

10th International Conference on Learning Representations

(ICLR 2022)

25-29 April 2022, Online

ISBN: 979-8-3313-0009-8

Title and Authors	Page#
<i>Improving Federated Learning Face Recognition Via Privacy-Agnostic Clusters</i> Qiang Meng, Feng Zhou, Hainan Ren, Tianshu Feng, Guochao Liu, Yuanqing Lin	1
<i>Probabilistic Implicit Scene Completion</i> Dongsu Zhang, Changwoon Choi, Inbum Park, Young Min Kim	22
<i>Learning Efficient Image Super-Resolution Networks Via Structure-Regularized Pruning</i> Yulun Zhang, Huan Wang, Can Qin, Yun Fu	54
<i>Towards Building a Group-Based Unsupervised Representation Disentanglement Framework</i> Tao Yang, Xuanchi Ren, Yuwang Wang, Wenjun Zeng, Nanning Zheng	66
<i>Learning Disentangled Representation by Exploiting Pretrained Generative Models: A Contrastive Learning View</i> Xuanchi Ren, Tao Yang, Yuwang Wang, Wenjun Zeng	107
<i>Clean Images Are Hard to Reblur: Exploiting the Ill-Posed Inverse Task for Dynamic Scene Deblurring</i> Seungjun Nah, Sanghyun Son, Jaerin Lee, Kyoung Mu Lee	147
<i>Multi-Critic Actor Learning: Teaching RL Policies to Act with Style</i> Siddharth Mysore, George Cheng, Yunqi Zhao, Kate Saenko, Meng Wu	166
<i>Incremental False Negative Detection for Contrastive Learning</i> Tsai-Shien Chen, Wei-Chih Hung, Hung-Yu Tseng, Shao-Yi Chien, Ming-Hsuan Yang	187
<i>Who is Your Right Mixup Partner in Positive and Unlabeled Learning</i> Changchun Li, Ximing Li, Lei Feng, Jihong Ouyang	205
<i>Learning Synthetic Environments and Reward Networks for Reinforcement Learning</i> Fabio Ferreira, Thomas Nierhoff, Andreas S��rlinger, Frank Hutter	219
<i>Learning the Dynamics of Physical Systems from Sparse Observations with Finite Element Networks</i> Marten Lienen, Stephan G��nnemann	256
<i>Representation-Agnostic Shape Fields</i> Xiaoyang Huang, Jiancheng Yang, Yanjun Wang, Ziyu Chen, Linguo Li, Teng Li, Bingbing Ni, Wenjun Zhang	277

Feature Kernel Distillation Bobby He, Mete Ozay	295
BiBERT: Accurate Fully Binarized BERT Haotong Qin, Yifu Ding, Mingyuan Zhang, Qinghua Yan, Aishan Liu, Qingqing Dang, Ziwei Liu, Xianglong Liu	333
ViDT: An Efficient and Effective Fully Transformer-Based Object Detector Hwanjun Song, Deqing Sun, Sanghyuk Chun, Varun Jampani, Dongyoon Han, Byeongho Heo, Wonjae Kim, Ming-Hsuan Yang	357
Boosted Curriculum Reinforcement Learning Pascal Klink, Carlo D'Eramo, Jan Peters, Joni Pajarinen	375
Increasing the Cost of Model Extraction with Calibrated Proof of Work Adam Dziedzic, Muhammad Ahmad Kaleem, Yu Shen Lu, Nicolas Papernot	397
RelViT: Concept-Guided Vision Transformer for Visual Relational Reasoning Xiaojuan Ma, Weili Nie, Zhiding Yu, Huaizu Jiang, Chaowei Xiao, Yuke Zhu, Song-Chun Zhu, Anima Anandkumar	426
MAML is a Noisy Contrastive Learner in Classification Chia Hsiang Kao, Wei-Chen Chiu, Pin-Yu Chen	445
Memory Replay with Data Compression for Continual Learning Liyuan Wang, Xingxing Zhang, Kuo Yang, Longhui Yu, Chongxuan Li, Lanqing Hong, Shifeng Zhang, Zhenguo Li, Yi Zhong, Jun Zhu	467
Language Modeling Via Stochastic Processes Rose E Wang, Esin Durmus, Noah Goodman, Tatsunori Hashimoto	492
Denoising Likelihood Score Matching for Conditional Score-Based Data Generation Chen-Hao Chao, Wei-Fang Sun, Bo-Wun Cheng, Yi-Chen Lo, Chia-Che Chang, Yu-Lun Liu, Yu-Lin Chang, Chia-Ping Chen, Chun-Yi Lee	528
Learning Features with Parameter-Free Layers Dongyoon Han, Youngjoon Yoo, Beomyoung Kim, Byeongho Heo	552
Learning Discrete Structured Variational Auto-Encoder Using Natural Evolution Strategies Alon Berliner, Guy Rotman, Yossi Adi, Roi Reichart, Tamir Hazan	571
Transformers Can Do Bayesian Inference Samuel Maller, Noah Hollmann, Sebastian Pineda Arango, Josif Grabocka, Frank Hutter	589
Rethinking Adversarial Transferability from a Data Distribution Perspective Yao Zhu, Jiacheng Sun, Zhenguo Li	612
Global Convergence of Multi-Agent Policy Gradient in Markov Potential Games Stefanos Leonardos, Will Overman, Ioannis Panageas, Georgios Piliouras	635

Random Matrices in Service of ML Footprint: Ternary Random Features with No Performance Loss <i>Hafiz Tiomoko Ali, Zhenyu Liao, Romain Couillet</i>	664
Zero-CL: Instance and Feature Decorrelation for Negative-Free Symmetric Contrastive Learning <i>Shaofeng Zhang, Feng Zhu, Junchi Yan, Rui Zhao, Xiaokang Yang</i>	690
Switch to Generalize: Domain-Switch Learning for Cross-Domain Few-Shot Classification <i>Zhengdong Hu, Yifan Sun, Yi Yang</i>	703
Online Coreset Selection for Rehearsal-Based Continual Learning <i>Jaehong Yoon, Divyam Madaan, Eunho Yang, Sung Ju Hwang</i>	718
GradMax: Growing Neural Networks Using Gradient Information <i>Utku Evci, Bart Van Merriënboer, Thomas Unterthiner, Fabian Pedregosa, Max Vladymyrov</i>	734
Measuring the Interpretability of Unsupervised Representations Via Quantized Reversed Probing <i>Iro Laina, Yuki M Asano, Andrea Vedaldi</i>	751
SketchODE: Learning Neural Sketch Representation in Continuous Time <i>Ayan Das, Yongxin Yang, Timothy Hospedales, Tao Xiang, Yi-Zhe Song</i>	777
Learning to Schedule Learning Rate with Graph Neural Networks <i>Yuanhao Xiong, Li-Cheng Lan, Xiangning Chen, Ruochen Wang, Cho-Jui Hsieh</i>	793
Beyond ImageNet Attack: Towards Crafting Adversarial Examples for Black-Box Domains <i>Qilong Zhang, Xiaodan Li, Yuefeng Chen, Jingkuan Song, Lianli Gao, Yuan He, Hui Xue'</i>	814
CDTrans: Cross-Domain Transformer for Unsupervised Domain Adaptation <i>Tongkun Xu, Weihua Chen, Pichao Wang, Fan Wang, Hao Li, Rong Jin</i>	832
Variational Oracle Guiding for Reinforcement Learning <i>Dongqi Han, Tadashi Kozuno, Xufang Luo, Zhao-Yun Chen, Kenji Doya, Yuqing Yang, Dongsheng Li</i>	846
Machine Learning for Elliptic PDEs: Fast Rate Generalization Bound, Neural Scaling Law and Minimax Optimality <i>Yiping Lu, Haoxuan Chen, Jianfeng Lu, Lexing Ying, Jose Blanchet</i>	868
NAS-Bench-Suite: NAS Evaluation is (Now) Surprisingly Easy <i>Yash Mehta, Colin White, Arber Zela, Arjun Krishnakumar, Guri Zabergja, Shakiba Moradian, Mahmoud Safari, Kaicheng Yu, Frank Hutter</i>	933
The Information Geometry of Unsupervised Reinforcement Learning <i>Benjamin Eysenbach, Ruslan Salakhutdinov, Sergey Levine</i>	957
C-Planning: An Automatic Curriculum for Learning Goal-Reaching Tasks <i>Tianjun Zhang, Benjamin Eysenbach, Ruslan Salakhutdinov, Sergey Levine, Joseph E. Gonzalez</i>	975

<p>Hybrid Local SGD for Federated Learning with Heterogeneous Communications</p> <p>Yuanxiong Guo, Ying Sun, Rui Hu, Yanmin Gong</p>	996
<p>Neural Graphical Modelling in Continuous-Time: Consistency Guarantees and Algorithms</p> <p>Alexis Bellot, Kim Branson, Mihaela Van Der Schaar</p>	1038
<p>VAT-Mart: Learning Visual Action Trajectory Proposals for Manipulating 3D ARTiculated Objects</p> <p>Ruihai Wu, Yan Zhao, Kaichun Mo, Zizheng Guo, Yian Wang, Tianhao Wu, Qingnan Fan, Xuelin Chen, Leonidas Guibas, Hao Dong</p>	1066
<p>IFR-Explore: Learning Inter-Object Functional Relationships in 3D Indoor Scenes</p> <p>Qi Li, Kaichun Mo, Yanchao Yang, Hang Zhao, Leonidas Guibas</p>	1089
<p>GNN is a Counter? Revisiting GNN for Question Answering</p> <p>Kuan Wang, Yuyu Zhang, Diyi Yang, Le Song, Tao Qin</p>	1109
<p>SGD Can Converge to Local Maxima</p> <p>Liu Ziyin, Botao Li, James B Simon, Masahito Ueda</p>	1128
<p>Dual Lottery Ticket Hypothesis</p> <p>Yue Bai, Huan Wang, Zhiqiang Tao, Kunpeng Li, Yun Fu</p>	1153
<p>Entroformer: A Transformer-Based Entropy Model for Learned Image Compression</p> <p>Yichen Qian, Xiuyu Sun, Ming Lin, Zhiyu Tan, Rong Jin</p>	1169
<p>On the Importance of Difficulty Calibration in Membership Inference Attacks</p> <p>Lauren Watson, Chuan Guo, Graham Cormode, Alexandre Sablayrolles</p>	1184
<p>Equivariant Transformers for Neural Network Based Molecular Potentials</p> <p>Philipp Thälke, Gianni De Fabritiis</p>	1200
<p>Know Your Action Set: Learning Action Relations for Reinforcement Learning</p> <p>Ayush Jain, Norio Kosaka, Kyung-Min Kim, Joseph J Lim</p>	1220
<p>RISP: Rendering-Invariant State Predictor with Differentiable Simulation and Rendering for Cross-Domain Parameter Estimation</p> <p>Pingchuan Ma, Tao Du, Joshua B. Tenenbaum, Wojciech Matusik, Chuang Gan</p>	1248
<p>Shuffle Private Stochastic Convex Optimization</p> <p>Albert Cheu, Matthew Joseph, Jieming Mao, Binghui Peng</p>	1266
<p>Label Encoding for Regression Networks</p> <p>Deval Shah, Zi Yu Xue, Tor Aamodt</p>	1294
<p>InfinityGAN: Towards Infinite-Pixel Image Synthesis</p> <p>Chieh Hubert Lin, Hsin-Ying Lee, Yen-Chi Cheng, Sergey Tulyakov, Ming-Hsuan Yang</p>	1329

<p>Analytic-DPM: An Analytic Estimate of the Optimal Reverse Variance in Diffusion Probabilistic Models</p> <p>Fan Bao, Chongxuan Li, Jun Zhu, Bo Zhang</p>	1372
<p>Skill-Based Meta-Reinforcement Learning</p> <p>Taewook Nam, Shao-Hua Sun, Karl Pertsch, Sung Ju Hwang, Joseph J Lim</p>	1411
<p>CycleMLP: A MLP-Like Architecture for Dense Prediction</p> <p>Shoufa Chen, Enze Xie, Chongjian Ge, Runjian Chen, Ding Liang, Ping Luo</p>	1433
<p>Chemical-Reaction-Aware Molecule Representation Learning</p> <p>Hongwei Wang, Weijiang Li, Xiaomeng Jin, Kyunghyun Cho, Heng Ji, Jiawei Han, Martin D. Burke</p>	1454
<p>Understanding the Role of Self Attention for Efficient Speech Recognition</p> <p>Kyuhong Shim, Jungwook Choi, Wonyong Sung</p>	1474
<p>Visual Representation Learning Over Latent Domains</p> <p>Lucas Deecke, Timothy Hospedales, Hakan Bilen</p>	1493
<p>Hot-Refresh Model Upgrades with Regression-Free Compatible Training in Image Retrieval</p> <p>Binjie Zhang, Yixiao Ge, Yantao Shen, Yu Li, Chun Yuan, Xuyuan Xu, Yexin Wang, Ying Shan</p>	1511
<p>Adversarially Robust Conformal Prediction</p> <p>Asaf Gendler, Tsui-Wei Weng, Luca Daniel, Yaniv Romano</p>	1532
<p>CrossFormer: A Versatile Vision Transformer Hinging on Cross-Scale Attention</p> <p>Wenxiao Wang, Lu Yao, Long Chen, Binbin Lin, Deng Cai, Xiaofei He, Wei Liu</p>	1555
<p>Neural Markov Controlled SDE: Stochastic Optimization for Continuous-Time Data</p> <p>Sung Woo Park, Kyungjae Lee, Junseok Kwon</p>	1571
<p>Retriever: Learning Content-Style Representation as a Token-Level Bipartite Graph</p> <p>Dacheng Yin, Xuanchi Ren, Chong Luo, Yuwang Wang, Zhiwei Xiong, Wenjun Zeng</p>	1593
<p>Byzantine-Robust Learning on Heterogeneous Datasets Via Bucketing</p> <p>Sai Praneeth Karimireddy, Lie He, Martin Jaggi</p>	1620
<p>Promoting Saliency from Depth: Deep Unsupervised RGB-D Saliency Detection</p> <p>Wei Ji, Jingjing Li, Qi Bi, Chuan Guo, Jie Liu, Li Cheng</p>	1633
<p>β-Intact-VAE: Identifying and Estimating Causal Effects Under Limited Overlap</p> <p>Pengzhou Abel Wu, Kenji Fukumizu</p>	1653
<p>Deep Point Cloud Reconstruction</p> <p>Jaesung Choe, Byeongjin Joung, Francois Rameau, Jaesik Park, In So Kweon</p>	1683
<p>Few-Shot Learning Via Dirichlet Tessellation Ensemble</p> <p>Chunwei Ma, Ziyun Huang, Mingchen Gao, Jinhui Xu</p>	1700

<p>Permutation Compressors for Provably Faster Distributed Nonconvex Optimization</p> <p>Rafał, Szlendak, Alexander Tyurin, Peter Richtárik</p>	1733
<p>OntoProtein: Protein Pretraining with Gene Ontology Embedding</p> <p>Ningyu Zhang, Zhen Bi, Xiaozhuan Liang, Siyuan Cheng, Haosen Hong, Shumin Deng, Qiang Zhang, Jiazhang Lian, Huajun Chen</p>	1786
<p>Differentiable Prompt Makes Pre-Trained Language Models Better Few-Shot Learners</p> <p>Ningyu Zhang, Luoqiu Li, Xiang Chen, Shumin Deng, Zhen Bi, Chuanqi Tan, Fei Huang, Huajun Chen</p>	1804
<p>A Biologically Interpretable Graph Convolutional Network to Link Genetic Risk Pathways and Imaging Phenotypes of Disease</p> <p>Sayan Ghosal, Qiang Chen, Giulio Pergola, Aaron L Goldman, William Ulrich, Daniel R Weinberger, Archana Venkataraman</p>	1823
<p>How to Inject Backdoors with Better Consistency: Logit Anchoring on Clean Data</p> <p>Zhiyuan Zhang, Lingjuan Lyu, Weiqiang Wang, Lichao Sun, Xu Sun</p>	1841
<p>Neural Spectral Marked Point Processes</p> <p>Shixiang Zhu, Haoyun Wang, Zheng Dong, Xiuyuan Cheng, Yao Xie</p>	1863
<p>Real-Time Neural Voice Camouflage</p> <p>Mia Chiquier, Chengzhi Mao, Carl Vondrick</p>	1889
<p>Hyperparameter Tuning with Renyi Differential Privacy</p> <p>Nicolas Papernot, Thomas Steinke</p>	1902
<p>Divergence-Aware Federated Self-Supervised Learning</p> <p>Weiming Zhuang, Yonggang Wen, Shuai Zhang</p>	1930
<p>PAC-Bayes Information Bottleneck</p> <p>Zifeng Wang, Shao-Lun Huang, Ercan Engin Kuruoglu, Jimeng Sun, Xi Chen, Yefeng Zheng</p>	1949
<p>IntSGD: Adaptive Floatless Compression of Stochastic Gradients</p> <p>Konstantin Mishchenko, Bokun Wang, Dmitry Kovalev, Peter Richtárik</p>	1965
<p>Steerable Partial Differential Operators for Equivariant Neural Networks</p> <p>Erik Jenner, Maurice Weiler</p>	1992
<p>Fixed Neural Network Steganography: Train the Images, Not the Network</p> <p>Varsha Kishore, Xiangyu Chen, Yan Wang, Boyi Li, Kilian Q Weinberger</p>	2039
<p>DKM: Differentiable k-Means Clustering Layer for Neural Network Compression</p> <p>Minsik Cho, Keivan Alizadeh-Vahid, Saurabh Adya, Mohammad Rastegari</p>	2061
<p>Closed-Form Sample Probing for Learning Generative Models in Zero-Shot Learning</p> <p>Samet Cetin, Orhun Buğra Baran, Ramazan Gokberk Cinbis</p>	2080

<p>Generalized Demographic Parity for Group Fairness</p> <p>Zhimeng Jiang, Xiaotian Han, Chao Fan, Fan Yang, Ali Mostafavi, Xia Hu</p>	2094
<p>Task Relatedness-Based Generalization Bounds for Meta Learning</p> <p>Jiechao Guan, Zhiwu Lu</p>	2119
<p>Automatic Loss Function Search for Predict-Then-Optimize Problems with Strong Ranking Property</p> <p>Boshi Wang, Jialin Yi, Hang Dong, Bo Qiao, Chuan Luo, Qingwei Lin</p>	2138
<p>Handling Distribution Shifts on Graphs: An Invariance Perspective</p> <p>Qitian Wu, Hengrui Zhang, Junchi Yan, David Wipf</p>	2153
<p>Distilling GANs with Style-Mixed Triplets for X2I Translation with Limited Data</p> <p>Yaxing Wang, Joost Van De Weijer, Lu Yu, Shangling Jui</p>	2183
<p>Towards Better Understanding and Better Generalization of Low-Shot Classification in Histology Images with Contrastive Learning</p> <p>Jiawei Yang, Hanbo Chen, Jiangpeng Yan, Xiaoyu Chen, Jianhua Yao</p>	2219
<p>Neural Variational Dropout Processes</p> <p>Insu Jeon, Youngjin Park, Gunhee Kim</p>	2248
<p>SQuant: On-The-Fly Data-Free Quantization Via Diagonal Hessian Approximation</p> <p>Cong Guo, Yuxian Qiu, Jingwen Leng, Xiaotian Gao, Chen Zhang, Yunxin Liu, Fan Yang, Yuhao Zhu, Minyi Guo</p>	2269
<p>Fast Differentiable Matrix Square Root</p> <p>Yue Song, Nicu Sebe, Wei Wang</p>	2287
<p>Group Equivariant Neural Posterior Estimation</p> <p>Maximilian Dax, Stephen R Green, Jonathan Gair, Michael Deistler, Bernhard Schälkopf, Jakob H. Macke</p>	2306
<p>Trivial Or Impossible --- Dichotomous Data Difficulty Masks Model Differences (on ImageNet and Beyond)</p> <p>Kristof Meding, Luca M. Schulze Buschoff, Robert Geirhos, Felix A. Wichmann</p>	2330
<p>ARTEMIS: Attention-Based Retrieval with Text-Explicit Matching and Implicit Similarity</p> <p>Ginger Delmas, Rafael S. Rezende, Gabriela Csurka, Diane Larlus</p>	2368
<p>Deep Attentive Variational Inference</p> <p>Ifigeneia Apostolopoulou, Ian Char, Elan Rosenfeld, Artur Dubrawski</p>	2392
<p>What's Wrong with Deep Learning in Tree Search for Combinatorial Optimization</p> <p>Maximilian Bähler, Otto Kiäyig, Martin Taraz, Sarel Cohen, Karen Seidel, Tobias Friedrich</p>	2405
<p>Visual Correspondence Hallucination</p> <p>Hugo Germain, Vincent Lepetit, Guillaume Bourmaud</p>	2430

<p>Quadtree Attention for Vision Transformers</p> <p><i>Shitao Tang, Jiahui Zhang, Siyu Zhu, Ping Tan</i></p>	2462
<p>RegionViT: Regional-To-Local Attention for Vision Transformers</p> <p><i>Chun-Fu Chen, Rameswar Panda, Quanfu Fan</i></p>	2478
<p>FedPara: Low-Rank Hadamard Product for Communication-Efficient Federated Learning</p> <p><i>Nam Hyeon-Woo, Moon Ye-Bin, Tae-Hyun Oh</i></p>	2496
<p>Understanding Dimensional Collapse in Contrastive Self-Supervised Learning</p> <p><i>Li Jing, Pascal Vincent, Yann Lecun, Yuandong Tian</i></p>	2519
<p>R4D: Utilizing Reference Objects for Long-Range Distance Estimation</p> <p><i>Yingwei Li, Tiffany Chen, Maya Kabkab, Ruichi Yu, Longlong Jing, Yurong You, Hang Zhao</i></p>	2536
<p>Surrogate Gap Minimization Improves Sharpness-Aware Training</p> <p><i>Juntang Zhuang, Boqing Gong, Liangzhe Yuan, Yin Cui, Hartwig Adam, Nicha C Dvornek, Sekhar Tatikonda, James S Duncan, Ting Liu</i></p>	2553
<p>Gradient Step Denoiser for Convergent Plug-And-Play</p> <p><i>Samuel Hurault, Arthur Leclaire, Nicolas Papadakis</i></p>	2577
<p>Unsupervised Discovery of Object Radiance Fields</p> <p><i>Hong-Xing Yu, Leonidas Guibas, Jiajun Wu</i></p>	2607
<p>Wisdom of Committees: An Overlooked Approach to Faster and More Accurate Models</p> <p><i>Xiaofang Wang, Dan Kondratyuk, Eric Christiansen, Kris M. Kitani, Yair Movshovitz-Attias, Elad Eban</i></p>	2639
<p>A Reduction-Based Framework for Conservative Bandits and Reinforcement Learning</p> <p><i>Yunchang Yang, Tianhao Wu, Han Zhong, Evrard Garcelon, Matteo Pirota, Alessandro Lazaric, Liwei Wang, Simon Shaolei Du</i></p>	2661
<p>Declarative Nets that Are Equilibrium Models</p> <p><i>Russell Tsuchida, Suk Yee Yong, Mohammad Ali Armin, Lars Petersson, Cheng Soon Ong</i></p>	2672
<p>Anytime Dense Prediction with Confidence Adaptivity</p> <p><i>Zhuang Liu, Zhiqiu Xu, Hung-Ju Wang, Trevor Darrell, Evan Shelhamer</i></p>	2705
<p>Online Adversarial Attacks</p> <p><i>Andjela Mladenovic, Joey Bose, Hugo Berard, William L. Hamilton, Simon Lacoste-Julien, Pascal Vincent, Gauthier Gidel</i></p>	2724
<p>Uncertainty Modeling for Out-Of-Distribution Generalization</p> <p><i>Xiaotong Li, Yongxing Dai, Yixiao Ge, Jun Liu, Ying Shan, Lingyu Duan</i></p>	2737
<p>Towards Deployment-Efficient Reinforcement Learning: Lower Bound and Optimality</p> <p><i>Jiawei Huang, Jinglin Chen, Li Zhao, Tao Qin, Nan Jiang, Tie-Yan Liu</i></p>	2753

Learning Efficient Online 3D Bin Packing on Packing Configuration Trees <i>Hang Zhao, Yang Yu, Kai Xu</i>	2766
SHINE: SHaring the INverse Estimate from the Forward Pass for Bi-Level Optimization and Implicit Models <i>Zaccharie Ramzi, Florian Mannel, Shaojie Bai, Jean-Luc Starck, Philippe Ciuciu, Thomas Moreau</i>	2784
Rethinking Class-Prior Estimation for Positive-Unlabeled Learning <i>Yu Yao, Tongliang Liu, Bo Han, Mingming Gong, Gang Niu, Masashi Sugiyama, Dacheng Tao</i>	2804
You Mostly Walk Alone: Analyzing Feature Attribution in Trajectory Prediction <i>Osama Makansi, Julius Von Kägelgen, Francesco Locatello, Peter Vincent Gehler, Dominik Janzing, Thomas Brox, Bernhard Schölkopf</i>	2816
QDrop: Randomly Dropping Quantization for Extremely Low-Bit Post-Training Quantization <i>Xiuying Wei, Ruihao Gong, Yuhang Li, Xianglong Liu, Fengwei Yu</i>	2831
Dropout Q-Functions for Doubly Efficient Reinforcement Learning <i>Takuya Hiraoka, Takahisa Imagawa, Taisei Hashimoto, Takashi Onishi, Yoshimasa Tsuruoka</i>	2850
A Generalization of the Randomized Singular Value Decomposition <i>Nicolas Boulle, Alex Townsend</i>	2883
Bregman Gradient Policy Optimization <i>Feihu Huang, Shangqian Gao, Heng Huang</i>	2899
Neural Relational Inference with Node-Specific Information <i>Ershad Banijamali</i>	2921
The Rich Get Richer: Disparate Impact of Semi-Supervised Learning <i>Zhaowei Zhu, Tianyi Luo, Yang Liu</i>	2936
Learning Temporally Causal Latent Processes from General Temporal Data <i>Weiran Yao, Yuewen Sun, Alex Ho, Changyin Sun, Kun Zhang</i>	2957
Learning Representation from Neural Fisher Kernel with Low-Rank Approximation <i>Ruixiang Zhang, Shuangfei Zhai, Etai Littwin, Joshua M. Susskind</i>	2988
Train Short, Test Long: Attention with Linear Biases Enables Input Length Extrapolation <i>Ofir Press, Noah Smith, Mike Lewis</i>	3012
Vector-Quantized Image Modeling with Improved VQGAN <i>Jiahui Yu, Xin Li, Jing Yu Koh, Han Zhang, Ruoming Pang, James Qin, Alexander Ku, Yuanzhong Xu, Jason Baldridge, Yonghui Wu</i>	3037
Learning to Annotate Part Segmentation with Gradient Matching <i>Yu Yang, Xiaotian Cheng, Hakan Bilen, Xiangyang Ji</i>	3054

<p>Optimization Inspired Multi-Branch Equilibrium Models</p> <p><i>Mingjie Li, Yisen Wang, Xingyu Xie, Zhouchen Lin</i></p>	3080
<p>Superclass-Conditional Gaussian Mixture Model for Learning Fine-Grained Embeddings</p> <p><i>Jingchao Ni, Wei Cheng, Zhengzhang Chen, Takayoshi Asakura, Tomoya Soma, Sho Kato, Haifeng Chen</i></p>	3101
<p>Learning with Noisy Labels Revisited: A Study Using Real-World Human Annotations</p> <p><i>Jiaheng Wei, Zhaowei Zhu, Hao Cheng, Tongliang Liu, Gang Niu, Yang Liu</i></p>	3121
<p>Online Continual Learning on Class Incremental Blurry Task Configuration with Anytime Inference</p> <p><i>Hyunseo Koh, Dahyun Kim, Jung-Woo Ha, Jonghyun Choi</i></p>	3144
<p>AS-MLP: An Axial Shifted MLP Architecture for Vision</p> <p><i>Dongze Lian, Zehao Yu, Xing Sun, Shenghua Gao</i></p>	3165
<p>Emergent Communication at Scale</p> <p><i>Rahma Chaabouni, Florian Strub, Florent Alth�, Eugene Tarassov, Corentin Tallec, Elnaz Davoodi, Kory Wallace Mathewson, Olivier Tieleman, Angeliki Lazaridou, Bilal Piot</i></p>	3184
<p>Optimal ANN-SNN Conversion for High-Accuracy and Ultra-Low-Latency Spiking Neural Networks</p> <p><i>Tong Bu, Wei Fang, Jianhao Ding, Penglin Dai, Zhaofei Yu, Tiejun Huang</i></p>	3214
<p>Modular Lifelong Reinforcement Learning Via Neural Composition</p> <p><i>Jorge A Mendez, Harm Van Seijen, Eric Eaton</i></p>	3233
<p>Learned Simulators for Turbulence</p> <p><i>Kim Stachenfeld, Drummond Buschman Fielding, Dmitrii Kochkov, Miles Cranmer, Tobias Pfaff, Jonathan Godwin, Can Cui, Shirley Ho, Peter Battaglia, Alvaro Sanchez-Gonzalez</i></p>	3255
<p>The Inductive Bias of In-Context Learning: Rethinking Pretraining Example Design</p> <p><i>Yoav Levine, Noam Wies, Daniel Jannai, Dan Navon, Yedid Hoshen, Amnon Shashua</i></p>	3282
<p>Multiset-Equivariant Set Prediction with Approximate Implicit Differentiation</p> <p><i>Yan Zhang, David W Zhang, Simon Lacoste-Julien, Gertjan J. Burghouts, Cees G. M. Snoek</i></p>	3322
<p>ProtoRes: Proto-Residual Network for Pose Authoring Via Learned Inverse Kinematics</p> <p><i>Boris N. Oreshkin, Florent Bocquelet, Felix G. Harvey, Bay Raitt, Dominic Laflamme</i></p>	3356
<p>Source-Free Adaptation to Measurement Shift Via Bottom-Up Feature Restoration</p> <p><i>Cian Eastwood, Ian Mason, Chris Williams, Bernhard Sch�lkopf</i></p>	3388
<p>Concurrent Adversarial Learning for Large-Batch Training</p> <p><i>Yong Liu, Xiangning Chen, Minhao Cheng, Cho-Jui Hsieh, Yang You</i></p>	3425
<p>Understanding and Preventing Capacity Loss in Reinforcement Learning</p> <p><i>Clare Lyle, Mark Rowland, Will Dabney</i></p>	3442

<p>On the Convergence of the Monte Carlo Exploring Starts Algorithm for Reinforcement Learning</p> <p><i>Che Wang, Shuhan Yuan, Kai Shao, Keith W. Ross</i></p>	3469
<p>Imbedding Deep Neural Networks</p> <p><i>Andrew Corbett, Dmitry Kangin</i></p>	3502
<p>Discrete Representations Strengthen Vision Transformer Robustness</p> <p><i>Chengzhi Mao, Lu Jiang, Mostafa Dehghani, Carl Vondrick, Rahul Sukthankar, Irfan Essa</i></p>	3524
<p>Transform2Act: Learning a Transform-And-Control Policy for Efficient Agent Design</p> <p><i>Ye Yuan, Yuda Song, Zhengyi Luo, Wen Sun, Kris M. Kitani</i></p>	3550
<p>In a Nutshell, the Human Asked for This: Latent Goals for Following Temporal Specifications</p> <p><i>Borja G. Leñn, Murray Shanahan, Francesco Belardinelli</i></p>	3567
<p>F8Net: Fixed-Point 8-Bit Only Multiplication for Network Quantization</p> <p><i>Qing Jin, Jian Ren, Richard Zhuang, Sumant Hanumante, Zhengang Li, Zhiyu Chen, Yanzhi Wang, Kaiyuan Yang, Sergey Tulyakov</i></p>	3591
<p>How Low Can We Go: Trading Memory for Error in Low-Precision Training</p> <p><i>Chengrun Yang, Ziyang Wu, Jerry Chee, Christopher De Sa, Madeleine Udell</i></p>	3610
<p>Differentially Private Fractional Frequency Moments Estimation with Polylogarithmic Space</p> <p><i>Lun Wang, Iosif Pinelis, Dawn Song</i></p>	3631
<p>Online Target Q-Learning with Reverse Experience Replay: Efficiently Finding the Optimal Policy for Linear MDPs</p> <p><i>Naman Agarwal, Syomantak Chaudhuri, Prateek Jain, Dheeraj Mysore Nagaraj, Praneeth Netrapalli</i></p>	3644
<p>Tackling the Generative Learning Trilemma with Denoising Diffusion GANs</p> <p><i>Zhisheng Xiao, Karsten Kreis, Arash Vahdat</i></p>	3680
<p>Sparse DETR: Efficient End-To-End Object Detection with Learnable Sparsity</p> <p><i>Byungseok Roh, Jaewoong Shin, Wuhyun Shin, Saehoon Kim</i></p>	3708
<p>Stein Latent Optimization for Generative Adversarial Networks</p> <p><i>Uiwon Hwang, Heeseung Kim, Dahuin Jung, Hyemi Jang, Hyungyu Lee, Sungroh Yoon</i></p>	3731
<p>Finetuned Language Models Are Zero-Shot Learners</p> <p><i>Jason Wei, Maarten Bosma, Vincent Zhao, Kelvin Guu, Adams Wei Yu, Brian Lester, Nan Du, Andrew M. Dai, Quoc V Le</i></p>	3766
<p>X-Model: Improving Data Efficiency in Deep Learning with a Minimax Model</p> <p><i>Ximei Wang, Xinyang Chen, Jianmin Wang, Mingsheng Long</i></p>	3812
<p>Constructing Orthogonal Convolutions in an Explicit Manner</p> <p><i>Tan Yu, Jun Li, Yunfeng Cai, Ping Li</i></p>	3826

Information Gain Propagation: A New Way to Graph Active Learning with Soft Labels Wentao Zhang, Yexin Wang, Zhenbang You, Meng Cao, Ping Huang, Jiulong Shan, Zhi Yang, Bin Cui	3840
Multimeasurement Generative Models Saeed Saremi, Rupesh Kumar Srivastava	3857
Contrastive Clustering to Mine Pseudo Parallel Data for Unsupervised Translation Xuan-Phi Nguyen, Hongyu Gong, Yun Tang, Changhan Wang, Philipp Koehn, Shafiq Joty	3888
An Unconstrained Layer-Peeled Perspective on Neural Collapse Wenlong Ji, Yiping Lu, Yiliang Zhang, Zhun Deng, Weijie J Su	3906
Filling the Gaps: Multivariate Time Series Imputation by Graph Neural Networks Andrea Cini, Ivan Marisca, Cesare Alippi	3934
Pareto Policy Pool for Model-Based Offline Reinforcement Learning Yijun Yang, Jing Jiang, Tianyi Zhou, Jie Ma, Yuhui Shi	3954
Generative Principal Component Analysis Zhaoqiang Liu, Jiulong Liu, Subhroshekhar Ghosh, Jun Han, Jonathan Scarlett	3976
Bootstrapping Semantic Segmentation with Regional Contrast Shikun Liu, Shuaifeng Zhi, Edward Johns, Andrew Davison	4001
Signing the Supermask: Keep, Hide, Invert Nils Koster, Oliver Grothe, Achim Rettinger	4018
Pseudo-Labeled Auto-Curriculum Learning for Semi-Supervised Keypoint Localization Can Wang, Sheng Jin, Yingda Guan, Wentao Liu, Chen Qian, Ping Luo, Wanli Ouyang	4055
Efficient Neural Causal Discovery Without Acyclicity Constraints Phillip Lippe, Taco Cohen, Efstratios Gavves	4071
On the Certified Robustness for Ensemble Models and Beyond Zhuolin Yang, Linyi Li, Xiaojun Xu, Bhavya Kailkhura, Tao Xie, Bo Li	4122
DAB-DETR: Dynamic Anchor Boxes Are Better Queries for DETR Shilong Liu, Feng Li, Hao Zhang, Xiao Yang, Xianbiao Qi, Hang Su, Jun Zhu, Lei Zhang	4173
Reverse Engineering of Imperceptible Adversarial Image Perturbations Yifan Gong, Yuguang Yao, Yize Li, Yimeng Zhang, Xiaoming Liu, Xue Lin, Sijia Liu	4193
Learning to Complete Code with Sketches Daya Guo, Alexey Svyatkovskiy, Jian Yin, Nan Duan, Marc Brockschmidt, Miltiadis Allamanis	4213
Variational Methods for Simulation-Based Inference Manuel Gläcker, Michael Deistler, Jakob H. Macke	4236

<p>Backdoor Defense Via Decoupling the Training Process</p> <p><i>Kunzhe Huang, Yiming Li, Baoyuan Wu, Zhan Qin, Kui Ren</i></p>	4262
<p>How Do Vision Transformers Work?</p> <p><i>Namuk Park, Songkuk Kim</i></p>	4287
<p>Few-Shot Backdoor Attacks on Visual Object Tracking</p> <p><i>Yiming Li, Haoxiang Zhong, Xingjun Ma, Yong Jiang, Shu-Tao Xia</i></p>	4313
<p>Online Facility Location with Predictions</p> <p><i>Shaofeng H.-C. Jiang, Erzhi Liu, You Lyu, Zhihao Gavin Tang, Yubo Zhang</i></p>	4334
<p>Sparse Communication Via Mixed Distributions</p> <p><i>Antônio Farinhas, Wilker Aziz, Vlad Niculae, Andre Martins</i></p>	4352
<p>Equivariant Subgraph Aggregation Networks</p> <p><i>Beatrice Bevilacqua, Fabrizio Frasca, Derek Lim, Balasubramaniam Srinivasan, Chen Cai, Gopinath Balamurugan, Michael M. Bronstein, Haggai Maron</i></p>	4379
<p>Pessimistic Bootstrapping for Uncertainty-Driven Offline Reinforcement Learning</p> <p><i>Chenjia Bai, Lingxiao Wang, Zhuoran Yang, Zhi-Hong Deng, Animesh Garg, Peng Liu, Zhaoran Wang</i></p>	4425
<p>Learning Super-Features for Image Retrieval</p> <p><i>Philippe Weinzaepfel, Thomas Lucas, Diane Larlus, Yannis Kalantidis</i></p>	4454
<p>Should We Be Pre-Training? an Argument for End-Task Aware Training as an Alternative</p> <p><i>Lucio M. Dery, Paul Michel, Ameet Talwalkar, Graham Neubig</i></p>	4473
<p>Simple GNN Regularisation for 3D Molecular Property Prediction and Beyond</p> <p><i>Jonathan Godwin, Michael Schaarschmidt, Alexander L Gaunt, Alvaro Sanchez-Gonzalez, Yulia Rubanova, Petar Velčkovič, James Kirkpatrick, Peter Battaglia</i></p>	4491
<p>Is High Variance Unavoidable in RL? a Case Study in Continuous Control</p> <p><i>Johan Bjorck, Carla P Gomes, Kilian Q Weinberger</i></p>	4514
<p>Do We Need Anisotropic Graph Neural Networks?</p> <p><i>Shyam A. Tailor, Felix Opolka, Pietro Lio, Nicholas Donald Lane</i></p>	4546
<p>Accelerated Policy Learning with Parallel Differentiable Simulation</p> <p><i>Jie Xu, Viktor Makoviychuk, Yashraj Narang, Fabio Ramos, Wojciech Matusik, Animesh Garg, Miles Macklin</i></p>	4565
<p>Generative Planning for Temporally Coordinated Exploration in Reinforcement Learning</p> <p><i>Haichao Zhang, Wei Xu, Haonan Yu</i></p>	4591
<p>Assessing Generalization of SGD Via Disagreement</p> <p><i>Yiding Jiang, Vaishnavh Nagarajan, Christina Baek, J Zico Kolter</i></p>	4611

<p>Image BERT Pre-Training with Online Tokenizer</p> <p>Jinghao Zhou, Chen Wei, Huiyu Wang, Wei Shen, Cihang Xie, Alan Yuille, Tao Kong</p>	4639
<p>Language-Driven Semantic Segmentation</p> <p>Boyi Li, Kilian Q Weinberger, Serge Belongie, Vladlen Koltun, Rene Ranftl</p>	4668
<p>Hindsight is 20/20: Leveraging Past Traversals to Aid 3D Perception</p> <p>Yurong You, Katie Z Luo, Xiangyu Chen, Junan Chen, Wei-Lun Chao, Wen Sun, Bharath Hariharan, Mark Campbell, Kilian Q Weinberger</p>	4681
<p>On the Role of Population Heterogeneity in Emergent Communication</p> <p>Mathieu Rita, Florian Strub, Jean-Bastien Grill, Olivier Pietquin, Emmanuel Dupoux</p>	4698
<p>Privacy Implications of Shuffling</p> <p>Casey Meehan, Amrita Roy Chowdhury, Kamalika Chaudhuri, Somesh Jha</p>	4717
<p>Bundle Networks: Fiber Bundles, Local Trivializations, and a Generative Approach to Exploring Many-To-One Maps</p> <p>Nico Courts, Henry Kvinge</p>	4747
<p>SimVLM: Simple Visual Language Model Pretraining with Weak Supervision</p> <p>Zirui Wang, Jiahui Yu, Adams Wei Yu, Zihang Dai, Yulia Tsvetkov, Yuan Cao</p>	4767
<p>A General Analysis of Example-Selection for Stochastic Gradient Descent</p> <p>Yucheng Lu, Si Yi Meng, Christopher De Sa</p>	4784
<p>Safe Neurosymbolic Learning with Differentiable Symbolic Execution</p> <p>Chenxi Yang, Swarat Chaudhuri</p>	4828
<p>Efficient and Differentiable Conformal Prediction with General Function Classes</p> <p>Yu Bai, Song Mei, Huan Wang, Yingbo Zhou, Caiming Xiong</p>	4856
<p>FP-DETR: Detection Transformer Advanced by Fully Pre-Training</p> <p>Wen Wang, Yang Cao, Jing Zhang, Dacheng Tao</p>	4884
<p>MetaShift: A Dataset of Datasets for Evaluating Contextual Distribution Shifts and Training Conflicts</p> <p>Weixin Liang, James Zou</p>	4898
<p>Associated Learning: An Alternative to End-To-End Backpropagation that Works on CNN, RNN, and Transformer</p> <p>Dennis Y. H. Wu, Dinan Lin, Vincent Chen, Hung-Hsuan Chen</p>	4917
<p>Demystifying Batch Normalization in ReLU Networks: Equivalent Convex Optimization Models and Implicit Regularization</p> <p>Tolga Ergen, Arda Sahiner, Batu Ozturkler, John M. Pauly, Morteza Mardani, Mert Pilanci</p>	4935

<p><i>New Insights on Reducing Abrupt Representation Change in Online Continual Learning</i></p> <p><i>Lucas Caccia, Rahaf Aljundi, Nader Asadi, Tinne Tuytelaars, Joelle Pineau, Eugene Belilovsky</i></p>	4971
<p><i>Progressive Distillation for Fast Sampling of Diffusion Models</i></p> <p><i>Tim Salimans, Jonathan Ho</i></p>	4998
<p><i>Rethinking Network Design and Local Geometry in Point Cloud: A Simple Residual MLP Framework</i></p> <p><i>Xu Ma, Can Qin, Haoxuan You, Haoxi Ran, Yun Fu</i></p>	5019
<p><i>Decoupled Adaptation for Cross-Domain Object Detection</i></p> <p><i>Junguang Jiang, Baixu Chen, Jianmin Wang, Mingsheng Long</i></p>	5034
<p><i>Ada-NETS: Face Clustering Via Adaptive Neighbour Discovery in the Structure Space</i></p> <p><i>Yaohua Wang, Yaobin Zhang, Fangyi Zhang, Senzhang Wang, Ming Lin, Yuqi Zhang, Xiuyu Sun</i></p>	5050
<p><i>Omni-Scale CNNs: A Simple and Effective Kernel Size Configuration for Time Series Classification</i></p> <p><i>Wensi Tang, Guodong Long, Lu Liu, Tianyi Zhou, Michael Blumenstein, Jing Jiang</i></p>	5066
<p><i>Focus on the Common Good: Group Distributional Robustness Follows</i></p> <p><i>Vihari Piratla, Praneeth Netrapalli, Sunita Sarawagi</i></p>	5083
<p><i>Generative Modeling with Optimal Transport Maps</i></p> <p><i>Litu Rout, Alexander Korotin, Evgeny Burnaev</i></p>	5100
<p><i>Representing Mixtures of Word Embeddings with Mixtures of Topic Embeddings</i></p> <p><i>Dongsheng Wang, Dan Dan Guo, He Zhao, Huangjie Zheng, Korawat Tanwisuth, Bo Chen, Mingyuan Zhou</i></p>	5122
<p><i>Discovering Invariant Rationales for Graph Neural Networks</i></p> <p><i>Yingxin Wu, Xiang Wang, An Zhang, Xiangnan He, Tat-Seng Chua</i></p>	5139
<p><i>Acceleration of Federated Learning with Alleviated Forgetting in Local Training</i></p> <p><i>Chencheng Xu, Zhiwei Hong, Minlie Huang, Tao Jiang</i></p>	5161
<p><i>Towards Model Agnostic Federated Learning Using Knowledge Distillation</i></p> <p><i>Andrei Afonin, Sai Praneeth Karimireddy</i></p>	5180
<p><i>Anomaly Transformer: Time Series Anomaly Detection with Association Discrepancy</i></p> <p><i>Jiehui Xu, Haixu Wu, Jianmin Wang, Mingsheng Long</i></p>	5203
<p><i>MiniF2F: A Cross-System Benchmark for Formal Olympiad-Level Mathematics</i></p> <p><i>Kunhao Zheng, Jesse Michael Han, Stanislas Polu</i></p>	5223
<p><i>A Unified Wasserstein Distributional Robustness Framework for Adversarial Training</i></p> <p><i>Anh Tuan Bui, Trung Le, Quan Hung Tran, He Zhao, Dinh Phung</i></p>	5234

GiraffeDet: A Heavy-Neck Paradigm for Object Detection Yiqi Jiang, Zhiyu Tan, Junyan Wang, Xiuyu Sun, Ming Lin, Hao Li	5259
Generative Models as a Data Source for Multiview Representation Learning Ali Jahanian, Xavier Puig, Yonglong Tian, Phillip Isola	5276
DISCOVERING AND EXPLAINING THE REPRESENTATION BOTTLENECK OF DNNS Huiqi Deng, Qihan Ren, Hao Zhang, Quanshi Zhang	5298
Graph-Based Nearest Neighbor Search in Hyperbolic Spaces Liudmila Prokhorenkova, Dmitry Baranchuk, Nikolay Bogachev, Yury Demidovich, Alexander Kolpakov	5323
Long Expressive Memory for Sequence Modeling T. Konstantin Rusch, Siddhartha Mishra, N. Benjamin Erichson, Michael W. Mahoney	5353
Task-Induced Representation Learning Jun Yamada, Karl Pertsch, Anisha Gunjal, Joseph J Lim	5386
Tighter Sparse Approximation Bounds for ReLU Neural Networks Carles Domingo-Enrich, Youssef Mroueh	5404
Learning Versatile Neural Architectures by Propagating Network Codes Mingyu Ding, Yuqi Huo, Haoyu Lu, Linjie Yang, Zhe Wang, Zhiwu Lu, Jingdong Wang, Ping Luo	5415
LEARNING GUARANTEES FOR GRAPH CONVOLUTIONAL NETWORKS ON THE STOCHASTIC BLOCK MODEL Wei Lu	5439
Online Hyperparameter Meta-Learning with Hypergradient Distillation Hae Beom Lee, Hayeon Lee, Jaewoong Shin, Eunho Yang, Timothy Hospedales, Sung Ju Hwang	5467
RvS: What is Essential for Offline RL Via Supervised Learning? Scott Emmons, Benjamin Eysenbach, Ilya Kostrikov, Sergey Levine	5484
On Covariate Shift of Latent Confounders in Imitation and Reinforcement Learning Guy Tennenholtz, Assaf Hallak, Gal Dalal, Shie Mannor, Gal Chechik, Uri Shalit	5498
Meta-Learning with Fewer Tasks Through Task Interpolation Huaxiu Yao, Linjun Zhang, Chelsea Finn	5528
When Vision Transformers Outperform ResNets Without Pre-Training Or Strong Data Augmentations Xiangning Chen, Cho-Jui Hsieh, Boqing Gong	5558
Vision-Based Manipulators Need to Also See from Their Hands Kyle Hsu, Moo Jin Kim, Rafael Rafailov, Jiajun Wu, Chelsea Finn	5578
Cross-Trajectory Representation Learning for Zero-Shot Generalization in RL Bogdan Mazouze, Ahmed M Ahmed, R Devon Hjelm, Andrey Kolobov, Patrick Macalpine	5608

<p>Do Deep Networks Transfer Invariances Across Classes?</p> <p>Allan Zhou, Fahim Tajwar, Alexander Robey, Tom Knowles, George J. Pappas, Hamed Hassani, Chelsea Finn</p>	5632
<p>Self-Ensemble Adversarial Training for Improved Robustness</p> <p>Hongjun Wang, Yisen Wang</p>	5652
<p>Anti-Oversmoothing in Deep Vision Transformers Via the Fourier Domain Analysis: From Theory to Practice</p> <p>Peihao Wang, Wenqing Zheng, Tianlong Chen, Zhangyang Wang</p>	5670
<p>Temporal Alignment Prediction for Supervised Representation Learning and Few-Shot Sequence Classification</p> <p>Bing Su, Ji-Rong Wen</p>	5694
<p>Deep AutoAugment</p> <p>Yu Zheng, Zhi Zhang, Shen Yan, Mi Zhang</p>	5713
<p>Active Hierarchical Exploration with Stable Subgoal Representation Learning</p> <p>Siyuan Li, Jin Zhang, Jianhao Wang, Yang Yu, Chongjie Zhang</p>	5729
<p>Programmatic Reinforcement Learning Without Oracles</p> <p>Wenjie Qiu, He Zhu</p>	5747
<p>A NON-PARAMETRIC REGRESSION VIEWPOINT : GENERALIZATION OF OVERPARAMETRIZED DEEP RELU NETWORK UNDER NOISY OBSERVATIONS</p> <p>Namjoon Suh, Hyunouk Ko, Xiaoming Huo</p>	5781
<p>Robust Unlearnable Examples: Protecting Data Privacy Against Adversarial Learning</p> <p>Shaopeng Fu, Fengxiang He, Yang Liu, Li Shen, Dacheng Tao</p>	5826
<p>CrossMatch: Cross-Classifer Consistency Regularization for Open-Set Single Domain Generalization</p> <p>Ronghang Zhu, Sheng Li</p>	5848
<p>Sparsity Winning Twice: Better Robust Generalization from More Efficient Training</p> <p>Tianlong Chen, Zhenyu Zhang, Pengjun Wang, Santosh Balachandra, Haoyu Ma, Zehao Wang, Zhangyang Wang</p>	5865
<p>Topological Experience Replay</p> <p>Zhang-Wei Hong, Tao Chen, Yen-Chen Lin, Joni Pajarinen, Pulkit Agrawal</p>	5890
<p>Environment Predictive Coding for Visual Navigation</p> <p>Santhosh Kumar Ramakrishnan, Tushar Nagarajan, Ziad Al-Halah, Kristen Grauman</p>	5914

<p>Supervision Exists Everywhere: A Data Efficient Contrastive Language-Image Pre-Training Paradigm</p> <p>Yangguang Li, Feng Liang, Lichen Zhao, Yufeng Cui, Wanli Ouyang, Jing Shao, Fengwei Yu, Junjie Yan</p>	5937
<p>Pseudo Numerical Methods for Diffusion Models on Manifolds</p> <p>Luping Liu, Yi Ren, Zhijie Lin, Zhou Zhao</p>	5954
<p>Representational Continuity for Unsupervised Continual Learning</p> <p>Divyam Madaan, Jaehong Yoon, Yuanchun Li, Yunxin Liu, Sung Ju Hwang</p>	5977
<p>Sequential Reptile: Inter-Task Gradient Alignment for Multilingual Learning</p> <p>Seanie Lee, Hae Beom Lee, Juho Lee, Sung Ju Hwang</p>	5995
<p>Resolving Training Biases Via Influence-Based Data Relabeling</p> <p>Shuming Kong, Yanyan Shen, Linpeng Huang</p>	6013
<p>Symbolic Learning to Optimize: Towards Interpretability and Scalability</p> <p>Wenqing Zheng, Tianlong Chen, Ting-Kuei Hu, Zhangyang Wang</p>	6034
<p>Dynamic Token Normalization Improves Vision Transformers</p> <p>Wenqi Shao, Yixiao Ge, Zhaoyang Zhang, Xuyuan Xu, Xiaogang Wang, Ying Shan, Ping Luo</p>	6053
<p>Multi-Task Processes</p> <p>Donggyun Kim, Seongwoong Cho, Wonkwang Lee, Seunghoon Hong</p>	6071
<p>Post-Training Detection of Backdoor Attacks for Two-Class and Multi-Attack Scenarios</p> <p>Zhen Xiang, David Miller, George Kesidis</p>	6120
<p>Model-Augmented Prioritized Experience Replay</p> <p>Youngmin Oh, Jinwoo Shin, Eunho Yang, Sung Ju Hwang</p>	6154
<p>Huber Additive Models for Non-Stationary Time Series Analysis</p> <p>Yingjie Wang, Xianrui Zhong, Fengxiang He, Hong Chen, Dacheng Tao</p>	6180
<p>PoNet: Pooling Network for Efficient Token Mixing in Long Sequences</p> <p>Chao-Hong Tan, Qian Chen, Wen Wang, Qinglin Zhang, Siqi Zheng, Zhen-Hua Ling</p>	6210
<p>Auto-Transfer: Learning to Route Transferable Representations</p> <p>Keerthiram Murugesan, Vijay Sadashivaiah, Ronny Luss, Karthikeyan Shanmugam, Pin-Yu Chen, Amit Dhurandhar</p>	6233
<p>A Conditional Point Diffusion-Refinement Paradigm for 3D Point Cloud Completion</p> <p>Zhaoyang Lyu, Zhifeng Kong, Xudong Xu, Liang Pan, Dahua Lin</p>	6254
<p>Zero Pixel Directional Boundary by Vector Transform</p> <p>Edoardo Mello Rella, Ajad Chhatkuli, Yun Liu, Ender Konukoglu, Luc Van Gool</p>	6278

<p>FlexConv: Continuous Kernel Convolutions with Differentiable Kernel Sizes</p> <p>David W. Romero, Robert-Jan Bruintjes, Jakub Mikolaj Tomczak, Erik J Bekkers, Mark Hoogendoorn, Jan Van Gemert</p>	6298
<p>IFlood: A Stable and Effective Regularizer</p> <p>Yuexiang Xie, Zhen Wang, Yaliang Li, Ce Zhang, Jingren Zhou, Bolin Ding</p>	6322
<p>On the Optimal Memorization Power of ReLU Neural Networks</p> <p>Gal Vardi, Gilad Yehudai, Ohad Shamir</p>	6338
<p>Unraveling Model-Agnostic Meta-Learning Via the Adaptation Learning Rate</p> <p>Yingtian Zou, Fusheng Liu, Qianxiao Li</p>	6356
<p>HyperDQN: A Randomized Exploration Method for Deep Reinforcement Learning</p> <p>Ziniu Li, Yingru Li, Yushun Zhang, Tong Zhang, Zhi-Quan Luo</p>	6394
<p>Implicit Bias of Projected Subgradient Method Gives Provable Robust Recovery of Subspaces of Unknown Codimension</p> <p>Paris Giampouras, Benjamin David Haeffele, Rene Vidal</p>	6430
<p>Deep Ensembling with No Overhead for Either Training Or Testing: The All-Round Blessings of Dynamic Sparsity</p> <p>Shiwei Liu, Tianlong Chen, Zahra Atashgahi, Xiaohan Chen, Ghada Sokar, Elena Mocanu, Mykola Pechenizkiy, Zhangyang Wang, Decebal Constantin Mocanu</p>	6448
<p>On the Relation Between Statistical Learning and Perceptual Distances</p> <p>Alexander Hepburn, Valero Laparra, Raul Santos-Rodriguez, Johannes Ballā, Jesus Malo</p>	6473
<p>Bayesian Modeling and Uncertainty Quantification for Learning to Optimize: What, Why, and How</p> <p>Yuning You, Yue Cao, Tianlong Chen, Zhangyang Wang, Yang Shen</p>	6499
<p>LIGS: Learnable Intrinsic-Reward Generation Selection for Multi-Agent Learning</p> <p>David Henry Mguni, Taher Jafferjee, Jianhong Wang, Nicolas Perez-Nieves, Oliver Slumbers, Feifei Tong, Yang Li, Jiangcheng Zhu, Yaodong Yang, Jun Wang</p>	6518
<p>Anomaly Detection for Tabular Data with Internal Contrastive Learning</p> <p>Tom Shenkar, Lior Wolf</p>	6531
<p>Unsupervised Disentanglement with Tensor Product Representations on the Torus</p> <p>Michael Rotman, Amit Dekel, Shir Gur, Yaron Oz, Lior Wolf</p>	6557
<p>Trust Region Policy Optimisation in Multi-Agent Reinforcement Learning</p> <p>Jakub Grudzien Kuba, Ruiqing Chen, Muning Wen, Ying Wen, Fanglei Sun, Jun Wang, Yaodong Yang</p>	6569
<p>VOS: Learning What You Don't Know by Virtual Outlier Synthesis</p> <p>Xuefeng Du, Zhaoning Wang, Mu Cai, Yixuan Li</p>	6596

<p>Critical Points in Quantum Generative Models</p> <p><i>Eric Ricardo Anschuetz</i></p>	6617
<p>Continuous-Time Meta-Learning with Forward Mode Differentiation</p> <p><i>Tristan Deleu, David Kanaa, Leo Feng, Giancarlo Kerg, Yoshua Bengio, Guillaume Lajoie, Pierre-Luc Bacon</i></p>	6644
<p>A Relational Intervention Approach for Unsupervised Dynamics Generalization in Model-Based Reinforcement Learning</p> <p><i>Jiaxian Guo, Mingming Gong, Dacheng Tao</i></p>	6676
<p>Multitask Prompted Training Enables Zero-Shot Task Generalization</p> <p><i>Victor Sanh, Albert Webson, Colin Raffel, Stephen Bach, Lintang Sutawika, Zaid Alyafeai, Antoine Chaffin, Arnaud Stiegler, Arun Raja, Manan Dey, M Saiful Bari, Canwen Xu, Urmish Thakker, Shanya Sharma Sharma, Eliza Szczechla, Taewoon Kim, Gunjan Chhablani, Nihal Nayak, Debajyoti Datta, Jonathan Chang, Mike Tian-Jian Jiang, Han Wang, Matteo Manica, Sheng Shen, Zheng Xin Yong, Harshit Pandey, Rachel Bawden, Thomas Wang, Trishala Neeraj, Jos Rozen, Abheesht Sharma, Andrea Santilli, Thibault Fevry, Jason Alan Fries, Ryan Teehan, Teven Le Scao, Stella Biderman, Leo Gao, Thomas Wolf, Alexander M Rush</i></p>	6700
<p>Embedded-Model Flows: Combining the Inductive Biases of Model-Free Deep Learning and Explicit Probabilistic Modeling</p> <p><i>Gianluigi Silvestri, Emily Fertig, Dave Moore, Luca Ambrogioni</i></p>	6916
<p>Learning Prototype-Oriented Set Representations for Meta-Learning</p> <p><i>Dan Dan Guo, Long Tian, Minghe Zhang, Mingyuan Zhou, Hongyuan Zha</i></p>	6946
<p>Learning Optimal Conformal Classifiers</p> <p><i>David Stutz, Krishnamurthy Dj Dvijotham, Ali Taylan Cemgil, Arnaud Doucet</i></p>	6966
<p>On the Learning and Learnability of Quasimetrics</p> <p><i>Tongzhou Wang, Phillip Isola</i></p>	6993
<p>Fooling Explanations in Text Classifiers</p> <p><i>Adam Ivankay, Ivan Girardi, Chiara Marchiori, Pascal Frossard</i></p>	7050
<p>Policy Improvement by Planning with Gumbel</p> <p><i>Ivo Danihelka, Arthur Guez, Julian Schrittwieser, David Silver</i></p>	7063
<p>Spanning Tree-Based Graph Generation for Molecules</p> <p><i>Sungsoo Ahn, Binghong Chen, Tianzhe Wang, Le Song</i></p>	7085
<p>Fairness Guarantees Under Demographic Shift</p> <p><i>Stephen Giguere, Blossom Metevier, Bruno Castro Da Silva, Yuriy Brun, Philip S. Thomas, Scott Niekum</i></p>	7107
<p>Spherical Message Passing for 3D Molecular Graphs</p> <p><i>Yi Liu, Limei Wang, Meng Liu, Yuchao Lin, Xuan Zhang, Bora Oztekin, Shuiwang Ji</i></p>	7131

<p>Node Feature Extraction by Self-Supervised Multi-Scale Neighborhood Prediction</p> <p><i>Eli Chien, Wei-Cheng Chang, Cho-Jui Hsieh, Hsiang-Fu Yu, Jiong Zhang, Olgica Milenkovic, Inderjit S Dhillon</i></p>	7151
<p>Know Thyself: Transferable Visual Control Policies Through Robot-Awareness</p> <p><i>Edward S. Hu, Kun Huang, Oleh Rybkin, Dinesh Jayaraman</i></p>	7171
<p>D-CODE: Discovering Closed-Form ODEs from Observed Trajectories</p> <p><i>Zhaozhi Qian, Krzysztof Kacprzyk, Mihaela Van Der Schaar</i></p>	7191
<p>Latent Image Animator: Learning to Animate Images Via Latent Space Navigation</p> <p><i>Yaohui Wang, Di Yang, Francois Bremond, Antitza Dantcheva</i></p>	7219
<p>Task Affinity with Maximum Bipartite Matching in Few-Shot Learning</p> <p><i>Cat Phuoc Le, Juncheng Dong, Mohammadreza Soltani, Vahid Tarokh</i></p>	7236
<p>Crystal Diffusion Variational Autoencoder for Periodic Material Generation</p> <p><i>Tian Xie, Xiang Fu, Octavian-Eugen Ganea, Regina Barzilay, Tommi S. Jaakkola</i></p>	7257
<p>Certified Robustness for Deep Equilibrium Models Via Interval Bound Propagation</p> <p><i>Colin Wei, J Zico Kolter</i></p>	7277
<p>Surrogate NAS Benchmarks: Going Beyond the Limited Search Spaces of Tabular NAS Benchmarks</p> <p><i>Arber Zela, Julien Niklas Siems, Lucas Zimmer, Jovita Lukasik, Margret Keuper, Frank Hutter</i></p>	7294
<p>Spatial Graph Attention and Curiosity-Driven Policy for Antiviral Drug Discovery</p> <p><i>Yulun Wu, Nicholas Choma, Andrew Deru Chen, Mikaela Cashman, Erica Teixeira Prates, Veronica G Melesse Vergara, Manesh B Shah, Austin Clyde, Thomas Brettin, Wibe Albert De Jong, Neeraj Kumar, Martha S Head, Rick L. Stevens, Peter Nugent, Daniel A Jacobson, James B Brown</i></p>	7330
<p>MobileViT: Light-Weight, General-Purpose, and Mobile-Friendly Vision Transformer</p> <p><i>Sachin Mehta, Mohammad Rastegari</i></p>	7346
<p>Communication-Efficient Actor-Critic Methods for Homogeneous Markov Games</p> <p><i>Dingyang Chen, Yile Li, Qi Zhang</i></p>	7372
<p>How to Train Your MAML to Excel in Few-Shot Classification</p> <p><i>Han-Jia Ye, Wei-Lun Chao</i></p>	7395
<p>NAS: Label- And Data-Agnostic Neural Architecture Search at Initialization</p> <p><i>Yao Shu, Shaofeng Cai, Zhongxiang Dai, Beng Chin Ooi, Bryan Kian Hsiang Low</i></p>	7416
<p>Cold Brew: Distilling Graph Node Representations with Incomplete Or Missing Neighborhoods</p> <p><i>Wenqing Zheng, Edward W Huang, Nikhil Rao, Sumeet Katariya, Zhangyang Wang, Karthik Subbian</i></p>	7441
<p>Pessimistic Model-Based Offline Reinforcement Learning Under Partial Coverage</p> <p><i>Masatoshi Uehara, Wen Sun</i></p>	7457

<p>Energy-Based Learning for Cooperative Games, with Applications to Valuation Problems in Machine Learning</p> <p><i>Yatao Bian, Yu Rong, Tingyang Xu, Jiaxiang Wu, Andreas Krause, Junzhou Huang</i></p>	7492
<p>Learning Curves for Continual Learning in Neural Networks: Self-Knowledge Transfer and Forgetting</p> <p><i>Ryo Karakida, Shotaro Akaho</i></p>	7515
<p>Synchromesh: Reliable Code Generation from Pre-Trained Language Models</p> <p><i>Gabriel Poesia, Alex Polozov, Vu Le, Ashish Tiwari, Gustavo Soares, Christopher Meek, Sumit Gulwani</i></p>	7542
<p>You Are AllSet: A Multiset Function Framework for Hypergraph Neural Networks</p> <p><i>Eli Chien, Chao Pan, Jianhao Peng, Olgica Milenkovic</i></p>	7561
<p>EViT: Expediting Vision Transformers Via Token Reorganizations</p> <p><i>Youwei Liang, Chongjian Ge, Zhan Tong, Yibing Song, Jue Wang, Pengtao Xie</i></p>	7585
<p>GradSign: Model Performance Inference with Theoretical Insights</p> <p><i>Zhihao Zhang, Zhihao Jia</i></p>	7606
<p>Do Not Escape from the Manifold: Discovering the Local Coordinates on the Latent Space of GANs</p> <p><i>Jaewoong Choi, Junho Lee, Changyeon Yoon, Jung Ho Park, Geonho Hwang, Myungjoo Kang</i></p>	7627
<p>Omni-Dimensional Dynamic Convolution</p> <p><i>Chao Li, Aojun Zhou, Anbang Yao</i></p>	7651
<p>Creating Training Sets Via Weak Indirect Supervision</p> <p><i>Jieyu Zhang, Bohan Wang, Xiangchen Song, Yujing Wang, Yaming Yang, Jing Bai, Alexander Ratner</i></p>	7671
<p>Automated Self-Supervised Learning for Graphs</p> <p><i>Wei Jin, Xiaorui Liu, Xiangyu Zhao, Yao Ma, Neil Shah, Jiliang Tang</i></p>	7710
<p>Blaschke Product Neural Networks (BPNN): A Physics-Infused Neural Network for Phase Retrieval of Meromorphic Functions</p> <p><i>Juncheng Dong, Simiao Ren, Yang Deng, Omar Khatib, Jordan Malof, Mohammadreza Soltani, Willie Padilla, Vahid Tarokh</i></p>	7730
<p>Sound and Complete Neural Network Repair with Minimality and Locality Guarantees</p> <p><i>Feisi Fu, Wenchao Li</i></p>	7747
<p>Finding Biological Plausibility for Adversarially Robust Features Via Metameric Tasks</p> <p><i>Anne Harrington, Arturo Deza</i></p>	7765
<p>Causal Contextual Bandits with Targeted Interventions</p> <p><i>Chandrasekar Subramanian, Balaraman Ravindran</i></p>	7799

<p>Learning Continuous Environment Fields Via Implicit Functions</p> <p><i>Xueting Li, Shalini De Mello, Xiaolong Wang, Ming-Hsuan Yang, Jan Kautz, Sifei Liu</i></p>	7821
<p>Differentiable Gradient Sampling for Learning Implicit 3D Scene Reconstructions from a Single Image</p> <p><i>Shizhan Zhu, Sayna Ebrahimi, Angjoo Kanazawa, Trevor Darrell</i></p>	7839
<p>Disentanglement Analysis with Partial Information Decomposition</p> <p><i>Seiya Tokui, Issei Sato</i></p>	7868
<p>8-Bit Optimizers Via Block-Wise Quantization</p> <p><i>Tim Dettmers, Mike Lewis, Sam Shleifer, Luke Zettlemoyer</i></p>	7891
<p>Exploring Memorization in Adversarial Training</p> <p><i>Yinpeng Dong, Ke Xu, Xiao Yang, Tianyu Pang, Zhijie Deng, Hang Su, Jun Zhu</i></p>	7910
<p>WeakM3D: Towards Weakly Supervised Monocular 3D Object Detection</p> <p><i>Liang Peng, Senbo Yan, Boxi Wu, Zheng Yang, Xiaofei He, Deng Cai</i></p>	7934
<p>The Boltzmann Policy Distribution: Accounting for Systematic Suboptimality in Human Models</p> <p><i>Cassidy Laidlaw, Anca Dragan</i></p>	7954
<p>Diffusion-Based Voice Conversion with Fast Maximum Likelihood Sampling Scheme</p> <p><i>Vadim Popov, Ivan Vovk, Vladimir Gogoryan, Tasnima Sadekova, Mikhail Sergeevich Kudinov, Jiansheng Wei</i></p>	7975
<p>Learning Value Functions from Undirected State-Only Experience</p> <p><i>Matthew Chang, Arjun Gupta, Saurabh Gupta</i></p>	7999
<p>Topological Graph Neural Networks</p> <p><i>Max Horn, Edward De Brouwer, Michael Moor, Yves Moreau, Bastian Rieck, Karsten Borgwardt</i></p>	8025
<p>Pixelated Butterfly: Simple and Efficient Sparse Training for Neural Network Models</p> <p><i>Beidi Chen, Tri Dao, Kaizhao Liang, Jiaming Yang, Zhao Song, Atri Rudra, Christopher Re</i></p>	8052
<p>On the Approximation Properties of Recurrent Encoder-Decoder Architectures</p> <p><i>Zhong Li, Haotian Jiang, Qianxiao Li</i></p>	8095
<p>Context-Aware Sparse Deep Coordination Graphs</p> <p><i>Tonghan Wang, Liang Zeng, Weijun Dong, Qianlan Yang, Yang Yu, Chongjie Zhang</i></p>	8126
<p>How to Deal with Missing Data in Supervised Deep Learning?</p> <p><i>Niels Bruun Ipsen, Pierre-Alexandre Mattei, Jes Frellsen</i></p>	8149
<p>Transferable Adversarial Attack Based on Integrated Gradients</p> <p><i>Yi Huang, Adams Wai-Kin Kong</i></p>	8179
<p>Learning Altruistic Behaviours in Reinforcement Learning Without External Rewards</p> <p><i>Tim Franzmeyer, Mateusz Malinowski, Joao F. Henriques</i></p>	8205

<p>HyAR: Addressing Discrete-Continuous Action Reinforcement Learning Via Hybrid Action Representation</p> <p>Boyan Li, Hongyao Tang, Yan Zheng, Jianye Hao, Pengyi Li, Zhen Wang, Zhaopeng Meng, Li Wang</p>	8225
<p>Understanding Over-Squashing and Bottlenecks on Graphs Via Curvature</p> <p>Jake Topping, Francesco Di Giovanni, Benjamin Paul Chamberlain, Xiaowen Dong, Michael M. Bronstein</p>	8247
<p>FALCON: Fast Visual Concept Learning by Integrating Images, Linguistic Descriptions, and Conceptual Relations</p> <p>Lingjie Mei, Jiayuan Mao, Ziqi Wang, Chuang Gan, Joshua B. Tenenbaum</p>	8277
<p>CosFormer: Rethinking Softmax in Attention</p> <p>Zhen Qin, Weixuan Sun, Hui Deng, Dongxu Li, Yunshen Wei, Baohong Lv, Junjie Yan, Lingpeng Kong, Yiran Zhong</p>	8304
<p>Givens Coordinate Descent Methods for Rotation Matrix Learning in Trainable Embedding Indexes</p> <p>Yunjiang Jiang, Han Zhang, Yiming Qiu, Yun Xiao, Bo Long, Wen-Yun Yang</p>	8319
<p>Natural Posterior Network: Deep Bayesian Predictive Uncertainty for Exponential Family Distributions</p> <p>Bertrand Charpentier, Oliver Borchert, Daniel Zagner, Simon Geisler, Stephan Gannemann</p>	8334
<p>Exploiting Class Activation Value for Partial-Label Learning</p> <p>Fei Zhang, Lei Feng, Bo Han, Tongliang Liu, Gang Niu, Tao Qin, Masashi Sugiyama</p>	8365
<p>Evaluation Metrics for Graph Generative Models: Problems, Pitfalls, and Practical Solutions</p> <p>Leslie O'Bray, Max Horn, Bastian Rieck, Karsten Borgwardt</p>	8382
<p>Measuring CLEVRness: Black-Box Testing of Visual Reasoning Models</p> <p>Spyridon Mouselinos, Henryk Michalewski, Mateusz Malinowski</p>	8405
<p>Learning Graphon Mean Field Games and Approximate Nash Equilibria</p> <p>Kai Cui, Heinz Koepl</p>	8452
<p>On Robust Prefix-Tuning for Text Classification</p> <p>Zonghan Yang, Yang Liu</p>	8483
<p>DEPTS: Deep Expansion Learning for Periodic Time Series Forecasting</p> <p>Wei Fan, Shun Zheng, Xiaohan Yi, Wei Cao, Yanjie Fu, Jiang Bian, Tie-Yan Liu</p>	8520
<p>Unifying Likelihood-Free Inference with Black-Box Optimization and Beyond</p> <p>Dinghuai Zhang, Jie Fu, Yoshua Bengio, Aaron Courville</p>	8544
<p>Explaining Point Processes by Learning Interpretable Temporal Logic Rules</p> <p>Shuang Li, Mingquan Feng, Lu Wang, Abdelmajid Essofi, Yufeng Cao, Junchi Yan, Le Song</p>	8564

DARA: Dynamics-Aware Reward Augmentation in Offline Reinforcement Learning <i>Jinxin Liu, Zhang Hongyin, Donglin Wang</i>	8587
TAPEX: Table Pre-Training Via Learning a Neural SQL Executor <i>Qian Liu, Bei Chen, Jiaqi Guo, Morteza Ziyadi, Zeqi Lin, Weizhu Chen, Jian-Guang Lou</i>	8618
On-Policy Model Errors in Reinforcement Learning <i>Lukas Froehlich, Maksym Lefarov, Melanie Zeilinger, Felix Berkenkamp</i>	8637
Imitation Learning by Reinforcement Learning <i>Kamil Ciosek</i>	8680
Quantitative Performance Assessment of CNN Units Via Topological Entropy Calculation <i>Yang Zhao, Hao Zhang</i>	8695
Generalized Decision Transformer for Offline Hindsight Information Matching <i>Hiroki Furuta, Yutaka Matsuo, Shixiang Shane Gu</i>	8715
Partial Wasserstein Adversarial Network for Non-Rigid Point Set Registration <i>Ziming Wang, Nan Xue, Ling Lei, Gui-Song Xia</i>	8743
Visual Hyperacuity with Moving Sensor and Recurrent Neural Computations <i>Alexander Rivkind, Or Ram, Eldad Assa, Michael Kreiserman, Ehud Ahissar</i>	8778
Delaunay Component Analysis for Evaluation of Data Representations <i>Petra Poklukar, Vladislav Polianskii, Anastasiia Varava, Florian T. Pokorny, Danica Kragic Jensfelt</i>	8802
Neural Program Synthesis with Query <i>Di Huang, Rui Zhang, Xing Hu, Xishan Zhang, Pengwei Jin, Nan Li, Zidong Du, Qi Guo, Yunji Chen</i>	8825
Reliable Adversarial Distillation with Unreliable Teachers <i>Jianing Zhu, Jiangchao Yao, Bo Han, Jingfeng Zhang, Tongliang Liu, Gang Niu, Jingren Zhou, Jianliang Xu, Hongxia Yang</i>	8849
Temporal Efficient Training of Spiking Neural Network Via Gradient Re-Weighting <i>Shikuang Deng, Yuhang Li, Shanghang Zhang, Shi Gu</i>	8864
SUMNAS: Supernet with Unbiased Meta-Features for Neural Architecture Search <i>Hyeonmin Ha, Ji-Hoon Kim, Semin Park, Byung-Gon Chun</i>	8879
High Probability Generalization Bounds with Fast Rates for Minimax Problems <i>Shaojie Li, Yong Liu</i>	8895
VICReg: Variance-Invariance-Covariance Regularization for Self-Supervised Learning <i>Adrien Bardes, Jean Ponce, Yann Lecun</i>	8941
Gradient Information Matters in Policy Optimization by Back-Propagating Through Model <i>Chongchong Li, Yue Wang, Wei Chen, Yuting Liu, Zhi-Ming Ma, Tie-Yan Liu</i>	8964

<p>VAE Approximation Error: ELBO and Exponential Families</p> <p>Alexander Shekhovtsov, Dmitrij Schlesinger, Boris Flach</p>	8988
<p>Policy Gradients Incorporating the Future</p> <p>David Venuto, Elaine Lau, Doina Precup, Ofir Nachum</p>	9006
<p>Actor-Critic Policy Optimization in a Large-Scale Imperfect-Information Game</p> <p>Haobo Fu, Weiming Liu, Shuang Wu, Yijia Wang, Tao Yang, Kai Li, Junliang Xing, Bin Li, Bo Ma, Qiang Fu, Yang Wei</p>	9019
<p>Using Graph Representation Learning with Schema Encoders to Measure the Severity of Depressive Symptoms</p> <p>Simin Hong, Anthony Cohn, David Crossland Hogg</p>	9047
<p>Subspace Regularizers for Few-Shot Class Incremental Learning</p> <p>Afra Feyza Aky�rek, Ekin Aky�rek, Dery Wijaya, Jacob Andreas</p>	9059
<p>Constrained Policy Optimization Via Bayesian World Models</p> <p>Yarden As, Ilhura Usmanova, Sebastian Curi, Andreas Krause</p>	9078
<p>How Well Does Self-Supervised Pre-Training Perform with Streaming Data?</p> <p>Dapeng Hu, Shipeng Yan, Qizhengqiu Lu, Lanqing Hong, Hailin Hu, Yifan Zhang, Zhenguo Li, Xinchao Wang, Jiashi Feng</p>	9102
<p>Meta Discovery: Learning to Discover Novel Classes Given Very Limited Data</p> <p>Haoang Chi, Feng Liu, Wenjing Yang, Long Lan, Tongliang Liu, Bo Han, Gang Niu, Mingyuan Zhou, Masashi Sugiyama</p>	9125
<p>Peek-A-Boo: What (More) is Disguised in a Randomly Weighted Neural Network, and How to Find it Efficiently</p> <p>Xiaohan Chen, Jason Zhang, Zhangyang Wang</p>	9145
<p>Noisy Feature Mixup</p> <p>Soon Hoe Lim, N. Benjamin Erichson, Francisco Utrera, Winnie Xu, Michael W. Mahoney</p>	9160
<p>Evidential Turing Processes</p> <p>Melih Kandemir, Abdullah Akg�l, Manuel Hausmann, Gozde Unal</p>	9196
<p>Near-Optimal Reward-Free Exploration for Linear Mixture MDPs with Plug-In Solver</p> <p>Xiaoyu Chen, Jiachen Hu, Lin Yang, Liwei Wang</p>	9212
<p>Dive Deeper into Integral Pose Regression</p> <p>Kerui Gu, Linlin Yang, Angela Yao</p>	9247
<p>Constraining Linear-Chain CRFs to Regular Languages</p> <p>Sean Papay, Roman Klinger, Sebastian Pado</p>	9262
<p>Understanding Approximate and Unrolled Dictionary Learning for Pattern Recovery</p> <p>Beno�t Mal�zieux, Thomas Moreau, Matthieu Kowalski</p>	9279

<p>Doubly Adaptive Scaled Algorithm for Machine Learning Using Second-Order Information</p> <p>Majid Jahani, Sergey Rusakov, Zheng Shi, Peter Richtárik, Michael W. Mahoney, Martin Takac</p>	9300
<p>Constrained Physical-Statistics Models for Dynamical System Identification and Prediction</p> <p>Járámie Dona, Marie Dächelle, Patrick Gallinari, Marina Levy</p>	9348
<p>Distributionally Robust Models with Parametric Likelihood Ratios</p> <p>Paul Michel, Tatsunori Hashimoto, Graham Neubig</p>	9377
<p>ADAVI: Automatic Dual Amortized Variational Inference Applied to Pyramidal Bayesian Models</p> <p>Louis Rouillard, Demian Wassermann</p>	9394
<p>Recursive Disentanglement Network</p> <p>Yixuan Chen, Yubin Shi, Dongsheng Li, Yujiang Wang, Mingzhi Dong, Yingying Zhao, Robert P. Dick, Qin Lv, Fan Yang, Li Shang</p>	9431
<p>On the Limitations of Multimodal VAEs</p> <p>Imant Daunhawer, Thomas M. Sutter, Kieran Chin-Cheong, Emanuele Palumbo, Julia E Vogt</p>	9454
<p>The Neural Data Router: Adaptive Control Flow in Transformers Improves Systematic Generalization</p> <p>Rábert Csordás, Kazuki Irie, Jürgen Schmidhuber</p>	9481
<p>Training Structured Neural Networks Through Manifold Identification and Variance Reduction</p> <p>Zih-Syuan Huang, Ching-Pei Lee</p>	9509
<p>Ab-Initio Potential Energy Surfaces by Pairing GNNs with Neural Wave Functions</p> <p>Nicholas Gao, Stephan Gännemann</p>	9538
<p>Ancestral Protein Sequence Reconstruction Using a Tree-Structured Ornstein-Uhlenbeck Variational Autoencoder</p> <p>Lys Sanz Moreta, Ola Rånning, Ahmad Salim Al-Sibahi, Jotun Hein, Douglas Theobald, Thomas Hamelryck</p>	9556
<p>Efficient Active Search for Combinatorial Optimization Problems</p> <p>Andrá Hottung, Yeong-Dae Kwon, Kevin Tierney</p>	9578
<p>Information Prioritization Through Empowerment in Visual Model-Based RL</p> <p>Homanga Bharadhwaj, Mohammad Babaeizadeh, Dumitru Erhan, Sergey Levine</p>	9591
<p>FILIP: Fine-Grained Interactive Language-Image Pre-Training</p> <p>Lewei Yao, Runhui Huang, Lu Hou, Guansong Lu, Minzhe Niu, Hang Xu, Xiaodan Liang, Zhenguo Li, Xin Jiang, Chunjing Xu</p>	9613
<p>Effective Model Sparsification by Scheduled Grow-And-Prune Methods</p> <p>Xiaolong Ma, Minghai Qin, Fei Sun, Zejiang Hou, Kun Yuan, Yi Xu, Yanzhi Wang, Yen-Kuang Chen, Rong Jin, Yuan Xie</p>	9634

<p>Learning Generalizable Representations for Reinforcement Learning Via Adaptive Meta-Learner of Behavioral Similarities</p> <p><i>Jianda Chen, Sinno Pan</i></p>	9659
<p>Lipschitz-Constrained Unsupervised Skill Discovery</p> <p><i>Seohong Park, Jongwook Choi, Jaekyeom Kim, Honglak Lee, Gunhee Kim</i></p>	9679
<p>Efficient Sharpness-Aware Minimization for Improved Training of Neural Networks</p> <p><i>Jiawei Du, Hanshu Yan, Jiashi Feng, Joey Tianyi Zhou, Liangli Zhen, Rick Siow Mong Goh, Vincent Tan</i></p>	9703
<p>The Geometry of Memoryless Stochastic Policy Optimization in Infinite-Horizon POMDPs</p> <p><i>Johannes Maller, Guido Montufar</i></p>	9721
<p>A Class of Short-Term Recurrence Anderson Mixing Methods and Their Applications</p> <p><i>Fuchao Wei, Chenglong Bao, Yang Liu</i></p>	9766
<p>Coherence-Based Label Propagation Over Time Series for Accelerated Active Learning</p> <p><i>Yooju Shin, Susik Yoon, Sundong Kim, Hwanjun Song, Jae-Gil Lee, Byung Suk Lee</i></p>	9815
<p>Plant 'n' Seek: Can You Find the Winning Ticket?</p> <p><i>Jonas Fischer, Rebekka Burkholz</i></p>	9833
<p>CoordX: Accelerating Implicit Neural Representation with a Split MLP Architecture</p> <p><i>Ruofan Liang, Hongyi Sun, Nandita Vijaykumar</i></p>	9861
<p>CoST: Contrastive Learning of Disentangled Seasonal-Trend Representations for Time Series Forecasting</p> <p><i>Gerald Woo, Chenghao Liu, Doyen Sahoo, Akshat Kumar, Steven Hoi</i></p>	9883
<p>Self-Supervised Learning is More Robust to Dataset Imbalance</p> <p><i>Hong Liu, Jeff Z. Haochen, Adrien Gaidon, Tengyu Ma</i></p>	9901
<p>Stiffness-Aware Neural Network for Learning Hamiltonian Systems</p> <p><i>Senwei Liang, Zhongzhan Huang, Hong Zhang</i></p>	9925
<p>DriPP: Driven Point Processes to Model Stimuli Induced Patterns in M/EEG Signals</p> <p><i>Cadric Allain, Alexandre Gramfort, Thomas Moreau</i></p>	9948
<p>On Lottery Tickets and Minimal Task Representations in Deep Reinforcement Learning</p> <p><i>Marc Vischer, Robert Tjarko Lange, Henning Sprekeler</i></p>	9973
<p>Learning Strides in Convolutional Neural Networks</p> <p><i>Rachid Riad, Olivier Teboul, David Grangier, Neil Zeghidour</i></p>	9991
<p>A Generalized Weighted Optimization Method for Computational Learning and Inversion</p> <p><i>Kui Ren, Yunan Yang, Bjarn Engquist</i></p>	10007

<p>πBO: Augmenting Acquisition Functions with User Beliefs for Bayesian Optimization</p> <p>Carl Hvarfner, Danny Stoll, Artur Souza, Marius Lindauer, Frank Hutter, Luigi Nardi</p>	10032
<p>Mastering Visual Continuous Control: Improved Data-Augmented Reinforcement Learning</p> <p>Denis Yarats, Rob Fergus, Alessandro Lazaric, Lerrel Pinto</p>	10062
<p>Explanations of Black-Box Models Based on Directional Feature Interactions</p> <p>Aria Masoomi, Davin Hill, Zhonghui Xu, Craig P Hersh, Edwin K. Silverman, Peter J. Castaldi, Stratis Ioannidis, Jennifer Dy</p>	10075
<p>Compositional Training for End-To-End Deep AUC Maximization</p> <p>Zhuoning Yuan, Zhishuai Guo, Nitesh Chawla, Tianbao Yang</p>	10106
<p>Implicit Bias of Adversarial Training for Deep Neural Networks</p> <p>Bochen Lv, Zhanxing Zhu</p>	10132
<p>The Uncanny Similarity of Recurrence and Depth</p> <p>Avi Schwarzschild, Arjun Gupta, Amin Ghiasi, Micah Goldblum, Tom Goldstein</p>	10167
<p>Learning Transferable Motor Skills with Hierarchical Latent Mixture Policies</p> <p>Dushyant Rao, Fereshteh Sadeghi, Leonard Hasenclever, Markus Wulfmeier, Martina Zambelli, Giulia Vezzani, Dhruva Tirumala, Yusuf Aytar, Josh Merel, Nicolas Heess, Raia Hadsell</p>	10184
<p>Sample Efficient Stochastic Policy Extragradient Algorithm for Zero-Sum Markov Game</p> <p>Ziyi Chen, Shaocong Ma, Yi Zhou</p>	10206
<p>BDDM: Bilateral Denoising Diffusion Models for Fast and High-Quality Speech Synthesis</p> <p>Max W. Y. Lam, Jun Wang, Dan Su, Dong Yu</p>	10245
<p>Learning Pruning-Friendly Networks Via Frank-Wolfe: One-Shot, Any-Sparsity, and No Retraining</p> <p>Miao Lu, Xiaolong Luo, Tianlong Chen, Wuyang Chen, Dong Liu, Zhangyang Wang</p>	10267
<p>Sample Efficient Deep Reinforcement Learning Via Uncertainty Estimation</p> <p>Vincent Mai, Kaustubh Mani, Liam Paull</p>	10283
<p>Scalable One-Pass Optimisation of High-Dimensional Weight-Update Hyperparameters by Implicit Differentiation</p> <p>Ross M Clarke, Elre Talea Oldewage, Josã Miguel Hernã;ndez-Lobato</p>	10307
<p>Solving Inverse Problems in Medical Imaging with Score-Based Generative Models</p> <p>Yang Song, Liyue Shen, Lei Xing, Stefano Ermon</p>	10348
<p>Finding an Unsupervised Image Segmenter in Each of Your Deep Generative Models</p> <p>Luke Melas-Kyriazi, Christian Rupprecht, Iro Laina, Andrea Vedaldi</p>	10366
<p>Open-Set Recognition: A Good Closed-Set Classifier is All You Need</p> <p>Sagar Vaze, Kai Han, Andrea Vedaldi, Andrew Zisserman</p>	10398

<p>Controlling the Complexity and Lipschitz Constant Improves Polynomial Nets</p> <p>Zhenyu Zhu, Fabian Latorre, Grigorios Chrysos, Volkan Cevher</p>	10424
<p>Sparse Attention with Learning to Hash</p> <p>Zhiqing Sun, Yiming Yang, Shinjae Yoo</p>	10468
<p>A Fine-Grained Analysis on Distribution Shift</p> <p>Olivia Wiles, Sven Gowal, Florian Stimberg, Sylvestre-Alvise Rebuffi, Ira Ktena, Krishnamurthy Dj Dvijotham, Ali Taylan Cemgil</p>	10488
<p>Bridging Recommendation and Marketing Via Recurrent Intensity Modeling</p> <p>Yifei Ma, Ge Liu, Anoop Deoras</p>	10503
<p>CodeTrek: Flexible Modeling of Code Using an Extensible Relational Representation</p> <p>Pardis Pashakhanloo, Aaditya Naik, Yuepeng Wang, Hanjun Dai, Petros Maniatis, Mayur Naik</p>	10523
<p>Semi-Relaxed Gromov-Wasserstein Divergence and Applications on Graphs</p> <p>Cădric Vincent-Cuaz, Rămi Flamary, Marco Corneli, Titouan Vayer, Nicolas Courty</p>	10546
<p>Label Leakage and Protection in Two-Party Split Learning</p> <p>Oscar Li, Jiankai Sun, Xin Yang, Weihao Gao, Hongyi Zhang, Junyuan Xie, Virginia Smith, Chong Wang</p>	10574
<p>CADDA: Class-Wise Automatic Differentiable Data Augmentation for EEG Signals</p> <p>Cădric Rommel, Thomas Moreau, Joseph Paillard, Alexandre Gramfort</p>	10601
<p>A Theoretical Analysis on Feature Learning in Neural Networks: Emergence from Inputs and Advantage Over Fixed Features</p> <p>Zhenmei Shi, Junyi Wei, Yingyu Liang</p>	10625
<p>Neural Deep Equilibrium Solvers</p> <p>Shaojie Bai, Vladlen Koltun, J Zico Kolter</p>	10706
<p>Case-Based Reasoning for Better Generalization in Textual Reinforcement Learning</p> <p>Mattia Atzeni, Shehzaad Zuzar Dhuliawala, Keerthiram Murugesan, Mrinmaya Sachan</p>	10725
<p>GRAND++: Graph Neural Diffusion with a Source Term</p> <p>Matthew Thorpe, Tan Minh Nguyen, Hedi Xia, Thomas Strohmer, Andrea Bertozzi, Stanley Osher, Bao Wang</p>	10746
<p>Autonomous Reinforcement Learning: Formalism and Benchmarking</p> <p>Archit Sharma, Kelvin Xu, Nikhil Sardana, Abhishek Gupta, Karol Hausman, Sergey Levine, Chelsea Finn</p>	10767
<p>Learning Causal Models from Conditional Moment Restrictions by Importance Weighting</p> <p>Masahiro Kato, Masaaki Imaizumi, Kenichiro McAlinn, Shota Yasui, Haruo Kakehi</p>	10784
<p>Churn Reduction Via Distillation</p> <p>Heinrich Jiang, Harikrishna Narasimhan, Dara Bahri, Andrew Cotter, Afshin Rostamizadeh</p>	10810

<p>Fast Topological Clustering with Wasserstein Distance</p> <p>Tananun Songdechakraiwt, Bryan M Krause, Matthew I Banks, Kirill V Nourski, Barry D Van Veen</p>	10863
<p>Learning Curves for Gaussian Process Regression with Power-Law Priors and Targets</p> <p>Hui Jin, Pradeep Kr. Banerjee, Guido Montufar</p>	10881
<p>Iterative Refinement Graph Neural Network for Antibody Sequence-Structure Co-Design</p> <p>Wengong Jin, Jeremy Wohlwend, Regina Barzilay, Tommi S. Jaakkola</p>	10934
<p>Frequency-Aware SGD for Efficient Embedding Learning with Provable Benefits</p> <p>Yan Li, Dhruv Choudhary, Xiaohan Wei, Baichuan Yuan, Bhargav Bhushanam, Tuo Zhao, Guanghui Lan</p>	10948
<p>What Do We Mean by Generalization in Federated Learning?</p> <p>Honglin Yuan, Warren Richard Morningstar, Lin Ning, Karan Singhal</p>	10973
<p>FedChain: Chained Algorithms for Near-Optimal Communication Cost in Federated Learning</p> <p>Charlie Hou, Kiran Koshy Thekumparampil, Giulia Fanti, Sewoong Oh</p>	10997
<p>Amortized Tree Generation for Bottom-Up Synthesis Planning and Synthesizable Molecular Design</p> <p>Wenhao Gao, Roc�o Mercado, Connor W. Coley</p>	11046
<p>Learning Fast, Learning Slow: A General Continual Learning Method Based on Complementary Learning System</p> <p>Elahe Arani, Fahad Sarfraz, Bahram Zonooz</p>	11071
<p>Eliminating Sharp Minima from SGD with Truncated Heavy-Tailed Noise</p> <p>Xingyu Wang, Sewoong Oh, Chang-Han Rhee</p>	11093
<p>Differentiable Scaffolding Tree for Molecule Optimization</p> <p>Tianfan Fu, Wenhao Gao, Cao Xiao, Jacob Yasonik, Connor W. Coley, Jimeng Sun</p>	11250
<p>An Explanation of In-Context Learning as Implicit Bayesian Inference</p> <p>Sang Michael Xie, Aditi Raghunathan, Percy Liang, Tengyu Ma</p>	11282
<p>Learning Audio-Visual Speech Representation by Masked Multimodal Cluster Prediction</p> <p>Bowen Shi, Wei-Ning Hsu, Kushal Lakhotia, Abdelrahman Mohamed</p>	11307
<p>Hierarchical Variational Memory for Few-Shot Learning Across Domains</p> <p>Yingjun Du, Xiantong Zhen, Ling Shao, Cees G. M. Snoek</p>	11331
<p>Taming Sparsely Activated Transformer with Stochastic Experts</p> <p>Simiao Zuo, Xiaodong Liu, Jian Jiao, Young Jin Kim, Hany Hassan, Ruofei Zhang, Jianfeng Gao, Tuo Zhao</p>	11348
<p>PER-ETD: A Polynomially Efficient Emphatic Temporal Difference Learning Method</p> <p>Ziwei Guan, Tengyu Xu, Yingbin Liang</p>	11365

<p>Pre-Training Molecular Graph Representation with 3D Geometry</p> <p><i>Shengchao Liu, Hanchen Wang, Weiyang Liu, Joan Lasenby, Hongyu Guo, Jian Tang</i></p>	11402
<p>On the Existence of Universal Lottery Tickets</p> <p><i>Rebekka Burkholz, Nilanjana Laha, Rajarshi Mukherjee, Alkis Gotovos</i></p>	11434
<p>DiffSkill: Skill Abstraction from Differentiable Physics for Deformable Object Manipulations with Tools</p> <p><i>Xingyu Lin, Zhiao Huang, Yunzhu Li, Joshua B. Tenenbaum, David Held, Chuang Gan</i></p>	11459
<p>On Incorporating Inductive Biases into VAEs</p> <p><i>Ning Miao, Emile Mathieu, Siddharth N, Yee Whye Teh, Tom Rainforth</i></p>	11473
<p>An Autoregressive Flow Model for 3D Molecular Geometry Generation from Scratch</p> <p><i>Youzhi Luo, Shuiwang Ji</i></p>	11496
<p>Scalable Sampling for Nonsymmetric Determinantal Point Processes</p> <p><i>Insu Han, Mike Gartrell, Jennifer Gillenwater, Elvis Dohmatob, Amin Karbasi</i></p>	11514
<p>Eigencurve: Optimal Learning Rate Schedule for SGD on Quadratic Objectives with Skewed Hessian Spectrums</p> <p><i>Rui Pan, Haishan Ye, Tong Zhang</i></p>	11533
<p>Fast Model Editing at Scale</p> <p><i>Eric Mitchell, Charles Lin, Antoine Bosselut, Chelsea Finn, Christopher D Manning</i></p>	11576
<p>Transformer Embeddings of Irregularly Spaced Events and Their Participants</p> <p><i>Hongyuan Mei, Chenghao Yang, Jason Eisner</i></p>	11597
<p>Leveraging Automated Unit Tests for Unsupervised Code Translation</p> <p><i>Baptiste Roziere, Jie Zhang, Francois Charton, Mark Harman, Gabriel Synnaeve, Guillaume Lample</i></p>	11618
<p>Contact Points Discovery for Soft-Body Manipulations with Differentiable Physics</p> <p><i>Sizhe Li, Zhiao Huang, Tao Du, Hao Su, Joshua B. Tenenbaum, Chuang Gan</i></p>	11638
<p>Federated Learning from Only Unlabeled Data with Class-Conditional-Sharing Clients</p> <p><i>Nan Lu, Zhao Wang, Xiaoxiao Li, Gang Niu, Qi Dou, Masashi Sugiyama</i></p>	11656
<p>Analyzing and Improving the Optimization Landscape of Noise-Contrastive Estimation</p> <p><i>Bingbin Liu, Elan Rosenfeld, Pradeep Kumar Ravikumar, Andrej Risteski</i></p>	11678
<p>Learnability Lock: Authorized Learnability Control Through Adversarial Invertible Transformations</p> <p><i>Weiqi Peng, Jinghui Chen</i></p>	11716
<p>A Global Convergence Theory for Deep ReLU Implicit Networks Via Over-Parameterization</p> <p><i>Tianxiang Gao, Hailiang Liu, Jia Liu, Hriday Rajan, Hongyang Gao</i></p>	11735

<p>When, Why, and Which Pretrained GANs Are Useful?</p> <p><i>Timofey Grigoryev, Andrey Voynov, Artem Babenko</i></p>	11763
<p>FastSHAP: Real-Time Shapley Value Estimation</p> <p><i>Neil Jethani, Mukund Sudarshan, Ian Connick Covert, Su-In Lee, Rajesh Ranganath</i></p>	11781
<p>RelaxLoss: Defending Membership Inference Attacks Without Losing Utility</p> <p><i>Dingfan Chen, Ning Yu, Mario Fritz</i></p>	11804
<p>Model Agnostic Interpretability for Multiple Instance Learning</p> <p><i>Joseph Early, Christine Evers, Sarvapali Ramchurn</i></p>	11832
<p>CrowdPlay: Crowdsourcing Human Demonstrations for Offline Learning</p> <p><i>Matthias Gerstgrasser, Rakshit Trivedi, David C. Parkes</i></p>	11857
<p>Monotonic Differentiable Sorting Networks</p> <p><i>Felix Petersen, Christian Borgelt, Hilde Kuehne, Oliver Deussen</i></p>	11881
<p>Training Data Generating Networks: Shape Reconstruction Via Bi-Level Optimization</p> <p><i>Biao Zhang, Peter Wonka</i></p>	11898
<p>Mention Memory: Incorporating Textual Knowledge into Transformers Through Entity Mention Attention</p> <p><i>Michiel De Jong, Yury Zemlyanskiy, Nicholas Fitzgerald, Fei Sha, William W. Cohen</i></p>	11918
<p>How to Robustify Black-Box ML Models? a Zeroth-Order Optimization Perspective</p> <p><i>Yimeng Zhang, Yuguang Yao, Jinghan Jia, Jinfeng Yi, Mingyi Hong, Shiyu Chang, Sijia Liu</i></p>	11933
<p>Explainable GNN-Based Models Over Knowledge Graphs</p> <p><i>David Jaime Tena Cucala, Bernardo Cuenca Grau, Egor V. Kostylev, Boris Motik</i></p>	11948
<p>Einops: Clear and Reliable Tensor Manipulations with Einstein-Like Notation</p> <p><i>Alex Rogozhnikov</i></p>	11967
<p>Spike-Inspired Rank Coding for Fast and Accurate Recurrent Neural Networks</p> <p><i>Alan Jeffares, Qinghai Guo, Pontus Stenetorp, Timoleon Moraitis</i></p>	11988
<p>The Evolution of Uncertainty of Learning in Games</p> <p><i>Yun Kuen Cheung, Georgios Piliouras, Yixin Tao</i></p>	12005
<p>FILM: Following Instructions in Language with Modular Methods</p> <p><i>So Yeon Min, Devendra Singh Chaplot, Pradeep Kumar Ravikumar, Yonatan Bisk, Ruslan Salakhutdinov</i></p>	12021
<p>Graph-Guided Network for Irregularly Sampled Multivariate Time Series</p> <p><i>Xiang Zhang, Marko Zeman, Theodoros Tsiligkaridis, Marinka Zitnik</i></p>	12038
<p>One After Another: Learning Incremental Skills for a Changing World</p> <p><i>Nur Muhammad Mahi Shafiullah, Lerrel Pinto</i></p>	12059

Generalization of Neural Combinatorial Solvers Through the Lens of Adversarial Robustness <i>Simon Geisler, Johanna Sommer, Jan Schuchardt, Aleksandar Bojchevski, Stephan Gännemann</i>	12087
EE-Net: Exploitation-Exploration Neural Networks in Contextual Bandits <i>Yikun Ban, Yuchen Yan, Arindam Banerjee, Jingrui He</i>	12108
PAC Prediction Sets Under Covariate Shift <i>Sangdon Park, Edgar Dobriban, Insup Lee, Osbert Bastani</i>	12131
Object Pursuit: Building a Space of Objects Via Discriminative Weight Generation <i>Chuanyu Pan, Yanchao Yang, Kaichun Mo, Yueqi Duan, Leonidas Guibas</i>	12164
Transformer-Based Transform Coding <i>Yinhao Zhu, Yang Yang, Taco Cohen</i>	12188
Half-Inverse Gradients for Physical Deep Learning <i>Patrick Schnell, Philipp Holl, Nils Thuerey</i>	12223
Unified Visual Transformer Compression <i>Shixing Yu, Tianlong Chen, Jiayi Shen, Huan Yuan, Jianchao Tan, Sen Yang, Ji Liu, Zhangyang Wang</i>	12245
Consistent Counterfactuals for Deep Models <i>Emily Black, Zifan Wang, Matt Fredrikson</i>	12262
Neural Methods for Logical Reasoning Over Knowledge Graphs <i>Alfonso Amayuelas, Shuai Zhang, Xi Susie Rao, Ce Zhang</i>	12284
Learning 3D Representations of Molecular Chirality with Invariance to Bond Rotations <i>Keir Adams, Lagnajit Pattanaik, Connor W. Coley</i>	12298
Conditional Image Generation by Conditioning Variational Auto-Encoders <i>William Harvey, Saeid Naderiparizi, Frank Wood</i>	12335
Learn Locally, Correct Globally: A Distributed Algorithm for Training Graph Neural Networks <i>Morteza Ramezani, Weilin Cong, Mehrdad Mahdavi, Mahmut Kandemir, Anand Sivasubramaniam</i>	12372
Effect of Scale on Catastrophic Forgetting in Neural Networks <i>Vinay Venkatesh Ramasesh, Aitor Lewkowycz, Ethan Dyer</i>	12385
On the Benefits of Maximum Likelihood Estimation for Regression and Forecasting <i>Pranjal Awasthi, Abhimanyu Das, Rajat Sen, Ananda Theertha Suresh</i>	12418
TRAIL: Near-Optimal Imitation Learning with Suboptimal Data <i>Mengjiao Yang, Sergey Levine, Ofir Nachum</i>	12450
Efficient Computation of Deep Nonlinear Infinite-Width Neural Networks that Learn Features <i>Greg Yang, Michael Santacroce, Edward J Hu</i>	12475

<p>Which Shortcut Cues Will DNNs Choose? a Study from the Parameter-Space Perspective</p> <p><i>Luca Scimeca, Seong Joon Oh, Sanghyuk Chun, Michael Poli, Sangdoon Yun</i></p>	12499
<p>LoRA: Low-Rank Adaptation of Large Language Models</p> <p><i>Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen</i></p>	12513
<p>Interpretable Unsupervised Diversity Denoising and Artefact Removal</p> <p><i>Mangal Prakash, Mauricio Delbracio, Peyman Milanfar, Florian Jug</i></p>	12526
<p>Hindsight Foresight Relabeling for Meta-Reinforcement Learning</p> <p><i>Michael Wan, Jian Peng, Tanmay Gangwani</i></p>	12552
<p>Step-Unrolled Denoising Autoencoders for Text Generation</p> <p><i>Nikolay Savinov, Junyoung Chung, Mikolaj Binkowski, Erich Elsen, Aaron Van Den Oord</i></p>	12570
<p>Possibility Before Utility: Learning and Using Hierarchical Affordances</p> <p><i>Robby Costales, Shariq Iqbal, Fei Sha</i></p>	12590
<p>Iterated Reasoning with Mutual Information in Cooperative and Byzantine Decentralized Teaming</p> <p><i>Sachin G Konan, Esmaeil Seraj, Matthew Gombolay</i></p>	12610
<p>P-Adapters: Robustly Extracting Factual Information from Language Models with Diverse Prompts</p> <p><i>Benjamin Newman, Prafulla Kumar Choubey, Nazneen Rajani</i></p>	12639
<p>Differentially Private Fine-Tuning of Language Models</p> <p><i>Da Yu, Saurabh Naik, Arturs Backurs, Sivakanth Gopi, Huseyin A Inan, Gautam Kamath, Janardhan Kulkarni, Yin Tat Lee, Andre Manoel, Lukas Wutschitz, Sergey Yekhanin, Huishuai Zhang</i></p>	12656
<p>A Fine-Tuning Approach to Belief State Modeling</p> <p><i>Samuel Sokota, Hengyuan Hu, David J Wu, J Zico Kolter, Jakob Nicolaus Foerster, Noam Brown</i></p>	12675
<p>An Agnostic Approach to Federated Learning with Class Imbalance</p> <p><i>Zebang Shen, Juan Cervino, Hamed Hassani, Alejandro Ribeiro</i></p>	12690
<p>Path Auxiliary Proposal for MCMC in Discrete Space</p> <p><i>Haoran Sun, Hanjun Dai, Wei Xia, Arun Ramamurthy</i></p>	12702
<p>Optimizer Amalgamation</p> <p><i>Tianshu Huang, Tianlong Chen, Sijia Liu, Shiyu Chang, Lisa Amini, Zhangyang Wang</i></p>	12724
<p>Curriculum Learning as a Tool to Uncover Learning Principles in the Brain</p> <p><i>Daniel R. Kepple, Rainer Engelken, Kanaka Rajan</i></p>	12745
<p>Learning to Dequantise with Truncated Flows</p> <p><i>Shawn Tan, Chin-Wei Huang, Alessandro Sordoni, Aaron Courville</i></p>	12757

<p>How Unlabeled Data Improve Generalization in Self-Training? a One-Hidden-Layer Theoretical Analysis</p> <p><i>Shuai Zhang, Meng Wang, Sijia Liu, Pin-Yu Chen, Jinjun Xiong</i></p>	12773
<p>Predicting Physics in Mesh-Reduced Space with Temporal Attention</p> <p><i>Xu Han, Han Gao, Tobias Pfaff, Jian-Xun Wang, Liping Liu</i></p>	12809
<p>Model Zoo: A Growing Brain that Learns Continually</p> <p><i>Rahul Ramesh, Pratik Chaudhari</i></p>	12831
<p>Responsible Disclosure of Generative Models Using Scalable Fingerprinting</p> <p><i>Ning Yu, Vladislav Skripniuk, Dingfan Chen, Larry S. Davis, Mario Fritz</i></p>	12858
<p>Path Integral Sampler: A Stochastic Control Approach for Sampling</p> <p><i>Qinsheng Zhang, Yongxin Chen</i></p>	12874
<p>Trans-Encoder: Unsupervised Sentence-Pair Modelling Through Self- And Mutual-Distillations</p> <p><i>Fangyu Liu, Yunlong Jiao, Jordan Massiah, Emine Yilmaz, Serhii Havrylov</i></p>	12900
<p>Scarf: Self-Supervised Contrastive Learning Using Random Feature Corruption</p> <p><i>Dara Bahri, Heinrich Jiang, Yi Tay, Donald Metzler</i></p>	12918
<p>$\mathcal{SO}(2)$-Equivariant Reinforcement Learning</p> <p><i>Dian Wang, Robin Walters, Robert Platt</i></p>	12942
<p>T-WaveNet: A Tree-Structured Wavelet Neural Network for Time Series Signal Analysis</p> <p><i>Minhao Liu, Ailing Zeng, Qiuxia Lai, Ruiyuan Gao, Min Li, Jing Qin, Qiang Xu</i></p>	12961
<p>Connectome-Constrained Latent Variable Model of Whole-Brain Neural Activity</p> <p><i>Lu Mi, Richard Xu, Sridhama Prakhya, Albert Lin, Nir Shavit, Aravinthan Samuel, Srinivas C Turaga</i></p>	12973
<p>Towards Deepening Graph Neural Networks: A GNTK-Based Optimization Perspective</p> <p><i>Wei Huang, Yayong Li, Weitao Du, Richard Xu, Jie Yin, Ling Chen, Miao Zhang</i></p>	13009
<p>Energy-Inspired Molecular Conformation Optimization</p> <p><i>Jiaqi Guan, Wesley Wei Qian, Qiang Liu, Wei-Ying Ma, Jianzhu Ma, Jian Peng</i></p>	13035
<p>Understanding Domain Randomization for Sim-To-Real Transfer</p> <p><i>Xiaoyu Chen, Jiachen Hu, Chi Jin, Lihong Li, Liwei Wang</i></p>	13059
<p>Granger Causal Inference on DAGs Identifies Genomic Loci Regulating Transcription</p> <p><i>Alexander P Wu, Rohit Singh, Bonnie Berger</i></p>	13087
<p>ConFeSS: A Framework for Single Source Cross-Domain Few-Shot Learning</p> <p><i>Debasmit Das, Sungrack Yun, Fatih Porikli</i></p>	13106
<p>Frame Averaging for Invariant and Equivariant Network Design</p> <p><i>Omri Puny, Matan Atzmon, Edward J. Smith, Ishan Misra, Aditya Grover, Heli Ben-Hamu, Yaron Lipman</i></p>	13124

<p>Conditional Contrastive Learning with Kernel</p> <p><i>Yao-Hung Hubert Tsai, Tianqin Li, Martin Q. Ma, Han Zhao, Kun Zhang, Louis-Philippe Morency, Ruslan Salakhutdinov</i></p>	13145
<p>Unsupervised Learning of Full-Waveform Inversion: Connecting CNN and Partial Differential Equation in a Loop</p> <p><i>Peng Jin, Xitong Zhang, Yinpeng Chen, Sharon X Huang, Zicheng Liu, Youzuo Lin</i></p>	13162
<p>BAM: Bayes with Adaptive Memory</p> <p><i>Josue Nassar, Jennifer Rogers Brennan, Ben Evans, Kendall Lowrey</i></p>	13182
<p>An Experimental Design Perspective on Model-Based Reinforcement Learning</p> <p><i>Viraj Mehta, Biswajit Paria, Jeff Schneider, Stefano Ermon, Willie Neiswanger</i></p>	13199
<p>Encoding Weights of Irregular Sparsity for Fixed-To-Fixed Model Compression</p> <p><i>Bae Seong Park, Se Jung Kwon, Daehwan Oh, Byeongwook Kim, Dongsoo Lee</i></p>	13216
<p>Learning Weakly-Supervised Contrastive Representations</p> <p><i>Yao-Hung Hubert Tsai, Tianqin Li, Weixin Liu, Peiyuan Liao, Ruslan Salakhutdinov, Louis-Philippe Morency</i></p>	13238
<p>Tuformer: Data-Driven Design of Transformers for Improved Generalization Or Efficiency</p> <p><i>Xiaoyu Liu, Jiahao Su, Furong Huang</i></p>	13261
<p>Capacity of Group-Invariant Linear Readouts from Equivariant Representations: How Many Objects Can Be Linearly Classified Under All Possible Views?</p> <p><i>Matthew Farrell, Blake Bordelon, Shubhendu Trivedi, Cengiz Pehlevan</i></p>	13282
<p>Lossless Compression with Probabilistic Circuits</p> <p><i>Anji Liu, Stephan Mandt, Guy Van Den Broeck</i></p>	13304
<p>An Operator Theoretic View on Pruning Deep Neural Networks</p> <p><i>William T Redman, Maria Fonoberova, Ryan Mohr, Yannis Kevrekidis, Igor Mezic</i></p>	13324
<p>Learning Object-Oriented Dynamics for Planning from Text</p> <p><i>Guiliang Liu, Ashutosh Adhikari, Amir-Massoud Farahmand, Pascal Poupart</i></p>	13338
<p>Neural Networks as Kernel Learners: The Silent Alignment Effect</p> <p><i>Alexander Atanasov, Blake Bordelon, Cengiz Pehlevan</i></p>	13358
<p>Representation Learning for Online and Offline RL in Low-Rank MDPs</p> <p><i>Masatoshi Uehara, Xuezhou Zhang, Wen Sun</i></p>	13387
<p>When Can We Learn General-Sum Markov Games with a Large Number of Players Sample-Efficiently?</p> <p><i>Ziang Song, Song Mei, Yu Bai</i></p>	13414
<p>The Three Stages of Learning Dynamics in High-Dimensional Kernel Methods</p> <p><i>Nikhil Ghosh, Song Mei, Bin Yu</i></p>	13459

<p>On Feature Learning in Neural Networks with Global Convergence Guarantees</p> <p>Zhengdao Chen, Eric Vanden-Eijnden, Joan Bruna</p>	13501
<p>GDA-AM: ON THE EFFECTIVENESS OF SOLVING MIN-IMAX OPTIMIZATION VIA ANDERSON MIXING</p> <p>Huan He, Shifan Zhao, Yuanzhe Xi, Joyce Ho, Yousef Saad</p>	13527
<p>R5: Rule Discovery with Reinforced and Recurrent Relational Reasoning</p> <p>Shengyao Lu, Bang Liu, Keith G Mills, Shangling Jui, Di Niu</p>	13558
<p>Training Transition Policies Via Distribution Matching for Complex Tasks</p> <p>Ju-Seung Byun, Andrew Perrault</p>	13577
<p>DictFormer: Tiny Transformer with Shared Dictionary</p> <p>Qian Lou, Ting Hua, Yen-Chang Hsu, Yilin Shen, Hongxia Jin</p>	13591
<p>Controlling Directions Orthogonal to a Classifier</p> <p>Yilun Xu, Hao He, Tianxiao Shen, Tommi S. Jaakkola</p>	13607
<p>Switch-GLAT: Multilingual Parallel Machine Translation Via Code-Switch Decoder</p> <p>Zhenqiao Song, Hao Zhou, Lihua Qian, Jingjing Xu, Shanbo Cheng, Mingxuan Wang, Lei Li</p>	13628
<p>Score-Based Generative Modeling with Critically-Damped Langevin Diffusion</p> <p>Tim Dockhorn, Arash Vahdat, Karsten Kreis</p>	13643
<p>Self-Supervision Enhanced Feature Selection with Correlated Gates</p> <p>Changhee Lee, Fergus Imrie, Mihaela Van Der Schaar</p>	13697
<p>The Unreasonable Effectiveness of Random Pruning: Return of the Most Naive Baseline for Sparse Training</p> <p>Shiwei Liu, Tianlong Chen, Xiaohan Chen, Li Shen, Decebal Constantin Mocanu, Zhangyang Wang, Mykola Pechenizkiy</p>	13723
<p>Generalizing Few-Shot NAS with Gradient Matching</p> <p>Shoukang Hu, Ruochen Wang, Lanqing Hong, Zhenguo Li, Cho-Jui Hsieh, Jiashi Feng</p>	13745
<p>Post Hoc Explanations May Be Ineffective for Detecting Unknown Spurious Correlation</p> <p>Julius Adebayo, Michael Muelly, Harold Abelson, Been Kim</p>	13768
<p>SDEdit: Guided Image Synthesis and Editing with Stochastic Differential Equations</p> <p>Chenlin Meng, Yutong He, Yang Song, Jiaming Song, Jiajun Wu, Jun-Yan Zhu, Stefano Ermon</p>	13781
<p>No Parameters Left Behind: Sensitivity Guided Adaptive Learning Rate for Training Large Transformer Models</p> <p>Chen Liang, Haoming Jiang, Simiao Zuo, Pengcheng He, Xiaodong Liu, Jianfeng Gao, Weizhu Chen, Tuo Zhao</p>	13814

<p>Variational Autoencoders in the Presence of Low-Dimensional Data: Landscape and Implicit Bias</p> <p><i>Frederic Koehler, Viraj Mehta, Chenghui Zhou, Andrej Risteski</i></p>	13832
<p>Towards Understanding the Data Dependency of Mixup-Style Training</p> <p><i>Muthu Chidambaram, Xiang Wang, Yuzheng Hu, Chenwei Wu, Rong Ge</i></p>	13851
<p>Deep Learning Without Shortcuts: Shaping the Kernel with Tailored Rectifiers</p> <p><i>Guodong Zhang, Aleksandar Botev, James Martens</i></p>	13877
<p>On Redundancy and Diversity in Cell-Based Neural Architecture Search</p> <p><i>Xingchen Wan, Binxin Ru, Pedro M Esperansa, Zhenguo Li</i></p>	13902
<p>An Information Fusion Approach to Learning with Instance-Dependent Label Noise</p> <p><i>Zhimeng Jiang, Kaixiong Zhou, Zirui Liu, Li Li, Rui Chen, Soo-Hyun Choi, Xia Hu</i></p>	13927
<p>ComPhy: Compositional Physical Reasoning of Objects and Events from Videos</p> <p><i>Zhenfang Chen, Kexin Yi, Yunzhu Li, Mingyu Ding, Antonio Torralba, Joshua B. Tenenbaum, Chuang Gan</i></p>	13947
<p>Cross-Lingual Transfer with Class-Weighted Language-Invariant Representations</p> <p><i>Ruicheng Xian, Heng Ji, Han Zhao</i></p>	13966
<p>FairCal: Fairness Calibration for Face Verification</p> <p><i>Tiago Salvador, Stephanie Cairns, Vikram Voleti, Noah Marshall, Adam M Oberman</i></p>	13988
<p>Finite-Time Convergence and Sample Complexity of Multi-Agent Actor-Critic Reinforcement Learning with Average Reward</p> <p><i>Fnu Hairi, Jia Liu, Songtao Lu</i></p>	14018
<p>Graph Auto-Encoder Via Neighborhood Wasserstein Reconstruction</p> <p><i>Mingyue Tang, Pan Li, Carl Yang</i></p>	14042
<p>On the Importance of Firth Bias Reduction in Few-Shot Classification</p> <p><i>Saba Ghaffari, Ehsan Saleh, David Forsyth, Yu-Xiong Wang</i></p>	14059
<p>Towards Understanding Generalization Via Decomposing Excess Risk Dynamics</p> <p><i>Jiaye Teng, Jianhao Ma, Yang Yuan</i></p>	14088
<p>The Close Relationship Between Contrastive Learning and Meta-Learning</p> <p><i>Renkun Ni, Manli Shu, Hossein Souri, Micah Goldblum, Tom Goldstein</i></p>	14119
<p>Policy Smoothing for Provably Robust Reinforcement Learning</p> <p><i>Aounon Kumar, Alexander Levine, Soheil Feizi</i></p>	14134
<p>GeoDiff: A Geometric Diffusion Model for Molecular Conformation Generation</p> <p><i>Minkai Xu, Lantao Yu, Yang Song, Chence Shi, Stefano Ermon, Jian Tang</i></p>	14163

Zero-Shot Self-Supervised Learning for MRI Reconstruction <i>Burhaneddin Yaman, Seyed Amir Hossein Hosseini, Mehmet Akcakaya</i>	14182
Large Language Models Can Be Strong Differentially Private Learners <i>Xuechen Li, Florian Tramer, Percy Liang, Tatsunori Hashimoto</i>	14198
Autoregressive Quantile Flows for Predictive Uncertainty Estimation <i>Phillip Si, Allan Bishop, Volodymyr Kuleshov</i>	14228
Graph Neural Networks with Learnable Structural and Positional Representations <i>Vijay Prakash Dwivedi, Anh Tuan Luu, Thomas Laurent, Yoshua Bengio, Xavier Bresson</i>	14245
Maximizing Ensemble Diversity in Deep Reinforcement Learning <i>Hassam Sheikh, Mariano Phielipp, Ladislau Boloni</i>	14270
Adversarial Unlearning of Backdoors Via Implicit Hypergradient <i>Yi Zeng, Si Chen, Won Park, Zhuoqing Mao, Ming Jin, Ruoxi Jia</i>	14295
Efficiently Modeling Long Sequences with Structured State Spaces <i>Albert Gu, Karan Goel, Christopher Re</i>	14323
Graph-Augmented Normalizing Flows for Anomaly Detection of Multiple Time Series <i>Enyan Dai, Jie Chen</i>	14350
Learning Neural Contextual Bandits Through Perturbed Rewards <i>Yiling Jia, Weitong Zhang, Dongruo Zhou, Quanquan Gu, Hongning Wang</i>	14366
PipeGCN: Efficient Full-Graph Training of Graph Convolutional Networks with Pipelined Feature Communication <i>Cheng Wan, Youjie Li, Cameron R. Wolfe, Anastasios Kyrillidis, Nam Sung Kim, Yingyan Lin</i>	14385
Demystifying Limited Adversarial Transferability in Automatic Speech Recognition Systems <i>Hadi Abdullah, Aditya Karlekar, Vincent Bindschaedler, Patrick Traynor</i>	14409
Large Learning Rate Tames Homogeneity: Convergence and Balancing Effect <i>Yuqing Wang, Minshuo Chen, Tuo Zhao, Molei Tao</i>	14426
How Much Can CLIP Benefit Vision-And-Language Tasks? <i>Sheng Shen, Liunian Harold Li, Hao Tan, Mohit Bansal, Anna Rohrbach, Kai-Wei Chang, Zhewei Yao, Kurt Keutzer</i>	14471
Generalized Kernel Thinning <i>Raaz Dwivedi, Lester Mackey</i>	14489
What Happens After SGD Reaches Zero Loss? --A Mathematical Framework <i>Zhiyuan Li, Tianhao Wang, Sanjeev Arora</i>	14519
No One Representation to Rule Them All: Overlapping Features of Training Methods <i>Raphael Gontijo-Lopes, Yann Dauphin, Ekin Dogus Cubuk</i>	14566

Bi-Linear Value Networks for Multi-Goal Reinforcement Learning Zhang-Wei Hong, Ge Yang, Pulkit Agrawal	14585
Revisit Kernel Pruning with Lottery Regulated Grouped Convolutions Shaochen Zhong, Guanqun Zhang, Ningjia Huang, Shuai Xu	14600
Learning Long-Term Reward Redistribution Via Randomized Return Decomposition Zhizhou Ren, Ruihan Guo, Yuan Zhou, Jian Peng	14623
Training Invariances and the Low-Rank Phenomenon: Beyond Linear Networks Thien Le, Stefanie Jegelka	14644
Graph-Relational Domain Adaptation Zihao Xu, Hao He, Guang-He Lee, Bernie Wang, Hao Wang	14670
Axiomatic Explanations for Visual Search, Retrieval, and Similarity Learning Mark Hamilton, Scott Lundberg, Stephanie Fu, Lei Zhang, William T. Freeman	14690
Unsupervised Semantic Segmentation by Distilling Feature Correspondences Mark Hamilton, Zhoutong Zhang, Bharath Hariharan, Noah Snavely, William T. Freeman	14716
AdaAug: Learning Class- And Instance-Adaptive Data Augmentation Policies Tsz-Him Cheung, Dit-Yan Yeung	14742
Bandit Learning with Joint Effect of Incentivized Sampling, Delayed Sampling Feedback, and Self-Reinforcing User Preferences Tianchen Zhou, Jia Liu, Chaosheng Dong, Yi Sun	14758
W-CTC: A Connectionist Temporal Classification Loss with Wild Cards Xingyu Cai, Jiahong Yuan, Yuchen Bian, Guangxu Xun, Jiaji Huang, Kenneth Church	14769
Provably Convergent Quasistatic Dynamics for Mean-Field Two-Player Zero-Sum Games Chao Ma, Lexing Ying	14791
EXACT: Scalable Graph Neural Networks Training Via Extreme Activation Compression Zirui Liu, Kaixiong Zhou, Fan Yang, Li Li, Rui Chen, Xia Hu	14808
Revisiting Over-Smoothing in BERT from the Perspective of Graph Han Shi, Jiahui Gao, Hang Xu, Xiaodan Liang, Zhenguo Li, Lingpeng Kong, Stephen M. S. Lee, James Kwok	14840
Properties from Mechanisms: An Equivariance Perspective on Identifiable Representation Learning Kartik Ahuja, Jason Hartford, Yoshua Bengio	14855
MonoDistill: Learning Spatial Features for Monocular 3D Object Detection Zhiyu Chong, Xinzhu Ma, Hong Zhang, Yuxin Yue, Haojie Li, Zhihui Wang, Wanli Ouyang	14878

<p>Offline Reinforcement Learning with Value-Based Episodic Memory</p> <p>Xiaoteng Ma, Yiqin Yang, Hao Hu, Jun Yang, Chongjie Zhang, Qianchuan Zhao, Bin Liang, Qihan Liu</p>	14895
<p>Variational Inference for Discriminative Learning with Generative Modeling of Feature Incompletion</p> <p>Kohei Miyaguchi, Takayuki Katsuki, Akira Koseki, Toshiya Iwamori</p>	14915
<p>Minimax Optimality (Probably) Doesn't Imply Distribution Learning for GANs</p> <p>Sitan Chen, Jerry Li, Yuanzhi Li, Raghu Meka</p>	14933
<p>Group-Based Interleaved Pipeline Parallelism for Large-Scale DNN Training</p> <p>Pengcheng Yang, Xiaoming Zhang, Wenpeng Zhang, Ming Yang, Hong Wei</p>	14961
<p>Self-Supervised Graph Neural Networks for Improved Electroencephalographic Seizure Analysis</p> <p>Siyi Tang, Jared Dunnmon, Khaled Kamal Saab, Xuan Zhang, Qianying Huang, Florian Dubost, Daniel Rubin, Christopher Lee-Messer</p>	14976
<p>Hindsight: Posterior-Guided Training of Retrievers for Improved Open-Ended Generation</p> <p>Ashwin Paranjape, Omar Khattab, Christopher Potts, Matei Zaharia, Christopher D Manning</p>	14999
<p>A Johnson-Lindenstrauss Framework for Randomly Initialized CNNs</p> <p>Ido Nachum, Jan Hazla, Michael Gastpar, Anatoly Khina</p>	15019
<p>Scale Mixtures of Neural Network Gaussian Processes</p> <p>Hyungi Lee, Eunggu Yun, Hongseok Yang, Juho Lee</p>	15037
<p>Model-Based Offline Meta-Reinforcement Learning with Regularization</p> <p>Sen Lin, Jialin Wan, Tengyu Xu, Yingbin Liang, Junshan Zhang</p>	15061
<p>TRGP: Trust Region Gradient Projection for Continual Learning</p> <p>Sen Lin, Li Yang, Deliang Fan, Junshan Zhang</p>	15083
<p>Open-Vocabulary Object Detection Via Vision and Language Knowledge Distillation</p> <p>Xiuye Gu, Tsung-Yi Lin, Weicheng Kuo, Yin Cui</p>	15099
<p>Local Feature Swapping for Generalization in Reinforcement Learning</p> <p>David Bertoin, Emmanuel Rachelson</p>	15119
<p>Exploring Extreme Parameter Compression for Pre-Trained Language Models</p> <p>Benyou Wang, Yuxin Ren, Lifeng Shang, Xin Jiang, Qun Liu</p>	15144
<p>Near-Optimal Offline Reinforcement Learning with Linear Representation: Leveraging Variance Information with Pessimism</p> <p>Ming Yin, Yaqi Duan, Mengdi Wang, Yu-Xiang Wang</p>	15168
<p>CURVATURE-GUIDED DYNAMIC SCALE NETWORKS FOR MULTI-VIEW STEREO</p> <p>Khang Truong Giang, Soohwan Song, Sungho Jo</p>	15207

<p>Evolutionary Diversity Optimization with Clustering-Based Selection for Reinforcement Learning</p> <p><i>Yutong Wang, Ke Xue, Chao Qian</i></p>	15226
<p>Enhancing Cross-Lingual Transfer by Manifold Mixup</p> <p><i>Huiyun Yang, Huadong Chen, Hao Zhou, Lei Li</i></p>	15243
<p>Efficient Learning of Safe Driving Policy Via Human-AI Copilot Optimization</p> <p><i>Quanyi Li, Zhenghao Peng, Bolei Zhou</i></p>	15262
<p>Generating Videos with Dynamics-Aware Implicit Generative Adversarial Networks</p> <p><i>Sihyun Yu, Jihoon Tack, Sangwoo Mo, Hyunsu Kim, Junho Kim, Jung-Woo Ha, Jinwoo Shin</i></p>	15281
<p>POETREE: Interpretable Policy Learning with Adaptive Decision Trees</p> <p><i>Aliz�e Pace, Alex Chan, Mihaela Van Der Schaar</i></p>	15306
<p>Improving the Accuracy of Learning Example Weights for Imbalance Classification</p> <p><i>Yuqi Liu, Bin Cao, Jing Fan</i></p>	15335
<p>SPIRAL: Self-Supervised Perturbation-Invariant Representation Learning for Speech Pre-Training</p> <p><i>Wenyong Huang, Zhenhe Zhang, Yu Ting Yeung, Xin Jiang, Qun Liu</i></p>	15350
<p>Phase Collapse in Neural Networks</p> <p><i>Florentin Guth, John Zarka, St�aphane Mallat</i></p>	15365
<p>Learning to Guide and to Be Guided in the Architect-Builder Problem</p> <p><i>Paul Barde, Tristan Karch, Derek Nowrouzezahrai, Cl�ement Moulin-Frier, Christopher Pal, Pierre-Yves Oudeyer</i></p>	15382
<p>Understanding and Improving Graph Injection Attack by Promoting Unnoticeability</p> <p><i>Yongqiang Chen, Han Yang, Yonggang Zhang, Ma Kaili, Tongliang Liu, Bo Han, James Cheng</i></p>	15404
<p>StyleAlign: Analysis and Applications of Aligned StyleGAN Models</p> <p><i>Zongze Wu, Yotam Nitzan, Eli Shechtman, Dani Lischinski</i></p>	15446
<p>Distributionally Robust Fair Principal Components Via Geodesic Descents</p> <p><i>Hieu Vu, Toan Tran, Man-Chung Yue, Viet Anh Nguyen</i></p>	15490
<p>ViTGAN: Training GANs with Vision Transformers</p> <p><i>Kwonjoon Lee, Huiwen Chang, Lu Jiang, Han Zhang, Zhuowen Tu, Ce Liu</i></p>	15513
<p>Adversarial Robustness Through the Lens of Causality</p> <p><i>Yonggang Zhang, Mingming Gong, Tongliang Liu, Gang Niu, Xinmei Tian, Bo Han, Bernhard Sch�alkopf, Kun Zhang</i></p>	15531
<p>Prototypical Contrastive Predictive Coding</p> <p><i>Kyungmin Lee</i></p>	15551

Pareto Set Learning for Neural Multi-Objective Combinatorial Optimization <i>Xi Lin, Zhiyuan Yang, Qingfu Zhang</i>	15566
Learning Vision-Guided Quadrupedal Locomotion End-To-End with Cross-Modal Transformers <i>Ruihan Yang, Minghao Zhang, Nicklas Hansen, Huazhe Xu, Xiaolong Wang</i>	15597
Language Model Compression with Weighted Low-Rank Factorization <i>Yen-Chang Hsu, Ting Hua, Sungen Chang, Qian Lou, Yilin Shen, Hongxia Jin</i>	15618
Label-Efficient Semantic Segmentation with Diffusion Models <i>Dmitry Baranchuk, Andrey Voynov, Ivan Rubachev, Valentin Khruikov, Artem Babenko</i>	15633
On the Pitfalls of Heteroscedastic Uncertainty Estimation with Probabilistic Neural Networks <i>Maximilian Seitzer, Arash Tavakoli, Dimitrije Antic, Georg Martius</i>	15648
Graph Neural Network Guided Local Search for the Traveling Salesperson Problem <i>Benjamin Hudson, Qingbiao Li, Matthew Malencia, Amanda Prorok</i>	15672
Hidden Parameter Recurrent State Space Models for Changing Dynamics Scenarios <i>Vaisakh Shaj, Dieter Bächler, Rohit Sonker, Philipp Becker, Gerhard Neumann</i>	15692
Objects in Semantic Topology <i>Shuo Yang, Peize Sun, Yi Jiang, Xiaobo Xia, Ruiheng Zhang, Zehuan Yuan, Changhu Wang, Ping Luo, Min Xu</i>	15715
COptiDICE: Offline Constrained Reinforcement Learning Via Stationary Distribution Correction Estimation <i>Jongmin Lee, Cosmin Paduraru, Daniel J Mankowitz, Nicolas Heess, Doina Precup, Kee-Eung Kim, Arthur Guez</i>	15727
NASPY: Automated Extraction of Automated Machine Learning Models <i>Xiaoxuan Lou, Shangwei Guo, Jiwei Li, Yaoxin Wu, Tianwei Zhang</i>	15751
Revisiting Design Choices in Offline Model Based Reinforcement Learning <i>Cong Lu, Philip Ball, Jack Parker-Holder, Michael Osborne, Stephen J. Roberts</i>	15766
Gradient Matching for Domain Generalization <i>Yuge Shi, Jeffrey Seely, Philip Torr, Siddharth N, Awni Hannun, Nicolas Usunier, Gabriel Synnaeve</i>	15803
Discrepancy-Based Active Learning for Domain Adaptation <i>Antoine De Mathelin, Franois Deheeger, Mathilde Mougéot, Nicolas Vayatis</i>	15831
Learning to Extend Molecular Scaffolds with Structural Motifs <i>Krzysztof Maziarz, Henry Richard Jackson-Flux, Pashmina Cameron, Finton Sirockin, Nadine Schneider, Nikolaus Stiefl, Marwin Segler, Marc Brockschmidt</i>	15863
When Should Agents Explore? <i>Miruna Pislari, David Szepesvari, Georg Ostrovski, Diana L Borsa, Tom Schaul</i>	15885

<p>DemoDICE: Offline Imitation Learning with Supplementary Imperfect Demonstrations</p> <p>Geon-Hyeong Kim, Seokin Seo, Jongmin Lee, Wonseok Jeon, Hyeongjoo Hwang, Hongseok Yang, Kee-Eung Kim</p>	15908
<p>Learning Multimodal VAEs Through Mutual Supervision</p> <p>Tom Joy, Yuge Shi, Philip Torr, Tom Rainforth, Sebastian M Schmon, Siddharth N</p>	15934
<p>Neural Stochastic Dual Dynamic Programming</p> <p>Hanjun Dai, Yuan Xue, Zia Syed, Dale Schuurmans, Bo Dai</p>	15957
<p>CLEVA-Compass: A Continual Learning Evaluation Assessment Compass to Promote Research Transparency and Comparability</p> <p>Martin Mundt, Steven Lang, Quentin Delfosse, Kristian Kersting</p>	15981
<p>THOMAS: Trajectory Heatmap Output with Learned Multi-Agent Sampling</p> <p>Thomas Gilles, Stefano Sabatini, Dzmitry Tsishkou, Bogdan Stanciulescu, Fabien Moutarde</p>	16005
<p>Offline Neural Contextual Bandits: Pessimism, Optimization and Generalization</p> <p>Thanh Nguyen-Tang, Sunil Gupta, A. Tuan Nguyen, Svetha Venkatesh</p>	16023
<p>Generalized Natural Gradient Flows in Hidden Convex-Concave Games and GANs</p> <p>Andjela Mladenovic, Iosif Sakos, Gauthier Gidel, Georgios Piliouras</p>	16048
<p>LORD: Lower-Dimensional Embedding of Log-Signature in Neural Rough Differential Equations</p> <p>Jaehoon Lee, Jinsung Jeon, Sheo Yon Jhin, Jihyeon Hyeong, Jayoung Kim, Minju Jo, Kook Seungji, Noseong Park</p>	16077
<p>Multi-Mode Deep Matrix and Tensor Factorization</p> <p>Jicong Fan</p>	16098
<p>Inverse Online Learning: Understanding Non-Stationary and Reactionary Policies</p> <p>Alex Chan, Alicia Curth, Mihaela Van Der Schaar</p>	16123
<p>UniFormer: Unified Transformer for Efficient Spatial-Temporal Representation Learning</p> <p>Kunchang Li, Yali Wang, Gao Peng, Guanglu Song, Yu Liu, Hongsheng Li, Yu Qiao</p>	16139
<p>Target-Side Input Augmentation for Sequence to Sequence Generation</p> <p>Shufang Xie, Ang Lv, Yingce Xia, Lijun Wu, Tao Qin, Tie-Yan Liu, Rui Yan</p>	16158
<p>PriorGrad: Improving Conditional Denoising Diffusion Models with Data-Dependent Adaptive Prior</p> <p>Sang-Gil Lee, Heeseung Kim, Chaehun Shin, Xu Tan, Chang Liu, Qi Meng, Tao Qin, Wei Chen, Sungroh Yoon, Tie-Yan Liu</p>	16176
<p>Optimal Transport for Long-Tailed Recognition with Learnable Cost Matrix</p> <p>Hanyu Peng, Mingming Sun, Ping Li</p>	16195
<p>Reinforcement Learning in Presence of Discrete Markovian Context Evolution</p> <p>Hang Ren, Aivar Sootla, Taher Jafferjee, Junxiao Shen, Jun Wang, Haitham Bou Ammar</p>	16210

<p>Inductive Relation Prediction Using Analogy Subgraph Embeddings</p> <p>Jiarui Jin, Yangkun Wang, Kounianhua Du, Weinan Zhang, Zheng Zhang, David Wipf, Yong Yu, Quan Gan</p>	16250
<p>Attention-Based Interpretability with Concept Transformers</p> <p>Mattia Rigotti, Christoph Miksovics, Ioana Giurgiu, Thomas Gschwind, Paolo Scotton</p>	16285
<p>How Does SimSiam Avoid Collapse Without Negative Samples? a Unified Understanding with Self-Supervised Contrastive Learning</p> <p>Chaoning Zhang, Kang Zhang, Chenshuang Zhang, Trung X. Pham, Chang D. Yoo, In So Kweon</p>	16301
<p>Gaussian Mixture Convolution Networks</p> <p>Adam Celarek, Pedro Hermosilla, Bernhard Kerbl, Timo Ropinski, Michael Wimmer</p>	16319
<p>ZeroFL: Efficient On-Device Training for Federated Learning with Local Sparsity</p> <p>Xinchi Qiu, Javier Fernandez-Marques, Pedro Pb Gusmao, Yan Gao, Titouan Parcollet, Nicholas Donald Lane</p>	16342
<p>Fine-Tuning Can Distort Pretrained Features and Underperform Out-Of-Distribution</p> <p>Ananya Kumar, Aditi Raghunathan, Robbie Matthew Jones, Tengyu Ma, Percy Liang</p>	16358
<p>Learning a Subspace of Policies for Online Adaptation in Reinforcement Learning</p> <p>Jean-Baptiste Gaya, Laure Soulier, Ludovic Denoyer</p>	16400
<p>BEiT: BERT Pre-Training of Image Transformers</p> <p>Hangbo Bao, Li Dong, Songhao Piao, Furu Wei</p>	16433
<p>TAda! Temporally-Adaptive Convolutions for Video Understanding</p> <p>Ziyuan Huang, Shiwei Zhang, Liang Pan, Zhiwu Qing, Mingqian Tang, Ziwei Liu, Marcelo H Ang Jr</p>	16451
<p>Better Supervisory Signals by Observing Learning Paths</p> <p>Yi Ren, Shangmin Guo, Danica J. Sutherland</p>	16474
<p>Contrastive Fine-Grained Class Clustering Via Generative Adversarial Networks</p> <p>Yunji Kim, Jung-Woo Ha</p>	16498
<p>ToM2C: Target-Oriented Multi-Agent Communication and Cooperation with Theory of Mind</p> <p>Yuanfei Wang, Fangwei Zhong, Jing Xu, Yizhou Wang</p>	16516
<p>PSA-GAN: Progressive Self Attention GANs for Synthetic Time Series</p> <p>Paul Jeha, Michael Bohlke-Schneider, Pedro Mercado, Shubham Kapoor, Rajbir Singh Nirwan, Valentin Flunkert, Jan Gasthaus, Tim Januschowski</p>	16533
<p>Compositional Attention: Disentangling Search and Retrieval</p> <p>Sarthak Mittal, Sharath Chandra Raparthy, Irina Rish, Yoshua Bengio, Guillaume Lajoie</p>	16553
<p>Neural Solvers for Fast and Accurate Numerical Optimal Control</p> <p>Federico Berto, Stefano Massaroli, Michael Poli, Jinkyoo Park</p>	16579

<p>Query Embedding on Hyper-Relational Knowledge Graphs</p> <p><i>Dimitrios Alivanistos, Max Berrendorf, Michael Cochez, Mikhail Galkin</i></p>	16598
<p>On the Pitfalls of Analyzing Individual Neurons in Language Models</p> <p><i>Omer Antverg, Yonatan Belinkov</i></p>	16621
<p>Reversible Instance Normalization for Accurate Time-Series Forecasting Against Distribution Shift</p> <p><i>Taesung Kim, Jinhee Kim, Yunwon Tae, Cheonbok Park, Jang-Ho Choi, Jaegul Choo</i></p>	16644
<p>Fast Generic Interaction Detection for Model Interpretability and Compression</p> <p><i>Tianjian Zhang, Feng Yin, Zhi-Quan Luo</i></p>	16669
<p>Unrolling PALM for Sparse Semi-Blind Source Separation</p> <p><i>Mohammad Fahes, Christophe Kervazo, J�r�me Bobin, Florence Tupin</i></p>	16698
<p>Stability Regularization for Discrete Representation Learning</p> <p><i>Adeel Pervez, Efstratios Gavves</i></p>	16719
<p>Bag of Instances Aggregation Boosts Self-Supervised Distillation</p> <p><i>Haohang Xu, Jiemin Fang, Xiaopeng Zhang, Lingxi Xie, Xinggang Wang, Wenrui Dai, Hongkai Xiong, Qi Tian</i></p>	16733
<p>Igeood: An Information Geometry Approach to Out-Of-Distribution Detection</p> <p><i>Eduardo Dadalto Camara Gomes, Florence Alberge, Pierre Duhamel, Pablo Piantanida</i></p>	16746
<p>Actor-Critic is Implicitly Biased Towards High Entropy Optimal Policies</p> <p><i>Yuzheng Hu, Ziwei Ji, Matus Telgarsky</i></p>	16783
<p>Knowledge Removal in Sampling-Based Bayesian Inference</p> <p><i>Shaopeng Fu, Fengxiang He, Dacheng Tao</i></p>	16811
<p>Regularized Autoencoders for Isometric Representation Learning</p> <p><i>Yonghyeon Lee, Sangwoong Yoon, Minjun Son, Frank C. Park</i></p>	16833
<p>Charformer: Fast Character Transformers Via Gradient-Based Subword Tokenization</p> <p><i>Yi Tay, Vinh Q. Tran, Sebastian Ruder, Jai Gupta, Hyung Won Chung, Dara Bahri, Zhen Qin, Simon Baumgartner, Cong Yu, Donald Metzler</i></p>	16856
<p>GPT-Critic: Offline Reinforcement Learning for End-To-End Task-Oriented Dialogue Systems</p> <p><i>Youngsoo Jang, Jongmin Lee, Kee-Eung Kim</i></p>	16876
<p>A Deep Variational Approach to Clustering Survival Data</p> <p><i>Laura Manduchi, Ri�ards Marcinkevi�s, Michela C. Massi, Thomas Weikert, Alexander Sauter, Verena Gotta, Timothy M�ller, Flavio Vasella, Marian C. Neidert, Marc Pfister, Bram Stieltjes, Julia E Vogt</i></p>	16892
<p>Generative Pseudo-Inverse Memory</p> <p><i>Kha Pham, Hung Le, Man Ngo, Truyen Tran, Bao Ho, Svetha Venkatesh</i></p>	16927

<p>Escaping Limit Cycles: Global Convergence for Constrained Nonconvex-Nonconcave Minimax Problems</p> <p><i>Thomas Pethick, Puya Latafat, Panos Patrinos, Olivier Fercoq, Volkan Cevher</i></p>	16945
<p>What Makes Better Augmentation Strategies? Augment Difficult but Not Too Different</p> <p><i>Jaehyung Kim, Dongyeop Kang, Sungsoo Ahn, Jinwoo Shin</i></p>	16964
<p>Shallow and Deep Networks Are Near-Optimal Approximators of Korobov Functions</p> <p><i>Moise Blanchard, Mohammed Amine Bennouna</i></p>	16989
<p>BadPre: Task-Agnostic Backdoor Attacks to Pre-Trained NLP Foundation Models</p> <p><i>Kangjie Chen, Yuxian Meng, Xiaofei Sun, Shangwei Guo, Tianwei Zhang, Jiwei Li, Chun Fan</i></p>	17012
<p>Equivariant and Stable Positional Encoding for More Powerful Graph Neural Networks</p> <p><i>Haorui Wang, Haoteng Yin, Muhan Zhang, Pan Li</i></p>	17029
<p>A Unified Contrastive Energy-Based Model for Understanding the Generative Ability of Adversarial Training</p> <p><i>Yifei Wang, Yisen Wang, Jiansheng Yang, Zhouchen Lin</i></p>	17053
<p>Maximum Entropy RL (Provably) Solves Some Robust RL Problems</p> <p><i>Benjamin Eysenbach, Sergey Levine</i></p>	17071
<p>Learning by Directional Gradient Descent</p> <p><i>David Silver, Anirudh Goyal, Ivo Danihelka, Matteo Hessel, Hado Van Hasselt</i></p>	17095
<p>Why Propagate Alone? Parallel Use of Labels and Features on Graphs</p> <p><i>Yangkun Wang, Jiarui Jin, Weinan Zhang, Yang Yongyi, Jiu hai Chen, Quan Gan, Yong Yu, Zheng Zhang, Zengfeng Huang, David Wipf</i></p>	17107
<p>Learning to Remember Patterns: Pattern Matching Memory Networks for Traffic Forecasting</p> <p><i>Hyunwook Lee, Seungmin Jin, Hyeshin Chu, Hongkyu Lim, Sungahn Ko</i></p>	17127
<p>Filtered-CoPhy: Unsupervised Learning of Counterfactual Physics in Pixel Space</p> <p><i>Steeven Janny, Fabien Baradel, Natalia Neverova, Madiha Nadri, Greg Mori, Christian Wolf</i></p>	17142
<p>Continuously Discovering Novel Strategies Via Reward-Switching Policy Optimization</p> <p><i>Zihan Zhou, Wei Fu, Bingliang Zhang, Yi Wu</i></p>	17170
<p>GraphENS: Neighbor-Aware Ego Network Synthesis for Class-Imbalanced Node Classification</p> <p><i>Joonhyung Park, Jaeyun Song, Eunho Yang</i></p>	17200
<p>GeneDisco: A Benchmark for Experimental Design in Drug Discovery</p> <p><i>Arash Mehrjou, Ashkan Soleymani, Andrew Jesson, Pascal Notin, Yarin Gal, Stefan Bauer, Patrick Schwab</i></p>	17220
<p>End-To-End Learning of Probabilistic Hierarchies on Graphs</p> <p><i>Daniel Zagner, Bertrand Charpentier, Morgane Ayle, Sascha Geringer, Stephan Gannemann</i></p>	17253

<p>Hierarchical Few-Shot Imitation with Skill Transition Models</p> <p><i>Kourosh Hakhamaneshi, Ruihan Zhao, Albert Zhan, Pieter Abbeel, Michael Laskin</i></p>	17276
<p>Evaluating Model-Based Planning and Planner Amortization for Continuous Control</p> <p><i>Arunkumar Byravan, Leonard Hasenclever, Piotr Trochim, Mehdi Mirza, Alessandro Davide Ialongo, Yuval Tassa, Jost Tobias Springenberg, Abbas Abdolmaleki, Nicolas Heess, Josh Merel, Martin Riedmiller</i></p>	17295
<p>EntQA: Entity Linking as Question Answering</p> <p><i>Wenzheng Zhang, Wenyue Hua, Karl Stratos</i></p>	17327
<p>Differentiable DAG Sampling</p> <p><i>Bertrand Charpentier, Simon Kibler, Stephan Gännemann</i></p>	17340
<p>Salient ImageNet: How to Discover Spurious Features in Deep Learning?</p> <p><i>Sahil Singla, Soheil Feizi</i></p>	17365
<p>Pretrained Language Model in Continual Learning: A Comparative Study</p> <p><i>Tongtong Wu, Massimo Caccia, Zhuang Li, Yuan-Fang Li, Guilin Qi, Gholamreza Haffari</i></p>	17441
<p>Hybrid Random Features</p> <p><i>Krzysztof Marcin Choromanski, Han Lin, Haoxian Chen, Arijit Sehanobish, Yuanzhe Ma, Deepali Jain, Jake Varley, Andy Zeng, Michael S Ryoo, Valerii Likhoshesterov, Dmitry Kalashnikov, Vikas Sindhwani, Adrian Weller</i></p>	17458
<p>Goal-Directed Planning Via Hindsight Experience Replay</p> <p><i>Lorenzo Moro, Amarildo Likmeta, Enrico Prati, Marcello Restelli</i></p>	17489
<p>Bayesian Neural Network Priors Revisited</p> <p><i>Vincent Fortuin, Adrià Garriga-Alonso, Sebastian W. Ober, Florian Wenzel, Gunnar Ratsch, Richard E Turner, Mark Van Der Wilk, Laurence Aitchison</i></p>	17505
<p>Looking Back on Learned Experiences for Class/Task Incremental Learning</p> <p><i>Mozhgan Pourkeshavarzi, Guoying Zhao, Mohammad Sabokrou</i></p>	17540
<p>Expressiveness and Approximation Properties of Graph Neural Networks</p> <p><i>Floris Geerts, Juan L Reutter</i></p>	17552
<p>Scattering Networks on the Sphere for Scalable and Rotationally Equivariant Spherical CNNs</p> <p><i>Jason McEwen, Christopher Wallis, Augustine N. Mavor-Parker</i></p>	17595
<p>Space-Time Graph Neural Networks</p> <p><i>Samar Hadou, Charilaos I Kanatsoulis, Alejandro Ribeiro</i></p>	17613
<p>Vitruvion: A Generative Model of Parametric CAD Sketches</p> <p><i>Ari Seff, Wenda Zhou, Nick Richardson, Ryan P Adams</i></p>	17640

<p>Scale Efficiently: Insights from Pretraining and Finetuning Transformers</p> <p><i>Yi Tay, Mostafa Dehghani, Jinfeng Rao, William Fedus, Samira Abnar, Hyung Won Chung, Sharan Narang, Dani Yogatama, Ashish Vaswani, Donald Metzler</i></p>	17658
<p>Conditional Object-Centric Learning from Video</p> <p><i>Thomas Kipf, Gamaleldin Fathy Elsayed, Aravindh Mahendran, Austin Stone, Sara Sabour, Georg Heigold, Rico Jonschkowski, Alexey Dosovitskiy, Klaus Greff</i></p>	17676
<p>Meta Learning Low Rank Covariance Factors for Energy Based Deterministic Uncertainty</p> <p><i>Jeffrey Ryan Willette, Hae Beom Lee, Juho Lee, Sung Ju Hwang</i></p>	17708
<p>Transition to Linearity of Wide Neural Networks is an Emerging Property of Assembling Weak Models</p> <p><i>Chaoyue Liu, Libin Zhu, Misha Belkin</i></p>	17735
<p>Improved Deterministic L2 Robustness on CIFAR-10 and CIFAR-100</p> <p><i>Sahil Singla, Surbhi Singla, Soheil Feizi</i></p>	17751
<p>Missingness Bias in Model Debugging</p> <p><i>Saachi Jain, Hadi Salman, Eric Wong, Pengchuan Zhang, Vibhav Vineet, Sai Vemprala, Aleksander Madry</i></p>	17781
<p>Revisiting Flow Generative Models for Out-Of-Distribution Detection</p> <p><i>Dihong Jiang, Sun Sun, Yaoliang Yu</i></p>	17810
<p>Fine-Grained Differentiable Physics: A Yarn-Level Model for Fabrics</p> <p><i>Deshan Gong, Zhanxing Zhu, Andrew J. Bulpitt, He Wang</i></p>	17834
<p>Auto-Scaling Vision Transformers Without Training</p> <p><i>Wuyang Chen, Wei Huang, Xianzhi Du, Xiaodan Song, Zhangyang Wang, Denny Zhou</i></p>	17847
<p>Autoregressive Diffusion Models</p> <p><i>Emiel Hoogetboom, Alexey A. Gritsenko, Jasmijn Bastings, Ben Poole, Rianne Van Den Berg, Tim Salimans</i></p>	17861
<p>Attacking Deep Networks with Surrogate-Based Adversarial Black-Box Methods is Easy</p> <p><i>Nicholas A. Lord, Romain Mueller, Luca Bertinetto</i></p>	17884
<p>Defending Against Image Corruptions Through Adversarial Augmentations</p> <p><i>Dan Andrei Calian, Florian Stimberg, Olivia Wiles, Sylvestre-Alvise Rebuffi, Andr�as Gy�argy, Timothy A Mann, Sven Gowal</i></p>	17901
<p>Provable Learning-Based Algorithm for Sparse Recovery</p> <p><i>Xinshi Chen, Haoran Sun, Le Song</i></p>	17928
<p>MoReL: Multi-Omics Relational Learning</p> <p><i>Arman Hasanzadeh, Ehsan Hajiramezanali, Nick Duffield, Xiaoning Qian</i></p>	17975

GLASS: GNN with Labeling Tricks for Subgraph Representation Learning <i>Xiyuan Wang, Muhan Zhang</i>	17990
Learnability of Convolutional Neural Networks for Infinite Dimensional Input Via Mixed and Anisotropic Smoothness <i>Sho Okumoto, Taiji Suzuki</i>	18010
Information Bottleneck: Exact Analysis of (Quantized) Neural Networks <i>Stephan Sloth Lorenzen, Christian Igel, Mads Nielsen</i>	18042
Dealing with Non-Stationarity in MARL Via Trust-Region Decomposition <i>Wenhao Li, Xiangfeng Wang, Bo Jin, Junjie Sheng, Hongyuan Zha</i>	18061
Information-Theoretic Online Memory Selection for Continual Learning <i>Shengyang Sun, Daniele Calandriello, Huiyi Hu, Ang Li, Michalis Titsias</i>	18100
On the Role of Neural Collapse in Transfer Learning <i>Tomer Galanti, Andr�as Gy�argy, Marcus Hutter</i>	18125
Self-Supervised Inference in State-Space Models <i>David Ruhe, Patrick Forr�</i>	18160
Generalization Through the Lens of Leave-One-Out Error <i>Gregor Bachmann, Thomas Hofmann, Aurelien Lucchi</i>	18181
Learning Scenario Representation for Solving Two-Stage Stochastic Integer Programs <i>Yaoxin Wu, Wen Song, Zhiguang Cao, Jie Zhang</i>	18203
Spread Spurious Attribute: Improving Worst-Group Accuracy with Spurious Attribute Estimation <i>Junhyun Nam, Jaehyung Kim, Jaeho Lee, Jinwoo Shin</i>	18222
Convergent Graph Solvers <i>Junyoung Park, Jinhyun Choo, Jinkyoo Park</i>	18238
SURF: Semi-Supervised Reward Learning with Data Augmentation for Feedback-Efficient Preference-Based Reinforcement Learning <i>Jongjin Park, Younggyo Seo, Jinwoo Shin, Honglak Lee, Pieter Abbeel, Kimin Lee</i>	18262
Surreal-GAN: Semi-Supervised Representation Learning Via GAN for Uncovering Heterogeneous Disease-Related Imaging Patterns <i>Zhijian Yang, Junhao Wen, Christos Davatzikos</i>	18279
NODE-GAM: Neural Generalized Additive Model for Interpretable Deep Learning <i>Chun-Hao Chang, Rich Caruana, Anna Goldenberg</i>	18297

<p>Towards Continual Knowledge Learning of Language Models</p> <p>Joel Jang, Seonghyeon Ye, Sohee Yang, Joongbo Shin, Janghoon Han, Gyeonghun Kim, Stanley Jungkyu Choi, Minjoon Seo</p>	18323
<p>Equivariant Graph Mechanics Networks with Constraints</p> <p>Wenbing Huang, Jiaqi Han, Yu Rong, Tingyang Xu, Fuchun Sun, Junzhou Huang</p>	18350
<p>Continual Normalization: Rethinking Batch Normalization for Online Continual Learning</p> <p>Quang Pham, Chenghao Liu, Steven Hoi</p>	18370
<p>Continual Learning with Recursive Gradient Optimization</p> <p>Hao Liu, Huaping Liu</p>	18390
<p>Online Ad Hoc Teamwork Under Partial Observability</p> <p>Pengjie Gu, Mengchen Zhao, Jianye Hao, Bo An</p>	18408
<p>Nonlinear ICA Using Volume-Preserving Transformations</p> <p>Xiaojiang Yang, Yi Wang, Jiacheng Sun, Xing Zhang, Shifeng Zhang, Zhenguo Li, Junchi Yan</p>	18425
<p>PF-GNN: Differentiable Particle Filtering Based Approximation of Universal Graph Representations</p> <p>Mohammed Haroon Dupty, Yanfei Dong, Wee Sun Lee</p>	18444
<p>Topologically Regularized Data Embeddings</p> <p>Robin Vandaele, Bo Kang, Jefrey Lijffijt, Tijl De Bie, Yvan Saeys</p>	18466
<p>Rethinking Goal-Conditioned Supervised Learning and Its Connection to Offline RL</p> <p>Rui Yang, Yiming Lu, Wenzhe Li, Hao Sun, Meng Fang, Yali Du, Xiu Li, Lei Han, Chongjie Zhang</p>	18492
<p>Discovering Nonlinear PDEs from Scarce Data with Physics-Encoded Learning</p> <p>Chengping Rao, Pu Ren, Yang Liu, Hao Sun</p>	18521
<p>Proving the Lottery Ticket Hypothesis for Convolutional Neural Networks</p> <p>Arthur Da Cunha, Emanuele Natale, Laurent Viennot</p>	18540
<p>Almost Tight L0-Norm Certified Robustness of Top-K Predictions Against Adversarial Perturbations</p> <p>Jinyuan Jia, Binghui Wang, Xiaoyu Cao, Hongbin Liu, Neil Zhenqiang Gong</p>	18555
<p>QUERY EFFICIENT DECISION BASED SPARSE ATTACKS AGAINST BLACK-BOX DEEP LEARNING MODELS</p> <p>Viet Vo, Ehsan M Abbasnejad, Damith Ranasinghe</p>	18579
<p>Parallel Training of GRU Networks with a Multi-Grid Solver for Long Sequences</p> <p>Euhyun Moon, Eric C Cyr</p>	18598
<p>Knowledge Infused Decoding</p> <p>Ruibo Liu, Guoqing Zheng, Shashank Gupta, Radhika Gaonkar, Chongyang Gao, Soroush Vosoughi, Milad Shokouhi, Ahmed Hassan Awadallah</p>	18616

<p><i>Pyraformer: Low-Complexity Pyramidal Attention for Long-Range Time Series Modeling and Forecasting</i></p> <p><i>Shizhan Liu, Hang Yu, Cong Liao, Jianguo Li, Weiyao Lin, Alex X. Liu, Schahram Dustdar</i></p>	18639
<p><i>Toward Efficient Low-Precision Training: Data Format Optimization and Hysteresis Quantization</i></p> <p><i>Sunwoo Lee, Jeongwoo Park, Dongsuk Jeon</i></p>	18659
<p><i>Structure-Aware Transformer Policy for Inhomogeneous Multi-Task Reinforcement Learning</i></p> <p><i>Sunghoon Hong, Deunsol Yoon, Kee-Eung Kim</i></p>	18679
<p><i>Relational Surrogate Loss Learning</i></p> <p><i>Tao Huang, Zekang Li, Hua Lu, Yong Shan, Shusheng Yang, Yang Feng, Fei Wang, Shan You, Chang Xu</i></p>	18698
<p><i>Continual Learning with Filter Atom Swapping</i></p> <p><i>Zichen Miao, Ze Wang, Wei Chen, Qiang Qiu</i></p>	18714
<p><i>Sqrt(d) Dimension Dependence of Langevin Monte Carlo</i></p> <p><i>Ruilin Li, Hongyuan Zha, Molei Tao</i></p>	18734
<p><i>Anti-Concentrated Confidence Bonuses for Scalable Exploration</i></p> <p><i>Jordan T. Ash, Cyril Zhang, Surbhi Goel, Akshay Krishnamurthy, Sham M. Kakade</i></p>	18756
<p><i>Efficient Split-Mix Federated Learning for On-Demand and In-Situ Customization</i></p> <p><i>Junyuan Hong, Haotao Wang, Zhangyang Wang, Jiayu Zhou</i></p>	18774
<p><i>Understanding Intrinsic Robustness Using Label Uncertainty</i></p> <p><i>Xiao Zhang, David Evans</i></p>	18792
<p><i>Meta-Imitation Learning by Watching Video Demonstrations</i></p> <p><i>Jiayi Li, Tao Lu, Xiaoge Cao, Yinghao Cai, Shuo Wang</i></p>	18815
<p><i>A Zest of LIME: Towards Architecture-Independent Model Distances</i></p> <p><i>Hengrui Jia, Hongyu Chen, Jonas Guan, Ali Shahin Shamsabadi, Nicolas Papernot</i></p>	18832
<p><i>PiCO: Contrastive Label Disambiguation for Partial Label Learning</i></p> <p><i>Haobo Wang, Ruixuan Xiao, Yixuan Li, Lei Feng, Gang Niu, Gang Chen, Junbo Zhao</i></p>	18849
<p><i>Rethinking Supervised Pre-Training for Better Downstream Transferring</i></p> <p><i>Yutong Feng, Jianwen Jiang, Mingqian Tang, Rong Jin, Yue Gao</i></p>	18867
<p><i>Self-Joint Supervised Learning</i></p> <p><i>Navid Kardan, Mubarak Shah, Mitch Hill</i></p>	18889
<p><i>Practical Integration Via Separable Bijective Networks</i></p> <p><i>Christopher M Bender, Patrick Emmanuel, Michael K. Reiter, Junier Oliva</i></p>	18908

OBJECT DYNAMICS DISTILLATION FOR SCENE DECOMPOSITION AND REPRESENTATION Qu Tang, Xiangyu Zhu, Zhen Lei, Zhaoxiang Zhang	18928
Towards Evaluating the Robustness of Neural Networks Learned by Transduction Jiefeng Chen, Xi Wu, Yang Guo, Yingyu Liang, Somesh Jha	18942
Generalized Rectifier Wavelet Covariance Models for Texture Synthesis Antoine Brochard, Sixin Zhang, Stephane Mallat	18960
CoMPS: Continual Meta Policy Search Glen Berseth, Zhiwei Zhang, Grace Zhang, Chelsea Finn, Sergey Levine	18983
Prospect Pruning: Finding Trainable Weights at Initialization Using Meta-Gradients Milad Alizadeh, Shyam A. Taylor, Luisa M Zintgraf, Joost Van Amersfoort, Sebastian Farquhar, Nicholas Donald Lane, Yarin Gal	19006
GNN-LM: Language Modeling Based on Global Contexts Via GNN Yuxian Meng, Shi Zong, Xiaoya Li, Xiaofei Sun, Tianwei Zhang, Fei Wu, Jiwei Li	19020
Overcoming the Spectral Bias of Neural Value Approximation Ge Yang, Anurag Ajay, Pulkit Agrawal	19033
Differentiable Expectation-Maximization for Set Representation Learning Minyoung Kim	19052
Domain Adversarial Training: A Game Perspective David Acuna, Marc T Law, Guojun Zhang, Sanja Fidler	19074
GATSBI: Generative Adversarial Training for Simulation-Based Inference Poornima Ramesh, Jan-Matthis Lueckmann, Jan Boelts, Ivarro Tejero-Cantero, David S. Greenberg, Pedro J. Goncalves, Jakob H. Macke	19100
Transfer RL Across Observation Feature Spaces Via Model-Based Regularization Yanchao Sun, Ruijie Zheng, Xiyao Wang, Andrew E Cohen, Furong Huang	19129
Stochastic Training is Not Necessary for Generalization Jonas Geiping, Micah Goldblum, Phil Pope, Michael Moeller, Tom Goldstein	19152
Boosting the Certified Robustness of L-Infinity Distance Nets Bohang Zhang, Du Jiang, Di He, Liwei Wang	19176
Normalization of Language Embeddings for Cross-Lingual Alignment Prince Osei Aboagye, Yan Zheng, Chin-Chia Michael Yeh, Junpeng Wang, Wei Zhang, Liang Wang, Hao Yang, Jeff Phillips	19197
Towards a Unified View of Parameter-Efficient Transfer Learning Junxian He, Chunting Zhou, Xuezhe Ma, Taylor Berg-Kirkpatrick, Graham Neubig	19229

<p>Direct Then Diffuse: Incremental Unsupervised Skill Discovery for State Covering and Goal Reaching</p> <p><i>Pierre-Alexandre Kamienny, Jean Tarbouriech, Sylvain Lamprier, Alessandro Lazaric, Ludovic Denoyer</i></p>	19244
<p>Equivariant Self-Supervised Learning: Encouraging Equivariance in Representations</p> <p><i>Rumen Dangovski, Li Jing, Charlotte Loh, Seungwook Han, Akash Srivastava, Brian Cheung, Pulkit Agrawal, Marin Soljacic</i></p>	19269
<p>Unsupervised Vision-Language Grammar Induction with Shared Structure Modeling</p> <p><i>Bo Wan, Wenjuan Han, Zilong Zheng, Tinne Tuytelaars</i></p>	19290
<p>Independent SE(3)-Equivariant Models for End-To-End Rigid Protein Docking</p> <p><i>Octavian-Eugen Ganea, Xinyuan Huang, Charlotte Bunne, Yatao Bian, Regina Barzilay, Tommi S. Jaakkola, Andreas Krause</i></p>	19309
<p>LOSSY COMPRESSION WITH DISTRIBUTION SHIFT AS ENTROPY CONSTRAINED OPTIMAL TRANSPORT</p> <p><i>Huan Liu, George Zhang, Jun Chen, Ashish J Khisti</i></p>	19330
<p>Conditioning Sequence-To-Sequence Networks with Learned Activations</p> <p><i>Alberto Gil Couto Pimentel Ramos, Abhinav Mehrotra, Nicholas Donald Lane, Sourav Bhattacharya</i></p>	19364
<p>Adversarial Retriever-Ranker for Dense Text Retrieval</p> <p><i>Hang Zhang, Yeyun Gong, Yelong Shen, Jiancheng Lv, Nan Duan, Weizhu Chen</i></p>	19380
<p>Hybrid Memoised Wake-Sleep: Approximate Inference at the Discrete-Continuous Interface</p> <p><i>Tuan Anh Le, Katherine M. Collins, Luke Hewitt, Kevin Ellis, Siddharth N, Samuel Gershman, Joshua B. Tenenbaum</i></p>	19396
<p>The Efficiency Misnomer</p> <p><i>Mostafa Dehghani, Yi Tay, Anurag Arnab, Lucas Beyer, Ashish Vaswani</i></p>	19413
<p>Mapping Language Models to Grounded Conceptual Spaces</p> <p><i>Roma Patel, Ellie Pavlick</i></p>	19429
<p>Multi-Objective Optimization by Learning Space Partition</p> <p><i>Yiyang Zhao, Linnan Wang, Kevin Yang, Tianjun Zhang, Tian Guo, Yuandong Tian</i></p>	19450
<p>Amortized Implicit Differentiation for Stochastic Bilevel Optimization</p> <p><i>Michael Arbel, Julien Mairal</i></p>	19475
<p>On the Generalization of Models Trained with SGD: Information-Theoretic Bounds and Implications</p> <p><i>Ziqiao Wang, Yongyi Mao</i></p>	19514
<p>Mapping Conditional Distributions for Domain Adaptation Under Generalized Target Shift</p> <p><i>Matthieu Kirchmeyer, Alain Rakotomamonjy, Emmanuel De Bezenac, Patrick Gallinari</i></p>	19541

<p>On Non-Random Missing Labels in Semi-Supervised Learning</p> <p><i>Xinting Hu, Yulei Niu, Chunyan Miao, Xian-Sheng Hua, Hanwang Zhang</i></p>	19563
<p>LFPT5: A Unified Framework for Lifelong Few-Shot Language Learning Based on Prompt Tuning of T5</p> <p><i>Chengwei Qin, Shafiq Joty</i></p>	19575
<p>AdaRL: What, Where, and How to Adapt in Transfer Reinforcement Learning</p> <p><i>Biwei Huang, Fan Feng, Chaochao Lu, Sara Magliacane, Kun Zhang</i></p>	19590
<p>Invariant Causal Representation Learning for Out-Of-Distribution Generalization</p> <p><i>Chaochao Lu, Yuhuai Wu, Josã Miguel Hernã;ndez-Lobato, Bernhard Schãlkopf</i></p>	19631
<p>The Spectral Bias of Polynomial Neural Networks</p> <p><i>Moulik Choraria, Leello Tadesse Dadi, Grigorios Chrysos, Julien Mairal, Volkan Cevher</i></p>	19663
<p>Top-N: Equivariant Set and Graph Generation Without Exchangeability</p> <p><i>Clement Vignac, Pascal Frossard</i></p>	19693
<p>Procedural Generalization by Planning with Self-Supervised World Models</p> <p><i>Ankesh Anand, Jacob C Walker, Yazhe Li, Eszter Vãrtes, Julian Schrittwieser, Sherjil Ozair, Theophane Weber, Jessica B Hamrick</i></p>	19710
<p>Scaling Laws for Neural Machine Translation</p> <p><i>Behrooz Ghorbani, Orhan Firat, Markus Freitag, Ankur Bapna, Maxim Krikun, Xavier Garcia, Ciprian Chelba, Colin Cherry</i></p>	19732
<p>IGLU: Efficient GCN Training Via Lazy Updates</p> <p><i>S Deepak Narayanan, Aditya Sinha, Prateek Jain, Purushottam Kar, Sundararajan Sellamanickam</i></p>	19774
<p>Diurnal Or Nocturnal? Federated Learning of Multi-Branch Networks from Periodically Shifting Distributions</p> <p><i>Chen Zhu, Zheng Xu, Mingqing Chen, Jakub Koneãnã½, Andrew Hard, Tom Goldstein</i></p>	19810
<p>Geometric Transformers for Protein Interface Contact Prediction</p> <p><i>Alex Morehead, Chen Chen, Jianlin Cheng</i></p>	19829
<p>Evaluating Disentanglement of Structured Representations</p> <p><i>Raphaã«l Dang-Nhu</i></p>	19847
<p>Neural Processes with Stochastic Attention: Paying More Attention to the Context Dataset</p> <p><i>Mingyu Kim, Kyeong Ryeol Go, Se-Young Yun</i></p>	19876
<p>Robust and Scalable SDE Learning: A Functional Perspective</p> <p><i>Scott Alexander Cameron, Tyron Luke Cameron, Arnû Pretorius, Stephen J. Roberts</i></p>	19917
<p>Large-Scale Representation Learning on Graphs Via Bootstrapping</p> <p><i>Shantanu Thakoor, Corentin Tallec, Mohammad Gheshlaghi Azar, Mehdi Azabou, Eva L Dyer, Remi Munos, Petar Veliãkoviã, Michal Valko</i></p>	19932

<p>Cross-Domain Imitation Learning Via Optimal Transport</p> <p><i>Arnaud Fickinger, Samuel Cohen, Stuart Russell, Brandon Amos</i></p>	19953
<p>A Loss Curvature Perspective on Training Instabilities of Deep Learning Models</p> <p><i>Justin Gilmer, Behrooz Ghorbani, Ankush Garg, Sneha Kudugunta, Behnam Neyshabur, David Cardoze, George Edward Dahl, Zachary Nado, Orhan Firat</i></p>	19968
<p>Universal Approximation Under Constraints is Possible with Transformers</p> <p><i>Anastasis Kratsios, Behnoosh Zamanlooy, Tianlin Liu, Ivan Dokmaniä</i></p>	19990
<p>Exploring the Limits of Large Scale Pre-Training</p> <p><i>Samira Abnar, Mostafa Dehghani, Behnam Neyshabur, Hanie Sedghi</i></p>	20004
<p>MIDI-DDSP: Detailed Control of Musical Performance Via Hierarchical Modeling</p> <p><i>Yusong Wu, Ethan Manilow, Yi Deng, Rigel Swavely, Kyle Kastner, Tim Cooijmans, Aaron Courville, Cheng-Zhi Anna Huang, Jesse Engel</i></p>	20050
<p>Mirror Descent Policy Optimization</p> <p><i>Manan Tomar, Lior Shani, Yonathan Efroni, Mohammad Ghavamzadeh</i></p>	20077
<p>Proof Artifact Co-Training for Theorem Proving with Language Models</p> <p><i>Jesse Michael Han, Jason Rute, Yuhuai Wu, Edward Ayers, Stanislas Polu</i></p>	20101
<p>Multi-Stage Episodic Control for Strategic Exploration in Text Games</p> <p><i>Jens Tuyls, Shunyu Yao, Sham M. Kakade, Karthik R Narasimhan</i></p>	20127
<p>Enabling Arbitrary Translation Objectives with Adaptive Tree Search</p> <p><i>Wang Ling, Wojciech Stokowiec, Domenic Donato, Chris Dyer, Lei Yu, Laurent Sartran, Austin Matthews</i></p>	20149
<p>The Role of Pretrained Representations for the OOD Generalization of RL Agents</p> <p><i>Frederik Trãuble, Andrea Dittadi, Manuel Wuthrich, Felix Widmaier, Peter Vincent Gehler, Ole Winther, Francesco Locatello, Olivier Bachem, Bernhard Schãlkopf, Stefan Bauer</i></p>	20166
<p>StyleNeRF: A Style-Based 3D Aware Generator for High-Resolution Image Synthesis</p> <p><i>Jiatao Gu, Lingjie Liu, Peng Wang, Christian Theobalt</i></p>	20186
<p>Neural Models for Output-Space Invariance in Combinatorial Problems</p> <p><i>Yatin Nandwani, Vidit Jain, Mausam ., Parag Singla</i></p>	20211
<p>Is Importance Weighting Incompatible with Interpolating Classifiers?</p> <p><i>Ke Alexander Wang, Niladri Shekhar Chatterji, Saminul Haque, Tatsunori Hashimoto</i></p>	20231
<p>How Many Degrees of Freedom Do We Need to Train Deep Networks: A Loss Landscape Perspective</p> <p><i>Brett W Larsen, Stanislav Fort, Nic Becker, Surya Ganguli</i></p>	20265
<p>Relating Transformers to Models and Neural Representations of the Hippocampal Formation</p> <p><i>James C. R. Whittington, Joseph Warren, Tim E. J. Behrens</i></p>	20285

<p>Graph-Less Neural Networks: Teaching Old MLPs New Tricks Via Distillation</p> <p><i>Shichang Zhang, Yozen Liu, Yizhou Sun, Neil Shah</i></p>	20305
<p>Message Passing Neural PDE Solvers</p> <p><i>Johannes Brandstetter, Daniel E. Worrall, Max Welling</i></p>	20326
<p>Comparing Distributions by Measuring Differences that Affect Decision Making</p> <p><i>Shengjia Zhao, Abhishek Sinha, Yutong He, Aidan Perreault, Jiaming Song, Stefano Ermon</i></p>	20353
<p>Permutation-Based SGD: Is Random Optimal?</p> <p><i>Shashank Rajput, Kangwook Lee, Dimitris Papailiopoulos</i></p>	20373
<p>Asymmetry Learning for Counterfactually-Invariant Classification in OOD Tasks</p> <p><i>S Chandra Mouli, Bruno Ribeiro</i></p>	20422
<p>Decentralized Learning for Overparameterized Problems: A Multi-Agent Kernel Approximation Approach</p> <p><i>Prashant Khanduri, Haibo Yang, Mingyi Hong, Jia Liu, Hoi To Wai, Sijia Liu</i></p>	20441
<p>Reward Uncertainty for Exploration in Preference-Based Reinforcement Learning</p> <p><i>Xinran Liang, Katherine Shu, Kimin Lee, Pieter Abbeel</i></p>	20454
<p>Practical Conditional Neural Process Via Tractable Dependent Predictions</p> <p><i>Stratis Markou, James Requeima, Wessel Bruinsma, Anna Vaughan, Richard E Turner</i></p>	20471
<p>Robbing the Fed: Directly Obtaining Private Data in Federated Learning with Modified Models</p> <p><i>Liam H Fowl, Jonas Geiping, Wojciech Czaja, Micah Goldblum, Tom Goldstein</i></p>	20494
<p>Language-Biased Image Classification: Evaluation Based on Semantic Representations</p> <p><i>Yoann Lemesle, Masataka Sawayama, Guillermo Valle-Perez, Maxime Adolphe, Hãlã`ne Sauzãon, Pierre-Yves Oudeyer</i></p>	20519
<p>Chaos is a Ladder: A New Theoretical Understanding of Contrastive Learning Via Augmentation Overlap</p> <p><i>Yifei Wang, Qi Zhang, Yisen Wang, Jiansheng Yang, Zhouchen Lin</i></p>	20539
<p>Robust Learning Meets Generative Models: Can Proxy Distributions Improve Adversarial Robustness?</p> <p><i>Vikash Sehwal, Saeed Mahloujifar, Tinashe Handina, Sihui Dai, Chong Xiang, Mung Chiang, Prateek Mittal</i></p>	20560
<p>A Tale of Two Flows: Cooperative Learning of Langevin Flow and Normalizing Flow Toward Energy-Based Model</p> <p><i>Jianwen Xie, Yaxuan Zhu, Jun Li, Ping Li</i></p>	20590
<p>How Did the Model Change? Efficiently Assessing Machine Learning API Shifts</p> <p><i>Lingjiao Chen, Matei Zaharia, James Zou</i></p>	20613

<p>Map Induction: Compositional Spatial Submap Learning for Efficient Exploration in Novel Environments</p> <p><i>Sugandha Sharma, Aidan Curtis, Marta Kryven, Joshua B. Tenenbaum, Ila R Fiete</i></p>	20643
<p>High Probability Bounds for a Class of Nonconvex Algorithms with AdaGrad Stepsize</p> <p><i>Ali Kavis, Kfir Yehuda Levy, Volkan Cevher</i></p>	20665
<p>The MultiBERTs: BERT Reproductions for Robustness Analysis</p> <p><i>Thibault Sellam, Steve Yadlowsky, Ian Tenney, Jason Wei, Naomi Saphra, Alexander D'Amour, Tal Linzen, Jasmijn Bastings, Iulia Raluca Turc, Jacob Eisenstein, Dipanjan Das, Ellie Pavlick</i></p>	20692
<p>Collapse by Conditioning: Training Class-Conditional GANs with Limited Data</p> <p><i>Mohamad Shahbazi, Martin Danelljan, Danda Pani Paudel, Luc Van Gool</i></p>	20722
<p>Constructing a Good Behavior Basis for Transfer Using Generalized Policy Updates</p> <p><i>Safa Alver, Doina Precup</i></p>	20743
<p>Non-Linear Operator Approximations for Initial Value Problems</p> <p><i>Gaurav Gupta, Xiongye Xiao, Radu Balan, Paul Bogdan</i></p>	20760
<p>TAMP-S2GCNets: Coupling Time-Aware Multipersistence Knowledge Representation with Spatio-Supra Graph Convolutional Networks for Time-Series Forecasting</p> <p><i>Yuzhou Chen, Ignacio Segovia-Dominguez, Baris Coskunuzer, Yulia Gel</i></p>	20782
<p>Neural Parameter Allocation Search</p> <p><i>Bryan A. Plummer, Nikoli Dryden, Julius Frost, Torsten Hoefler, Kate Saenko</i></p>	20810
<p>Counterfactual Plans Under Distributional Ambiguity</p> <p><i>Ngoc Bui, Duy Nguyen, Viet Anh Nguyen</i></p>	20831
<p>Learning Distributionally Robust Models at Scale Via Composite Optimization</p> <p><i>Farzin Haddadpour, Mohammad Mahdi Kamani, Mehrdad Mahdavi, Amin Karbasi</i></p>	20850
<p>Offline Reinforcement Learning with Implicit Q-Learning</p> <p><i>Ilya Kostrikov, Ashvin Nair, Sergey Levine</i></p>	20876
<p>Linking Emergent and Natural Languages Via Corpus Transfer</p> <p><i>Shunyu Yao, Mo Yu, Yang Zhang, Karthik R Narasimhan, Joshua B. Tenenbaum, Chuang Gan</i></p>	20887
<p>Reinforcement Learning with Sparse Rewards Using Guidance from Offline Demonstration</p> <p><i>Desik Rengarajan, Gargi Vaidya, Akshay Sarvesh, Dileep Kalathil, Srinivas Shakkottai</i></p>	20905
<p>Optimizing Neural Networks with Gradient Lexicase Selection</p> <p><i>Li Ding, Lee Spector</i></p>	20926
<p>The Effects of Reward Misspecification: Mapping and Mitigating Misaligned Models</p> <p><i>Alexander Pan, Kush Bhatia, Jacob Steinhardt</i></p>	20941

<p>Illiterate DALL-E Learns to Compose</p> <p>Gautam Singh, Fei Deng, Sungjin Ahn</p>	20960
<p>Extending the WILDS Benchmark for Unsupervised Adaptation</p> <p>Shiori Sagawa, Pang Wei Koh, Tony Lee, Irena Gao, Sang Michael Xie, Kendrick Shen, Ananya Kumar, Weihua Hu, Michihiro Yasunaga, Henrik Marklund, Sara Beery, Etienne David, Ian Stavness, Wei Guo, Jure Leskovec, Kate Saenko, Tatsunori Hashimoto, Sergey Levine, Chelsea Finn, Percy Liang</p>	21006
<p>HTLM: Hyper-Text Pre-Training and Prompting of Language Models</p> <p>Armen Aghajanyan, Dmytro Okhonko, Mike Lewis, Mandar Joshi, Hu Xu, Gargi Ghosh, Luke Zettlemoyer</p>	21057
<p>MetaMorph: Learning Universal Controllers with Transformers</p> <p>Agrim Gupta, Linxi Fan, Surya Ganguli, Li Fei-Fei</p>	21070
<p>Data Poisoning Won't Save You from Facial Recognition</p> <p>Evani Radiya-Dixit, Sanghyun Hong, Nicholas Carlini, Florian Tramèr</p>	21088
<p>Fairness in Representation for Multilingual NLP: Insights from Controlled Experiments on Conditional Language Modeling</p> <p>Ada Wan</p>	21109
<p>The Role of Permutation Invariance in Linear Mode Connectivity of Neural Networks</p> <p>Rahim Entezari, Hanie Sedghi, Olga Saukh, Behnam Neyshabur</p>	21149
<p>Discovering Latent Concepts Learned in BERT</p> <p>Fahim Dalvi, Abdul Rafee Khan, Firoj Alam, Nadir Durrani, Jia Xu, Hassan Sajjad</p>	21173
<p>Implicit Bias of MSE Gradient Optimization in Underparameterized Neural Networks</p> <p>Benjamin Bowman, Guido Montufar</p>	21198
<p>A Fast and Accurate Splitting Method for Optimal Transport: Analysis and Implementation</p> <p>Vien V. Mai, Jacob Lindbäck, Mikael Johansson</p>	21255
<p>Autonomous Learning of Object-Centric Abstractions for High-Level Planning</p> <p>Steven James, Benjamin Rosman, George Konidaris</p>	21276
<p>MCMC Should Mix: Learning Energy-Based Model with Neural Transport Latent Space MCMC</p> <p>Erik Nijkamp, Ruiqi Gao, Pavel Sountsov, Srinivas Vasudevan, Bo Pang, Song-Chun Zhu, Ying Nian Wu</p>	21306
<p>Value Gradient Weighted Model-Based Reinforcement Learning</p> <p>Claas A Voelcker, Victor Liao, Animesh Garg, Amir-Massoud Farahmand</p>	21324
<p>Imitation Learning from Observations Under Transition Model Disparity</p> <p>Tanmay Gangwani, Yuan Zhou, Jian Peng</p>	21343

Likelihood Training of Schrödinger Bridge Using Forward-Backward SDEs Theory <i>Tianrong Chen, Guan-Horng Liu, Evangelos Theodorou</i>	21358
ILQR-VAE : Control-Based Learning of Input-Driven Dynamics with Applications to Neural Data <i>Marine Schimel, Ta-Chu Kao, Kristopher T Jensen, Guillaume Hennequin</i>	21385
On the Convergence of mSGD and AdaGrad for Stochastic Optimization <i>Ruinan Jin, Yu Xing, Xingkang He</i>	21415
Adaptive Wavelet Transformer Network for 3D Shape Representation Learning <i>Hao Huang, Yi Fang</i>	21471
The Convex Geometry of Backpropagation: Neural Network Gradient Flows Converge to Extreme Points of the Dual Convex Program <i>Yifei Wang, Mert Pilanci</i>	21487
Fair Normalizing Flows <i>Mislav Balunovic, Anian Ruoss, Martin Vechev</i>	21515
Pareto Policy Adaptation <i>Panagiotis Kyriakis, Jyotirmoy Deshmukh, Paul Bogdan</i>	21536
Towards Empirical Sandwich Bounds on the Rate-Distortion Function <i>Yibo Yang, Stephan Mandt</i>	21551
On Bridging Generic and Personalized Federated Learning for Image Classification <i>Hong-You Chen, Wei-Lun Chao</i>	21579
CKConv: Continuous Kernel Convolution for Sequential Data <i>David W. Romero, Anna Kuzina, Erik J Bekkers, Jakub Mikolaj Tomczak, Mark Hoogendoorn</i>	21611
Learning Transferable Reward for Query Object Localization with Policy Adaptation <i>Tingfeng Li, Shaobo Han, Martin Renqiang Min, Dimitris N. Metaxas</i>	21636
How Attentive Are Graph Attention Networks? <i>Shaked Brody, Uri Alon, Eran Yahav</i>	21659
Optimal Transport for Causal Discovery <i>Ruibo Tu, Kun Zhang, Hedvig Kjellstrom, Cheng Zhang</i>	21685
Phenomenology of Double Descent in Finite-Width Neural Networks <i>Sidak Pal Singh, Aurelien Lucchi, Thomas Hofmann, Bernhard Schölkopf</i>	21710
Deep ReLU Networks Preserve Expected Length <i>Boris Hanin, Ryan Jeong, David Rolnick</i>	21748
A First-Occupancy Representation for Reinforcement Learning <i>Ted Moskowitz, Spencer R Wilson, Maneesh Sahani</i>	21767

<p>TPU-GAN: Learning Temporal Coherence from Dynamic Point Cloud Sequences</p> <p><i>Zijie Li, Tianqin Li, Amir Barati Farimani</i></p>	21794
<p>CoBERL: Contrastive BERT for Reinforcement Learning</p> <p><i>Andrea Banino, Adria Puigdomenech Badia, Jacob C Walker, Tim Scholtes, Jovana Mitrovic, Charles Blundell</i></p>	21814
<p>Relational Multi-Task Learning: Modeling Relations Between Data and Tasks</p> <p><i>Kaidi Cao, Jiaxuan You, Jure Leskovec</i></p>	21849
<p>Prototype Memory and Attention Mechanisms for Few Shot Image Generation</p> <p><i>Tianqin Li, Zijie Li, Andrew Luo, Harold Rockwell, Amir Barati Farimani, Tai Sing Lee</i></p>	21863
<p>SOSP: Efficiently Capturing Global Correlations by Second-Order Structured Pruning</p> <p><i>Manuel Nonnenmacher, Thomas Pfeil, Ingo Steinwart, David Reeb</i></p>	21882
<p>Learning to Generalize Across Domains on Single Test Samples</p> <p><i>Zehao Xiao, Xiantong Zhen, Ling Shao, Cees G. M. Snoek</i></p>	21907
<p>Orchestrated Value Mapping for Reinforcement Learning</p> <p><i>Mehdi Fatemi, Arash Tavakoli</i></p>	21927
<p>Weighted Training for Cross-Task Learning</p> <p><i>Shuxiao Chen, Koby Crammer, Hangfeng He, Dan Roth, Weijie J Su</i></p>	21950
<p>Memory Augmented Optimizers for Deep Learning</p> <p><i>Paul-Aymeric Martin McRae, Prasanna Parthasarathi, Mido Assran, Sarath Chandar</i></p>	21974
<p>Hidden Convexity of Wasserstein GANs: Interpretable Generative Models with Closed-Form Solutions</p> <p><i>Arda Sahiner, Tolga Ergen, Batu Ozturkler, Burak Bartan, John M. Pauly, Morteza Mardani, Mert Pilanci</i></p>	21998
<p>Neural Collapse Under MSE Loss: Proximity to and Dynamics on the Central Path</p> <p><i>X. Y. Han, Vardan Papyan, David L. Donoho</i></p>	22032
<p>Visual Representation Learning Does Not Generalize Strongly Within the Same Domain</p> <p><i>Lukas Schott, Julius Von Kagelgen, Frederik Trauble, Peter Vincent Gehler, Chris Russell, Matthias Bethge, Bernhard Schalkopf, Francesco Locatello, Wieland Brendel</i></p>	22068
<p>Efficient Self-Supervised Vision Transformers for Representation Learning</p> <p><i>Chunyuan Li, Jianwei Yang, Pengchuan Zhang, Mei Gao, Bin Xiao, Xiyang Dai, Lu Yuan, Jianfeng Gao</i></p>	22102
<p>Low-Budget Active Learning Via Wasserstein Distance: An Integer Programming Approach</p> <p><i>Rafid Mahmood, Sanja Fidler, Marc T Law</i></p>	22129
<p>Joint Shapley Values: A Measure of Joint Feature Importance</p> <p><i>Chris Harris, Richard Pymar, Colin Rowat</i></p>	22160

<p>Provably Robust Adversarial Examples</p> <p><i>Dimitar Iliev Dimitrov, Gagandeep Singh, Timon Gehr, Martin Vechev</i></p>	22176
<p>Boosting Randomized Smoothing with Variance Reduced Classifiers</p> <p><i>Mikls Z. Horvth, Mark Niklas Mueller, Marc Fischer, Martin Vechev</i></p>	22201
<p>Relational Learning with Variational Bayes</p> <p><i>Kuang-Hung Liu</i></p>	22234
<p>Augmented Sliced Wasserstein Distances</p> <p><i>Xiongjie Chen, Yongxin Yang, Yunpeng Li</i></p>	22258
<p>Toward Faithful Case-Based Reasoning Through Learning Prototypes in a Nearest Neighbor-Friendly Space.</p> <p><i>Seyed Omid Davoudi, Majid Komeili</i></p>	22295
<p>Distributional Reinforcement Learning with Monotonic Splines</p> <p><i>Yudong Luo, Guiliang Liu, Haonan Duan, Oliver Schulte, Pascal Poupart</i></p>	22308
<p>Bridging the Gap: Providing Post-Hoc Symbolic Explanations for Sequential Decision-Making Problems with Inscrutable Representations</p> <p><i>Sarath Sreedharan, Utkarsh Soni, Mudit Verma, Siddharth Srivastava, Subbarao Kambhampati</i></p>	22327
<p>Network Augmentation for Tiny Deep Learning</p> <p><i>Han Cai, Chuang Gan, Ji Lin, Song Han</i></p>	22354
<p>Graphon Based Clustering and Testing of Networks: Algorithms and Theory</p> <p><i>Mahalakshmi Sabanayagam, Leena Chennuru Vankadara, Debarghya Ghoshdastidar</i></p>	22367
<p>NASViT: Neural Architecture Search for Efficient Vision Transformers with Gradient Conflict Aware Supernet Training</p> <p><i>Chengyue Gong, Dilin Wang, Meng Li, Xinlei Chen, Zhicheng Yan, Yuandong Tian, Qiang Liu, Vikas Chandra</i></p>	22393
<p>Learning Curves for SGD on Structured Features</p> <p><i>Blake Bordelon, Cengiz Pehlevan</i></p>	22411
<p>Towards Understanding the Robustness Against Evasion Attack on Categorical Data</p> <p><i>Hongyan Bao, Yufei Han, Yujun Zhou, Yun Shen, Xiangliang Zhang</i></p>	22432
<p>Tracking the Risk of a Deployed Model and Detecting Harmful Distribution Shifts</p> <p><i>Aleksandr Podkopaev, Aaditya Ramdas</i></p>	22456
<p>The Effects of Invertibility on the Representational Complexity of Encoders in Variational Autoencoders</p> <p><i>Divyansh Pareek, Andrej Risteski</i></p>	22482

<p>SphereFace2: Binary Classification is All You Need for Deep Face Recognition</p> <p><i>Yandong Wen, Weiyang Liu, Adrian Weller, Bhiksha Raj, Rita Singh</i></p>	22514
<p>Particle Stochastic Dual Coordinate Ascent: Exponential Convergent Algorithm for Mean Field Neural Network Optimization</p> <p><i>Kazusato Oko, Taiji Suzuki, Atsushi Nitanda, Denny Wu</i></p>	22529
<p>Geometric and Physical Quantities Improve E(3) Equivariant Message Passing</p> <p><i>Johannes Brandstetter, Rob Hesselink, Elise Van Der Pol, Erik J Bekkers, Max Welling</i></p>	22572
<p>Pix2seq: A Language Modeling Framework for Object Detection</p> <p><i>Ting Chen, Saurabh Saxena, Lala Li, David J. Fleet, Geoffrey Hinton</i></p>	22604
<p>NodePiece: Compositional and Parameter-Efficient Representations of Large Knowledge Graphs</p> <p><i>Mikhail Galkin, Etienne Denis, Jiapeng Wu, William L. Hamilton</i></p>	22621
<p>Does Your Graph Need a Confidence Boost? Convergent Boosted Smoothing on Graphs with Tabular Node Features</p> <p><i>Jiuhai Chen, Jonas Mueller, Vassilis N. Ioannidis, Soji Adeshina, Yangkun Wang, Tom Goldstein, David Wipf</i></p>	22646
<p>Is Fairness Only Metric Deep? Evaluating and Addressing Subgroup Gaps in Deep Metric Learning</p> <p><i>Natalie Dullerud, Karsten Roth, Kimia Hamidieh, Nicolas Papernot, Marzyeh Ghassemi</i></p>	22667
<p>Triangle and Four Cycle Counting with Predictions in Graph Streams</p> <p><i>Justin Y Chen, Talya Eden, Piotr Indyk, Honghao Lin, Shyam Narayanan, Ronitt Rubinfeld, Sandeep Silwal, Tal Wagner, David Woodruff, Michael Zhang</i></p>	22700
<p>Poisoning and Backdooring Contrastive Learning</p> <p><i>Nicholas Carlini, Andreas Terzis</i></p>	22738
<p>Fast AdvProp</p> <p><i>Jieru Mei, Yucheng Han, Yutong Bai, Yixiao Zhang, Yingwei Li, Xianhang Li, Alan Yuille, Cihang Xie</i></p>	22751
<p>Expressivity of Emergent Languages is a Trade-Off Between Contextual Complexity and Unpredictability</p> <p><i>Shangmin Guo, Yi Ren, Kory Wallace Mathewson, Simon Kirby, Stefano V Albrecht, Kenny Smith</i></p>	22763
<p>Reducing Excessive Margin to Achieve a Better Accuracy Vs. Robustness Trade-Off</p> <p><i>Rahul Rade, Seyed-Mohsen Moosavi-Dezfooli</i></p>	22785
<p>Evading Adversarial Example Detection Defenses with Orthogonal Projected Gradient Descent</p> <p><i>Oliver Bryniarski, Nabeel Hingun, Pedro Pachuca, Vincent Wang, Nicholas Carlini</i></p>	22806
<p>Data-Efficient Graph Grammar Learning for Molecular Generation</p> <p><i>Minghao Guo, Veronika Thost, Beichen Li, Payel Das, Jie Chen, Wojciech Matusik</i></p>	22818

<p>Open-World Semi-Supervised Learning</p> <p><i>Kaidi Cao, Maria Brbic, Jure Leskovec</i></p>	22837
<p>Learning State Representations Via Retracing in Reinforcement Learning</p> <p><i>Changmin Yu, Dong Li, Jianye Hao, Jun Wang, Neil Burgess</i></p>	22856
<p>MT3: Multi-Task Multitrack Music Transcription</p> <p><i>Joshua P Gardner, Ian Simon, Ethan Manilow, Curtis Hawthorne, Jesse Engel</i></p>	22876
<p>Should I Run Offline Reinforcement Learning Or Behavioral Cloning?</p> <p><i>Aviral Kumar, Joey Hong, Anikait Singh, Sergey Levine</i></p>	22897
<p>FedBABU: Toward Enhanced Representation for Federated Image Classification</p> <p><i>Jaehoon Oh, Sangmook Kim, Se-Young Yun</i></p>	22933
<p>A Statistical Framework for Efficient Out of Distribution Detection in Deep Neural Networks</p> <p><i>Matan Haroush, Tzviel Frostig, Ruth Heller, Daniel Soudry</i></p>	22962
<p>Data Efficient Language-Supervised Zero-Shot Recognition with Optimal Transport Distillation</p> <p><i>Bichen Wu, Ruizhe Cheng, Peizhao Zhang, Tianren Gao, Joseph E. Gonzalez, Peter Vajda</i></p>	22993
<p>It Takes Two to Tango: Mixup for Deep Metric Learning</p> <p><i>Shashanka Venkataramanan, Bill Psomas, Ewa Kijak, Laurent Amsaleg, Konstantinos Karantzas, Yannis Avrithis</i></p>	23012
<p>Modeling Label Space Interactions in Multi-Label Classification Using Box Embeddings</p> <p><i>Dhruvesh Patel, Pavitra Dangati, Jay-Yoon Lee, Michael Boratko, Andrew McCallum</i></p>	23033
<p>DR3: Value-Based Deep Reinforcement Learning Requires Explicit Regularization</p> <p><i>Aviral Kumar, Rishabh Agarwal, Tengyu Ma, Aaron Courville, George Tucker, Sergey Levine</i></p>	23054
<p>Geometry-Consistent Neural Shape Representation with Implicit Displacement Fields</p> <p><i>Wang Yifan, Lukas Rahmann, Olga Sorkine-Hornung</i></p>	23095
<p>Multi-Agent MDP Homomorphic Networks</p> <p><i>Elise Van Der Pol, Herke Van Hoof, Frans A Oliehoek, Max Welling</i></p>	23115
<p>Coordination Among Neural Modules Through a Shared Global Workspace</p> <p><i>Anirudh Goyal, Aniket Rajiv Didolkar, Alex Lamb, Kartikeya Badola, Nan Rosemary Ke, Nasim Rahaman, Jonathan Binas, Charles Blundell, Michael Curtis Mozer, Yoshua Bengio</i></p>	23134
<p>Data-Driven Offline Optimization for Architecting Hardware Accelerators</p> <p><i>Aviral Kumar, Amir Yazdanbakhsh, Milad Hashemi, Kevin Swersky, Sergey Levine</i></p>	23157
<p>Perceiver IO: A General Architecture for Structured Inputs & Outputs</p> <p><i>Andrew Jaegle, Sebastian Borgeaud, Jean-Baptiste Alayrac, Carl Doersch, Catalin Ionescu, David Ding, Skanda Koppula, Daniel Zoran, Andrew Brock, Evan Shelhamer, Olivier J Henaff, Matthew Botvinick, Andrew Zisserman, Oriol Vinyals, Joao Carreira</i></p>	23185

<p>Sample Selection with Uncertainty of Losses for Learning with Noisy Labels</p> <p><i>Xiaobo Xia, Tongliang Liu, Bo Han, Mingming Gong, Jun Yu, Gang Niu, Masashi Sugiyama</i></p>	23214
<p>Efficient Token Mixing for Transformers Via Adaptive Fourier Neural Operators</p> <p><i>John Guibas, Morteza Mardani, Zongyi Li, Andrew Tao, Anima Anandkumar, Bryan Catanzaro</i></p>	23237
<p>Provable Adaptation Across Multiway Domains Via Representation Learning</p> <p><i>Zhili Feng, Shaobo Han, Simon Shaolei Du</i></p>	23252
<p>ExT5: Towards Extreme Multi-Task Scaling for Transfer Learning</p> <p><i>Vamsi Aribandi, Yi Tay, Tal Schuster, Jinfeng Rao, Huaixiu Steven Zheng, Sanket Vaibhav Mehta, Honglei Zhuang, Vinh Q. Tran, Dara Bahri, Jianmo Ni, Jai Gupta, Kai Hui, Sebastian Ruder, Donald Metzler</i></p>	23269
<p>On the Uncomputability of Partition Functions in Energy-Based Sequence Models</p> <p><i>Chu-Cheng Lin, Arya D. McCarthy</i></p>	23295
<p>COPA: Certifying Robust Policies for Offline Reinforcement Learning Against Poisoning Attacks</p> <p><i>Fan Wu, Linyi Li, Huan Zhang, Bhavya Kailkhura, Krishnaram Kenthapadi, Ding Zhao, Bo Li</i></p>	23317
<p>Chunked Autoregressive GAN for Conditional Waveform Synthesis</p> <p><i>Max Morrison, Rithesh Kumar, Kundan Kumar, Prem Seetharaman, Aaron Courville, Yoshua Bengio</i></p>	23358
<p>Who is the Strongest Enemy? Towards Optimal and Efficient Evasion Attacks in Deep RL</p> <p><i>Yanchao Sun, Ruijie Zheng, Yongyuan Liang, Furong Huang</i></p>	23377
<p>Trigger Hunting with a Topological Prior for Trojan Detection</p> <p><i>Xiaoling Hu, Xiao Lin, Michael Cogswell, Yi Yao, Susmit Jha, Chao Chen</i></p>	23417
<p>Convergent and Efficient Deep Q Learning Algorithm</p> <p><i>Zhikang T. Wang, Masahito Ueda</i></p>	23434
<p>A Theory of Tournament Representations</p> <p><i>Arun Rajkumar, Vishnu Veerathu, Abdul Bakey Mir</i></p>	23461
<p>Improving Non-Autoregressive Translation Models Without Distillation</p> <p><i>Xiao Shi Huang, Felipe Perez, Maksims Volkovs</i></p>	23477
<p>PolyLoss: A Polynomial Expansion Perspective of Classification Loss Functions</p> <p><i>Zhaoqi Leng, Mingxing Tan, Chenxi Liu, Ekin Dogus Cubuk, Jay Shi, Shuyang Cheng, Dragomir Anguelov</i></p>	23489
<p>Learning-Augmented k-Means Clustering</p> <p><i>Jon C. Ergun, Zhili Feng, Sandeep Silwal, David Woodruff, Samson Zhou</i></p>	23505
<p>Generalisation in Lifelong Reinforcement Learning Through Logical Composition</p> <p><i>Geraud Nangue Tasse, Steven James, Benjamin Rosman</i></p>	23535

<p>Understanding the Variance Collapse of SVGD in High Dimensions</p> <p><i>Jimmy Ba, Murat A Erdogdu, Marzyeh Ghassemi, Shengyang Sun, Taiji Suzuki, Denny Wu, Tianzong Zhang</i></p>	23556
<p>Do Users Benefit from Interpretable Vision? a User Study, Baseline, and Dataset</p> <p><i>Leon Sixt, Martin Schuessler, Oana-Iuliana Popescu, Philipp Wei��y, Tim Landgraf</i></p>	23582
<p>Bootstrapped Meta-Learning</p> <p><i>Sebastian Flennerhag, Yannick Schroecker, Tom Zahavy, Hado Van Hasselt, David Silver, Satinder Singh</i></p>	23605
<p>Memorizing Transformers</p> <p><i>Yuhuai Wu, Markus Norman Rabe, Delesley Hutchins, Christian Szegedy</i></p>	23642
<p>Gradient Importance Learning for Incomplete Observations</p> <p><i>Qitong Gao, Dong Wang, Joshua David Amason, Siyang Yuan, Chenyang Tao, Ricardo Henao, Majda Hadziahmetovic, Lawrence Carin, Miroslav Pajic</i></p>	23661
<p>Deconstructing the Inductive Biases of Hamiltonian Neural Networks</p> <p><i>Nate Gruver, Marc Anton Finzi, Samuel Don Stanton, Andrew Gordon Wilson</i></p>	23688
<p>NETWORK INSENSITIVITY TO PARAMETER NOISE VIA PARAMETER ATTACK DURING TRAINING</p> <p><i>Julian B��chel, Fynn Firouz Faber, Dylan Richard Muir</i></p>	23704
<p>From Stars to Subgraphs: Uplifting Any GNN with Local Structure Awareness</p> <p><i>Lingxiao Zhao, Wei Jin, Leman Akoglu, Neil Shah</i></p>	23733
<p>KL Guided Domain Adaptation</p> <p><i>A. Tuan Nguyen, Toan Tran, Yarin Gal, Philip Torr, Atilim Gunes Baydin</i></p>	23755
<p>Maximum N-Times Coverage for Vaccine Design</p> <p><i>Ge Liu, Alexander Dimitrakakis, Brandon Carter, David Gifford</i></p>	23775
<p>Universalizing Weak Supervision</p> <p><i>Changho Shin, Winfred Li, Harit Vishwakarma, Nicholas Carl Roberts, Frederic Sala</i></p>	23795
<p>Bayesian Framework for Gradient Leakage</p> <p><i>Mislav Balunovic, Dimitar Iliev Dimitrov, Robin Staab, Martin Vechev</i></p>	23827
<p>AlphaZero-Based Proof Cost Network to Aid Game Solving</p> <p><i>Ti-Rong Wu, Chung-Chin Shih, Ting Han Wei, Meng-Yu Tsai, Wei-Yuan Hsu, I-Chen Wu</i></p>	23843
<p>A Neural Tangent Kernel Perspective of Infinite Tree Ensembles</p> <p><i>Ryuichi Kanoh, Mahito Sugiyama</i></p>	23873
<p>Heteroscedastic Temporal Variational Autoencoder for Irregularly Sampled Time Series</p> <p><i>Satya Narayan Shukla, Benjamin Marlin</i></p>	23906

<p>DISSECT: Disentangled Simultaneous Explanations Via Concept Traversals</p> <p><i>Asma Ghandeharioun, Been Kim, Chun-Liang Li, Brendan Jou, Brian Eoff, Rosalind Picard</i></p>	23926
<p>Scene Transformer: A Unified Architecture for Predicting Future Trajectories of Multiple Agents</p> <p><i>Jiquan Ngiam, Vijay Vasudevan, Benjamin Caine, Zhengdong Zhang, Hao-Tien Lewis Chiang, Jeffrey Ling, Rebecca Roelofs, Alex Bewley, Chenxi Liu, Ashish Venugopal, David J Weiss, Benjamin Sapp, Zhifeng Chen, Jonathon Shlens</i></p>	23950
<p>Contextualized Scene Imagination for Generative Commonsense Reasoning</p> <p><i>Peifeng Wang, Jonathan Zamora, Junfeng Liu, Filip Ilievski, Muhao Chen, Xiang Ren</i></p>	23975
<p>EigenGame Unloaded: When Playing Games is Better than Optimizing</p> <p><i>Ian Gemp, Brian McWilliams, Claire Vernade, Thore Graepel</i></p>	23992
<p>Understanding Latent Correlation-Based Multiview Learning and Self-Supervision: An Identifiability Perspective</p> <p><i>Qi Lyu, Xiao Fu, Weiran Wang, Songtao Lu</i></p>	24016
<p>Fortuitous Forgetting in Connectionist Networks</p> <p><i>Hattie Zhou, Ankit Vani, Hugo Larochelle, Aaron Courville</i></p>	24053
<p>Optimal Representations for Covariate Shift</p> <p><i>Yangjun Ruan, Yann Dubois, Chris J. Maddison</i></p>	24077
<p>VC Dimension of Partially Quantized Neural Networks in the Overparametrized Regime</p> <p><i>Yutong Wang, Clayton Scott</i></p>	24116
<p>Leveraging Unlabeled Data to Predict Out-Of-Distribution Performance</p> <p><i>Saurabh Garg, Sivaraman Balakrishnan, Zachary Chase Lipton, Behnam Neyshabur, Hanie Sedghi</i></p>	24138
<p>On Distributed Adaptive Optimization with Gradient Compression</p> <p><i>Xiaoyun Li, Belhal Karimi, Ping Li</i></p>	24166
<p>Provably Filtering Exogenous Distractors Using Multistep Inverse Dynamics</p> <p><i>Yonathan Efroni, Dipendra Misra, Akshay Krishnamurthy, Alekh Agarwal, John Langford</i></p>	24188
<p>Minimax Optimization with Smooth Algorithmic Adversaries</p> <p><i>Tanner Fiez, Chi Jin, Praneeth Netrapalli, Lillian J Ratliff</i></p>	24199
<p>A Program to Build $E(N)$-Equivariant Steerable CNNs</p> <p><i>Gabriele Cesa, Leon Lang, Maurice Weiler</i></p>	24230
<p>Latent Variable Sequential Set Transformers for Joint Multi-Agent Motion Prediction</p> <p><i>Roger Girgis, Florian Golemo, Felipe Codevilla, Martin Weiss, Jim Aldon D'Souza, Samira Ebrahimi Kahou, Felix Heide, Christopher Pal</i></p>	24292
<p>A Comparison of Hamming Errors of Representative Variable Selection Methods</p> <p><i>Tracy Ke, Longlin Wang</i></p>	24318

<p>ClimateGAN: Raising Climate Change Awareness by Generating Images of Floods</p> <p>Victor Schmidt, Alexandra Luccioni, Mălisande Teng, Tianyu Zhang, Alexia Reynaud, Sunand Raghupathi, Gautier Cosne, Adrien Juraver, Vahe Vardanyan, Alex Hernández-García, Yoshua Bengio</p>	24329
<p>Exposing the Implicit Energy Networks Behind Masked Language Models Via Metropolis--Hastings</p> <p>Kartik Goyal, Chris Dyer, Taylor Berg-Kirkpatrick</p>	24356
<p>Towards General Function Approximation in Zero-Sum Markov Games</p> <p>Baihe Huang, Jason D. Lee, Zhaoran Wang, Zhuoran Yang</p>	24372
<p>DIVA: Dataset Derivative of a Learning Task</p> <p>Yonatan Dukler, Alessandro Achille, Giovanni Paolini, Avinash Ravichandran, Marzia Polito, Stefano Soatto</p>	24404
<p>Graph Condensation for Graph Neural Networks</p> <p>Wei Jin, Lingxiao Zhao, Shichang Zhang, Yozen Liu, Jiliang Tang, Neil Shah</p>	24418
<p>Selective Ensembles for Consistent Predictions</p> <p>Emily Black, Klas Leino, Matt Fredrikson</p>	24437
<p>On Evaluation Metrics for Graph Generative Models</p> <p>Rylee Thompson, Boris Knyazev, Elahe Ghalebi, Jungtaek Kim, Graham W. Taylor</p>	24462
<p>The Hidden Convex Optimization Landscape of Regularized Two-Layer ReLU Networks: An Exact Characterization of Optimal Solutions</p> <p>Yifei Wang, Jonathan Lacotte, Mert Pilanci</p>	24492
<p>Mind the Gap: Domain Gap Control for Single Shot Domain Adaptation for Generative Adversarial Networks</p> <p>Peihao Zhu, Rameen Abdal, John Femiani, Peter Wonka</p>	24518
<p>Benchmarking the Spectrum of Agent Capabilities</p> <p>Danijar Hafner</p>	24538
<p>Learning to Map for Active Semantic Goal Navigation</p> <p>Georgios Georgakis, Bernadette Bucher, Karl Schmeckpeper, Siddharth Singh, Kostas Daniilidis</p>	24559
<p>Minibatch Vs Local SGD with Shuffling: Tight Convergence Bounds and Beyond</p> <p>Chulhee Yun, Shashank Rajput, Suvrit Sra</p>	24578
<p>Learning Meta-Features for AutoML</p> <p>Herilalaina Rakotoarison, Louisot Milijaona, Andry Rasoanaivo, Michele Sebag, Marc Schoenauer</p>	24654
<p>Audio Lottery: Speech Recognition Made Ultra-Lightweight, Noise-Robust, and Transferable</p> <p>Shaojin Ding, Tianlong Chen, Zhangyang Wang</p>	24687

<p>Capturing Structural Locality in Non-Parametric Language Models</p> <p><i>Frank F. Xu, Junxian He, Graham Neubig, Vincent Josua Hellendoorn</i></p>	24708
<p>Distribution Compression in Near-Linear Time</p> <p><i>Abhishek Shetty, Raaz Dwivedi, Lester Mackey</i></p>	24724
<p>Learning Fast Samplers for Diffusion Models by Differentiating Through Sample Quality</p> <p><i>Daniel Watson, William Chan, Jonathan Ho, Mohammad Norouzi</i></p>	24748
<p>GreaseLM: Graph REASONing Enhanced Language Models</p> <p><i>Xikun Zhang, Antoine Bosselut, Michihiro Yasunaga, Hongyu Ren, Percy Liang, Christopher D Manning, Jure Leskovec</i></p>	24765
<p>Complete Verification Via Multi-Neuron Relaxation Guided Branch-And-Bound</p> <p><i>Claudio Ferrari, Mark Niklas Mueller, Nikola Jovanovia, Martin Vechev</i></p>	24781
<p>AdaMatch: A Unified Approach to Semi-Supervised Learning and Domain Adaptation</p> <p><i>David Berthelot, Rebecca Roelofs, Kihyuk Sohn, Nicholas Carlini, Alexey Kurakin</i></p>	24796
<p>Patch-Fool: Are Vision Transformers Always Robust Against Adversarial Perturbations?</p> <p><i>Yonggan Fu, Shun Yao Zhang, Shang Wu, Cheng Wan, Yingyan Lin</i></p>	24846
<p>Adversarial Support Alignment</p> <p><i>Shangyuan Tong, Timur Garipov, Yang Zhang, Shiyu Chang, Tommi S. Jaakkola</i></p>	24864
<p>Learning Towards the Largest Margins</p> <p><i>Xiong Zhou, Xianming Liu, Deming Zhai, Junjun Jiang, Xin Gao, Xiangyang Ji</i></p>	24891
<p>Neural Network Approximation Based on Hausdorff Distance of Tropical Zonotopes</p> <p><i>Panagiotis Misiakos, Georgios Smyrnis, George Retsinas, Petros Maragos</i></p>	24921
<p>DeSKO: Stability-Assured Robust Control with a Deep Stochastic Koopman Operator</p> <p><i>Minghao Han, Jacob Euler-Rolle, Robert K. Katzschmann</i></p>	24939
<p>A New Perspective on "How Graph Neural Networks Go Beyond Weisfeiler-Lehman?"</p> <p><i>Asiri Wijesinghe, Qing Wang</i></p>	24961
<p>Reinforcement Learning Under a Multi-Agent Predictive State Representation Model: Method and Theory</p> <p><i>Zhi Zhang, Zhuoran Yang, Han Liu, Pratap Tokekar, Furong Huang</i></p>	24984
<p>NeuPL: Neural Population Learning</p> <p><i>Siqi Liu, Luke Marris, Daniel Hennes, Josh Merel, Nicolas Heess, Thore Graepel</i></p>	25013
<p>Interacting Contour Stochastic Gradient Langevin Dynamics</p> <p><i>Wei Deng, Siqi Liang, Botao Hao, Guang Lin, Faming Liang</i></p>	25032
<p>Learning More Skills Through Optimistic Exploration</p> <p><i>Dj Strouse, Kate Baumli, David Warde-Farley, Volodymyr Mnih, Steven Stenberg Hansen</i></p>	25061

Strength of Minibatch Noise in SGD <i>Liu Ziyin, Kangqiao Liu, Takashi Mori, Masahito Ueda</i>	25088
On the Connection Between Local Attention and Dynamic Depth-Wise Convolution <i>Qi Han, Zejia Fan, Qi Dai, Lei Sun, Ming-Ming Cheng, Jiaying Liu, Jingdong Wang</i>	25121
Can an Image Classifier Suffice for Action Recognition? <i>Quanfu Fan, Chun-Fu Chen, Rameswar Panda</i>	25146
Non-Parallel Text Style Transfer with Self-Parallel Supervision <i>Ruibo Liu, Chongyang Gao, Chenyan Jia, Guangxuan Xu, Soroush Vosoughi</i>	25159
Optimization and Adaptive Generalization of Three Layer Neural Networks <i>Khashayar Gatmiry, Stefanie Jegelka, Jonathan Kelner</i>	25181
Understanding and Leveraging Overparameterization in Recursive Value Estimation <i>Chenjun Xiao, Bo Dai, Jincheng Mei, Oscar A Ramirez, Ramki Gummadi, Chris Harris, Dale Schuurmans</i>	25296
On Predicting Generalization Using GANs <i>Yi Zhang, Arushi Gupta, Nikunj Saunshi, Sanjeev Arora</i>	25318
On Improving Adversarial Transferability of Vision Transformers <i>Muzammal Naseer, Kanchana Ranasinghe, Salman Khan, Fahad Khan, Fatih Porikli</i>	25328
RotoGrad: Gradient Homogenization in Multitask Learning <i>Adri�jn Javaloy, Isabel Valera</i>	25352
Towards Training Billion Parameter Graph Neural Networks for Atomic Simulations <i>Anuroop Sriram, Abhishek Das, Brandon M Wood, Siddharth Goyal, C. Lawrence Zitnick</i>	25376
Neural Structured Prediction for Inductive Node Classification <i>Meng Qu, Huiyu Cai, Jian Tang</i>	25391
Pretraining Text Encoders with Adversarial Mixture of Training Signal Generators <i>Yu Meng, Chenyan Xiong, Payal Bajaj, Saurabh Tiwary, Paul N. Bennett, Jiawei Han, Xia Song</i>	25412
On the Convergence of Certified Robust Training with Interval Bound Propagation <i>Yihan Wang, Zhouxing Shi, Quanquan Gu, Cho-Jui Hsieh</i>	25430
Neural Link Prediction with Walk Pooling <i>Liming Pan, Cheng Shi, Ivan Dokmani�</i>	25449
CROP: Certifying Robust Policies for Reinforcement Learning Through Functional Smoothing <i>Fan Wu, Linyi Li, Zijian Huang, Yevgeniy Vorobeychik, Ding Zhao, Bo Li</i>	25467
It Takes Four to Tango: Multiagent Self Play for Automatic Curriculum Generation <i>Yuqing Du, Pieter Abbeel, Aditya Grover</i>	25501

Discriminative Similarity for Data Clustering Yingzhen Yang, Ping Li	25522
Non-Transferable Learning: A New Approach for Model Ownership Verification and Applicability Authorization Lixu Wang, Shichao Xu, Ruiqi Xu, Xiao Wang, Qi Zhu	25545
PI3NN: Out-Of-Distribution-Aware Prediction Intervals from Three Neural Networks Siyao Liu, Pei Zhang, Dan Lu, Guannan Zhang	25569
Neural Contextual Bandits with Deep Representation and Shallow Exploration Pan Xu, Zheng Wen, Handong Zhao, Quanquan Gu	25585
Planning in Stochastic Environments with a Learned Model Ioannis Antonoglou, Julian Schrittwieser, Sherjil Ozair, Thomas K Hubert, David Silver	25610
Sampling with Mirrored Stein Operators Jiaxin Shi, Chang Liu, Lester Mackey	25640
MaGNET: Uniform Sampling from Deep Generative Network Manifolds Without Retraining Ahmed Imtiaz Humayun, Randall Balestriero, Richard Baraniuk	25666
Evaluating Distributional Distortion in Neural Language Modeling Benjamin Lebrun, Alessandro Sordani, Timothy J. O'Donnell	25694
Divisive Feature Normalization Improves Image Recognition Performance in AlexNet Michelle Miller, Sueyeon Chung, Kenneth D. Miller	25712
PEARL: Data Synthesis Via Private Embeddings and Adversarial Reconstruction Learning Seng Pei Liew, Tsubasa Takahashi, Michihiko Ueno	25752
CrossBeam: Learning to Search in Bottom-Up Program Synthesis Kensen Shi, Hanjun Dai, Kevin Ellis, Charles Sutton	25774
Fast Regression for Structured Inputs Raphael A Meyer, Cameron N Musco, Christopher P Musco, David Woodruff, Samson Zhou	25791
Learning Hierarchical Structures with Differentiable Nondeterministic Stacks Brian Dusell, David Chiang	25813
Natural Language Descriptions of Deep Visual Features Evan Hernandez, Sarah Schwettmann, David Bau, Teona Bagashvili, Antonio Torralba, Jacob Andreas	25830
Value Function Spaces: Skill-Centric State Abstractions for Long-Horizon Reasoning Dhruv Shah, Peng Xu, Yao Lu, Ted Xiao, Alexander T Toshev, Sergey Levine, Brian Ichter	25851
Approximation and Learning with Deep Convolutional Models: A Kernel Perspective Alberto Bietti	25867

<p>Back2Future: Leveraging Backfill Dynamics for Improving Real-Time Predictions in Future</p> <p>Harshavardhan Kamarthi, Alexander Rodr�guez, B. Aditya Prakash</p>	25895
<p>Anisotropic Random Feature Regression in High Dimensions</p> <p>Gabriel Mel, Jeffrey Pennington</p>	25909
<p>Top-Label Calibration and Multiclass-To-Binary Reductions</p> <p>Chirag Gupta, Aaditya Ramdas</p>	25953
<p>Domino: Discovering Systematic Errors with Cross-Modal Embeddings</p> <p>Sabri Eyuboglu, Maya Varma, Khaled Kamal Saab, Jean-Benoit Delbrouck, Christopher Lee-Messer, Jared Dunnmon, James Zou, Christopher Re</p>	25990
<p>Resonance in Weight Space: Covariate Shift Can Drive Divergence of SGD with Momentum</p> <p>Kirby Banman, Garnet Liam Peet-Pare, Nidhi Hegde, Alona Fyshe, Martha White</p>	26018
<p>AEVA: Black-Box Backdoor Detection Using Adversarial Extreme Value Analysis</p> <p>Junfeng Guo, Ang Li, Cong Liu</p>	26050
<p>Dynamics-Aware Comparison of Learned Reward Functions</p> <p>Blake Wulfe, Logan Michael Ellis, Jean Mercat, Rowan Thomas McAllister, Adrien Gaidon</p>	26074
<p>Out-Of-Distribution Generalization in the Presence of Nuisance-Induced Spurious Correlations</p> <p>Aahlad Manas Puli, Lily H Zhang, Eric Karl Oermann, Rajesh Ranganath</p>	26097
<p>Sound Adversarial Audio-Visual Navigation</p> <p>Yinfeng Yu, Wenbing Huang, Fuchun Sun, Changan Chen, Yikai Wang, Xiaohong Liu</p>	26129
<p>Sample and Computation Redistribution for Efficient Face Detection</p> <p>Jia Guo, Jiankang Deng, Alexandros Lattas, Stefanos Zafeiriou</p>	26153
<p>Variational Predictive Routing with Nested Subjective Timescales</p> <p>Alexey Zakharov, Qinghai Guo, Zafeirios Fountas</p>	26170
<p>From Intervention to Domain Transportation: A Novel Perspective to Optimize Recommendation</p> <p>Da Xu, Yuting Ye, Chuanwei Ruan, Evren Korpeoglu, Sushant Kumar, Kannan Achan</p>	26192
<p>Diverse Client Selection for Federated Learning Via Submodular Maximization</p> <p>Ravikumar Balakrishnan, Tian Li, Tianyi Zhou, Nageen Himayat, Virginia Smith, Jeff Bilmes</p>	26226
<p>Sequence Approximation Using Feedforward Spiking Neural Network for Spatiotemporal Learning: Theory and Optimization Methods</p> <p>Xueyuan She, Saurabh Dash, Saibal Mukhopadhyay</p>	26244
<p>Improving Mutual Information Estimation with Annealed and Energy-Based Bounds</p> <p>Rob Brekelmans, Sicong Huang, Marzyeh Ghassemi, Greg Ver Steeg, Roger Baker Grosse, Alireza Makhzani</p>	26266

<p>DEGREE: Decomposition Based Explanation for Graph Neural Networks</p> <p><i>Qizhang Feng, Ninghao Liu, Fan Yang, Ruixiang Tang, Mengnan Du, Xia Hu</i></p>	26323
<p>Is Homophily a Necessity for Graph Neural Networks?</p> <p><i>Yao Ma, Xiaorui Liu, Neil Shah, Jiliang Tang</i></p>	26342
<p>Recycling Model Updates in Federated Learning: Are Gradient Subspaces Low-Rank?</p> <p><i>Sheikh Shams Azam, Seyyedali Hosseinalipour, Qiang Qiu, Christopher Brinton</i></p>	26370
<p>L0-Sparse Canonical Correlation Analysis</p> <p><i>Ofir Lindenbaum, Moshe Salhov, Amir Averbuch, Yuval Kluger</i></p>	26440
<p>Wish You Were Here: Hindsight Goal Selection for Long-Horizon Dexterous Manipulation</p> <p><i>Todor Davchev, Oleg Olegovich Sushkov, Jean-Baptiste Regli, Stefan Schaal, Yusuf Aytar, Markus Wulfmeier, Jon Scholz</i></p>	26459
<p>Variational Neural Cellular Automata</p> <p><i>Rasmus Berg Palm, Miguel Gonzalez Duque, Shyam Sudhakaran, Sebastian Risi</i></p>	26477
<p>Learning to Downsample for Segmentation of Ultra-High Resolution Images</p> <p><i>Chen Jin, Ryutaro Tanno, Thomy Mertzanidou, Eleftheria Panagiotaki, Daniel C. Alexander</i></p>	26498
<p>Wiring Up Vision: Minimizing Supervised Synaptic Updates Needed to Produce a Primate Ventral Stream</p> <p><i>Franziska Geiger, Martin Schrimpf, Tiago Marques, James J. Dicarlo</i></p>	26515