication Efficient Learning for Heterogeneous Federated Systems .....	27493
<i>Zhuangdi Zhu, Junyuan Hong, Steve Drew, Jiayu Zhou</i>	
On Numerical Integration in Neural Ordinary Differential Equations .....	27516
<i>Aiqing Zhu, Pengzhan Jin, Beibei Zhu, Yifa Tang</i>	
When AUC Meets DRO: Optimizing Partial AUC for Deep Learning with Non-Convex Convergence Guarantee.....	27537
<i>Dixian Zhu, Gang Li, Bokun Wang, Xiaodong Wu, Tianbao Yang</i>	
Contextual Bandits with Smooth Regret: Efficient Learning in Continuous Action Spaces .....	27563
<i>Yinglun Zhu, Paul Mineiro</i>	
Residual-Based Sampling for Online Outlier-Robust PCA.....	27580
<i>Tianhao Zhu, Jie Shen</i>	
Region-Based Semantic Factorization in GANs .....	27601
<i>Jiapeng Zhu, Yujun Shen, Yinghao Xu, Deli Zhao, Qifeng Chen</i>	
Beyond Images: Label Noise Transition Matrix Estimation for Tasks with Lower-Quality Features .....	27622
<i>Zhaowei Zhu, Jialu Wang, Yang Liu</i>	
Towards Uniformly Superhuman Autonomy Via Subdominance Minimization .....	27643
<i>Brian Ziebart, Sanjiban Choudhury, Xinyan Yan, Paul Vernaza</i>	
Inductive Matrix Completion: No Bad Local Minima and a Fast Algorithm.....	27660
<i>Pini Zilber, Boaz Nadler</i>	
Counterfactual Prediction for Outcome-Oriented Treatments.....	27682
<i>Hao Zou, Bo Li, Jiangan Han, Shuiping Chen, Xuetao Ding, Peng Cui</i>	
SpaceMAP: Visualizing High-Dimensional Data by Space Expansion .....	27696
<i>Xinrui Zu, Qian Tao</i>	

**Author Index**