

11th International Conference on Learning Representations

(ICLR 2023)

1-5 May 2023, Kigali, Rwanda

ISBN: 978-1-7138-9925-9

Title and Authors	Page#
<i>Continuous-Discrete Convolution for Geometry-Sequence Modeling in Proteins</i> <i>Hehe Fan, Zhangyang Wang, Yi Yang, Mohan Kankanhalli</i>	1
<i>Patch-Level Contrasting without Patch Correspondence for Accurate and Dense Contrastive Representation Learning</i> <i>Shaofeng Zhang, Feng Zhu, Rui Zhao, Junchi Yan</i>	27
<i>Contextual Image Masking Modeling via Synergized Contrasting without View Augmentation for Faster and Better Visual Pretraining</i> <i>Shaofeng Zhang, Feng Zhu, Rui Zhao, Junchi Yan</i>	41
<i>DAG Matters! GFlowNets Enhanced Explainer for Graph Neural Networks</i> <i>Wenqian Li, Yinchuan Li, Zhigang Li, Jianye HAO, Yan Pang</i>	55
<i>Delving into Semantic Scale Imbalance</i> <i>Yanbiao Ma, Licheng Jiao, Fang Liu, Yuxin Li, Shuyuan Yang, Xu Liu</i>	80
<i>The Augmented Image Prior: Distilling 1000 Classes by Extrapolating from a Single Image</i> <i>Yuki M Asano, Aaqib Saeed</i>	121
<i>SIMPLE: Specialized Model-Sample Matching for Domain Generalization</i> <i>Ziyue Li, Kan Ren, XINYANG JIANG, Yifei Shen, Haipeng Zhang, Dongsheng Li</i>	151
<i>Molecular Geometry Pretraining with SE(3)-Invariant Denoising Distance Matching</i> <i>Shengchao Liu, Hongyu Guo, Jian Tang</i>	187
<i>DualAfford: Learning Collaborative Visual Affordance for Dual-gripper Manipulation</i> <i>Yan Zhao, Ruihai Wu, Zhehuan Chen, Yourong Zhang, Qingnan Fan, Kaichun Mo, Hao Dong</i>	214
<i>On amortizing convex conjugates for optimal transport</i> <i>Brandon Amos</i>	235
<i>Diversify and Disambiguate: Out-of-Distribution Robustness via Disagreement</i> <i>Yoonho Lee, Huaxiu Yao, Chelsea Finn</i>	261
<i>Surgical Fine-Tuning Improves Adaptation to Distribution Shifts</i> <i>Yoonho Lee, Annie S Chen, Fahim Tajwar, Ananya Kumar, Huaxiu Yao, Percy Liang, Chelsea Finn</i>	284

Representation Learning for Low-rank General-sum Markov Games Chengzhuo Ni, Yuda Song, Xuezhou Zhang, Zihan Ding, Chi Jin, Mengdi Wang	307
Pushing the Limits of Fewshot Anomaly Detection in Industry Vision: Graphcore Guoyang Xie, Jinbao Wang, Jiaqi Liu, Yaochu Jin, Feng Zheng	367
Designing BERT for Convolutional Networks: Sparse and Hierarchical Masked Modeling Keyu Tian, Yi Jiang, qishuai diao, Chen Lin, Liwei Wang, Zehuan Yuan	385
H2RBox: Horizontal Box Annotation is All You Need for Oriented Object Detection Xue Yang, Gefan Zhang, Wentong Li, Yue Zhou, Xuehui Wang, Junchi Yan	401
Achieve the Minimum Width of Neural Networks for Universal Approximation Yongqiang Cai	415
Suppressing the Heterogeneity: A Strong Feature Extractor for Few-shot Segmentation Zhengdong Hu, Yifan Sun, Yi Yang	430
A Multi-Grained Self-Interpretable Symbolic-Neural Model For Single/Multi-Labeled Text Classification Xiang Hu, XinYu KONG, Kewei Tu	445
BEVDistill: Cross-Modal BEV Distillation for Multi-View 3D Object Detection Zehui Chen, Zhenyu Li, Shiquan Zhang, Liangji Fang, Qinhong Jiang, Feng Zhao	471
Logical Entity Representation in Knowledge-Graphs for Differentiable Rule Learning Chi Han, Qizheng He, Charles Yu, Xinya Du, Hanghang Tong, Heng Ji	488
Minimax Optimal Kernel Operator Learning via Multilevel Training Jikai Jin, Yiping Lu, Jose Blanchet, Lexing Ying	508
Learning Domain-Agnostic Representation for Disease Diagnosis Churan Wang, Jing Li, Xinwei Sun, Fandong Zhang, Yizhou Yu, Yizhou Wang	538
TaskPrompter: Spatial-Channel Multi-Task Prompting for Dense Scene Understanding Hanrong Ye, Dan Xu	559
A Unified Framework for Soft Threshold Pruning Yanqi Chen, Zhengyu Ma, Wei Fang, Xiawu Zheng, Zhaofei Yu, Yonghong Tian	579
DropIT: Dropping Intermediate Tensors for Memory-Efficient DNN Training Joya Chen, Kai Xu, Yuhui Wang, Yifei Cheng, Angela Yao	608
Harnessing Out-Of-Distribution Examples via Augmenting Content and Style Zhuo Huang, Xiaobo Xia, Li Shen, Bo Han, Mingming Gong, Chen Gong, Tongliang Liu	625
Diagnosing and Rectifying Vision Models using Language Yuhui Zhang, Jeff Z. HaoChen, Shih-Cheng Huang, Kuan-Chieh Wang, James Zou, Serena Yeung	644

Trainability Preserving Neural Pruning Huan Wang, Yun Fu	671
Image as Set of Points Xu Ma, Yuqian Zhou, Huan Wang, Can Qin, Bin Sun, Chang Liu, Yun Fu	692
MapTR: Structured Modeling and Learning for Online Vectorized HD Map Construction Bencheng Liao, Shaoyu Chen, Xinggang Wang, Tianheng Cheng, Qian Zhang, Wenyu Liu, Chang Huang	710
IDEAL: Query-Efficient Data-Free Learning from Black-Box Models Jie Zhang, Chen Chen, Lingjuan Lyu	728
Decompose to Generalize: Species-Generalized Animal Pose Estimation Guangrui Li, Yifan Sun, Zongxin Yang, Yi Yang	743
Learnable Behavior Control: Breaking Atari Human World Records via Sample-Efficient Behavior Selection Jiajun Fan, Yuzheng Zhuang, Yuecheng Liu, Jianye HAO, Bin Wang, Jiangcheng Zhu, Hao Wang, Shu-Tao Xia	764
IS SYNTHETIC DATA FROM GENERATIVE MODELS READY FOR IMAGE RECOGNITION? Ruifei He, Shuyang Sun, Xin Yu, Chuhui Xue, Wenqing Zhang, Philip Torr, Song Bai, XIAOJUAN QI	776
kNN-Diffusion: Image Generation via Large-Scale Retrieval Shelly Sheynin, Oron Ashual, Adam Polyak, Uriel Singer, Oran Gafni, Eliya Nachmani, Yaniv Taigman	789
Bit-Pruning: A Sparse Multiplication-Less Dot-Product Yusuke Sekikawa, Shingo Yashima	817
DELTA: DEGRADATION-FREE FULLY TEST-TIME ADAPTATION Bowen Zhao, Chen Chen, Shu-Tao Xia	829
Cross-Level Distillation and Feature Denoising for Cross-Domain Few-Shot Classification Hao ZHENG, Runqi Wang, Jianzhuang Liu, Asako Kanezaki	854
Real-Time Image Demoirising on Mobile Devices Yuxin Zhang, Mingbao Lin, Xunchao Li, Han Liu, Guozhi Wang, Fei Chao, Ren Shuai, Yafei Wen, Xiaoxin Chen, Rongrong Ji	868
ChiroDiff: Modelling chirographic data with Diffusion Models Ayan Das, Yongxin Yang, Timothy Hospedales, Tao Xiang, Yi-Zhe Song	881
PatchDCT: Patch Refinement for High Quality Instance Segmentation Qinrou Wen, Jirui Yang, Xue Yang, Kewei Liang	896
Dilated convolution with learnable spacings Ismail Khalfaoui Hassani, Thomas Pellegrini, Timothée Masquelier	911

<p>Rotamer Density Estimator is an Unsupervised Learner of the Effect of Mutations on Protein-Protein Interaction</p> <p><i>Shitong Luo, Yufeng Su, Zuofan Wu, Chenpeng Su, Jian Peng, Jianzhu Ma</i></p>	940
<p>Fairness-aware Contrastive Learning with Partially Annotated Sensitive Attributes</p> <p><i>Fengda Zhang, Kun Kuang, Long Chen, Yuxuan Liu, Chao Wu, Jun Xiao</i></p>	960
<p>Learnable Topological Features For Phylogenetic Inference via Graph Neural Networks</p> <p><i>Cheng Zhang</i></p>	978
<p>MCAL: Minimum Cost Human-Machine Active Labeling</p> <p><i>Hang Qiu, Krishna Chintalapudi, Ramesh Govindan</i></p>	995
<p>The Devil is in the Wrongly-classified Samples: Towards Unified Open-set Recognition</p> <p><i>Jun CEN, Di Luan, Shiwei Zhang, Yixuan Pei, Yingya Zhang, Deli Zhao, Shaojie Shen, Qifeng Chen</i></p>	1009
<p>Function-Consistent Feature Distillation</p> <p><i>Dongyang Liu, Meina Kan, Shiguang Shan, Xilin CHEN</i></p>	1033
<p>Hierarchical Relational Learning for Few-Shot Knowledge Graph Completion</p> <p><i>Han Wu, Jie Yin, Bala Rajaratnam, Jianyuan Guo</i></p>	1054
<p>An Image is Worth One Word: Personalizing Text-to-Image Generation using Textual Inversion</p> <p><i>Rinon Gal, Yuval Alaluf, Yuval Atzmon, Or Patashnik, Amit Haim Bermano, Gal Chechik, Daniel Cohen-or</i></p>	1069
<p>Neural ePDOs: Spatially Adaptive Equivariant Partial Differential Operator Based Networks</p> <p><i>Lingshen He, Yuxuan Chen, Zhengyang Shen, Yibo Yang, Zhouchen Lin</i></p>	1100
<p>Neural Networks and the Chomsky Hierarchy</p> <p><i>Gregoire Deletang, Anian Ruoss, Jordi Grau-Moya, Tim Genewein, Li Kevin Wenliang, Elliot Catt, Chris Cundy, Marcus Hutter, Shane Legg, Joel Veness, Pedro A Ortega</i></p>	1111
<p>GFlowNets and variational inference</p> <p><i>Nikolay Malkin, Salem Lahlou, Tristan Deleu, Xu Ji, Edward J Hu, Katie E Everett, Dinghuai Zhang, Yoshua Bengio</i></p>	1143
<p>Volumetric Optimal Transportation by Fast Fourier Transform</p> <p><i>Na Lei, DONGSHENG An, Min Zhang, Xiaoyin Xu, David Gu</i></p>	1167
<p>Neural Collapse Inspired Feature-Classifer Alignment for Few-Shot Class-Incremental Learning</p> <p><i>Yibo Yang, Haobo Yuan, Xiangtai Li, Zhouchen Lin, Philip Torr, Dacheng Tao</i></p>	1189
<p>On The Inadequacy of Optimizing Alignment and Uniformity in Contrastive Learning of Sentence Representations</p> <p><i>Zhijie Nie, Richong Zhang, Yongyi Mao</i></p>	1207
<p>Stochastic Differentially Private and Fair Learning</p> <p><i>Andrew Lowy, Devansh Gupta, Meisam Razaviyayn</i></p>	1229

<p>Sample Complexity of Nonparametric Off-Policy Evaluation on Low-Dimensional Manifolds using Deep Networks</p> <p>Xiang Ji, Minshuo Chen, Mengdi Wang, Tuo Zhao</p>	1257
<p>E-CRF: Embedded Conditional Random Field for Boundary-caused Class Weights Confusion in Semantic Segmentation</p> <p>Jie Zhu, Huabin Huang, Banghuai Li, Leye Wang</p>	1294
<p>Neural Groundplans: Persistent Neural Scene Representations from a Single Image</p> <p>Prafull Sharma, Ayush Tewari, Yilun Du, Sergey Zakharov, Rares Andrei Ambrus, Adrien Gaidon, William T. Freeman, Fredo Durand, Joshua B. Tenenbaum, Vincent Sitzmann</p>	1314
<p>Extracting Robust Models with Uncertain Examples</p> <p>Guanlin Li, Guowen Xu, Shangwei Guo, Han Qiu, Jiwei Li, Tianwei Zhang</p>	1334
<p>Out-of-distribution Detection with Implicit Outlier Transformation</p> <p>Qizhou Wang, Junjie Ye, Feng Liu, Quanyu Dai, Marcus Kalander, Tongliang Liu, Jianye HAO, Bo Han</p>	1363
<p>Alternating Differentiation for Optimization Layers</p> <p>Haixiang Sun, Ye Shi, Jingya Wang, Hoang Duong Tuan, H. Vincent Poor, Dacheng Tao</p>	1385
<p>Measuring axiomatic soundness of counterfactual image models</p> <p>Miguel Monteiro, Fabio De Sousa Ribeiro, Nick Pawlowski, Daniel C. Castro, Ben Glocker</p>	1406
<p>Rethinking skip connection model as a learnable Markov chain</p> <p>Dengsheng Chen, Jie Hu, Wenwen Qiang, Xiaoming Wei, Enhua Wu</p>	1438
<p>Revisiting the Entropy Semiring for Neural Speech Recognition</p> <p>Oscar Chang, Dongseong Hwang, Olivier Siohan</p>	1453
<p>De Novo Molecular Generation via Connection-aware Motif Mining</p> <p>Zijie Geng, Shufang Xie, Yingce Xia, Lijun Wu, Tao Qin, Jie Wang, Yongdong Zhang, Feng Wu, Tie-Yan Liu</p>	1476
<p>Token Merging: Your ViT But Faster</p> <p>Daniel Bolya, Cheng-Yang Fu, Xiaoliang Dai, Peizhao Zhang, Christoph Feichtenhofer, Judy Hoffman</p>	1498
<p>Limitless Stability for Graph Convolutional Networks</p> <p>Christian Koke</p>	1518
<p>EAGLE: Large-scale Learning of Turbulent Fluid Dynamics with Mesh Transformers</p> <p>Steeven JANNY, Aurélien BÃ©nÃ©teau, Madiha Nadri, Julie Digne, Nicolas THOME, Christian Wolf</p>	1570
<p>Topology-aware Robust Optimization for Out-of-Distribution Generalization</p> <p>Fengchun Qiao, Xi Peng</p>	1595
<p>STREET: A MULTI-TASK STRUCTURED REASONING AND EXPLANATION BENCHMARK</p> <p>Daniilo Neves Ribeiro, Shen Wang, Xiaofei Ma, Henghui Zhu, Rui Dong, Deguang Kong, Juliette Burger, Anjelica Ramos, zhiheng huang, William Yang Wang, George Karypis, Bing Xiang, Dan Roth</p>	1612

<p>Robust and Controllable Object-Centric Learning through Energy-based Models</p> <p>Ruixiang ZHANG, Tong Che, Boris Ivanovic, Renhao Wang, Marco Pavone, Yoshua Bengio, Liam Paull</p>	1632
<p>ISS: Image as Stepping Stone for Text-Guided 3D Shape Generation</p> <p>Zhengzhe Liu, Peng Dai, Ruihui Li, XIAOJUAN QI, Chi-Wing Fu</p>	1654
<p>Empowering Networks With Scale and Rotation Equivariance Using A Similarity Convolution</p> <p>Zikai Sun, Thierry Blu</p>	1666
<p>PEER: A Collaborative Language Model</p> <p>Timo Schick, Jane A. Yu, Zhengbao Jiang, Fabio Petroni, Patrick Lewis, Gautier Izacard, Qingfei You, Christoforos Nalmpantis, Edouard Grave, Sebastian Riedel</p>	1682
<p>Sparse Mixture-of-Experts are Domain Generalizable Learners</p> <p>Bo Li, Yifei Shen, Jingkang Yang, Yezhen Wang, Jiawei Ren, Tong Che, Jun Zhang, Ziwei Liu</p>	1706
<p>SQA3D: Situated Question Answering in 3D Scenes</p> <p>Xiaojuan Ma, Silong Yong, Zilong Zheng, Qing Li, Yitao Liang, Song-Chun Zhu, Siyuan Huang</p>	1750
<p>Learning to Compose Soft Prompts for Compositional Zero-Shot Learning</p> <p>Nihal V. Nayak, Peilin Yu, Stephen Bach</p>	1773
<p>Mind the Gap: Offline Policy Optimization for Imperfect Rewards</p> <p>Jianxiong Li, Xiao Hu, Haoran Xu, Jingjing Liu, Xianyuan Zhan, Qing-Shan Jia, Ya-Qin Zhang</p>	1794
<p>Generating Diverse Cooperative Agents by Learning Incompatible Policies</p> <p>Rujikorn Charakorn, Poramate Manoonpong, Nat Dilokthanakul</p>	1822
<p>One Transformer Can Understand Both 2D & 3D Molecular Data</p> <p>Shengjie Luo, Tianlang Chen, Yixian Xu, Shuxin Zheng, Tie-Yan Liu, Liwei Wang, Di He</p>	1848
<p>Rethinking the Expressive Power of GNNs via Graph Biconnectivity</p> <p>Bohang Zhang, Shengjie Luo, Liwei Wang, Di He</p>	1868
<p>Revocable Deep Reinforcement Learning with Affinity Regularization for Outlier-Robust Graph Matching</p> <p>Chang Liu, Zetian Jiang, Runzhong Wang, Lingxiao Huang, Pinyan Lu, Junchi Yan</p>	1927
<p>3D generation on ImageNet</p> <p>Ivan Skorokhodov, Aliaksandr Siarohin, Yinghao Xu, Jian Ren, Hsin-Ying Lee, Peter Wonka, Sergey Tulyakov</p>	1952
<p>Distilling Cognitive Backdoor Patterns within an Image</p> <p>Hanxun Huang, Xingjun Ma, Sarah Monazam Erfani, James Bailey</p>	1981
<p>A Graph Neural Network Approach to Automated Model Building in Cryo-EM Maps</p> <p>Kiarash Jamali, Dari Kimanius, Sjors HW Scheres</p>	2012

<p>Deja Vu: Continual Model Generalization for Unseen Domains</p> <p>Chenxi Liu, Lixu Wang, Lingjuan Lyu, Chen Sun, Xiao Wang, Qi Zhu</p>	2039
<p>ManiSkill2: A Unified Benchmark for Generalizable Manipulation Skills</p> <p>Jiayuan Gu, Fanbo Xiang, Xuanlin Li, Zhan Ling, Xiqiang Liu, Tongzhou Mu, Yihe Tang, Stone Tao, Xinyue Wei, Yunchao Yao, Xiaodi Yuan, Pengwei Xie, Zhiao Huang, Rui Chen, Hao Su</p>	2058
<p>Conditional Positional Encodings for Vision Transformers</p> <p>Xiangxiang Chu, Zhi Tian, Bo Zhang, Xinlong Wang, Chunhua Shen</p>	2088
<p>Voxurf: Voxel-based Efficient and Accurate Neural Surface Reconstruction</p> <p>Tong Wu, Jiaqi Wang, Xingang Pan, Xudong XU, Christian Theobalt, Ziwei Liu, Dahua Lin</p>	2107
<p>DaxBench: Benchmarking Deformable Object Manipulation with Differentiable Physics</p> <p>Siwei Chen, Yiqing Xu, Cunjun Yu, Linfeng Li, Xiao Ma, Zhongwen Xu, David Hsu</p>	2120
<p>Nearly Minimax Optimal Offline Reinforcement Learning with Linear Function Approximation: Single-Agent MDP and Markov Game</p> <p>Wei Xiong, Han Zhong, Chengshuai Shi, Cong Shen, Liwei Wang, Tong Zhang</p>	2136
<p>EVA3D: Compositional 3D Human Generation from 2D Image Collections</p> <p>Fangzhou Hong, Zhaoxi Chen, Yushi LAN, Liang Pan, Ziwei Liu</p>	2168
<p>Linear Convergence of Natural Policy Gradient Methods with Log-Linear Policies</p> <p>Rui Yuan, Simon Shaolei Du, Robert M. Gower, Alessandro Lazaric, Lin Xiao</p>	2183
<p>No Reason for No Supervision: Improved Generalization in Supervised Models</p> <p>Mert Bakkas, Saravanan Sridharan, Yannis Kalantidis, Karteek Alahari, Diane Larlus</p>	2234
<p>What shapes the loss landscape of self supervised learning?</p> <p>Liu Ziyin, Ekdeep Singh Lubana, Masahito Ueda, Hidenori Tanaka</p>	2261
<p>Interpretability with full complexity by constraining feature information</p> <p>Kieran A Murphy, Danielle Bassett</p>	2279
<p>MocoSFL: enabling cross-client collaborative self-supervised learning</p> <p>Jingtao Li, Lingjuan Lyu, Daisuke Iso, Chaitali Chakrabarti, Michael Spranger</p>	2299
<p>MECTA: Memory-Economic Continual Test-Time Model Adaptation</p> <p>Junyuan Hong, Lingjuan Lyu, Jiayu Zhou, Michael Spranger</p>	2312
<p>Multimodal Analogical Reasoning over Knowledge Graphs</p> <p>Ningyu Zhang, Lei Li, Xiang Chen, Xiaozhuan Liang, Shumin Deng, Huajun Chen</p>	2329
<p>Spikformer: When Spiking Neural Network Meets Transformer</p> <p>Zhaokun Zhou, Yuesheng Zhu, Chao He, Yaowei Wang, Shuicheng YAN, Yonghong Tian, Li Yuan</p>	2349
<p>Explicit Box Detection Unifies End-to-End Multi-Person Pose Estimation</p> <p>Jie Yang, Ailing Zeng, Shilong Liu, Feng Li, Ruimao Zhang, Lei Zhang</p>	2366

<p>Graph Contrastive Learning for Skeleton-based Action Recognition</p> <p><i>Xiaohu Huang, Hao Zhou, Jian Wang, Haocheng Feng, Junyu Han, Errui Ding, Jingdong Wang, Xinggang Wang, Wenyu Liu, Bin Feng</i></p>	2383
<p>Knowledge Distillation based Degradation Estimation for Blind Super-Resolution</p> <p><i>Bin Xia, Yulun Zhang, Yitong Wang, Yapeng Tian, Wenming Yang, Radu Timofte, Luc Van Gool</i></p>	2398
<p>Towards Inferential Reproducibility of Machine Learning Research</p> <p><i>Michael Hagmann, Philipp Meier, Stefan Riezler</i></p>	2410
<p>DiffMimic: Efficient Motion Mimicking with Differentiable Physics</p> <p><i>Jiawei Ren, Cunjun Yu, Siwei Chen, Xiao Ma, Liang Pan, Ziwei Liu</i></p>	2425
<p>Continuous PDE Dynamics Forecasting with Implicit Neural Representations</p> <p><i>Yuan Yin, Matthieu Kirchmeyer, Jean-Yves Franceschi, Alain Rakotomamonjy, patrick gallinari</i></p>	2444
<p>Towards Smooth Video Composition</p> <p><i>Qihang Zhang, Ceyuan Yang, Yujun Shen, Yinghao Xu, Bolei Zhou</i></p>	2468
<p>Towards Stable Test-time Adaptation in Dynamic Wild World</p> <p><i>Shuaicheng Niu, Jiaxiang Wu, Yifan Zhang, Zhiquan Wen, Yaofu Chen, Peilin Zhao, Mingkui Tan</i></p>	2487
<p>DamoFD: Digging into Backbone Design on Face Detection</p> <p><i>Yang Liu, Jiankang Deng, Fei Wang, Lei Shang, Xuansong Xie, Baigui Sun</i></p>	2514
<p>LPT: Long-tailed Prompt Tuning for Image Classification</p> <p><i>Bowen Dong, Pan Zhou, Shuicheng Yan, Wangmeng Zuo</i></p>	2527
<p>Visual Recognition with Deep Nearest Centroids</p> <p><i>Wenguan Wang, Cheng Han, Tianfei Zhou, Dongfang Liu</i></p>	2547
<p>Easy Differentially Private Linear Regression</p> <p><i>Kareem Amin, Matthew Joseph, M^ánica Ribero, Sergei Vassilvitskii</i></p>	2577
<p>Scenario-based Question Answering with Interacting Contextual Properties</p> <p><i>Haitian Sun, William W. Cohen, Ruslan Salakhutdinov</i></p>	2598
<p>Statistical Inference for Fisher Market Equilibrium</p> <p><i>Luofeng Liao, Yuan Gao, Christian Kroer</i></p>	2612
<p>Contextual Convolutional Networks</p> <p><i>Shuxian Liang, Xu Shen, Tongliang Liu, Xian-Sheng Hua</i></p>	2626
<p>Self-Distillation for Further Pre-training of Transformers</p> <p><i>Seanie Lee, Minki Kang, Juho Lee, Sung Ju Hwang, Kenji Kawaguchi</i></p>	2638
<p>Interactive Portrait Harmonization</p> <p><i>Jeya Maria Jose Valanarasu, HE Zhang, Jianming Zhang, Yilin Wang, Zhe Lin, Jose Echevarria, Yinglan Ma, Zijun Wei, Kalyan Sunkavalli, Vishal Patel</i></p>	2660

<p>ESD: Expected Squared Difference as a Tuning-Free Trainable Calibration Measure</p> <p>Hee Suk Yoon, Joshua Tian Jin Tee, Eunseop Yoon, Sunjae Yoon, Gwangsu Kim, Yingzhen Li, Chang D. Yoo</p>	2675
<p>Win: Weight-Decay-Integrated Nesterov Acceleration for Adaptive Gradient Algorithms</p> <p>Pan Zhou, Xingyu Xie, Shuicheng YAN</p>	2694
<p>Offline RL with No OOD Actions: In-Sample Learning via Implicit Value Regularization</p> <p>Haoran Xu, Li Jiang, Jianxiong Li, Zhuoran Yang, Zhaoran Wang, Victor Wai Kin Chan, Xianyuan Zhan</p>	2722
<p>Bridging the Gap between ANNs and SNNs by Calibrating Offset Spikes</p> <p>Zecheng Hao, Jianhao Ding, Tong Bu, Tiejun Huang, Zhaofei Yu</p>	2742
<p>A view of mini-batch SGD via generating functions: conditions of convergence, phase transitions, benefit from negative momenta.</p> <p>Maksim Velikanov, Denis Kuznedelev, Dmitry Yarotsky</p>	2761
<p>Can discrete information extraction prompts generalize across language models?</p> <p>Nathanael Carraz Rakotonirina, Roberto Dessi, Fabio Petroni, Sebastian Riedel, Marco Baroni</p>	2812
<p>Filter-Recovery Network for Multi-Speaker Audio-Visual Speech Separation</p> <p>Haoyue Cheng, Zhaoyang Liu, Wayne Wu, Limin Wang</p>	2829
<p>Human Motion Diffusion Model</p> <p>Guy Tevet, Sigal Raab, Brian Gordon, Yoni Shafir, Daniel Cohen-or, Amit Haim Bermano</p>	2843
<p>TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Second</p> <p>Noah Hollmann, Samuel Müller, Katharina Eggenberger, Frank Hutter</p>	2859
<p>OTOv2: Automatic, Generic, User-Friendly</p> <p>Tianyi Chen, Luming Liang, Tianyu DING, Zhihui Zhu, Ilya Zharkov</p>	2895
<p>Disparate Impact in Differential Privacy from Gradient Misalignment</p> <p>Maria S. Esipova, Atiyeh Ashari Ghomi, Yaqiao Luo, Jesse C Cresswell</p>	2921
<p>MPCFORMER: FAST, PERFORMANT AND PRIVATE TRANSFORMER INFERENCE WITH MPC</p> <p>Dacheng Li, Hongyi Wang, Rulin Shao, Han Guo, Eric Xing, Hao Zhang</p>	2944
<p>CoRTX: Contrastive Framework for Real-time Explanation</p> <p>Yu-Neng Chuang, Guanchu Wang, Fan Yang, Quan Zhou, Pushkar Tripathi, Xuanting Cai, Xia Hu</p>	2960
<p>Effective Self-supervised Pre-training on Low-compute Networks without Distillation</p> <p>Fuwen Tan, Fatemeh Sadat Saleh, Brais Martinez</p>	2983
<p>NTK-SAP: Improving neural network pruning by aligning training dynamics</p> <p>Yite Wang, Dawei Li, Ruoyu Sun</p>	2998
<p>Truncated Diffusion Probabilistic Models and Diffusion-based Adversarial Auto-Encoders</p> <p>Huangjie Zheng, Pengcheng He, Weizhu Chen, Mingyuan Zhou</p>	3030

<p>Automating Nearest Neighbor Search Configuration with Constrained Optimization</p> <p><i>Philip Sun, Ruiqi Guo, Sanjiv Kumar</i></p>	3058
<p>Equiformer: Equivariant Graph Attention Transformer for 3D Atomistic Graphs</p> <p><i>Yi-Lun Liao, Tess Smidt</i></p>	3077
<p>Unified Detoxifying and Debiasing in Language Generation via Inference-time Adaptive Optimization</p> <p><i>Zonghan Yang, Xiaoyuan Yi, Peng Li, Yang Liu, Xing Xie</i></p>	3109
<p>GRACE-C: Generalized Rate Agnostic Causal Estimation via Constraints</p> <p><i>Mohammadsajad Abavisani, David Danks, Sergey Plis</i></p>	3151
<p>Multi-lingual Evaluation of Code Generation Models</p> <p><i>Ben Athiwaratkun, Sanjay Krishna Gouda, Zijian Wang, Xiaopeng Li, Yuchen Tian, Ming Tan, Wasi Uddin Ahmad, Shiqi Wang, Qing Sun, Mingyue Shang, Sujan Kumar Gonugondla, Hantian Ding, Varun Kumar, Nathan Fulton, Arash Farahani, Siddhartha Jain, Robert Giaquinto, Haifeng Qian, Murali Krishna Ramanathan, Ramesh Nallapati, Baishakhi Ray, Parminder Bhatia, Sudipta Sengupta, Dan Roth, Bing Xiang</i></p>	3169
<p>Socratic Models: Composing Zero-Shot Multimodal Reasoning with Language</p> <p><i>Andy Zeng, Maria Attarian, brian ichter, Krzysztof Marcin Choromanski, Adrian Wong, Stefan Welker, Federico Tombari, Aveek Purohit, Michael S Ryoo, Vikas Sindhwani, Johnny Lee, Vincent Vanhoucke, Pete Florence</i></p>	3248
<p>Multivariate Time-series Imputation with Disentangled Temporal Representations</p> <p><i>SHUAI LIU, Xiucheng Li, Gao Cong, Yile Chen, YUE JIANG</i></p>	3283
<p>DDM²: Self-Supervised Diffusion MRI Denoising with Generative Diffusion Models</p> <p><i>Tiange Xiang, Mahmut Yurt, Ali B Syed, Kawin Setsompop, Akshay Chaudhari</i></p>	3302
<p>Semi-supervised Community Detection via Structural Similarity Metrics</p> <p><i>Yicong Jiang, Tracy Ke</i></p>	3326
<p>Rethinking the Effect of Data Augmentation in Adversarial Contrastive Learning</p> <p><i>Rundong Luo, Yifei Wang, Yisen Wang</i></p>	3382
<p>Block and Subword-Scaling Floating-Point (BSFP) : An Efficient Non-Uniform Quantization For Low Precision Inference</p> <p><i>Yun-Chen Lo, Tse-Kuang Lee, Ren-Shuo Liu</i></p>	3399
<p>Towards One-shot Neural Combinatorial Solvers: Theoretical and Empirical Notes on the Cardinality-Constrained Case</p> <p><i>Runzhong Wang, Li Shen, Yiting Chen, Xiaokang Yang, Dacheng Tao, Junchi Yan</i></p>	3417
<p>CROM: Continuous Reduced-Order Modeling of PDEs Using Implicit Neural Representations</p> <p><i>Peter Yichen Chen, Jinxu Xiang, Dong Heon Cho, Yue Chang, G A Pershing, Henrique Teles Maia, Maurizio M Chieramonte, Kevin Thomas Carlberg, Eitan Grinspun</i></p>	3448

<p>Breaking Correlation Shift via Conditional Invariant Regularizer</p> <p>Mingyang Yi, Ruoyu Wang, Jiacheng Sun, Zhenguo Li, Zhi-Ming Ma</p>	3489
<p>Towards Addressing Label Skews in One-Shot Federated Learning</p> <p>Yiqun Diao, Qinbin Li, Bingsheng He</p>	3524
<p>Causal Balancing for Domain Generalization</p> <p>Xinyi Wang, Michael Saxon, Jiachen Li, Hongyang Zhang, Kun Zhang, William Yang Wang</p>	3546
<p>Exploring Low-Rank Property in Multiple Instance Learning for Whole Slide Image Classification</p> <p>Jinxi Xiang, Jun Zhang</p>	3570
<p>DBQ-SSD: Dynamic Ball Query for Efficient 3D Object Detection</p> <p>Jinrong Yang, Lin Song, Songtao Liu, Weixin Mao, Zeming Li, Xiaoping Li, Hongbin Sun, Jian Sun, Nanning Zheng</p>	3588
<p>Differentiable Gaussianization Layers for Inverse Problems Regularized by Deep Generative Models</p> <p>Dongzhuo Li</p>	3605
<p>CFlowNets: Continuous Control with Generative Flow Networks</p> <p>Yinchuan Li, Shuang Luo, Haozhi Wang, Jianye HAO</p>	3638
<p>Cycle-consistent Masked AutoEncoder for Unsupervised Domain Generalization</p> <p>Haiyang Yang, Xiaotong Li, SHIXIANG TANG, Feng Zhu, Yizhou Wang, Meilin Chen, LEI BAI, Rui Zhao, Wanli Ouyang</p>	3662
<p>S-NeRF: Neural Radiance Fields for Street Views</p> <p>Ziyang Xie, Junge Zhang, Wenye Li, Feihu Zhang, Li Zhang</p>	3679
<p>The Trade-off between Universality and Label Efficiency of Representations from Contrastive Learning</p> <p>Zhenmei Shi, Jiefeng Chen, Kunyang Li, Jayaram Raghuram, Xi Wu, Yingyu Liang, Somesh Jha</p>	3703
<p>Progressively Compressed Auto-Encoder for Self-supervised Representation Learning</p> <p>Jin Li, Yaoming Wang, XIAOPENG ZHANG, Yabo Chen, Dongsheng Jiang, Wenrui Dai, Chenglin Li, Hongkai Xiong, Qi Tian</p>	3745
<p>MLPInit: Embarrassingly Simple GNN Training Acceleration with MLP Initialization</p> <p>Xiaotian Han, Tong Zhao, Yozen Liu, Xia Hu, Neil Shah</p>	3761
<p>The Power of Regularization in Solving Extensive-Form Games</p> <p>Mingyang Liu, Asuman E. Ozdaglar, Tiancheng Yu, Kaiqing Zhang</p>	3792
<p>Bort: Towards Explainable Neural Networks with Bounded Orthogonal Constraint</p> <p>Borui Zhang, Wenzhao Zheng, Jie Zhou, Jiwen Lu</p>	3834

<p>Generative Augmented Flow Networks</p> <p>Ling Pan, Dinghuai Zhang, Aaron Courville, Longbo Huang, Yoshua Bengio</p>	3853
<p>TempCLR: Temporal Alignment Representation with Contrastive Learning</p> <p>Yuncong Yang, Jiawei Ma, Shiyuan Huang, Long Chen, Xudong Lin, Guangxing Han, Shih-Fu Chang</p>	3871
<p>On the Word Boundaries of Emergent Languages Based on Harris's Articulation Scheme</p> <p>Ryo Ueda, Taiga Ishii, Yusuke Miyao</p>	3897
<p>Phase transition for detecting a small community in a large network</p> <p>Jiashun Jin, Tracy Ke, Paxton Turner, Anru Zhang</p>	3918
<p>Learning Object-Language Alignments for Open-Vocabulary Object Detection</p> <p>Chuang Lin, Peize Sun, Yi Jiang, Ping Luo, Lizhen Qu, Gholamreza Haffari, Zehuan Yuan, Jianfei Cai</p>	3929
<p>Neuro-Symbolic Procedural Planning with Commonsense Prompting</p> <p>Yujie Lu, Weixi Feng, Wanrong Zhu, Wenda Xu, Xin Eric Wang, Miguel Eckstein, William Yang Wang</p>	3942
<p>Dynamic Prompt Learning via Policy Gradient for Semi-structured Mathematical Reasoning</p> <p>Pan Lu, Liang Qiu, Kai-Wei Chang, Ying Nian Wu, Song-Chun Zhu, Tanmay Rajpurohit, Peter Clark, Ashwin Kalyan</p>	3976
<p>Masked Frequency Modeling for Self-Supervised Visual Pre-Training</p> <p>Jiahao Xie, Wei Li, Xiaohang Zhan, Ziwei Liu, Yew-Soon Ong, Chen Change Loy</p>	4002
<p>Exploring Active 3D Object Detection from a Generalization Perspective</p> <p>Yadan Luo, Zhuoxiao Chen, Zijian Wang, Xin Yu, Zi Huang, Mahsa Baktashmotlagh</p>	4024
<p>Meta Knowledge Condensation for Federated Learning</p> <p>Ping Liu, Xin Yu, Joey Tianyi Zhou</p>	4037
<p>Denoising Diffusion Error Correction Codes</p> <p>Yoni Choukroun, Lior Wolf</p>	4053
<p>Learning to Linearize Deep Neural Networks for Secure and Efficient Private Inference</p> <p>Souvik Kundu, Shunlin Lu, Yuke Zhang, Jacqueline Tiffany Liu, Peter Anthony Bearel</p>	4071
<p>Meta-prediction Model for Distillation-Aware NAS on Unseen Datasets</p> <p>Hayeon Lee, Sohyun An, Minseon Kim, Sung Ju Hwang</p>	4086
<p>LAVA: Data Valuation without Pre-Specified Learning Algorithms</p> <p>Hoang Anh Just, Feiyang Kang, Tianhao Wang, Yi Zeng, Myeongseob Ko, Ming Jin, Ruoxi Jia</p>	4102
<p>Test-Time Robust Personalization for Federated Learning</p> <p>Liangze Jiang, Tao Lin</p>	4125
<p>Provably Efficient Risk-Sensitive Reinforcement Learning: Iterated CVaR and Worst Path</p> <p>Yihan Du, Siwei Wang, Longbo Huang</p>	4161

<p>Collaborative Pure Exploration in Kernel Bandit</p> <p><i>Yihan Du, Wei Chen, Yuko Kuroki, Longbo Huang</i></p>	4213
<p>Sound Randomized Smoothing in Floating-Point Arithmetic</p> <p><i>Vaclav Voracek, Matthias Hein</i></p>	4243
<p>LDMIC: Learning-based Distributed Multi-view Image Coding</p> <p><i>Xinjie Zhang, Jiawei Shao, Jun Zhang</i></p>	4264
<p>DASHA: Distributed Nonconvex Optimization with Communication Compression and Optimal Oracle Complexity</p> <p><i>Alexander Tyurin, Peter Richtárik</i></p>	4280
<p>An Additive Instance-Wise Approach to Multi-class Model Interpretation</p> <p><i>Vy Vo, Van Nguyen, Trung Le, Quan Hung Tran, Reza Haf, Seyit Camtepe, Dinh Phung</i></p>	4340
<p>PLOT: Prompt Learning with Optimal Transport for Vision-Language Models</p> <p><i>Guangyi Chen, Weiran Yao, Xiangchen Song, Xinyue Li, Yongming Rao, Kun Zhang</i></p>	4372
<p>SCoMoE: Efficient Mixtures of Experts with Structured Communication</p> <p><i>zhiyuan zeng, Deyi Xiong</i></p>	4393
<p>Approximate Bayesian Inference with Stein Functional Variational Gradient Descent</p> <p><i>Tobias Pielok, Bernd Bischl, David Röger</i></p>	4416
<p>Estimating individual treatment effects under unobserved confounding using binary instruments</p> <p><i>Dennis Frauen, Stefan Feuerriegel</i></p>	4432
<p>Denoising Masked Autoencoders Help Robust Classification</p> <p><i>QuanLin Wu, Hang Ye, Yuntian Gu, Huishuai Zhang, Liwei Wang, Di He</i></p>	4465
<p>Deep Variational Implicit Processes</p> <p><i>Luis A. Ortega, Simon Rodriguez Santana, Daniel Hernández-Lobato</i></p>	4479
<p>Thalamus: a brain-inspired algorithm for biologically-plausible continual learning and disentangled representations</p> <p><i>Ali Hummos</i></p>	4491
<p>Divide to Adapt: Mitigating Confirmation Bias for Domain Adaptation of Black-Box Predictors</p> <p><i>Jianfei Yang, Xiangyu Peng, Kai Wang, Zheng Zhu, Jiashi Feng, Lihua Xie, Yang You</i></p>	4512
<p>Self-Supervised Category-Level Articulated Object Pose Estimation with Part-Level SE(3) Equivariance</p> <p><i>Xueyi Liu, Ji Zhang, Ruizhen Hu, Haibin Huang, He Wang, Li Yi</i></p>	4533
<p>Towards Understanding Why Mask Reconstruction Pretraining Helps in Downstream Tasks</p> <p><i>Jiachun Pan, Pan Zhou, Shuicheng YAN</i></p>	4556

<p>Vision Transformer Adapter for Dense Predictions</p> <p>Zhe Chen, Yuchen Duan, Wenhai Wang, Junjun He, Tong Lu, Jifeng Dai, Yu Qiao</p>	4604
<p>Simplicial Embeddings in Self-Supervised Learning and Downstream Classification</p> <p>Samuel Lavoie, Christos Tsirigotis, Max Schwarzer, Ankit Vani, Michael Noukhovitch, Kenji Kawaguchi, Aaron Courville</p>	4624
<p>Deep Ensembles for Graphs with Higher-order Dependencies</p> <p>Steven Krieg, William Burgis, Patrick Soga, Nitesh Chawla</p>	4654
<p>Least-to-Most Prompting Enables Complex Reasoning in Large Language Models</p> <p>Denny Zhou, Nathanael Schärli, Le Hou, Jason Wei, Nathan Scales, Xuezhi Wang, Dale Schuurmans, Claire Cui, Olivier Bousquet, Quoc V Le, Ed H. Chi</p>	4667
<p>ERL-Re²: Efficient Evolutionary Reinforcement Learning with Shared State Representation and Individual Policy Representation</p> <p>Jianye HAO, Pengyi Li, Hongyao Tang, YAN ZHENG, Xian Fu, Zhaopeng Meng</p>	4728
<p>Meta-learning Adaptive Deep Kernel Gaussian Processes for Molecular Property Prediction</p> <p>Wenlin Chen, Austin Tripp, José Miguel Hernández-Lobato</p>	4755
<p>Weakly-supervised HOI Detection via Prior-guided Bi-level Representation Learning</p> <p>Bo Wan, Yongfei Liu, Desen Zhou, Tinne Tuytelaars, Xuming He</p>	4780
<p>Single-shot General Hyper-parameter Optimization for Federated Learning</p> <p>Yi Zhou, Parikshit Ram, Theodoros Salonidis, Nathalie Baracaldo, Horst Samulowitz, Heiko Ludwig</p>	4797
<p>Robustness to corruption in pre-trained Bayesian neural networks</p> <p>Xi Wang, Laurence Aitchison</p>	4828
<p>Multi-skill Mobile Manipulation for Object Rearrangement</p> <p>Jiayuan Gu, Devendra Singh Chaplot, Hao Su, Jitendra Malik</p>	4846
<p>Deep Ranking Ensembles for Hyperparameter Optimization</p> <p>Abdus Salam Khazi, Sebastian Pineda Arango, Josif Grabocka</p>	4871
<p>Fair Attribute Completion on Graph with Missing Attributes</p> <p>Dongliang Guo, Zhixuan Chu, Sheng Li</p>	4889
<p>Reliability of CKA as a Similarity Measure in Deep Learning</p> <p>MohammadReza Davari, Stefan Horoi, Amine Natick, Guillaume Lajoie, Guy Wolf, Eugene Belilovsky</p>	4904
<p>Riemannian Metric Learning via Optimal Transport</p> <p>Christopher Scovel, Justin Solomon</p>	4933
<p>Learning Label Encodings for Deep Regression</p> <p>Deval Shah, Tor M. Aamodt</p>	4952

<p>Automated Data Augmentations for Graph Classification</p> <p>Youzhi Luo, Michael Curtis McThrow, Wing Yee Au, Tao Komikado, Kanji Uchino, Koji Maruhashi, Shuiwang Ji</p>	4982
<p>Composing Ensembles of Pre-trained Models via Iterative Consensus</p> <p>Shuang Li, Yilun Du, Joshua B. Tenenbaum, Antonio Torralba, Igor Mordatch</p>	5009
<p>Gradient-Guided Importance Sampling for Learning Binary Energy-Based Models</p> <p>Meng Liu, Haoran Liu, Shuiwang Ji</p>	5030
<p>Linear Connectivity Reveals Generalization Strategies</p> <p>Jeevesh Juneja, Rachit Bansal, Kyunghyun Cho, JoÃ£o Sedoc, Naomi Saphra</p>	5048
<p>Betty: An Automatic Differentiation Library for Multilevel Optimization</p> <p>Sang Keun Choe, Willie Neiswanger, Pengtao Xie, Eric Xing</p>	5073
<p>Understanding Why Generalized Reweighting Does Not Improve Over ERM</p> <p>Runtian Zhai, Chen Dan, J Zico Kolter, Pradeep Kumar Ravikumar</p>	5099
<p>Discovering Informative and Robust Positives for Video Domain Adaptation</p> <p>Chang Liu, Kunpeng Li, Michael Stopa, Jun Amano, Yun Fu</p>	5139
<p>Generate rather than Retrieve: Large Language Models are Strong Context Generators</p> <p>Wenhao Yu, Dan Iter, Shuohang Wang, Yichong Xu, Mingxuan Ju, Soumya Sanyal, Chenguang Zhu, Michael Zeng, Meng Jiang</p>	5155
<p>Nonlinear Reconstruction for Operator Learning of PDEs with Discontinuities</p> <p>Samuel Lanthaler, Roberto Molinaro, Patrik Hadorn, Siddhartha Mishra</p>	5182
<p>General Neural Gauge Fields</p> <p>Fangneng Zhan, Lingjie Liu, Adam Kortylewski, Christian Theobalt</p>	5230
<p>BrainBERT: Self-supervised representation learning for intracranial recordings</p> <p>Christopher Wang, Vighnesh Subramaniam, Adam Uri Yaari, Gabriel Kreiman, Boris Katz, Ignacio Cases, Andrei Barbu</p>	5246
<p>Crossformer: Transformer Utilizing Cross-Dimension Dependency for Multivariate Time Series Forecasting</p> <p>Yunhao Zhang, Junchi Yan</p>	5268
<p>The KFIoU Loss for Rotated Object Detection</p> <p>Xue Yang, Yue Zhou, Gefan Zhang, Jirui Yang, Wentao Wang, Junchi Yan, XIAOPENG ZHANG, Qi Tian</p>	5289
<p>Plateau in Monotonic Linear Interpolation --- A "Biased" View of Loss Landscape for Deep Networks</p> <p>Xiang Wang, Annie N. Wang, Mo Zhou, Rong Ge</p>	5306

<p>StrucTexTv2: Masked Visual-Textual Prediction for Document Image Pre-training</p> <p>Yuechen Yu, Yulin Li, Chengquan Zhang, Xiaoqiang Zhang, Zengyuan Guo, Xiameng Qin, Kun Yao, Junyu Han, Errui Ding, Jingdong Wang</p>	5343
<p>Dataset Pruning: Reducing Training Data by Examining Generalization Influence</p> <p>Shuo Yang, Zeke Xie, Hanyu Peng, Min Xu, Mingming Sun, Ping Li</p>	5355
<p>ResAct: Reinforcing Long-term Engagement in Sequential Recommendation with Residual Actor</p> <p>Wanqi Xue, Qingpeng Cai, Ruohan Zhan, Dong Zheng, Peng Jiang, Kun Gai, Bo An</p>	5368
<p>Predicting Cellular Responses with Variational Causal Inference and Refined Relational Information</p> <p>Yulun Wu, Rob Barton, Zichen Wang, Vassilis N. Ioannidis, Carlo De Donno, Layne C Price, Luis F. Voloch, George Karypis</p>	5385
<p>Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model</p> <p>Yinhuai Wang, Jiwen Yu, Jian Zhang</p>	5402
<p>EVC: Towards Real-Time Neural Image Compression with Mask Decay</p> <p>Wang Guo-Hua, Jiahao Li, Bin Li, Yan Lu</p>	5433
<p>Diffusion Models Already Have A Semantic Latent Space</p> <p>Mingi Kwon, Jaeseok Jeong, Youngjung Uh</p>	5456
<p>The Role of ImageNet Classes in Fr�chet Inception Distance</p> <p>Tuomas Kynk�nniemi, Tero Karras, Miika Aittala, Timo Aila, Jaakko Lehtinen</p>	5491
<p>Learning to Generate Columns with Application to Vertex Coloring</p> <p>Yuan Sun, Andreas T Ernst, Xiaodong Li, Jake Weiner</p>	5517
<p>CodeT: Code Generation with Generated Tests</p> <p>Bei Chen, Fengji Zhang, Anh Nguyen, Daoguang Zan, Zeqi Lin, Jian-Guang Lou, Weizhu Chen</p>	5540
<p>Visual Imitation Learning with Patch Rewards</p> <p>Minghuan Liu, Tairan He, Weinan Zhang, Shuicheng YAN, Zhongwen Xu</p>	5559
<p>Causal Representation Learning for Instantaneous and Temporal Effects in Interactive Systems</p> <p>Phillip Lippe, Sara Magliacane, Sindy L�we, Yuki M Asano, Taco Cohen, Efstratios Gavves</p>	5578
<p>Neural Episodic Control with State Abstraction</p> <p>Zhuo Li, Derui Zhu, Yujing Hu, Xiaofei Xie, Lei Ma, YAN ZHENG, Yan Song, Yingfeng Chen, Jianjun Zhao</p>	5630
<p>Self-Supervised Set Representation Learning for Unsupervised Meta-Learning</p> <p>Dong Bok Lee, Seanie Lee, Kenji Kawaguchi, Yunji Kim, Jihwan Bang, Jung-Woo Ha, Sung Ju Hwang</p>	5648
<p>Priors, Hierarchy, and Information Asymmetry for Skill Transfer in Reinforcement Learning</p> <p>Sasha Salter, Kristian Hartikainen, Walter Goodwin, Ingmar Posner</p>	5672

<p>Accurate Image Restoration with Attention Retractable Transformer</p> <p><i>Jiale Zhang, Yulun Zhang, Jinjin Gu, Yongbing Zhang, Linghe Kong, Xin Yuan</i></p>	5696
<p>Dirichlet-based Uncertainty Calibration for Active Domain Adaptation</p> <p><i>Mixue Xie, Shuang Li, Rui Zhang, Chi Harold Liu</i></p>	5709
<p>Learning Symbolic Models for Graph-structured Physical Mechanism</p> <p><i>Hongzhi Shi, Jingtao Ding, Yufan Cao, quanming yao, Li Liu, Yong Li</i></p>	5731
<p>Continual Transformers: Redundancy-Free Attention for Online Inference</p> <p><i>Lukas Hedegaard, Arian Bakhtiarnia, Alexandros Iosifidis</i></p>	5750
<p>Strategic Classification with Graph Neural Networks</p> <p><i>Itay Eilat, Ben Finkelshtein, Chaim Baskin, Nir Rosenfeld</i></p>	5766
<p>Truthful Self-Play</p> <p><i>Shohei Ohsawa</i></p>	5793
<p>Maximizing Communication Efficiency for Large-scale Training via 0/1 Adam</p> <p><i>Yucheng Lu, Conglong Li, Minjia Zhang, Christopher De Sa, Yuxiong He</i></p>	5811
<p>InPL: Pseudo-labeling the Inliers First for Imbalanced Semi-supervised Learning</p> <p><i>Zhuoran Yu, Yin Li, Yong Jae Lee</i></p>	5843
<p>Spherical Sliced-Wasserstein</p> <p><i>Clément Bonet, Paul Berg, Nicolas Courty, François Septier, Lucas Drumetz, Minh Tan Pham</i></p>	5860
<p>Learning Sparse and Low-Rank Priors for Image Recovery via Iterative Reweighted Least Squares Minimization</p> <p><i>Stamatios Lefkimmiatis, Iaroslav Sergeevich Koshelev</i></p>	5898
<p>Joint Edge-Model Sparse Learning is Provably Efficient for Graph Neural Networks</p> <p><i>Shuai Zhang, Meng Wang, Pin-Yu Chen, Sijia Liu, Songtao Lu, Miao Liu</i></p>	5929
<p>SeaFormer: Squeeze-enhanced Axial Transformer for Mobile Semantic Segmentation</p> <p><i>Qiang Wan, Zilong Huang, Jiachen Lu, Gang YU, Li Zhang</i></p>	5974
<p>Neural Optimal Transport</p> <p><i>Alexander Korotin, Daniil Selikhanovych, Evgeny Burnaev</i></p>	5993
<p>Kernel Neural Optimal Transport</p> <p><i>Alexander Korotin, Daniil Selikhanovych, Evgeny Burnaev</i></p>	6027
<p>Robust Active Distillation</p> <p><i>Cenk Baykal, Khoa Trinh, Fotis Iliopoulos, Gaurav Menghani, Erik Vee</i></p>	6063
<p>Sparse Token Transformer with Attention Back Tracking</p> <p><i>Heejun Lee, Minki Kang, Youngwan Lee, Sung Ju Hwang</i></p>	6091

Deep Learning meets Nonparametric Regression: Are Weight-Decayed DNNs Locally Adaptive? Kaiqi Zhang, Yu-Xiang Wang	6110
Autoencoders as Cross-Modal Teachers: Can Pretrained 2D Image Transformers Help 3D Representation Learning? Runpei Dong, Zekun Qi, Linfeng Zhang, Junbo Zhang, Jianjian Sun, Zheng Ge, Li Yi, Kaisheng Ma	6145
Critic Sequential Monte Carlo Vasileios Lioutas, Jonathan Wilder Lavington, Justice Sefas, Matthew Niedoba, Yunpeng Liu, Berend Zwartsenberg, Setareh Dabiri, Frank Wood, Adam Scibior	6165
GPViT: A High Resolution Non-Hierarchical Vision Transformer with Group Propagation Chenhongyi Yang, Jiarui Xu, Shalini De Mello, Elliot J. Crowley, Xiaolong Wang	6188
NORM: Knowledge Distillation via N-to-One Representation Matching Xiaolong Liu, LUKING LI, Chao Li, Anbang Yao	6207
Near-Optimal Adversarial Reinforcement Learning with Switching Costs Ming Shi, Yingbin Liang, Ness Shroff	6230
Active Learning in Bayesian Neural Networks with Balanced Entropy Learning Principle Jae Oh Woo	6258
A Holistic View of Label Noise Transition Matrix in Deep Learning and Beyond LIN Yong, Renjie Pi, WEIZHONG ZHANG, Xiaobo Xia, Jiahui Gao, Xiao Zhou, Tongliang Liu, Bo Han	6298
Switch-NeRF: Learning Scene Decomposition with Mixture of Experts for Large-scale Neural Radiance Fields Zhenxing MI, Dan Xu	6322
Selective Annotation Makes Language Models Better Few-Shot Learners Hongjin SU, Jungo Kasai, Chen Henry Wu, Weijia Shi, Tianlu Wang, Jiayi Xin, Rui Zhang, Mari Ostendorf, Luke Zettlemoyer, Noah A. Smith, Tao Yu	6337
Approximate Vanishing Ideal Computations at Scale Eliás Samuel Wirth, Hiroshi Kera, Sebastian Pokutta	6361
FINDE: Neural Differential Equations for Finding and Preserving Invariant Quantities Takashi Matsubara, Takaharu Yaguchi	6380
Learning without Prejudices: Continual Unbiased Learning via Benign and Malignant Forgetting Myeongho Jeon, Hyoje Lee, Yedarm Seong, Myungjoo Kang	6405
TimesNet: Temporal 2D-Variation Modeling for General Time Series Analysis Haixu Wu, Tengge Hu, Yong Liu, Hang Zhou, Jianmin Wang, Mingsheng Long	6423
Policy Pre-training for Autonomous Driving via Self-supervised Geometric Modeling Penghao Wu, Li Chen, Hongyang Li, Xiaosong Jia, Junchi Yan, Yu Qiao	6446

Light Sampling Field and BRDF Representation for Physically-based Neural Rendering Jing Yang, Hanyuan Xiao, Wenbin Teng, Yunxuan Cai, Yajie Zhao	6459
Towards Robust Object Detection Invariant to Real-World Domain Shifts Qi Fan, Mattia Segu, Yu-Wing Tai, Fisher Yu, Chi-Keung Tang, Bernt Schiele, Dengxin Dai	6479
Combating Exacerbated Heterogeneity for Robust Models in Federated Learning Jianing Zhu, Jiangchao Yao, Tongliang Liu, quanming yao, Jianliang Xu, Bo Han	6495
AutoGT: Automated Graph Transformer Architecture Search Zizhao Zhang, Xin Wang, Chaoyu Guan, Ziwei Zhang, Haoyang Li, Wenwu Zhu	6529
Learning Locality and Isotropy in Dialogue Modeling Han Wu, Haochen Tan, Mingjie Zhan, Gangming Zhao, Shaoqing Lu, Ding Liang, Linqi Song	6543
Particle-based Variational Inference with Preconditioned Functional Gradient Flow Hanze Dong, Xi Wang, LIN Yong, Tong Zhang	6562
Everybody Needs Good Neighbours: An Unsupervised Locality-based Method for Bias Mitigation Xudong Han, Timothy Baldwin, Trevor Cohn	6588
POPGym: Benchmarking Partially Observable Reinforcement Learning Steven Morad, Ryan Kortvelesy, Matteo Bettini, Stephan Liwicki, Amanda Prorok	6612
Deterministic training of generative autoencoders using invertible layers Gianluigi Silvestri, Daan Roos, Luca Ambrogioni	6647
Deep Reinforcement Learning for Cost-Effective Medical Diagnosis Zheng Yu, Yikuan Li, Joseph Chahn Kim, Kaixuan Huang, Yuan Luo, Mengdi Wang	6669
GReTo: Remedying dynamic graph topology-task discordance via target homophily Zhengyang Zhou, Qihe Huang, Gengyu Lin, Kuo Yang, LEI BAI, Yang Wang	6694
Data-Free One-Shot Federated Learning Under Very High Statistical Heterogeneity Clare Elizabeth Heinbaugh, Emilio Luz-Ricca, Huajie Shao	6721
CLARE: Conservative Model-Based Reward Learning for Offline Inverse Reinforcement Learning Sheng Yue, Guanbo Wang, Wei Shao, Zhaofeng Zhang, Sen Lin, Ju Ren, Junshan Zhang	6738
Graph Domain Adaptation via Theory-Grounded Spectral Regularization Yuning You, Tianlong Chen, Zhangyang Wang, Yang Shen	6761
Spatial Attention Kinetic Networks with E(n)-Equivariance Yuanqing Wang, John Chodera	6786

<p>Learning to Jointly Share and Prune Weights for Grounding Based Vision and Language Models</p> <p>Shangqian Gao, Burak Uzkent, Yilin Shen, Heng Huang, Hongxia Jin</p>	6802
<p>Robust Fair Clustering: A Novel Fairness Attack and Defense Framework</p> <p>Anshuman Chhabra, Peizhao Li, Prasant Mohapatra, Hongfu Liu</p>	6818
<p>A General Framework For Proving The Equivariant Strong Lottery Ticket Hypothesis</p> <p>Damien Ferbach, Christos Tsirigotis, Gauthier Gidel, Joey Bose</p>	6839
<p>Sparse Random Networks for Communication-Efficient Federated Learning</p> <p>Berivan Isik, Francesco Pase, Deniz Gunduz, Tsachy Weissman, Zorzi Michele</p>	6867
<p>One-Pixel Shortcut: On the Learning Preference of Deep Neural Networks</p> <p>Shutong Wu, Sizhe Chen, Cihang Xie, Xiaolin Huang</p>	6889
<p>DFlow: Learning to Synthesize Better Optical Flow Datasets via a Differentiable Pipeline</p> <p>Kwon Byung-Ki, Nam Hyeon-Woo, Ji-Yun Kim, Tae-Hyun Oh</p>	6904
<p>Universal Vision-Language Dense Retrieval: Learning A Unified Representation Space for Multi-Modal Retrieval</p> <p>Zhenghao Liu, Chenyan Xiong, Yuanhuiyi Lv, Zhiyuan Liu, Ge Yu</p>	6927
<p>View Synthesis with Sculpted Neural Points</p> <p>Yiming Zuo, Jia Deng</p>	6944
<p>Domain-Indexing Variational Bayes: Interpretable Domain Index for Domain Adaptation</p> <p>Zihao Xu, Guang-Yuan Hao, Hao He, Hao Wang</p>	6970
<p>Dataless Knowledge Fusion by Merging Weights of Language Models</p> <p>Xisen Jin, Xiang Ren, Daniel Preotiuc-Pietro, Pengxiang Cheng</p>	6992
<p>Can We Faithfully Represent Absence States to Compute Shapley Values on a DNN?</p> <p>Jie Ren, Zhanpeng Zhou, Qirui Chen, Quanshi Zhang</p>	7011
<p>Continual evaluation for lifelong learning: Identifying the stability gap</p> <p>Matthias De Lange, Gido M van de Ven, Tinne Tuytelaars</p>	7037
<p>Understanding Zero-shot Adversarial Robustness for Large-Scale Models</p> <p>Chengzhi Mao, Scott Geng, Junfeng Yang, Xin Wang, Carl Vondrick</p>	7058
<p>Pessimism in the Face of Confounders: Provably Efficient Offline Reinforcement Learning in Partially Observable Markov Decision Processes</p> <p>Miao Lu, Yifei Min, Zhaoran Wang, Zhuoran Yang</p>	7077
<p>Information-Theoretic Analysis of Unsupervised Domain Adaptation</p> <p>Ziqiao Wang, Yongyi Mao</p>	7124

<p>Multi-domain image generation and translation with identifiability guarantees</p> <p>Shaoan Xie, Lingjing Kong, Mingming Gong, Kun Zhang</p>	7155
<p>Parallel Deep Neural Networks Have Zero Duality Gap</p> <p>Yifei Wang, Tolga Ergen, Mert Pilanci</p>	7190
<p>Mega: Moving Average Equipped Gated Attention</p> <p>Xuezhe Ma, Chunting Zhou, Xiang Kong, Junxian He, Liangke Gui, Graham Neubig, Jonathan May, Luke Zettlemoyer</p>	7216
<p>A General Rank Preserving Framework for Asymmetric Image Retrieval</p> <p>Hui Wu, Min Wang, Wengang Zhou, Houqiang Li</p>	7236
<p>Hidden Markov Transformer for Simultaneous Machine Translation</p> <p>Shaolei Zhang, Yang Feng</p>	7256
<p>Backstepping Temporal Difference Learning</p> <p>Han-Dong Lim, Donghwan Lee</p>	7278
<p>Human MotionFormer: Transferring Human Motions with Vision Transformers</p> <p>Hongyu Liu, Xintong Han, Chenbin Jin, Lihui Qian, Huawei Wei, Zhe Lin, Faqiang Wang, Haoye Dong, Yibing Song, Jia Xu, Qifeng Chen</p>	7289
<p>Become a Proficient Player with Limited Data through Watching Pure Videos</p> <p>Weirui Ye, Yunsheng Zhang, Pieter Abbeel, Yang Gao</p>	7304
<p>Proactive Multi-Camera Collaboration for 3D Human Pose Estimation</p> <p>Hai Ci, Mickel Liu, Xuehai Pan, fangwei zhong, Yizhou Wang</p>	7320
<p>Packed Ensembles for efficient uncertainty estimation</p> <p>Olivier Laurent, Adrien Lafage, Enzo Tartaglione, Geoffrey Daniel, Jean-marc Martinez, Andrei Bursuc, Gianni Franchi</p>	7348
<p>Evaluating Long-Term Memory in 3D Mazes</p> <p>Jurgis PaÅ¼ukonis, Timothy P Lillicrap, Danijar Hafner</p>	7374
<p>Sub-Task Decomposition Enables Learning in Sequence to Sequence Tasks</p> <p>Noam Wies, Yoav Levine, Amnon Shashua</p>	7390
<p>Multifactor Sequential Disentanglement via Structured Koopman Autoencoders</p> <p>Nimrod Berman, Ilan Naiman, Omri Azencot</p>	7418
<p>Exploring The Role of Mean Teachers in Self-supervised Masked Auto-Encoders</p> <p>Youngwan Lee, Jeffrey Ryan Willette, Jonghee Kim, Juho Lee, Sung Ju Hwang</p>	7442
<p>Analyzing Tree Architectures in Ensembles via Neural Tangent Kernel</p> <p>Ryuichi Kanoh, Mahito Sugiyama</p>	7463

<p>Multi-task Self-supervised Graph Neural Networks Enable Stronger Task Generalization</p> <p>Mingxuan Ju, Tong Zhao, Qianlong Wen, Wenhao Yu, Neil Shah, Yanfang Ye, Chuxu Zhang</p>	7481
<p>What Is Missing in IRM Training and Evaluation? Challenges and Solutions</p> <p>Yihua Zhang, Pranay Sharma, Parikshit Ram, Mingyi Hong, Kush R. Varshney, Sijia Liu</p>	7501
<p>Behind the Scenes of Gradient Descent: A Trajectory Analysis via Basis Function Decomposition</p> <p>Jianhao Ma, Lingjun Guo, Salar Fattahi</p>	7520
<p>On the Robustness of Safe Reinforcement Learning under Observational Perturbations</p> <p>Zuxin Liu, Zijian Guo, Zhepeng Cen, Huan Zhang, Jie Tan, Bo Li, Ding Zhao</p>	7562
<p>Modeling Multimodal Aleatoric Uncertainty in Segmentation with Mixture of Stochastic Experts</p> <p>Zhitong Gao, Yucong Chen, Chuyu Zhang, Xuming He</p>	7592
<p>Formal Mathematics Statement Curriculum Learning</p> <p>Stanislas Polu, Jesse Michael Han, Kunhao Zheng, Mantas Baksys, Igor Babuschkin, Ilya Sutskever</p>	7619
<p>Reversible Column Networks</p> <p>Yuxuan Cai, Yizhuang Zhou, Qi Han, Jianjian Sun, Xiangwen Kong, Jun Li, Xiangyu Zhang</p>	7650
<p>Cheap Talk Discovery and Utilization in Multi-Agent Reinforcement Learning</p> <p>Yat Long Lo, Christian Schroeder de Witt, Samuel Sokota, Jakob Nicolaus Foerster, Shimon Whiteson</p>	7674
<p>Private Federated Learning Without a Trusted Server: Optimal Algorithms for Convex Losses</p> <p>Andrew Lowy, Meisam Razaviyayn</p>	7691
<p>NeRN: Learning Neural Representations for Neural Networks</p> <p>Maor Ashkenazi, Zohar Rimon, Ron Vainshtein, Shir Levi, Elad Richardson, Pinchas Mintz, Eran Treister</p>	7745
<p>Rethinking Graph Lottery Tickets: Graph Sparsity Matters</p> <p>Bo Hui, Da Yan, Xiaolong Ma, Wei-Shinn Ku</p>	7761
<p>NeRF-SOS: Any-View Self-supervised Object Segmentation on Complex Scenes</p> <p>Zhiwen Fan, Peihao Wang, Yifan Jiang, Xinyu Gong, DeJia Xu, Zhangyang Wang</p>	7778
<p>Diffusion Probabilistic Modeling of Protein Backbones in 3D for the motif-scaffolding problem</p> <p>Brian L. Trippe, Jason Yim, Doug Tischer, David Baker, Tamara Broderick, Regina Barzilay, Tommi S. Jaakkola</p>	7798
<p>Discrete Contrastive Diffusion for Cross-Modal Music and Image Generation</p> <p>Ye Zhu, Yu Wu, Kyle Olszewski, Jian Ren, Sergey Tulyakov, Yan Yan</p>	7828
<p>AANG : Automating Auxiliary Learning</p> <p>Lucio M. Dery, Paul Michel, Mikhail Khodak, Graham Neubig, Ameet Talwalkar</p>	7853

<p><i>FIT: Parameter Efficient Few-shot Transfer Learning for Personalized and Federated Image Classification</i></p> <p><i>Aliaksandra Shysheya, John F Bronskill, Massimiliano Patacchiola, Sebastian Nowozin, Richard E Turner</i></p>	7875
<p><i>Generalized Precision Matrix for Scalable Estimation of Nonparametric Markov Networks</i></p> <p><i>Yujia Zheng, Ignavier Ng, Yewen Fan, Kun Zhang</i></p>	7903
<p><i>Neural Radiance Field Codebooks</i></p> <p><i>Matthew Wallingford, Aditya Kusupati, Alex Fang, Vivek Ramanujan, Aniruddha Kembhavi, Roozbeh Mottaghi, Ali Farhadi</i></p>	7925
<p><i>Guiding continuous operator learning through Physics-based boundary constraints</i></p> <p><i>Nadim Saad, Gaurav Gupta, Shima Alizadeh, Danielle C. Maddix</i></p>	7944
<p><i>Learning Proximal Operators to Discover Multiple Optima</i></p> <p><i>Lingxiao Li, Noam Aigerman, Vladimir Kim, Jiajin Li, Kristjan Greenewald, Mikhail Yurochkin, Justin Solomon</i></p>	7973
<p><i>Understanding Edge-of-Stability Training Dynamics with a Minimalist Example</i></p> <p><i>Xingyu Zhu, Zixuan Wang, Xiang Wang, Mo Zhou, Rong Ge</i></p>	8000
<p><i>Analog Bits: Generating Discrete Data using Diffusion Models with Self-Conditioning</i></p> <p><i>Ting Chen, Ruixiang ZHANG, Geoffrey Hinton</i></p>	8065
<p><i>Deep Declarative Dynamic Time Warping for End-to-End Learning of Alignment Paths</i></p> <p><i>Ming Xu, Sourav Garg, Michael Milford, Stephen Gould</i></p>	8088
<p><i>Trainable Weight Averaging: Efficient Training by Optimizing Historical Solutions</i></p> <p><i>Tao Li, Zhehao Huang, Qinghua Tao, Yingwen Wu, Xiaolin Huang</i></p>	8110
<p><i>Protein Representation Learning by Geometric Structure Pretraining</i></p> <p><i>Zuobai Zhang, Minghao Xu, Arian Rokkum Jamasb, Vijil Chenthamarakshan, Aurelie Lozano, Payel Das, Jian Tang</i></p>	8126
<p><i>Unsupervised Model Selection for Time Series Anomaly Detection</i></p> <p><i>Mononito Goswami, Cristian Ignacio Challu, Laurent Callot, Lenon Minorics, Andrey Kan</i></p>	8153
<p><i>On the Saturation Effect of Kernel Ridge Regression</i></p> <p><i>Yicheng Li, Haobo Zhang, Qian Lin</i></p>	8182
<p><i>Bi-level Physics-Informed Neural Networks for PDE Constrained Optimization using Broyden's Hypergradients</i></p> <p><i>Zhongkai Hao, Chengyang Ying, Hang Su, Jun Zhu, Jian Song, Ze Cheng</i></p>	8213
<p><i>A Minimalist Dataset for Systematic Generalization of Perception, Syntax, and Semantics</i></p> <p><i>Qing Li, Siyuan Huang, Yining Hong, Yixin Zhu, Ying Nian Wu, Song-Chun Zhu</i></p>	8245

<p>Risk-Aware Reinforcement Learning with Coherent Risk Measures and Non-linear Function Approximation</p> <p><i>Thanh Lam, Arun Verma, Bryan Kian Hsiang Low, Patrick Jaillet</i></p>	8263
<p>Can CNNs Be More Robust Than Transformers?</p> <p><i>Zeyu Wang, Yutong Bai, Yuyin Zhou, Cihang Xie</i></p>	8289
<p>Eva: Practical Second-order Optimization with Kronecker-vectorized Approximation</p> <p><i>Lin Zhang, Shaohuai Shi, Bo Li</i></p>	8305
<p>Multimodal Federated Learning via Contrastive Representation Ensemble</p> <p><i>Qiyang Yu, Yang Liu, Yimu Wang, Ke Xu, Jingjing Liu</i></p>	8329
<p>HiViT: A Simpler and More Efficient Design of Hierarchical Vision Transformer</p> <p><i>Xiaosong Zhang, Yunjie Tian, Lingxi Xie, Wei Huang, Qi Dai, Qixiang Ye, Qi Tian</i></p>	8344
<p>A Laplace-inspired Distribution on $SO(3)$ for Probabilistic Rotation Estimation</p> <p><i>Yingda Yin, Yang Wang, He Wang, Baoquan Chen</i></p>	8359
<p>Basic Binary Convolution Unit for Binarized Image Restoration Network</p> <p><i>Bin Xia, Yulun Zhang, Yitong Wang, Yapeng Tian, Wenming Yang, Radu Timofte, Luc Van Gool</i></p>	8378
<p>Deep Generative Modeling on Limited Data with Regularization by Nontransferable Pre-trained Models</p> <p><i>Yong Zhong, Hongtao Liu, Xiaodong Liu, Fan Bao, Weiran Shen, Chongxuan Li</i></p>	8390
<p>Towards Robustness Certification Against Universal Perturbations</p> <p><i>Yi Zeng, Zhouxing Shi, Ming Jin, Feiyang Kang, Lingjuan Lyu, Cho-Jui Hsieh, Ruoxi Jia</i></p>	8416
<p>Molecule Generation For Target Protein Binding with Structural Motifs</p> <p><i>ZAIXI ZHANG, Yaosen Min, Shuxin Zheng, Qi Liu</i></p>	8434
<p>Globally Optimal Training of Neural Networks with Threshold Activation Functions</p> <p><i>Tolga Ergen, Halil Ibrahim Gulluk, Jonathan Lacotte, Mert Pilanci</i></p>	8450
<p>Outcome-directed Reinforcement Learning by Uncertainty & Temporal Distance-Aware Curriculum Goal Generation</p> <p><i>Daesol Cho, Seungjae Lee, H. Jin Kim</i></p>	8476
<p>Provable Sim-to-real Transfer in Continuous Domain with Partial Observations</p> <p><i>Jiachen Hu, Han Zhong, Chi Jin, Liwei Wang</i></p>	8501
<p>Revisiting Graph Adversarial Attack and Defense From a Data Distribution Perspective</p> <p><i>Kuan Li, Yang Liu, Xiang Ao, Qing He</i></p>	8537
<p>Statistical Efficiency of Score Matching: The View from Isoperimetry</p> <p><i>Frederic Koehler, Alexander Heckett, Andrej Risteski</i></p>	8560

<p>DINO: DETR with Improved DeNoising Anchor Boxes for End-to-End Object Detection</p> <p>Hao Zhang, Feng Li, Shilong Liu, Lei Zhang, Hang Su, Jun Zhu, Lionel Ni, Heung-Yeung Shum</p>	8587
<p>Learning the Positions in CountSketch</p> <p>Yi Li, Honghao Lin, Simin Liu, Ali Vakilian, David Woodruff</p>	8606
<p>GeneFace: Generalized and High-Fidelity Audio-Driven 3D Talking Face Synthesis</p> <p>Zhenhui Ye, Ziyue Jiang, Yi Ren, Jinglin Liu, Jinzheng He, Zhou Zhao</p>	8645
<p>Masked Unsupervised Self-training for Label-free Image Classification</p> <p>Junnan Li, Silvio Savarese, Steven Hoi</p>	8661
<p>GoBigger: A Scalable Platform for Cooperative-Competitive Multi-Agent Interactive Simulation</p> <p>Ming Zhang, Shenghan Zhang, Zhenjie Yang, Lekai Chen, Jinliang Zheng, Chao Yang, Chuming Li, Hang Zhou, Yazhe Niu, Yu Liu</p>	8673
<p>Mass-Editing Memory in a Transformer</p> <p>Kevin Meng, Arnab Sen Sharma, Alex J Andonian, Yonatan Belinkov, David Bau</p>	8694
<p>Masked Image Modeling with Denoising Contrast</p> <p>Kun Yi, Yixiao Ge, Xiaotong Li, Shusheng Yang, Dian Li, Jianping Wu, Ying Shan, Xiaohu Qie</p>	8715
<p>DepthFL : Depthwise Federated Learning for Heterogeneous Clients</p> <p>Minjae Kim, Sangyoon Yu, Suhyun Kim, Soo-Mook Moon</p>	8733
<p>Diffusion Posterior Sampling for General Noisy Inverse Problems</p> <p>Hyungjin Chung, Jeongsol Kim, Michael Thompson Mccann, Marc Louis Klasky, Jong Chul Ye</p>	8753
<p>Modeling Sequential Sentence Relation to Improve Cross-lingual Dense Retrieval</p> <p>Shunyu Zhang, Yaobo Liang, MING GONG, Daxin Jiang, Nan Duan</p>	8783
<p>The Provable Benefit of Unsupervised Data Sharing for Offline Reinforcement Learning</p> <p>Hao Hu, Yiqin Yang, Qianchuan Zhao, Chongjie Zhang</p>	8801
<p>Protein Representation Learning via Knowledge Enhanced Primary Structure Reasoning</p> <p>Hong-Yu Zhou, Yunxiang Fu, Zhicheng Zhang, Bian Cheng, Yizhou Yu</p>	8826
<p>Re-parameterizing Your Optimizers rather than Architectures</p> <p>Xiaohan Ding, Honghao Chen, Xiangyu Zhang, Kaiqi Huang, Jungong Han, Guiguang Ding</p>	8840
<p>Instance-wise Batch Label Restoration via Gradients in Federated Learning</p> <p>Kailang Ma, Yu Sun, Jian Cui, Dawei Li, Zhenyu Guan, Jianwei Liu</p>	8858
<p>Advancing Radiograph Representation Learning with Masked Record Modeling</p> <p>Hong-Yu Zhou, Chenyu Lian, Liansheng Wang, Yizhou Yu</p>	8873
<p>FedSpeed: Larger Local Interval, Less Communication Round, and Higher Generalization Accuracy</p> <p>Yan Sun, Li Shen, Tiansheng Huang, Liang Ding, Dacheng Tao</p>	8889

Bag of Tricks for Unsupervised Text-to-Speech Yi Ren, Chen Zhang, Shuicheng YAN	8921
CO3: Cooperative Unsupervised 3D Representation Learning for Autonomous Driving Runjian Chen, Yao Mu, Runsen Xu, Wenqi Shao, Chenhan Jiang, Hang Xu, Yu Qiao, Zhenguo Li, Ping Luo	8941
On the complexity of nonsmooth automatic differentiation Jerome Bolte, Ryan Boustany, Edouard Pauwels, B�atrice Pesquet-Popescu	8963
Towards the Generalization of Contrastive Self-Supervised Learning Weiran Huang, Mingyang Yi, Xuyang Zhao, Zihao Jiang	8990
Generalize Learned Heuristics to Solve Large-scale Vehicle Routing Problems in Real-time Qingchun Hou, Jingwei Yang, Yiqiang Su, Xiaoqing Wang, Yuming Deng	9019
Towards Effective and Interpretable Human-Agent Collaboration in MOBA Games: A Communication Perspective Yiming Gao, Feiyu Liu, Liang Wang, Zhenjie Lian, Weixuan Wang, Siqin Li, Xianliang Wang, Xianhan Zeng, Rundong Wang, Jiawei Wang, QIANG FU, Yang Wei, Lanxiao Huang, Wei Liu	9056
Adaptive Budget Allocation for Parameter-Efficient Fine-Tuning Qingru Zhang, Minshuo Chen, Alexander Bukharin, Pengcheng He, Yu Cheng, Weizhu Chen, Tuo Zhao	9082
D4AM: A General Denoising Framework for Downstream Acoustic Models Chi-Chang Lee, Yu Tsao, Hsin-Min Wang, Chu-Song Chen	9099
From t-SNE to UMAP with contrastive learning Sebastian Damrich, Niklas B�hm, Fred A Hamprecht, Dmitry Kobak	9116
Learning Heterogeneous Interaction Strengths by Trajectory Prediction with Graph Neural Network Seungwoong Ha, Hawoong Jeong	9160
Monocular Scene Reconstruction with 3D SDF Transformers Weihao Yuan, Xiaodong Gu, Heng Li, Zilong Dong, Siyu Zhu	9171
DENSE RGB SLAM WITH NEURAL IMPLICIT MAPS Heng Li, Xiaodong Gu, Weihao Yuan, luwei yang, Zilong Dong, Ping Tan	9188
When to Make and Break Commitments? Alihan H�k, Zhaozhi Qian, Mihaela van der Schaar	9209
Approximate Nearest Neighbor Search through Modern Error-Correcting Codes Noam Touitou, Nissim Halabi	9236

<p>QuAnt: Quantum Annealing with Learnt Couplings</p> <p>Marcel Seelbach Benkner, Maximilian Krahn, Edith Tretschk, Zorah LÄhner, Michael Moeller, Vladislav Golyanik</p>	9256
<p>Voint Cloud: Multi-View Point Cloud Representation for 3D Understanding</p> <p>Abdullah Hamdi, Silvio Giancola, Bernard Ghanem</p>	9277
<p>GAMR: A Guided Attention Model for (visual) Reasoning</p> <p>Mohit Vaishnav, Thomas Serre</p>	9308
<p>Planckian Jitter: countering the color-crippling effects of color jitter on self-supervised training</p> <p>Simone Zini, Alex Gomez-Villa, Marco Buzzelli, BartÅmiej Twardowski, Andrew D. Bagdanov, Joost van de weijer</p>	9344
<p>Modelling Long Range Dependencies in \$N\$D: From Task-Specific to a General Purpose CNN</p> <p>David M Knigge, David W. Romero, Albert Gu, Efstratios Gavves, Erik J Bekkers, Jakub Mikolaj Tomczak, Mark Hoogendoorn, Jan-jakob Sonke</p>	9356
<p>Learning Zero-Shot Cooperation with Humans, Assuming Humans Are Biased</p> <p>Chao Yu, Jiaxuan Gao, Weilin Liu, Botian Xu, Hao Tang, Jiaqi Yang, Yu Wang, Yi Wu</p>	9376
<p>Gradient Boosting Performs Gaussian Process Inference</p> <p>Aleksei Ustimenko, Artem Beliaikov, Liudmila Prokhorenkova</p>	9407
<p>Contextual bandits with concave rewards, and an application to fair ranking</p> <p>Virginie Do, Elvis Dohmatob, Matteo Pirota, Alessandro Lazaric, Nicolas Usunier</p>	9436
<p>ImaginaryNet: Learning Object Detectors without Real Images and Annotations</p> <p>Minheng Ni, Zitong Huang, Kailai Feng, Wangmeng Zuo</p>	9478
<p>Deep Learning on Implicit Neural Representations of Shapes</p> <p>Luca De Luigi, Adriano Cardace, Riccardo Spezialetti, Pierluigi Zama Ramirez, Samuele Salti, Luigi di Stefano</p>	9495
<p>Temperature Schedules for self-supervised contrastive methods on long-tail data</p> <p>Anna Kukleva, Moritz BÄhle, Bernt Schiele, Hilde Kuehne, Christian Rupprecht</p>	9534
<p>Toeplitz Neural Network for Sequence Modeling</p> <p>Zhen Qin, Xiaodong Han, Weixuan Sun, Bowen He, Dong Li, Dongxu Li, Yuchao Dai, Lingpeng Kong, Yiran Zhong</p>	9549
<p>ArCL: Enhancing Contrastive Learning with Augmentation-Robust Representations</p> <p>Xuyang Zhao, Tianqi Du, Yisen Wang, Jun Yao, Weiran Huang</p>	9567
<p>Learning to Estimate Single-View Volumetric Flow Motions without 3D Supervision</p> <p>Erik Franz, Barbara Solenthaler, Nils Thuerey</p>	9585
<p>Improved Convergence of Differential Private SGD with Gradient Clipping</p> <p>Huang Fang, Xiaoyun Li, Chenglin Fan, Ping Li</p>	9607

<p>Scalable Subset Sampling with Neural Conditional Poisson Networks</p> <p>Adeel Pervez, Phillip Lippe, Efstratios Gavves</p>	9633
<p>Transformers are Sample-Efficient World Models</p> <p>Vincent Micheli, Eloi Alonso, François Fleuret</p>	9654
<p>Differentiable Mathematical Programming for Object-Centric Representation Learning</p> <p>Adeel Pervez, Phillip Lippe, Efstratios Gavves</p>	9675
<p>A Model or 603 Exemplars: Towards Memory-Efficient Class-Incremental Learning</p> <p>Da-Wei Zhou, Qi-Wei Wang, Han-Jia Ye, De-Chuan Zhan</p>	9689
<p>SYNC: SAFETY-AWARE NEURAL CONTROL FOR STABILIZING STOCHASTIC DELAY-DIFFERENTIAL EQUATIONS</p> <p>Jingdong Zhang, Qunxi Zhu, Wei Yang, Wei Lin</p>	9724
<p>Anisotropic Message Passing: Graph Neural Networks with Directional and Long-Range Interactions</p> <p>Moritz Thaler, Sereina Riniker</p>	9737
<p>Tailoring Language Generation Models under Total Variation Distance</p> <p>Haozhe Ji, Pei Ke, Zhipeng Hu, Rongsheng Zhang, Minlie Huang</p>	9756
<p>Evidential Uncertainty and Diversity Guided Active Learning for Scene Graph Generation</p> <p>Shuzhou Sun, Shuaifeng Zhi, Janne Heikkilä, Li Liu</p>	9781
<p>SketchKnitter: Vectorized Sketch Generation with Diffusion Models</p> <p>Qiang Wang, Haoge Deng, Yonggang Qi, Da Li, Yi-Zhe Song</p>	9803
<p>Towards Understanding and Mitigating Dimensional Collapse in Heterogeneous Federated Learning</p> <p>Yujun Shi, Jian Liang, Wenqing Zhang, Vincent Tan, Song Bai</p>	9820
<p>An Exact Poly-Time Membership-Queries Algorithm for Extracting a Three-Layer ReLU Network</p> <p>Amit Daniely, Elad Granot</p>	9844
<p>TabCaps: A Capsule Neural Network for Tabular Data Classification with BoW Routing</p> <p>Jintai Chen, KuanLun Liao, Yanwen Fang, Danny Chen, Jian Wu</p>	9866
<p>Programmatically Grounded, Compositionally Generalizable Robotic Manipulation</p> <p>Renhao Wang, Jiayuan Mao, Joy Hsu, Hang Zhao, Jiajun Wu, Yang Gao</p>	9881
<p>Emergent World Representations: Exploring a Sequence Model Trained on a Synthetic Task</p> <p>Kenneth Li, Aspen K Hopkins, David Bau, Fernanda Viégas, Hanspeter Pfister, Martin Wattenberg</p>	9901
<p>GOOD: Exploring geometric cues for detecting objects in an open world</p> <p>Haiwen Huang, Andreas Geiger, Dan Zhang</p>	9918

Clifford Neural Layers for PDE Modeling Johannes Brandstetter, Rianne van den Berg, Max Welling, Jayesh K Gupta	9941
3D Segmenter: 3D Transformer based Semantic Segmentation via 2D Panoramic Distillation ZHENNAN WU, YANG LI, Yifei Huang, Lin Gu, Tatsuya Harada, Hiroyuki Sato	9999
Efficient Federated Domain Translation Zeyu Zhou, Sheikh Shams Azam, Christopher Brinton, David I. Inouye	10014
Fuzzy Alignments in Directed Acyclic Graph for Non-Autoregressive Machine Translation Zhengrui Ma, Chenze Shao, Shangtong Gui, Min Zhang, Yang Feng	10045
Conditional Antibody Design as 3D Equivariant Graph Translation Xiangzhe Kong, Wenbing Huang, Yang Liu	10064
Iterative Patch Selection for High-Resolution Image Recognition Benjamin Bergner, Christoph Lippert, Aravindh Mahendran	10085
Unified Discrete Diffusion for Simultaneous Vision-Language Generation Minghui Hu, Chuanxia Zheng, Zuopeng Yang, Tat-Jen Cham, Heliang Zheng, Chaoyue Wang, Dacheng Tao, Ponnuthurai N. Suganthan	10105
Improving Deep Regression with Ordinal Entropy Shihao Zhang, Linlin Yang, Michael Bi Mi, Xiaoxu Zheng, Angela Yao	10128
AnyDA: Anytime Domain Adaptation Omprakash Chakraborty, Aadarsh Sahoo, Rameswar Panda, Abir Das	10144
How to Exploit Hyperspherical Embeddings for Out-of-Distribution Detection? Yifei Ming, Yiyu Sun, Ousmane Dia, Yixuan Li	10166
Long-Tailed Learning Requires Feature Learning Thomas Laurent, James von Brecht, Xavier Bresson	10185
Learning Human-Compatible Representations for Case-Based Decision Support Han Liu, Yizhou Tian, Chacha Chen, Shi Feng, Yuxin Chen, Chenhao Tan	10224
ISAAC Newton: Input-based Approximate Curvature for Newton's Method Felix Petersen, Tobias Sutter, Christian Borgelt, Dongsung Huh, Hilde Kuehne, Yuekai Sun, Oliver Deussen	10252
LMC: Fast Training of GNNs via Subgraph Sampling with Provable Convergence Zhihao Shi, Xize Liang, Jie Wang	10267
Trading Information between Latents in Hierarchical Variational Autoencoders Tim Z. Xiao, Robert Bamler	10299
DM-NeRF: 3D Scene Geometry Decomposition and Manipulation from 2D Images Bing WANG, Lu Chen, Bo Yang	10316

<p>Novel View Synthesis with Diffusion Models</p> <p><i>Daniel Watson, William Chan, Ricardo Martin Brualla, Jonathan Ho, Andrea Tagliasacchi, Mohammad Norouzi</i></p>	10328
<p>Asynchronous Gradient Play in Zero-Sum Multi-agent Games</p> <p><i>Ruicheng Ao, Shicong Cen, Yuejie Chi</i></p>	10350
<p>Fooling SHAP with Stealthily Biased Sampling</p> <p><i>gabriel laberge, Ulrich AÅvodji, Satoshi Hara, Mario Marchand, Foutse Khomh</i></p>	10398
<p>Continual Unsupervised Disentangling of Self-Organizing Representations</p> <p><i>Zhiyuan Li, Xiajun Jiang, Ryan Missel, Prashnna Kumar Gyawali, Nilesh Kumar, Linwei Wang</i></p>	10428
<p>Multi-Objective Reinforcement Learning: Convexity, Stationarity and Pareto Optimality</p> <p><i>Haoye Lu, Daniel Herman, Yaoliang Yu</i></p>	10450
<p>Causality Compensated Attention for Contextual Biased Visual Recognition</p> <p><i>Ruyang Liu, Jingjia Huang, Thomas H. Li, Ge Li</i></p>	10477
<p>NANSY++: Unified Voice Synthesis with Neural Analysis and Synthesis</p> <p><i>Hyeong-Seok Choi, Jinhyeok Yang, Juheon Lee, Hyeongju Kim</i></p>	10494
<p>DecAF: Joint Decoding of Answers and Logical Forms for Question Answering over Knowledge Bases</p> <p><i>Donghan Yu, Sheng Zhang, Patrick Ng, Henghui Zhu, Alexander Hanbo Li, Jun Wang, Yiqun Hu, William Yang Wang, Zhiguo Wang, Bing Xiang</i></p>	10518
<p>Flow Annealed Importance Sampling Bootstrap</p> <p><i>Laurence Illing Midgley, Vincent Stimper, Gregor N. C. Simm, Bernhard SchÅlkopf, JosÅ Miguel HernÅndez-Lobato</i></p>	10536
<p>ManyDG: Many-domain Generalization for Healthcare Applications</p> <p><i>Chaoqi Yang, M Brandon Westover, Jimeng Sun</i></p>	10580
<p>Omnigrok: Grokking Beyond Algorithmic Data</p> <p><i>Ziming Liu, Eric J Michaud, Max Tegmark</i></p>	10604
<p>GAIN: On the Generalization of Instructional Action Understanding</p> <p><i>Junlong Li, Guangyi Chen, Yansong Tang, Jinan Bao, Kun Zhang, Jie Zhou, Jiwen Lu</i></p>	10620
<p>Code Translation with Compiler Representations</p> <p><i>Marc Szafraniec, Baptiste Roziere, Hugh James Leather, Patrick Labatut, Francois Charton, Gabriel Synnaeve</i></p>	10642
<p>Sequential Latent Variable Models for Few-Shot High-Dimensional Time-Series Forecasting</p> <p><i>Xiajun Jiang, Ryan Missel, Zhiyuan Li, Linwei Wang</i></p>	10662

<p>Sign and Basis Invariant Networks for Spectral Graph Representation Learning</p> <p>Derek Lim, Joshua David Robinson, Lingxiao Zhao, Tess Smidt, Suvrit Sra, Haggai Maron, Stefanie Jegelka</p>	10683
<p>Model ensemble instead of prompt fusion: a sample-specific knowledge transfer method for few-shot prompt tuning</p> <p>XIANGYU PENG, Chen Xing, Prafulla Kumar Choubey, Chien-Sheng Wu, Caiming Xiong</p>	10727
<p>Hybrid RL: Using both offline and online data can make RL efficient</p> <p>Yuda Song, Yifei Zhou, Ayush Sekhari, Drew Bagnell, Akshay Krishnamurthy, Wen Sun</p>	10750
<p>Beyond calibration: estimating the grouping loss of modern neural networks</p> <p>Alexandre Perez-Lebel, Marine Le Morvan, Gael Varoquaux</p>	10784
<p>Stochastic Multi-Person 3D Motion Forecasting</p> <p>Sirui Xu, Yu-Xiong Wang, Liangyan Gui</p>	10832
<p>Diffusion Probabilistic Fields</p> <p>Peiye Zhuang, Samira Abnar, Jiatao Gu, Alex Schwing, Joshua M. Susskind, Miguel Ángel Bautista</p>	10853
<p>Unicom: Universal and Compact Representation Learning for Image Retrieval</p> <p>Xiang An, Jiankang Deng, Kaicheng Yang, Jaiwei Li, Ziyong Feng, Jia Guo, Jing Yang, Tongliang Liu</p>	10873
<p>Transformer Meets Boundary Value Inverse Problems</p> <p>Ruchi Guo, Shuhao Cao, Long Chen</p>	10889
<p>Distributionally Robust Post-hoc Classifiers under Prior Shifts</p> <p>Jiaheng Wei, Harikrishna Narasimhan, Ehsan Amid, Wen-Sheng Chu, Yang Liu, Abhishek Kumar</p>	10919
<p>Impossibly Good Experts and How to Follow Them</p> <p>Aaron Walsman, Muru Zhang, Sanjiban Choudhury, Dieter Fox, Ali Farhadi</p>	10942
<p>AIM: Adapting Image Models for Efficient Video Action Recognition</p> <p>Taojiannan Yang, Yi Zhu, Yusheng Xie, Aston Zhang, Chen Chen, Mu Li</p>	10959
<p>Predictor-corrector algorithms for stochastic optimization under gradual distribution shift</p> <p>Subha Maity, Debarghya Mukherjee, Moulinath Banerjee, Yuekai Sun</p>	10977
<p>What Can we Learn From The Selective Prediction And Uncertainty Estimation Performance Of 523 Imagenet Classifiers?</p> <p>Ido Galil, Mohammed Dabbah, Ran El-Yaniv</p>	10988
<p>Deep Generative Symbolic Regression</p> <p>Samuel Holt, Zhaozhi Qian, Mihaela van der Schaar</p>	11014
<p>Retrieval-based Controllable Molecule Generation</p> <p>Zichao Wang, Weili Nie, Zhuoran Qiao, Chaowei Xiao, Richard Baraniuk, Anima Anandkumar</p>	11056

<p>Simplifying Model-based RL: Learning Representations, Latent-space Models, and Policies with One Objective</p> <p><i>Raj Ghugare, Homanga Bharadhwaj, Benjamin Eysenbach, Sergey Levine, Russ Salakhutdinov</i></p>	11086
<p>SlotFormer: Unsupervised Visual Dynamics Simulation with Object-Centric Models</p> <p><i>Ziyi Wu, Nikita Dvornik, Klaus Greff, Thomas Kipf, Animesh Garg</i></p>	11117
<p>A framework for benchmarking Class-out-of-distribution detection and its application to ImageNet</p> <p><i>Ido Galil, Mohammed Dabbah, Ran El-Yaniv</i></p>	11143
<p>Learning to Estimate Shapley Values with Vision Transformers</p> <p><i>Ian Connick Covert, Chanwoo Kim, Su-In Lee</i></p>	11162
<p>FedDAR: Federated Domain-Aware Representation Learning</p> <p><i>Aoxiao Zhong, Hao He, Zhaolin Ren, Na Li, Quanzheng Li</i></p>	11210
<p>Learning differentiable solvers for systems with hard constraints</p> <p><i>Geoffrey NÅ@gjar, Michael W. Mahoney, Aditi Krishnapriyan</i></p>	11241
<p>Calibrating Sequence likelihood Improves Conditional Language Generation</p> <p><i>Yao Zhao, Mikhail Khalman, Rishabh Joshi, Shashi Narayan, Mohammad Saleh, Peter J Liu</i></p>	11260
<p>Understanding new tasks through the lens of training data via exponential tilting</p> <p><i>Subha Maity, Mikhail Yurochkin, Moulinath Banerjee, Yuekai Sun</i></p>	11281
<p>CUDA: Curriculum of Data Augmentation for Long-tailed Recognition</p> <p><i>Sumyeong Ahn, Jongwoo Ko, Se-Young Yun</i></p>	11295
<p>Benchmarking Offline Reinforcement Learning on Real-Robot Hardware</p> <p><i>Nico GÅ¼rtler, Sebastian Blaes, Pavel Kolev, Felix Widmaier, Manuel Wuthrich, Stefan Bauer, Bernhard SchÅ¼kopf, Georg Martius</i></p>	11319
<p>UNIFIED-IO: A Unified Model for Vision, Language, and Multi-modal Tasks</p> <p><i>Jiasen Lu, Christopher Clark, Rowan Zellers, Roozbeh Mottaghi, Aniruddha Kembhavi</i></p>	11354
<p>How Much Space Has Been Explored? Measuring the Chemical Space Covered by Databases and Machine-Generated Molecules</p> <p><i>Yutong Xie, Ziqiao Xu, Jiaqi Ma, Qiaozhu Mei</i></p>	11388
<p>Markup-to-Image Diffusion Models with Scheduled Sampling</p> <p><i>Yuntian Deng, Noriyuki Kojima, Alexander M Rush</i></p>	11422
<p>Certified Defences Against Adversarial Patch Attacks on Semantic Segmentation</p> <p><i>Maksym Yatsura, Kaspar Sakmann, N. Grace Hua, Matthias Hein, Jan Hendrik Metzen</i></p>	11443
<p>On the Performance of Temporal Difference Learning With Neural Networks</p> <p><i>HAOXING TIAN, Ioannis Paschalidis, Alex Olshevsky</i></p>	11474

<p>Agent-based Graph Neural Networks</p> <p><i>Karolis Martinkus, Pıl Andrıs Papp, Benedikt Schesch, Roger Wattenhofer</i></p>	11504
<p>Masked Vision and Language Modeling for Multi-modal Representation Learning</p> <p><i>Gukyeong Kwon, Zhaowei Cai, Avinash Ravichandran, Erhan Bas, Rahul Bhotika, Stefano Soatto</i></p>	11536
<p>Bridge the Inference Gaps of Neural Processes via Expectation Maximization</p> <p><i>Qi Wang, Marco Federici, Herke van Hoof</i></p>	11554
<p>InCoder: A Generative Model for Code Infilling and Synthesis</p> <p><i>Daniel Fried, Armen Aghajanyan, Jessy Lin, Sida Wang, Eric Wallace, Freda Shi, Ruiqi Zhong, Scott Yih, Luke Zettlemoyer, Mike Lewis</i></p>	11587
<p>Provable Memorization Capacity of Transformers</p> <p><i>Junghwan Kim, Michelle Kim, Barzan Mozafari</i></p>	11613
<p>Generative Modeling Helps Weak Supervision (and Vice Versa)</p> <p><i>Benedikt Boecking, Nicholas Roberts, Willie Neiswanger, Stefano Ermon, Frederic Sala, Artur Dubrawski</i></p>	11645
<p>Learning to Grow Pretrained Models for Efficient Transformer Training</p> <p><i>Peihao Wang, Rameswar Panda, Lucas Torroba Hennigen, Philip Greengard, Leonid Karlinsky, Rogerio Feris, David Daniel Cox, Zhangyang Wang, Yoon Kim</i></p>	11677
<p>FoSR: First-order spectral rewiring for addressing oversquashing in GNNs</p> <p><i>Kedar Karhadkar, Pradeep Kr. Banerjee, Guido Montufar</i></p>	11695
<p>Near-optimal Policy Identification in Active Reinforcement Learning</p> <p><i>Xiang Li, Viraj Mehta, Johannes Kirschner, Ian Char, Willie Neiswanger, Jeff Schneider, Andreas Krause, Ilija Bogunovic</i></p>	11716
<p>The Lazy Neuron Phenomenon: On Emergence of Activation Sparsity in Transformers</p> <p><i>Zonglin Li, Chong You, Srinadh Bhojanapalli, Daliang Li, Ankit Singh Rawat, Sashank J. Reddi, Ke Ye, Felix Chern, Felix Yu, Ruiqi Guo, Sanjiv Kumar</i></p>	11736
<p>VoGE: A Differentiable Volume Renderer using Gaussian Ellipsoids for Analysis-by-Synthesis</p> <p><i>Angtian Wang, Peng Wang, Jian Sun, Adam Kortylewski, Alan Yuille</i></p>	11768
<p>Learning Soft Constraints From Constrained Expert Demonstrations</p> <p><i>Ashish Gaurav, Kasra Rezaee, Guiliang Liu, Pascal Poupart</i></p>	11793
<p>Sparse tree-based Initialization for Neural Networks</p> <p><i>Patrick Lutz, Ludovic Arnould, Claire Boyer, Erwan Scornet</i></p>	11825
<p>Continuized Acceleration for Quasar Convex Functions in Non-Convex Optimization</p> <p><i>Jun-Kun Wang, Andre Wibisono</i></p>	11855

<p>Constructive TT-representation of the tensors given as index interaction functions with applications</p> <p><i>Gleb Ryzhakov, Ivan Oseledets</i></p>	11879
<p>DiffDock: Diffusion Steps, Twists, and Turns for Molecular Docking</p> <p><i>Gabriele Corso, Hannes StÅark, Bowen Jing, Regina Barzilay, Tommi S. Jaakkola</i></p>	11907
<p>Meta-Learning in Games</p> <p><i>Keegan Harris, Ioannis Anagnostides, Gabriele Farina, Mikhail Khodak, Steven Wu, Tuomas Sandholm</i></p>	11940
<p>Combinatorial-Probabilistic Trade-Off: P-Values of Community Properties Test in the Stochastic Block Models</p> <p><i>Shuting Shen, Junwei Lu</i></p>	11957
<p>Multi-Rate VAE: Train Once, Get the Full Rate-Distortion Curve</p> <p><i>Juhan Bae, Michael R. Zhang, Michael Ruan, Eric Wang, So Hasegawa, Jimmy Ba, Roger Baker Grosse</i></p>	11992
<p>GNNDelete: A General Strategy for Unlearning in Graph Neural Networks</p> <p><i>Jiali Cheng, George Dasoulas, Huan He, Chirag Agarwal, Marinka Zitnik</i></p>	12006
<p>Continuous pseudo-labeling from the start</p> <p><i>Dan Berrebbi, Ronan Collobert, Samy Bengio, Navdeep Jaitly, Tatiana Likhomanenko</i></p>	12029
<p>Stable Target Field for Reduced Variance Score Estimation in Diffusion Models</p> <p><i>Yilun Xu, Shangyuan Tong, Tommi S. Jaakkola</i></p>	12044
<p>Distilling Model Failures as Directions in Latent Space</p> <p><i>Saachi Jain, Hannah Lawrence, Ankur Moitra, Aleksander Madry</i></p>	12056
<p>Relational Attention: Generalizing Transformers for Graph-Structured Tasks</p> <p><i>Cameron Diao, Ricky Loynd</i></p>	12104
<p>Average Sensitivity of Decision Tree Learning</p> <p><i>Satoshi Hara, Yuichi Yoshida</i></p>	12128
<p>ACMP: Allen-Cahn Message Passing with Attractive and Repulsive Forces for Graph Neural Networks</p> <p><i>Yuelin Wang, Kai Yi, Xinliang Liu, Yu Guang Wang, Shi Jin</i></p>	12148
<p>Dual Diffusion Implicit Bridges for Image-to-Image Translation</p> <p><i>Xuan Su, Jiaming Song, Chenlin Meng, Stefano Ermon</i></p>	12175
<p>Hungry Hungry Hippos: Towards Language Modeling with State Space Models</p> <p><i>Daniel Y Fu, Tri Dao, Khaled Kamal Saab, Armin W Thomas, Atri Rudra, Christopher Re</i></p>	12193

<p>The Modality Focusing Hypothesis: Towards Understanding Crossmodal Knowledge Distillation</p> <p>Zihui Xue, Zhengqi Gao, Sucheng Ren, Hang Zhao</p>	12218
<p>Adaptive Robust Evidential Optimization For Open Set Detection from Imbalanced Data</p> <p>Hitesh Sapkota, Qi Yu</p>	12239
<p>Learning Math Reasoning from Self-Sampled Correct and Partially-Correct Solutions</p> <p>Ansong Ni, Jeevana Priya Inala, Chenglong Wang, Alex Polozov, Christopher Meek, Dragomir Radev, Jianfeng Gao</p>	12267
<p>MIMT: Masked Image Modeling Transformer for Video Compression</p> <p>Jinxi Xiang, Kuan Tian, Jun Zhang</p>	12284
<p>Visual Classification via Description from Large Language Models</p> <p>Sachit Menon, Carl Vondrick</p>	12301
<p>On the Perils of Cascading Robust Classifiers</p> <p>Ravi Mangal, Zifan Wang, Chi Zhang, Klas Leino, Corina Pasareanu, Matt Fredrikson</p>	12318
<p>SCALE-UP: An Efficient Black-box Input-level Backdoor Detection via Analyzing Scaled Prediction Consistency</p> <p>Junfeng Guo, Yiming Li, Xun Chen, Hanqing Guo, Lichao Sun, Cong Liu</p>	12337
<p>From Play to Policy: Conditional Behavior Generation from Uncurated Robot Data</p> <p>Zichen Jeff Cui, Yibin Wang, Nur Muhammad Mahi Shafiullah, Lerrel Pinto</p>	12361
<p>Behavior Prior Representation learning for Offline Reinforcement Learning</p> <p>Hongyu Zang, Xin Li, Jie Yu, Chen Liu, Riashat Islam, Remi Tachet des Combes, Romain Laroche</p>	12377
<p>RPM: Generalizable Multi-Agent Policies for Multi-Agent Reinforcement Learning</p> <p>Wei Qiu, Xiao Ma, Bo An, Svetlana Obraztsova, Shuicheng YAN, Zhongwen Xu</p>	12411
<p>HotProtein: A Novel Framework for Protein Thermostability Prediction and Editing</p> <p>Tianlong Chen, Chengyue Gong, Daniel Jesus Diaz, Xuxi Chen, Jordan Tyler Wells, qiang liu, Zhangyang Wang, Andrew Ellington, Alex Dimakis, Adam Klivans</p>	12426
<p>Data Valuation Without Training of a Model</p> <p>Ki Nohyun, Hoyong Choi, Hye Won Chung</p>	12451
<p>Progressive Voronoi Diagram Subdivision Enables Accurate Data-free Class-Incremental Learning</p> <p>Chunwei Ma, Zhanghexuan Ji, Ziyun Huang, Yan Shen, Mingchen Gao, Jinhui Xu</p>	12478
<p>Write and Paint: Generative Vision-Language Models are Unified Modal Learners</p> <p>Shizhe Diao, Wangchunshu Zhou, Xinsong Zhang, Jiawei Wang</p>	12528
<p>LightGCL: Simple Yet Effective Graph Contrastive Learning for Recommendation</p> <p>Xuheng Cai, Chao Huang, Lianghao Xia, Xubin Ren</p>	12553

Distributionally Robust Recourse Action Duy Nguyen, Ngoc Bui, Viet Anh Nguyen	12568
Personalized Federated Learning with Feature Alignment and Classifier Collaboration Jian Xu, Xinyi Tong, Shao-Lun Huang	12593
Spiking Convolutional Neural Networks for Text Classification Changze Lv, Jianhan Xu, Xiaoqing Zheng	12618
Ensuring DNN Solution Feasibility for Optimization Problems with Linear Constraints Tianyu Zhao, Xiang Pan, Minghua Chen, Steven Low	12635
Self-Consistency Improves Chain of Thought Reasoning in Language Models Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc V Le, Ed H. Chi, Sharan Narang, Aakanksha Chowdhery, Denny Zhou	12648
Harnessing Mixed Offline Reinforcement Learning Datasets via Trajectory Weighting Zhang-Wei Hong, Pulkit Agrawal, Remi Tachet des Combes, Romain Laroche	12672
Your Contrastive Learning Is Secretly Doing Stochastic Neighbor Embedding Tianyang Hu, Zhili LIU, Fengwei Zhou, Wenjia Wang, Weiran Huang	12698
Schema Inference for Interpretable Image Classification Haofei Zhang, Mengqi Xue, Xiaokang Liu, Kaixuan Chen, Jie Song, Mingli Song	12723
Out-of-distribution Representation Learning for Time Series Classification Wang Lu, Jindong Wang, Xinwei Sun, Yiqiang Chen, Xing Xie	12752
BEEF: Bi-Compatible Class-Incremental Learning via Energy-Based Expansion and Fusion Fu-Yun Wang, Da-Wei Zhou, Liu Liu, Han-Jia Ye, Yatao Bian, De-Chuan Zhan, Peilin Zhao	12773
Contrastive Alignment of Vision to Language Through Parameter-Efficient Transfer Learning Zaid Khan, Yun Fu	12798
ILA-DA: Improving Transferability of Intermediate Level Attack with Data Augmentation Chiu Wai Yan, Tsz-Him Cheung, Dit-Yan Yeung	12815
Coverage-centric Coreset Selection for High Pruning Rates Haizhong Zheng, Rui Liu, Fan Lai, Atul Prakash	12840
RGI: robust GAN-inversion for mask-free image inpainting and unsupervised pixel-wise anomaly detection Shancong Mou, Xiaoyi Gu, Meng Cao, Haoping Bai, Ping Huang, Jiulong Shan, Jianjun Shi	12859
Adversarial Attacks on Adversarial Bandits Yuzhe Ma, Zhijin Zhou	12887
Learning Hierarchical Protein Representations via Complete 3D Graph Networks Limei Wang, Haoran Liu, Yi Liu, Jerry Kurtin, Shuiwang Ji	12899

<p>NERDS: A General Framework to Train Camera Denoisers from Raw-RGB Noisy Image Pairs</p> <p>Heewon Kim, Kyoung Mu Lee</p>	12920
<p>Prototypical Calibration for Few-shot Learning of Language Models</p> <p>Zhixiong Han, Yaru Hao, Li Dong, Yutao Sun, Furu Wei</p>	12940
<p>Hierarchical Sliced Wasserstein Distance</p> <p>Khai Nguyen, Tongzheng Ren, Huy Nguyen, Litu Rout, Tan Minh Nguyen, Nhat Ho</p>	12956
<p>SWIFT: Rapid Decentralized Federated Learning via Wait-Free Model Communication</p> <p>Marco Bornstein, Tahseen Rabbani, Evan Z Wang, Amrit Bedi, Furong Huang</p>	12984
<p>MICN: Multi-scale Local and Global Context Modeling for Long-term Series Forecasting</p> <p>Huiqiang Wang, Jian Peng, Feihu Huang, Jince Wang, Junhui Chen, Yifei Xiao</p>	13014
<p>Characterizing the Influence of Graph Elements</p> <p>Zizhang Chen, Peizhao Li, Hongfu Liu, Pengyu Hong</p>	13036
<p>PV3D: A 3D Generative Model for Portrait Video Generation</p> <p>Zhongcong Xu, Jianfeng Zhang, Jun Hao Liew, Wenqing Zhang, Song Bai, Jiashi Feng, Mike Zheng Shou</p>	13058
<p>Image to Sphere: Learning Equivariant Features for Efficient Pose Prediction</p> <p>David Klee, Ondrej Biza, Robert Platt, Robin Walters</p>	13075
<p>wav2tok: Deep Sequence Tokenizer for Audio Retrieval</p> <p>Adhiraj Banerjee, Vipul Arora</p>	13093
<p>Boosting Multiagent Reinforcement Learning via Permutation Invariant and Permutation Equivariant Networks</p> <p>Jianye HAO, Xiaotian Hao, Hangyu Mao, Weixun Wang, Yaodong Yang, Dong Li, YAN ZHENG, Zhen Wang</p>	13109
<p>Learned Index with Dynamic ϵ</p> <p>Daoyuan Chen, Wuchao Li, Yaliang Li, Bolin Ding, Kai Zeng, Defu Lian, Jingren Zhou</p>	13133
<p>Rethinking Self-Supervised Visual Representation Learning in Pre-training for 3D Human Pose and Shape Estimation</p> <p>Hongsuk Choi, Hyeongjin Nam, Taeryung Lee, Gyeongsik Moon, Kyoung Mu Lee</p>	13158
<p>MaskFusion: Feature Augmentation for Click-Through Rate Prediction via Input-adaptive Mask Fusion</p> <p>Chao Liao, Jianchao Tan, Jiyuan Jia, Yi Guo, Chengru Song</p>	13176
<p>Generalizing and Decoupling Neural Collapse via Hyperspherical Uniformity Gap</p> <p>Weiyang Liu, Longhui Yu, Adrian Weller, Bernhard Schölkopf</p>	13196
<p>Latent State Marginalization as a Low-cost Approach for Improving Exploration</p> <p>Dinghui Zhang, Aaron Courville, Yoshua Bengio, Qinqing Zheng, Amy Zhang, Ricky T. Q. Chen</p>	13229

<p>Cycle to Clique (Cy2C) Graph Neural Network: A Sight to See beyond Neighborhood Aggregation</p> <p><i>Yun Young Choi, Sun Woo Park, Youngho Woo, U Jin Choi</i></p>	13253
<p>Maximizing Spatio-Temporal Entropy of Deep 3D CNNs for Efficient Video Recognition</p> <p><i>Junyan Wang, Zhenhong Sun, Yichen Qian, Dong Gong, Xiuyu Sun, Ming Lin, Maurice Pagnucco, Yang Song</i></p>	13290
<p>A General Framework for Sample-Efficient Function Approximation in Reinforcement Learning</p> <p><i>Zixiang Chen, Chris Junchi Li, Huizhuo Yuan, Quanquan Gu, Michael Jordan</i></p>	13310
<p>EUCLID: Towards Efficient Unsupervised Reinforcement Learning with Multi-choice Dynamics Model</p> <p><i>Yifu Yuan, Jianye HAO, Fei Ni, Yao Mu, YAN ZHENG, Yujing Hu, Jinyi Liu, Yingfeng Chen, Changjie Fan</i></p>	13352
<p>Feature Reconstruction From Outputs Can Mitigate Simplicity Bias in Neural Networks</p> <p><i>Sravanti Addepalli, Anshul Nasery, Venkatesh Babu Radhakrishnan, Praneeth Netrapalli, Prateek Jain</i></p>	13374
<p>Improving Object-centric Learning with Query Optimization</p> <p><i>Baoxiong Jia, Yu Liu, Siyuan Huang</i></p>	13398
<p>Optimal Activation Functions for the Random Features Regression Model</p> <p><i>Jianxin Wang, JosÃ© Bento</i></p>	13430
<p>Gromov-Wasserstein Autoencoders</p> <p><i>Nao Nakagawa, Ren Togo, Takahiro Ogawa, Miki Haseyama</i></p>	13471
<p>Toward Adversarial Training on Contextualized Language Representation</p> <p><i>Hongqiu Wu, Yongxiang Liu, Hanwen Shi, hai zhao, Min Zhang</i></p>	13509
<p>ODAM: Gradient-based Instance-Specific Visual Explanations for Object Detection</p> <p><i>Chenyang ZHAO, Antoni B. Chan</i></p>	13526
<p>Treeformer: Dense Gradient Trees for Efficient Attention Computation</p> <p><i>Lovish Madaan, Srinadh Bhojanapalli, Himanshu Jain, Prateek Jain</i></p>	13556
<p>RoPAWS: Robust Semi-supervised Representation Learning from Uncurated Data</p> <p><i>Sangwoo Mo, Jong-Chyi Su, Chih-Yao Ma, Mido Assran, Ishan Misra, Licheng Yu, Sean Bell</i></p>	13571
<p>LMSeg: Language-guided Multi-dataset Segmentation</p> <p><i>Qiang Zhou, Yuang Liu, Chaohui Yu, Jingliang Li, Zhibin Wang, Fan Wang</i></p>	13597
<p>Leveraging Future Relationship Reasoning for Vehicle Trajectory Prediction</p> <p><i>Daehee Park, Hobin Ryu, Yunseo Yang, Jegyeong Cho, Jiwon Kim, Kuk-Jin Yoon</i></p>	13609
<p>In-sample Actor Critic for Offline Reinforcement Learning</p> <p><i>Hongchang Zhang, Yixiu Mao, Boyuan Wang, Shuncheng He, Yi Xu, Xiangyang Ji</i></p>	13622

<p>CircNet: Meshing 3D Point Clouds with Circumcenter Detection</p> <p>Huan Lei, Ruitao Leng, Liang Zheng, Hongdong Li</p>	13645
<p>Variance-Aware Sparse Linear Bandits</p> <p>Yan Dai, Ruosong Wang, Simon Shaolei Du</p>	13657
<p>Q-Pensieve: Boosting Sample Efficiency of Multi-Objective RL Through Memory Sharing of Q-Snapshots</p> <p>Wei Hung, Bo Kai Huang, Ping-Chun Hsieh, Xi Liu</p>	13693
<p>Exploring Temporally Dynamic Data Augmentation for Video Recognition</p> <p>Taeoh Kim, Jinhjung Kim, Minho Shim, Sangdoon Yun, Myunggu Kang, Dongyoon Wee, Sangyoun Lee</p>	13713
<p>Towards Lightweight, Model-Agnostic and Diversity-Aware Active Anomaly Detection</p> <p>Xu Zhang, Yuan Zhao, Ziang Cui, Liqun Li, Shilin He, Qingwei Lin, Yingnong Dang, Saravan Rajmohan, Dongmei Zhang</p>	13735
<p>Time to augment self-supervised visual representation learning</p> <p>Arthur Aubret, Markus R. Ernst, Cline Teulire, Jochen Triesch</p>	13752
<p>Measure the Predictive Heterogeneity</p> <p>Jiashuo Liu, Jiayun Wu, Renjie Pi, Renzhe Xu, Xingxuan Zhang, Bo Li, Peng Cui</p>	13773
<p>Recon: Reducing Conflicting Gradients From the Root For Multi-Task Learning</p> <p>Guangyuan SHI, Qimai Li, Wenlong Zhang, Jiaxin Chen, Xiao-Ming Wu</p>	13798
<p>Predictive Inference with Feature Conformal Prediction</p> <p>Jiaye Teng, Chuan Wen, Dinghuai Zhang, Yoshua Bengio, Yang Gao, Yang Yuan</p>	13818
<p>Benign Overfitting in Classification: Provably Counter Label Noise with Larger Models</p> <p>Kaiyue Wen, Jiaye Teng, Jingzhao Zhang</p>	13846
<p>Sharper Bounds for Uniformly Stable Algorithms with Stationary Mixing Process</p> <p>Shi Fu, Yunwen Lei, Qiong Cao, Xinmei Tian, Dacheng Tao</p>	13871
<p>Semi-Implicit Variational Inference via Score Matching</p> <p>Longlin Yu, Cheng Zhang</p>	13892
<p>Corrupted Image Modeling for Self-Supervised Visual Pre-Training</p> <p>Yuxin Fang, Li Dong, Hangbo Bao, Xinggang Wang, Furu Wei</p>	13909
<p>Transformer-Patcher: One Mistake Worth One Neuron</p> <p>Zeyu Huang, Yikang Shen, Xiaofeng Zhang, Jie Zhou, Wenge Rong, Zhang Xiong</p>	13927
<p>Rarity Score : A New Metric to Evaluate the Uncommonness of Synthesized Images</p> <p>Jiyeon Han, Hwanil Choi, Yunjey Choi, Junho Kim, Jung-Woo Ha, Jaesik Choi</p>	13943

<p>Self-supervision through Random Segments with Autoregressive Coding (RandSAC)</p> <p><i>Tianyu Hua, Yonglong Tian, Sucheng Ren, Michalis Raptis, Hang Zhao, Leonid Sigal</i></p>	13970
<p>Generative Modelling with Inverse Heat Dissipation</p> <p><i>Severi Rissanen, Markus Heinonen, Arno Solin</i></p>	13992
<p>Better Generative Replay for Continual Federated Learning</p> <p><i>Daiqing Qi, Handong Zhao, Sheng Li</i></p>	14046
<p>Multitask Prompt Tuning Enables Parameter-Efficient Transfer Learning</p> <p><i>Zhen Wang, Rameswar Panda, Leonid Karlinsky, Rogerio Feris, Huan Sun, Yoon Kim</i></p>	14063
<p>Graph Neural Networks for Link Prediction with Subgraph Sketching</p> <p><i>Benjamin Paul Chamberlain, Sergey Shirobokov, Emanuele Rossi, Fabrizio Frasca, Thomas Markovich, Nils Yannick Hammerla, Michael M. Bronstein, Max Hansmire</i></p>	14079
<p>Variance Reduction is an Antidote to Byzantines: Better Rates, Weaker Assumptions and Communication Compression as a Cherry on the Top</p> <p><i>Eduard Gorbunov, Samuel Horváth, Peter Richtárik, Gauthier Gidel</i></p>	14106
<p>Clean-image Backdoor: Attacking Multi-label Models with Poisoned Labels Only</p> <p><i>Kangjie Chen, Xiaoxuan Lou, Guowen Xu, Jiwei Li, Tianwei Zhang</i></p>	14148
<p>VRMA²: Aggregation-Aware Quantization for Graph Neural Networks</p> <p><i>Zeyu Zhu, Fanrong Li, Zitao Mo, Qinghao Hu, Gang Li, Zejian Liu, Xiaoyao Liang, Jian Cheng</i></p>	14166
<p>A Differential Geometric View and Explainability of GNN on Evolving Graphs</p> <p><i>Yazheng Liu, Xi Zhang, Sihong Xie</i></p>	14195
<p>Distributional Meta-Gradient Reinforcement Learning</p> <p><i>Haiyan Yin, Shuicheng YAN, Zhongwen Xu</i></p>	14220
<p>Classically Approximating Variational Quantum Machine Learning with Random Fourier Features</p> <p><i>Jonas Landman, Slimane Thabet, Constantin Dalyac, Hela Mhiri, Elham Kashefi</i></p>	14232
<p>Versatile Neural Processes for Learning Implicit Neural Representations</p> <p><i>Zongyu Guo, Cuiling Lan, Zhizheng Zhang, Yan Lu, Zhibo Chen</i></p>	14256
<p>Simplicial Hopfield networks</p> <p><i>Thomas F Burns, Tomoki Fukai</i></p>	14268
<p>Are More Layers Beneficial to Graph Transformers?</p> <p><i>Haiteng Zhao, Shuming Ma, Dongdong Zhang, Zhi-Hong Deng, Furu Wei</i></p>	14304
<p>DiffEdit: Diffusion-based semantic image editing with mask guidance</p> <p><i>Guillaume Couairon, Jakob Verbeek, Holger Schwenk, Matthieu Cord</i></p>	14324

Re-weighting Based Group Fairness Regularization via Classwise Robust Optimization Sangwon Jung, Taeon Park, Sanghyuk Chun, Taesup Moon	14346
Prompt-to-Prompt Image Editing with Cross-Attention Control Amir Hertz, Ron Mokady, Jay Tenenbaum, Kfir Aberman, Yael Pritch, Daniel Cohen-or	14369
VA-DepthNet: A Variational Approach to Single Image Depth Prediction Ce Liu, Suryansh Kumar, Shuhang Gu, Radu Timofte, Luc Van Gool	14405
E3Bind: An End-to-End Equivariant Network for Protein-Ligand Docking Yangtian Zhang, Huiyu Cai, Chence Shi, Jian Tang	14426
Semi-supervised learning with a principled likelihood from a generative model of data curation Stoil Krasimirov Ganev, Laurence Aitchison	14446
A Convergent Single-Loop Algorithm for Relaxation of Gromov-Wasserstein in Graph Data Jiajin Li, Jianheng Tang, Lemin Kong, Huikang Liu, Jia Li, Anthony Man-Cho So, Jose Blanchet	14458
BAYES RISK CTC: CONTROLLABLE CTC ALIGNMENT IN SEQUENCE-TO-SEQUENCE TASKS Jinchuan Tian, Brian Yan, Jianwei Yu, CHAO WENG, Dong Yu, Shinji Watanabe	14471
Symbolic Physics Learner: Discovering governing equations via Monte Carlo tree search Fangzheng Sun, Yang Liu, Jian-Xun Wang, Hao Sun	14496
PowerQuant: Automorphism Search for Non-Uniform Quantization Edouard YVINEC, Arnaud Dapogny, Matthieu Cord, Kevin Bailly	14518
Unsupervised visualization of image datasets using contrastive learning Niklas BÅllm, Philipp Berens, Dmitry Kobak	14539
Video Scene Graph Generation from Single-Frame Weak Supervision Siqi Chen, Jun Xiao, Long Chen	14560
Searching Lottery Tickets in Graph Neural Networks: A Dual Perspective Kun Wang, Yuxuan Liang, Pengkun Wang, Xu Wang, Pengfei Gu, Junfeng Fang, Yang Wang	14576
StyleMorph: Disentangled 3D-Aware Image Synthesis with a 3D Morphable StyleGAN Eric-Tuan Le, Edward Bartrum, Iasonas Kokkinos	14596
MixPro: Data Augmentation with MaskMix and Progressive Attention Labeling for Vision Transformer Qihao Zhao, Yangyu Huang, Wei Hu, Fan Zhang, Jun Liu	14608
New Insights for the Stability-Plasticity Dilemma in Online Continual Learning Dahuin Jung, Dongjin Lee, Sunwon Hong, Hyemi Jang, Ho Bae, Sungroh Yoon	14624
Efficient Offline Policy Optimization with a Learned Model Zichen Liu, Siyi Li, Wee Sun Lee, Shuicheng YAN, Zhongwen Xu	14648

<p>Hyperbolic Self-paced Learning for Self-supervised Skeleton-based Action Representations</p> <p><i>Luca Franco, Paolo Mandica, Bharti Munjal, Fabio Galasso</i></p>	14670
<p>Blurring Diffusion Models</p> <p><i>Emiel Hoogeboom, Tim Salimans</i></p>	14689
<p>How Informative is the Approximation Error from Tensor Decomposition for Neural Network Compression?</p> <p><i>Jetze Schuurmans, kim batselier, Julian Kooij</i></p>	14703
<p>MultiViz: Towards Visualizing and Understanding Multimodal Models</p> <p><i>Paul Pu Liang, Yiwei Lyu, Gunjan Chhablani, Nihal Jain, Zihao Deng, Xingbo Wang, Louis-Philippe Morency, Ruslan Salakhutdinov</i></p>	14718
<p>Faster Gradient-Free Methods for Escaping Saddle Points</p> <p><i>Hualin Zhang, Bin Gu</i></p>	14767
<p>MARS: Meta-learning as Score Matching in the Function Space</p> <p><i>Krunoslav Lehman Pavasovic, Jonas Rothfuss, Andreas Krause</i></p>	14797
<p>Addressing Parameter Choice Issues in Unsupervised Domain Adaptation by Aggregation</p> <p><i>Marius-Constantin Dinu, Markus Holzleitner, Maximilian Beck, Hoan Duc Nguyen, Andrea Huber, Hamid Eghbal-zadeh, Bernhard A. Moser, Sergei Pereverzyev, Sepp Hochreiter, Werner Zellinger</i></p>	14823
<p>Finding the Global Semantic Representation in GAN through Fréchet Mean</p> <p><i>Jaewoong Choi, Geonho Hwang, Hyunsoo Cho, Myungjoo Kang</i></p>	14874
<p>KnowDA: All-in-One Knowledge Mixture Model for Data Augmentation in Low-Resource NLP</p> <p><i>Yufei Wang, Jiayi Zheng, Can Xu, Xiubo Geng, Tao Shen, Chongyang Tao, Daxin Jiang</i></p>	14899
<p>EA-HAS-Bench: Energy-aware Hyperparameter and Architecture Search Benchmark</p> <p><i>Shuguang Dou, XINYANG JIANG, Cai Rong Zhao, Dongsheng Li</i></p>	14924
<p>Massively Scaling Heteroscedastic Classifiers</p> <p><i>Mark Collier, Rodolphe Jenatton, Basil Mustafa, Neil Houlsby, Jesse Berent, Efrosyni Kokiopoulou</i></p>	14948
<p>Language Modelling with Pixels</p> <p><i>Phillip Rust, Jonas F. Lotz, Emanuele Bugliarello, Elizabeth Salesky, Miryam de Lhoneux, Desmond Elliott</i></p>	14978
<p>Learning with Auxiliary Activation for Memory-Efficient Training</p> <p><i>Sunghyeon Woo, Dongsuk Jeon</i></p>	15010
<p>The Symmetric Generalized Eigenvalue Problem as a Nash Equilibrium</p> <p><i>Ian Gemp, Charlie Chen, Brian McWilliams</i></p>	15034
<p>DEP-RL: Embodied Exploration for Reinforcement Learning in Overactuated and Musculoskeletal Systems</p> <p><i>Pierre Schumacher, Daniel Haeuffle, Dieter Bacher, Syn Schmitt, Georg Martius</i></p>	15047

<p>Uni-Mol: A Universal 3D Molecular Representation Learning Framework</p> <p>Gengmo Zhou, Zhifeng Gao, Qiankun Ding, Hang Zheng, Hongteng Xu, Zhewei Wei, Linfeng Zhang, Guolin Ke</p>	15076
<p>Dynamic Update-to-Data Ratio: Minimizing World Model Overfitting</p> <p>Nicolai Dorka, Tim Welschehold, Wolfram Burgard</p>	15107
<p>On the Effectiveness of Out-of-Distribution Data in Self-Supervised Long-Tail Learning.</p> <p>Jianhong Bai, Zuoqiu Liu, Hualiang Wang, Jin Hao, YANG FENG, Huanpeng Chu, Haoji Hu</p>	15126
<p>Simple and Scalable Nearest Neighbor Machine Translation</p> <p>Yuhan Dai, Zhirui Zhang, Qiuzhi Liu, Qu Cui, Weihua Li, Yichao Du, Tong Xu</p>	15145
<p>Unsupervised Manifold Alignment with Joint Multidimensional Scaling</p> <p>Dexiong Chen, Bowen Fan, Carlos Oliver, Karsten Borgwardt</p>	15162
<p>Relative representations enable zero-shot latent space communication</p> <p>Luca Moschella, Valentino Maiorca, Marco Fumero, Antonio Norelli, Francesco Locatello, Emanuele Rodolà</p>	15186
<p>Accurate Neural Training with 4-bit Matrix Multiplications at Standard Formats</p> <p>Brian Chmiel, Ron Banner, Elad Hoffer, Hilla Ben-Yaacov, Daniel Soudry</p>	15212
<p>Test-Time Adaptation via Self-Training with Nearest Neighbor Information</p> <p>Minguk Jang, Sae-Young Chung, Hye Won Chung</p>	15223
<p>Context-enriched molecule representations improve few-shot drug discovery</p> <p>Johannes Schimunek, Philipp Seidl, Lukas Friedrich, Daniel Kuhn, Friedrich Rippmann, Sepp Hochreiter, G�nter Klambauer</p>	15252
<p>Learning to Decompose Visual Features with Latent Textual Prompts</p> <p>Feng Wang, Manling Li, Xudong Lin, Hairong Lv, Alex Schwing, Heng Ji</p>	15282
<p>Accurate Bayesian Meta-Learning by Accurate Task Posterior Inference</p> <p>Michael Volpp, Philipp Dahlinger, Philipp Becker, Christian Daniel, Gerhard Neumann</p>	15294
<p>On the Usefulness of Embeddings, Clusters and Strings for Text Generation Evaluation</p> <p>Tiago Pimentel, Clara Isabel Meister, Ryan Cotterell</p>	15327
<p>Batch Multivald Conformal Prediction</p> <p>Christopher Jung, Georgy Noarov, Ramya Ramalingam, Aaron Roth</p>	15347
<p>Accelerating Guided Diffusion Sampling with Splitting Numerical Methods</p> <p>Suttisak Wizadwongsa, Supasorn Suwajanakorn</p>	15380
<p>Improving Out-of-distribution Generalization with Indirection Representations</p> <p>Kha Pham, Hung Le, Man Ngo, Truyen Tran</p>	15392

<p>Does Deep Learning Learn to Abstract? A Systematic Probing Framework</p> <p>Shengnan An, Zeqi Lin, Bei Chen, Qiang Fu, Nanning Zheng, Jian-Guang Lou</p>	15411
<p>Neural Architecture Design and Robustness: A Dataset</p> <p>Steffen Jung, Jovita Lukasik, Margret Keuper</p>	15437
<p>SpeedyZero: Mastering Atari with Limited Data and Time</p> <p>Yixuan Mei, Jiaxuan Gao, Weirui Ye, Shaohuai Liu, Yang Gao, Yi Wu</p>	15470
<p>Discovering Policies with DOMiNO: Diversity Optimization Maintaining Near Optimality</p> <p>Tom Zahavy, Yannick Schroecker, Feryal Behbahani, Kate Baumli, Sebastian Flennerhag, Shaobo Hou, Satinder Singh</p>	15486
<p>Memory Gym: Partially Observable Challenges to Memory-Based Agents</p> <p>Marco Pleines, Matthias Pallasch, Frank Zimmer, Mike Preuss</p>	15517
<p>Benchmarking Constraint Inference in Inverse Reinforcement Learning</p> <p>Guiliang Liu, Yudong Luo, Ashish Gaurav, Kasra Rezaee, Pascal Poupart</p>	15543
<p>ViewCo: Discovering Text-Supervised Segmentation Masks via Multi-View Semantic Consistency</p> <p>Pengzhen Ren, Changlin Li, Hang Xu, Yi Zhu, Guangrun Wang, Jianzhuang Liu, Xiaojun Chang, Xiaodan Liang</p>	15570
<p>Towards Open Temporal Graph Neural Networks</p> <p>Kaituo Feng, Changsheng Li, Xiaolu Zhang, JUN ZHOU</p>	15582
<p>Martingale Posterior Neural Processes</p> <p>Hyungi Lee, Eunggu Yun, Giung Nam, Edwin Fong, Juho Lee</p>	15605
<p>Decoupled Training for Long-Tailed Classification With Stochastic Representations</p> <p>Giung Nam, Sunguk Jang, Juho Lee</p>	15628
<p>A Simple Yet Powerful Deep Active Learning With Snapshots Ensembles</p> <p>Seohyeon Jung, Sanghyun Kim, Juho Lee</p>	15651
<p>Fast and Precise: Adjusting Planning Horizon with Adaptive Subgoal Search</p> <p>Michał Zawalski, Michał Tyrolski, Konrad Czechowski, Tomasz Odrzygóźdź, Damian Stachura, Piotr Piłkos, Yuhuai Wu, Łukasz Kuciński, Piotr Miłoś</p>	15669
<p>On the Feasibility of Cross-Task Transfer with Model-Based Reinforcement Learning</p> <p>Yifan Xu, Nicklas Hansen, Zirui Wang, Yung-Chieh Chan, Hao Su, Zhuowen Tu</p>	15701
<p>Equivariant Energy-Guided SDE for Inverse Molecular Design</p> <p>Fan Bao, Min Zhao, Zhongkai Hao, Peiyao Li, Chongxuan Li, Jun Zhu</p>	15725
<p>Pre-training via Denoising for Molecular Property Prediction</p> <p>Sheheryar Zaidi, Michael Schaarschmidt, James Martens, Hyunjik Kim, Yee Whye Teh, Alvaro Sanchez-Gonzalez, Peter Battaglia, Razvan Pascanu, Jonathan Godwin</p>	15748

<p>CLIP-ViP: Adapting Pre-trained Image-Text Model to Video-Language Alignment</p> <p>Hongwei Xue, Yuchong Sun, Bei Liu, Jianlong Fu, Ruihua Song, Houqiang Li, Jiebo Luo</p>	15774
<p>Diminishing Return of Value Expansion Methods in Model-Based Reinforcement Learning</p> <p>Daniel Palenicek, Michael Lutter, Joao Carvalho, Jan Peters</p>	15788
<p>Jointly Learning Visual and Auditory Speech Representations from Raw Data</p> <p>Alexandros Haliassos, Pingchuan Ma, Rodrigo Mira, Stavros Petridis, Maja Pantic</p>	15809
<p>Over-parameterized Model Optimization with Polyak-$\{L\}$ojasiewicz Condition</p> <p>Yixuan Chen, Yubin Shi, Mingzhi Dong, Xiaochen Yang, Dongsheng Li, Yujiang Wang, Robert P. Dick, Qin Lv, Yingying Zhao, Fan Yang, Ning Gu, Li Shang</p>	15831
<p>Warping the Space: Weight Space Rotation for Class-Incremental Few-Shot Learning</p> <p>Do-Yeon Kim, Dong-Jun Han, Jun Seo, Jaekyun Moon</p>	15857
<p>D4FT: A Deep Learning Approach to Kohn-Sham Density Functional Theory</p> <p>Tianbo Li, Min Lin, Zheyuan Hu, Kunhao Zheng, Giovanni Vignale, Kenji Kawaguchi, A.H. Castro Neto, Kostya S. Novoselov, Shuicheng YAN</p>	15876
<p>TransSpeech: Speech-to-Speech Translation With Bilateral Perturbation</p> <p>Rongjie Huang, Jinglin Liu, Huadai Liu, Yi Ren, Lichao Zhang, Jinzheng He, Zhou Zhao</p>	15900
<p>A Neural Mean Embedding Approach for Back-door and Front-door Adjustment</p> <p>Liyuan Xu, Arthur Gretton</p>	15917
<p>A Call to Reflect on Evaluation Practices for Failure Detection in Image Classification</p> <p>Paul F Jaeger, Carsten Tim L$\frac{1}{4}$th, Lukas Klein, Till J. Bungert</p>	15940
<p>Weakly Supervised Knowledge Transfer with Probabilistic Logical Reasoning for Object Detection</p> <p>Martijn Oldenhof, Adam Arany, Yves Moreau, Edward De Brouwer</p>	15978
<p>Learning Adversarial Linear Mixture Markov Decision Processes with Bandit Feedback and Unknown Transition</p> <p>Canzhe Zhao, Ruofeng Yang, Baoxiang Wang, Shuai Li</p>	15999
<p>Embedding Fourier for Ultra-High-Definition Low-Light Image Enhancement</p> <p>Chongyi Li, Chun-Le Guo, man zhou, Zhexin Liang, Shangchen Zhou, Ruicheng Feng, Chen Change Loy</p>	16013
<p>Approximation and non-parametric estimation of functions over high-dimensional spheres via deep ReLU networks</p> <p>Namjoon Suh, Tian-Yi Zhou, Xiaoming Huo</p>	16040
<p>Non-parametric Outlier Synthesis</p> <p>Leitian Tao, Xuefeng Du, Jerry Zhu, Yixuan Li</p>	16071

<p>Self-Supervised Geometric Correspondence for Category-Level 6D Object Pose Estimation in the Wild</p> <p><i>Kaifeng Zhang, Yang Fu, Shubhankar Borse, Hong Cai, Fatih Porikli, Xiaolong Wang</i></p>	16091
<p>Deep Transformers without Shortcuts: Modifying Self-attention for Faithful Signal Propagation</p> <p><i>Bobby He, James Martens, Guodong Zhang, Aleksandar Botev, Andrew Brock, Samuel L Smith, Yee Whye Teh</i></p>	16110
<p>A Learning Based Hypothesis Test for Harmful Covariate Shift</p> <p><i>Tom Ginsberg, Zhongyuan Liang, Rahul G Krishnan</i></p>	16152
<p>Temporal Coherent Test Time Optimization for Robust Video Classification</p> <p><i>Chenyu Yi, SIYUAN YANG, Yufei Wang, Haoliang Li, Yap-peng Tan, Alex Kot</i></p>	16186
<p>MEDICAL IMAGE UNDERSTANDING WITH PRETRAINED VISION LANGUAGE MODELS: A COMPREHENSIVE STUDY</p> <p><i>Ziyuan Qin, Hua Hui Yi, Qicheng Lao, Kang Li</i></p>	16201
<p>Improving Deep Policy Gradients with Value Function Search</p> <p><i>Enrico Marchesini, Christopher Amato</i></p>	16221
<p>Self-Guided Noise-Free Data Generation for Efficient Zero-Shot Learning</p> <p><i>Jiahui Gao, Renjie Pi, LIN Yong, Hang Xu, Jiacheng Ye, Zhiyong Wu, WEIZHONG ZHANG, Xiaodan Liang, Zhenguo Li, Lingpeng Kong</i></p>	16239
<p>BSST: A Bayesian Spatial-Temporal Transformer for Sleep Staging</p> <p><i>Yuchen Liu, Ziyu Jia</i></p>	16262
<p>Learning Cut Selection for Mixed-Integer Linear Programming via Hierarchical Sequence Model</p> <p><i>Zhihai Wang, Xijun Li, Jie Wang, Yufei Kuang, Mingxuan Yuan, Jia Zeng, Yongdong Zhang, Feng Wu</i></p>	16274
<p>Graph Neural Networks are Inherently Good Generalizers: Insights by Bridging GNNs and MLPs</p> <p><i>Chenxiao Yang, Qitian Wu, Jiahua Wang, Junchi Yan</i></p>	16306
<p>Mitigating Memorization of Noisy Labels via Regularization between Representations</p> <p><i>Hao Cheng, Zhaowei Zhu, Xing Sun, Yang Liu</i></p>	16335
<p>Policy Expansion for Bridging Offline-to-Online Reinforcement Learning</p> <p><i>Haichao Zhang, Wei Xu, Haonan Yu</i></p>	16360
<p>DiffuSeq: Sequence to Sequence Text Generation with Diffusion Models</p> <p><i>Shansan Gong, Mukai Li, Jiangtao Feng, Zhiyong Wu, Lingpeng Kong</i></p>	16379
<p>Loss Landscapes are All You Need: Neural Network Generalization Can Be Explained Without the Implicit Bias of Gradient Descent</p> <p><i>Ping-yeh Chiang, Renkun Ni, David Yu Miller, Arpit Bansal, Jonas Geiping, Micah Goldblum, Tom Goldstein</i></p>	16399

<p>How to prepare your task head for finetuning</p> <p><i>Yi Ren, Shangmin Guo, Wonho Bae, Danica J. Sutherland</i></p>	16415
<p>Neural Lagrangian Schrödinger Bridge: Diffusion Modeling for Population Dynamics</p> <p><i>Takeshi Koshizuka, Issei Sato</i></p>	16443
<p>DIFFormer: Scalable (Graph) Transformers Induced by Energy Constrained Diffusion</p> <p><i>Qitian Wu, Chenxiao Yang, Wentao Zhao, Yixuan He, David Wipf, Junchi Yan</i></p>	16480
<p>DiGress: Discrete Denoising diffusion for graph generation</p> <p><i>Clement Vignac, Igor Krawczuk, Antoine Siraudin, Bohan Wang, Volkan Cevher, Pascal Frossard</i></p>	16506
<p>Unbiased Stochastic Proximal Solver for Graph Neural Networks with Equilibrium States</p> <p><i>Mingjie Li, Yifei Wang, Yisen Wang, Zhouchen Lin</i></p>	16528
<p>Turning the Curse of Heterogeneity in Federated Learning into a Blessing for Out-of-Distribution Detection</p> <p><i>Shuyang Yu, Junyuan Hong, Haotao Wang, Zhangyang Wang, Jiayu Zhou</i></p>	16541
<p>Human alignment of neural network representations</p> <p><i>Lukas Muttenthaler, Jonas Dippel, Lorenz Linhardt, Robert A. Vandermeulen, Simon Kornblith</i></p>	16553
<p>Sparsity-Constrained Optimal Transport</p> <p><i>Tianlin Liu, Joan Puigcerver, Mathieu Blondel</i></p>	16591
<p>Adversarial Training of Self-supervised Monocular Depth Estimation against Physical-World Attacks</p> <p><i>Zhiyuan Cheng, James Chenhao Liang, Guanhong Tao, Dongfang Liu, Xiangyu Zhang</i></p>	16619
<p>Prompting GPT-3 To Be Reliable</p> <p><i>Chenglei Si, Zhe Gan, Zhengyuan Yang, Shuohang Wang, Jianfeng Wang, Jordan Lee Boyd-Graber, Lijuan Wang</i></p>	16638
<p>Neural Agents Struggle to Take Turns in Bidirectional Emergent Communication</p> <p><i>Valentin Taillandier, Dieuwke Hupkes, Benoît Sagot, Emmanuel Dupoux, Paul Michel</i></p>	16662
<p>Teacher Guided Training: An Efficient Framework for Knowledge Transfer</p> <p><i>Manzil Zaheer, Ankit Singh Rawat, Seungyeon Kim, Chong You, Himanshu Jain, Andreas Veit, Rob Fergus, Sanjiv Kumar</i></p>	16678
<p>Disentangling Learning Representations with Density Estimation</p> <p><i>Eric Yeats, Frank Y Liu, Hai Li</i></p>	16707
<p>Exponential Generalization Bounds with Near-Optimal Rates for L_q-Stable Algorithms</p> <p><i>Xiaotong Yuan, Ping Li</i></p>	16722
<p>FaiREE: fair classification with finite-sample and distribution-free guarantee</p> <p><i>Puheng Li, James Zou, Linjun Zhang</i></p>	16748

<p>TiAda: A Time-scale Adaptive Algorithm for Nonconvex Minimax Optimization</p> <p>Xiang Li, Junchi YANG, Niao He</p>	16783
<p>Compositional Semantic Parsing with Large Language Models</p> <p>Andrew Drozdov, Nathanael Schärli, Ekin Akyürek, Nathan Scales, Xinying Song, Xinyun Chen, Olivier Bousquet, Denny Zhou</p>	16818
<p>Sparse MoE as the New Dropout: Scaling Dense and Self-Slimmable Transformers</p> <p>Tianlong Chen, Zhenyu Zhang, AJAY KUMAR JAISWAL, Shiwei Liu, Zhangyang Wang</p>	16878
<p>REVISITING PRUNING AT INITIALIZATION THROUGH THE LENS OF RAMANUJAN GRAPH</p> <p>Duc N.M Hoang, Shiwei Liu, Radu Marculescu, Zhangyang Wang</p>	16895
<p>Hard-Meta-Dataset++: Towards Understanding Few-Shot Performance on Difficult Tasks</p> <p>Samyadeep Basu, Megan Stanley, John F Bronskill, Soheil Feizi, Daniela Massiceti</p>	16909
<p>Ollivier-Ricci Curvature for Hypergraphs: A Unified Framework</p> <p>Corinna Coupette, Sebastian Dalleiger, Bastian Rieck</p>	16945
<p>Equivariant Hypergraph Diffusion Neural Operators</p> <p>Peihao Wang, Shenghao Yang, Yunyu Liu, Zhangyang Wang, Pan Li</p>	16983
<p>Liquid Structural State-Space Models</p> <p>Ramin Hasani, Mathias Lechner, Tsun-Hsuan Wang, Makram Chahine, Alexander Amini, Daniela Rus</p>	17007
<p>Draft, Sketch, and Prove: Guiding Formal Theorem Provers with Informal Proofs</p> <p>Albert Qiaochu Jiang, Sean Welleck, Jin Peng Zhou, Timothee Lacroix, Jiacheng Liu, Wenda Li, Mateja Jamnik, Guillaume Lample, Yuhuai Wu</p>	17026
<p>Scaleformer: Iterative Multi-scale Refining Transformers for Time Series Forecasting</p> <p>Mohammad Amin Shabani, Amir H. Abdi, Lili Meng, Tristan Sylvain</p>	17045
<p>A Non-Asymptotic Analysis of Oversmoothing in Graph Neural Networks</p> <p>Xinyi Wu, Zhengdao Chen, William Wei Wang, Ali Jadbabaie</p>	17068
<p>Which Layer is Learning Faster? A Systematic Exploration of Layer-wise Convergence Rate for Deep Neural Networks</p> <p>Yixiong Chen, Alan Yuille, Zongwei Zhou</p>	17093
<p>Sparsity May Cry: Let Us Fail (Current) Sparse Neural Networks Together!</p> <p>Shiwei Liu, Tianlong Chen, Zhenyu Zhang, Xuxi Chen, Tianjin Huang, AJAY KUMAR JAISWAL, Zhangyang Wang</p>	17112
<p>RLx2: Training a Sparse Deep Reinforcement Learning Model from Scratch</p> <p>Yiqin Tan, Pihe Hu, Ling Pan, Jiatai Huang, Longbo Huang</p>	17135
<p>Few-shot Cross-domain Image Generation via Inference-time Latent-code Learning</p> <p>Arnab Kumar Mondal, Piyush Tiwary, Parag Singla, Prathosh AP</p>	17167

<p>More ConvNets in the 2020s: Scaling up Kernels Beyond 51x51 using Sparsity</p> <p>Shiwei Liu, Tianlong Chen, Xiaohan Chen, Xuxi Chen, Qiao Xiao, Boqian Wu, Tommi Kärkkäinen, Mykola Pechenizkiy, Decebal Constantin Mocanu, Zhangyang Wang</p>	17180
<p>Generalization Bounds for Federated Learning: Fast Rates, Unparticipating Clients and Unbounded Losses</p> <p>Xiaolin Hu, Shaojie Li, Yong Liu</p>	17203
<p>Quasi-optimal Reinforcement Learning with Continuous Actions</p> <p>Yuhan Li, Wenzhuo Zhou, Ruoqing Zhu</p>	17244
<p>Energy-based Out-of-Distribution Detection for Graph Neural Networks</p> <p>Qitian Wu, Yiting Chen, Chenxiao Yang, Junchi Yan</p>	17290
<p>On the Data-Efficiency with Contrastive Image Transformation in Reinforcement Learning</p> <p>Sicong Liu, Xi Sheryl Zhang, Yushuo Li, Yifan Zhang, Jian Cheng</p>	17307
<p>The Influence of Learning Rule on Representation Dynamics in Wide Neural Networks</p> <p>Blake Bordelon, Cengiz Pehlevan</p>	17320
<p>Towards Minimax Optimal Reward-free Reinforcement Learning in Linear MDPs</p> <p>Pihe Hu, Yu Chen, Longbo Huang</p>	17353
<p>Hebbian and Gradient-based Plasticity Enables Robust Memory and Rapid Learning in RNNs</p> <p>Yu Duan, Zhongfan Jia, Qian Li, Yi Zhong, Kaisheng Ma</p>	17389
<p>AudioGen: Textually Guided Audio Generation</p> <p>Felix Kreuk, Gabriel Synnaeve, Adam Polyak, Uriel Singer, Alexandre D'Áfossez, Jade Copet, Devi Parikh, Yaniv Taigman, Yossi Adi</p>	17409
<p>A Unified Algebraic Perspective on Lipschitz Neural Networks</p> <p>Alexandre Araujo, Aaron J Havens, Blaise Delattre, Alexandre Allauzen, Bin Hu</p>	17425
<p>Efficient Model Updates for Approximate Unlearning of Graph-Structured Data</p> <p>Eli Chien, Chao Pan, Olgica Milenkovic</p>	17440
<p>Mitigating Dataset Bias by Using Per-Sample Gradient</p> <p>Sumyeong Ahn, Seongyeon Kim, Se-Young Yun</p>	17473
<p>STaSy: Score-based Tabular data Synthesis</p> <p>Jayoung Kim, Chaejeong Lee, Noseong Park</p>	17509
<p>LilNetX: Lightweight Networks with EXtreme Model Compression and Structured Sparsification</p> <p>Sharath Girish, Kamal Gupta, Saurabh Singh, Abhinav Shrivastava</p>	17536
<p>Pink Noise Is All You Need: Colored Noise Exploration in Deep Reinforcement Learning</p> <p>Onno Eberhard, Jakob Hollenstein, Cristina Pinneri, Georg Martius</p>	17557

<p>Compositional Law Parsing with Latent Random Functions</p> <p><i>Fan Shi, Bin Li, Xiangyang Xue</i></p>	17584
<p>PandA: Unsupervised Learning of Parts and Appearances in the Feature Maps of GANs</p> <p><i>James Oldfield, Christos Tzelepis, Yannis Panagakis, Mihalis Nicolaou, Ioannis Patras</i></p>	17619
<p>Learning to CROSS exchange to solve min-max vehicle routing problems</p> <p><i>Minjun Kim, Junyoung Park, Jinkyoo Park</i></p>	17648
<p>Learning Simultaneous Navigation and Construction in Grid Worlds</p> <p><i>Wenyu Han, Haoran Wu, Eisuke Hirota, Alexander Gao, Lerrel Pinto, Ludovic Righetti, Chen Feng</i></p>	17677
<p>Active Image Indexing</p> <p><i>Pierre Fernandez, Matthijs Douze, Herve Jegou, Teddy Furon</i></p>	17698
<p>Expressive Monotonic Neural Networks</p> <p><i>Niklas Nolte, Ouail Kitouni, Mike Williams</i></p>	17718
<p>Statistical Guarantees for Consensus Clustering</p> <p><i>Zhixin Zhou, Gautam Dudeja, Arash A Amini</i></p>	17730
<p>ZiCo: Zero-shot NAS via inverse Coefficient of Variation on Gradients</p> <p><i>Guihong Li, Yuedong Yang, Kartikeya Bhardwaj, Radu Marculescu</i></p>	17754
<p>Extremely Simple Activation Shaping for Out-of-Distribution Detection</p> <p><i>Andrija Djuric, Nebojsa Bozanic, Arjun Ashok, Rosanne Liu</i></p>	17785
<p>PGrad: Learning Principal Gradients For Domain Generalization</p> <p><i>Zhe Wang, Jake Grigsby, Yanjun Qi</i></p>	17807
<p>Part-Based Models Improve Adversarial Robustness</p> <p><i>Chawin Sitawarin, Kornrapat Pongmala, Yizheng Chen, Nicholas Carlini, David Wagner</i></p>	17828
<p>MOAT: Alternating Mobile Convolution and Attention Brings Strong Vision Models</p> <p><i>Chenglin Yang, Siyuan Qiao, Qihang Yu, Xiaoding Yuan, Yukun Zhu, Alan Yuille, Hartwig Adam, Liang-Chieh Chen</i></p>	17857
<p>NTFields: Neural Time Fields for Physics-Informed Robot Motion Planning</p> <p><i>Ruiqi Ni, Ahmed H Qureshi</i></p>	17880
<p>SGDA with shuffling: faster convergence for nonconvex-P&A minimax optimization</p> <p><i>Hanseul Cho, Chulhee Yun</i></p>	17897
<p>Matching receptor to odorant with protein language and graph neural networks</p> <p><i>Matej Hladik, Maxence Lalis, Sebastien Fiorucci, J&A@r&A@mie Topin</i></p>	17943
<p>Static Prediction of Runtime Errors by Learning to Execute Programs with External Resource Descriptions</p> <p><i>David Bieber, Rishab Goel, Dan Zheng, Hugo Larochelle, Daniel Tarlow</i></p>	17968

<p>Boosting Adversarial Transferability using Dynamic Cues</p> <p>Muzammal Naseer, Ahmad Mahmood, Salman Khan, Fahad Khan</p>	17991
<p>ChordMixer: A Scalable Neural Attention Model for Sequences with Different Length</p> <p>Ruslan Khalitov, Tong Yu, Lei Cheng, Zhirong Yang</p>	18014
<p>Generalization and Estimation Error Bounds for Model-based Neural Networks</p> <p>Avner Shultzman, Eyar Azar, Miguel R. D. Rodrigues, Yonina C. Eldar</p>	18036
<p>The Dark Side of AutoML: Towards Architectural Backdoor Search</p> <p>Ren Pang, Changjiang Li, Zhaohan Xi, Shouling Ji, Ting Wang</p>	18045
<p>Stochastic No-regret Learning for General Games with Variance Reduction</p> <p>Yichi Zhou, Fang Kong, Shuai Li</p>	18061
<p>Is Attention All That NeRF Needs?</p> <p>Mukund Varma T, Peihao Wang, Xuxi Chen, Tianlong Chen, Subhashini Venugopalan, Zhangyang Wang</p>	18078
<p>Red PANDA: Disambiguating Image Anomaly Detection by Removing Nuisance Factors</p> <p>Niv Cohen, Jonathan Kahana, Yedid Hoshen</p>	18100
<p>Calibrating Transformers via Sparse Gaussian Processes</p> <p>Wenlong Chen, Yingzhen Li</p>	18120
<p>Perfectly Secure Steganography Using Minimum Entropy Coupling</p> <p>Christian Schroeder de Witt, Samuel Sokota, J Zico Kolter, Jakob Nicolaus Foerster, Martin Strohmeier</p>	18149
<p>Scaling Laws for a Multi-Agent Reinforcement Learning Model</p> <p>Oren Neumann, Claudius Gros</p>	18163
<p>Pitfalls of Gaussians as a noise distribution in NCE</p> <p>Holden Lee, Chirag Pabbaraju, Anish Prasad Sevekari, Andrej Risteski</p>	18182
<p>A Unified Approach to Reinforcement Learning, Quantal Response Equilibria, and Two-Player Zero-Sum Games</p> <p>Samuel Sokota, Ryan D'Orazio, J Zico Kolter, Nicolas Loizou, Marc Lanctot, Ioannis Mitliagkas, Noam Brown, Christian Kroer</p>	18195
<p>Incompatibility Clustering as a Defense Against Backdoor Poisoning Attacks</p> <p>Charles Jin, Melinda Sun, Martin Rinard</p>	18236
<p>Temporal Domain Generalization with Drift-Aware Dynamic Neural Networks</p> <p>Guangji Bai, Chen Ling, Liang Zhao</p>	18272
<p>Human-level Atari 200x faster</p> <p>Steven Kapturowski, VÁctor Campos, Ray Jiang, Nemanja Rakicevic, Hado van Hasselt, Charles Blundell, Adria Puigdomenech Badia</p>	18291

<p>Differentially Private \mathcal{L}_2-Heavy Hitters in the Sliding Window Model</p> <p>Jeremiah Blocki, Seunghoon Lee, Tamalika Mukherjee, Samson Zhou</p>	18324
<p>Efficient Edge Inference by Selective Query</p> <p>Anil Kag, Igor Fedorov, Aditya Gangrade, Paul Whatmough, Venkatesh Saligrama</p>	18337
<p>Noise Injection Node Regularization for Robust Learning</p> <p>Noam Itzhak Levi, Itay Mimouni Bloch, Marat Freytsis, Tomer Volansky</p>	18362
<p>Diffusion Adversarial Representation Learning for Self-supervised Vessel Segmentation</p> <p>Boah Kim, Yujin Oh, Jong Chul Ye</p>	18382
<p>Discovering Latent Knowledge in Language Models Without Supervision</p> <p>Collin Burns, Haotian Ye, Dan Klein, Jacob Steinhardt</p>	18404
<p>Layer Grafted Pre-training: Bridging Contrastive Learning And Masked Image Modeling For Label-Efficient Representations</p> <p>Ziyu Jiang, Yinpeng Chen, Mengchen Liu, Dongdong Chen, Xiyang Dai, Lu Yuan, Zicheng Liu, Zhangyang Wang</p>	18432
<p>Out-of-Distribution Detection and Selective Generation for Conditional Language Models</p> <p>Jie Ren, Jiaming Luo, Yao Zhao, Kundan Krishna, Mohammad Saleh, Balaji Lakshminarayanan, Peter J Liu</p>	18448
<p>Interpretable Debiasing of Vectorized Language Representations with Iterative Orthogonalization</p> <p>Prince Osei Aboagye, Yan Zheng, Jack Shunn, Chin-Chia Michael Yeh, Junpeng Wang, Zhongfang Zhuang, Huiyuan Chen, Liang Wang, Wei Zhang, Jeff Phillips</p>	18480
<p>How to Train your HIPPO: State Space Models with Generalized Orthogonal Basis Projections</p> <p>Albert Gu, Isys Johnson, Aman Timalsina, Atri Rudra, Christopher Re</p>	18513
<p>Continuous-time identification of dynamic state-space models by deep subspace encoding</p> <p>Gerben I. Beintema, Maarten Schoukens, Roland TÅ³th</p>	18548
<p>Text Summarization with Oracle Expectation</p> <p>Yumo Xu, Mirella Lapata</p>	18563
<p>MaskViT: Masked Visual Pre-Training for Video Prediction</p> <p>Agrim Gupta, Stephen Tian, Yunzhi Zhang, Jiajun Wu, Roberto MartÅn-MartÅn, Li Fei-Fei</p>	18583
<p>Measuring Forgetting of Memorized Training Examples</p> <p>Matthew Jagielski, Om Thakkar, Florian Tramer, Daphne Ippolito, Katherine Lee, Nicholas Carlini, Eric Wallace, Shuang Song, Abhradeep Guha Thakurta, Nicolas Papernot, Chiyuan Zhang</p>	18607
<p>Learning to Segment from Noisy Annotations: A Spatial Correction Approach</p> <p>Jiachen Yao, Yikai Zhang, Songzhu Zheng, Mayank Goswami, Prateek Prasanna, Chao Chen</p>	18629

<p>Powderworld: A Platform for Understanding Generalization via Rich Task Distributions</p> <p>Kevin Frans, Phillip Isola</p>	18649
<p>Quantifying Memorization Across Neural Language Models</p> <p>Nicholas Carlini, Daphne Ippolito, Matthew Jagielski, Katherine Lee, Florian Tramèr, Chiyuan Zhang</p>	18670
<p>A Non-monotonic Self-terminating Language Model</p> <p>Eugene Choi, Kyunghyun Cho, Cheolhyoung Lee</p>	18689
<p>Dr.Spider: A Diagnostic Evaluation Benchmark towards Text-to-SQL Robustness</p> <p>Shuaichen Chang, Jun Wang, Mingwen Dong, Lin Pan, Henghui Zhu, Alexander Hanbo Li, Wuwei Lan, Sheng Zhang, Jiarong Jiang, Joseph Lillien, Steve Ash, William Yang Wang, Zhiguo Wang, Vittorio Castelli, Patrick Ng, Bing Xiang</p>	18712
<p>A critical look at the evaluation of GNNs under heterophily: Are we really making progress?</p> <p>Oleg Platonov, Denis Kuznedelev, Michael Diskin, Artem Babenko, Liudmila Prokhorenkova</p>	18738
<p>Unmasking the Lottery Ticket Hypothesis: What's Encoded in a Winning Ticket's Mask?</p> <p>Mansheej Paul, Feng Chen, Brett W. Larsen, Jonathan Frankle, Surya Ganguli, Gintare Karolina Dziugaite</p>	18753
<p>Characterizing the spectrum of the NTK via a power series expansion</p> <p>Michael Murray, Hui Jin, Benjamin Bowman, Guido Montufar</p>	18781
<p>Temporal Dependencies in Feature Importance for Time Series Prediction</p> <p>Kin Kwan Leung, Clayton Rooke, Jonathan Smith, Saba Zuberi, Maksims Volkovs</p>	18839
<p>AutoTransfer: AutoML with Knowledge Transfer - An Application to Graph Neural Networks</p> <p>Kaidi Cao, Jiaxuan You, Jiaju Liu, Jure Leskovec</p>	18857
<p>Efficient Discrete Multi Marginal Optimal Transport Regularization</p> <p>Ronak Mehta, Jeffery Kline, Vishnu Suresh Lokhande, Glenn Fung, Vikas Singh</p>	18873
<p>Neural Design for Genetic Perturbation Experiments</p> <p>Aldo Pacchiano, Drausin Wulsin, Robert A Barton, Luis Voloch</p>	18896
<p>Anamnesic Neural Differential Equations with Orthogonal Polynomial Projections</p> <p>Edward De Brouwer, Rahul G Krishnan</p>	18937
<p>Tier Balancing: Towards Dynamic Fairness over Underlying Causal Factors</p> <p>Zeyu Tang, Yatong Chen, Yang Liu, Kun Zhang</p>	18948
<p>Sample-Efficient Reinforcement Learning by Breaking the Replay Ratio Barrier</p> <p>Pierluca D'Oro, Max Schwarzer, Evgenii Nikishin, Pierre-Luc Bacon, Marc G Bellemare, Aaron Courville</p>	18995
<p>SMART: Sentences as Basic Units for Text Evaluation</p> <p>Reinald Kim Amplayo, Peter J Liu, Yao Zhao, Shashi Narayan</p>	19018

<p>Exploring and Exploiting Decision Boundary Dynamics for Adversarial Robustness</p> <p><i>Yuancheng Xu, Yanchao Sun, Micah Goldblum, Tom Goldstein, Furong Huang</i></p>	19038
<p>Malign Overfitting: Interpolation and Invariance are Fundamentally at Odds</p> <p><i>Yoav Wald, Gal Yona, Uri Shalit, Yair Carmon</i></p>	19065
<p>Does Learning from Decentralized Non-IID Unlabeled Data Benefit from Self Supervision?</p> <p><i>Lirui Wang, Kaiqing Zhang, Yunzhu Li, Yonglong Tian, Russ Tedrake</i></p>	19080
<p>MEDFAIR: Benchmarking Fairness for Medical Imaging</p> <p><i>Yongshuo Zong, Yongxin Yang, Timothy Hospedales</i></p>	19094
<p>Learning with Stochastic Orders</p> <p><i>Carles Domingo-Enrich, Yair Schiff, Youssef Mroueh</i></p>	19132
<p>Offline Congestion Games: How Feedback Type Affects Data Coverage Requirement</p> <p><i>Haozhe Jiang, Qiwen Cui, Zhihan Xiong, Maryam Fazel, Simon Shaolei Du</i></p>	19161
<p>Is Adversarial Training Really a Silver Bullet for Mitigating Data Poisoning?</p> <p><i>Rui Wen, Zhengyu Zhao, Zhuoran Liu, Michael Backes, Tianhao Wang, Yang Zhang</i></p>	19185
<p>FedFA: Federated Feature Augmentation</p> <p><i>Tianfei Zhou, Ender Konukoglu</i></p>	19202
<p>Learning Hyper Label Model for Programmatic Weak Supervision</p> <p><i>Renzhi Wu, Shen-En Chen, Jieyu Zhang, Xu Chu</i></p>	19223
<p>Parametrizing Product Shape Manifolds by Composite Networks</p> <p><i>Josua Sassen, Klaus Hildebrandt, Martin Rumpf, Benedikt Wirth</i></p>	19244
<p>Agnostic Learning of General ReLU Activation Using Gradient Descent</p> <p><i>Pranjal Awasthi, Alex Tang, Aravindan Vijayaraghavan</i></p>	19262
<p>The Surprising Computational Power of Nondeterministic Stack RNNs</p> <p><i>Brian DuSell, David Chiang</i></p>	19273
<p>Causal Imitation Learning via Inverse Reinforcement Learning</p> <p><i>Kangrui Ruan, Junzhe Zhang, Xuan Di, Elias Bareinboim</i></p>	19294
<p>Achieving Sub-linear Regret in Infinite Horizon Average Reward Constrained MDP with Linear Function Approximation</p> <p><i>Arnob Ghosh, Xingyu Zhou, Ness Shroff</i></p>	19306
<p>PaLI: A Jointly-Scaled Multilingual Language-Image Model</p> <p><i>Xi Chen, Xiao Wang, Soravit Changpinyo, AJ Piergiovanni, Piotr Padlewski, Daniel Salz, Sebastian Goodman, Adam Grycner, Basil Mustafa, Lucas Beyer, Alexander Kolesnikov, Joan Puigcerver, Nan Ding, Keran Rong, Hassan Akbari, Gaurav Mishra, Linting Xue, Ashish V Thapliyal, James Bradbury, Weicheng Kuo, Mojtaba Seyedhosseini, Chao Jia, Burcu Karagol Ayan, Carlos Riquelme Ruiz, Andreas Peter Steiner, Anelia Angelova, Xiaohua Zhai, Neil Houlsby, Radu Soricut</i></p>	19353

<p>BigVGAN: A Universal Neural Vocoder with Large-Scale Training</p> <p>Sang-gil Lee, Wei Ping, Boris Ginsburg, Bryan Catanzaro, Sungroh Yoon</p>	19386
<p>Where to Begin? On the Impact of Pre-Training and Initialization in Federated Learning</p> <p>John Nguyen, Jianyu Wang, Kshitiz Malik, Maziar Sanjabi, Michael Rabbat</p>	19406
<p>Learning Multimodal Data Augmentation in Feature Space</p> <p>Zichang Liu, Zhiqiang Tang, Xingjian Shi, Aston Zhang, Mu Li, Anshumali Shrivastava, Andrew Gordon Wilson</p>	19418
<p>Tuning Frequency Bias in Neural Network Training with Nonuniform Data</p> <p>Annan Yu, Yunan Yang, Alex Townsend</p>	19434
<p>Variational Latent Branching Model for Off-Policy Evaluation</p> <p>Qitong Gao, Ge Gao, Min Chi, Miroslav Pajic</p>	19471
<p>PerFedMask: Personalized Federated Learning with Optimized Masking Vectors</p> <p>Mehdi Setayesh, Xiaoxiao Li, Vincent W.S. Wong</p>	19502
<p>Minimum Description Length Control</p> <p>Ted Moskovitz, Ta-Chu Kao, Maneesh Sahani, Matthew Botvinick</p>	19529
<p>MetaGL: Evaluation-Free Selection of Graph Learning Models via Meta-Learning</p> <p>Namyong Park, Ryan A. Rossi, Nesreen Ahmed, Christos Faloutsos</p>	19560
<p>Koopman Neural Operator Forecaster for Time-series with Temporal Distributional Shifts</p> <p>Rui Wang, Yihe Dong, Sercan O Arik, Rose Yu</p>	19587
<p>Restricted Strong Convexity of Deep Learning Models with Smooth Activations</p> <p>Arindam Banerjee, Pedro Cisneros-Velarde, Libin Zhu, Misha Belkin</p>	19602
<p>Revisiting the Assumption of Latent Separability for Backdoor Defenses</p> <p>Xiangyu Qi, Tinghao Xie, Yiming Li, Saeed Mahloujifar, Prateek Mittal</p>	19637
<p>Open-Vocabulary Object Detection upon Frozen Vision and Language Models</p> <p>Weicheng Kuo, Yin Cui, Xiuye Gu, AJ Piergiovanni, Anelia Angelova</p>	19657
<p>Learning Controllable Adaptive Simulation for Multi-resolution Physics</p> <p>Tailin Wu, Takashi Maruyama, Qingqing Zhao, Gordon Wetzstein, Jure Leskovec</p>	19677
<p>Hyperbolic Deep Reinforcement Learning</p> <p>Edoardo Cedin, Benjamin Paul Chamberlain, Michael M. Bronstein, Jonathan J Hunt</p>	19696
<p>TILP: Differentiable Learning of Temporal Logical Rules on Knowledge Graphs</p> <p>Siheng Xiong, Yuan Yang, Faramarz Fekri, James Clayton Kerce</p>	19726
<p>Self-Stabilization: The Implicit Bias of Gradient Descent at the Edge of Stability</p> <p>Alex Damian, Eshaan Nichani, Jason D. Lee</p>	19741

<p>Does Zero-Shot Reinforcement Learning Exist?</p> <p>Ahmed Touati, JÃ©rÃ©my Rapin, Yann Ollivier</p>	19781
<p>Sampling with Mollified Interaction Energy Descent</p> <p>Lingxiao Li, qiang liu, Anna Korba, Mikhail Yurochkin, Justin Solomon</p>	19819
<p>An Extensible Multi-modal Multi-task Object Dataset with Materials</p> <p>Trevor Scott Standley, Ruohan Gao, Dawn Chen, Jiajun Wu, Silvio Savarese</p>	19851
<p>Neural Networks Efficiently Learn Low-Dimensional Representations with SGD</p> <p>Alireza Mousavi-Hosseini, Sejun Park, Manuela Girotti, Ioannis Mitliagkas, Murat A Erdogdu</p>	19869
<p>Constraining Representations Yields Models That Know What They Don't Know</p> <p>Joao Monteiro, Pau Rodriguez, Pierre-Andre Noel, Issam H. Laradji, David Vazquez</p>	19906
<p>Concept Gradient: Concept-based Interpretation Without Linear Assumption</p> <p>Andrew Bai, Chih-Kuan Yeh, Neil Y.C. Lin, Pradeep Kumar Ravikumar, Cho-Jui Hsieh</p>	19931
<p>Parameter-Efficient Fine-Tuning Design Spaces</p> <p>Jiaao Chen, Aston Zhang, Xingjian Shi, Mu Li, Alex Smola, Diyi Yang</p>	19954
<p>Is a Caption Worth a Thousand Images? A Study on Representation Learning</p> <p>Shibani Santurkar, Yann Dubois, Rohan Taori, Percy Liang, Tatsunori Hashimoto</p>	19972
<p>When Source-Free Domain Adaptation Meets Learning with Noisy Labels</p> <p>Li Yi, Gezheng Xu, Pengcheng Xu, Jiaqi Li, Ruizhi Pu, Charles Ling, Ian McLeod, Boyu Wang</p>	19986
<p>Post-hoc Concept Bottleneck Models</p> <p>Mert Yuksekogul, Maggie Wang, James Zou</p>	20019
<p>SimPer: Simple Self-Supervised Learning of Periodic Targets</p> <p>Yuzhe Yang, Xin Liu, Jiang Wu, Silviu Borac, Dina Katabi, Ming-Zher Poh, Daniel McDuff</p>	20039
<p>Effects of Graph Convolutions in Multi-layer Networks</p> <p>Aseem Baranwal, Kimon Fountoulakis, Aukosh Jagannath</p>	20062
<p>A CMDP-within-online framework for Meta-Safe Reinforcement Learning</p> <p>Vanshaj Khattar, Yuhao Ding, Bilgehan Sel, Javad Lavaei, Ming Jin</p>	20093
<p>Guiding Safe Exploration with Weakest Preconditions</p> <p>Greg Anderson, Swarat Chaudhuri, Isil Dillig</p>	20142
<p>Moving Forward by Moving Backward: Embedding Action Impact over Action Semantics</p> <p>Kuo-Hao Zeng, Luca Weihs, Roozbeh Mottaghi, Ali Farhadi</p>	20162
<p>Domain Generalization via Heckman-type Selection Models</p> <p>Hyungu Kahng, Hyungrok Do, Judy Zhong</p>	20183

<p>Partially Observable RL with B-Stability: Unified Structural Condition and Sharp Sample-Efficient Algorithms</p> <p>Fan Chen, Yu Bai, Song Mei</p>	20200
<p>Weighted Ensemble Self-Supervised Learning</p> <p>Yangjun Ruan, Saurabh Singh, Warren Richard Morningstar, Alexander A Alemi, Sergey Ioffe, Ian Fischer, Joshua V. Dillon</p>	20247
<p>simpleKT: A Simple But Tough-to-Beat Baseline for Knowledge Tracing</p> <p>Zitao Liu, Qionggiong Liu, Jiahao Chen, Shuyan Huang, Weiqi Luo</p>	20276
<p>Partial Label Unsupervised Domain Adaptation with Class-Prototype Alignment</p> <p>Yan Yan, Yuhong Guo</p>	20288
<p>Mutual Partial Label Learning with Competitive Label Noise</p> <p>Yan Yan, Yuhong Guo</p>	20302
<p>Faster federated optimization under second-order similarity</p> <p>Ahmed Khaled, Chi Jin</p>	20316
<p>MeshDiffusion: Score-based Generative 3D Mesh Modeling</p> <p>Zhen Liu, Yao Feng, Michael J. Black, Derek Nowrouzezahrai, Liam Paull, Weiyang Liu</p>	20353
<p>DCI-ES: An Extended Disentanglement Framework with Connections to Identifiability</p> <p>Cian Eastwood, Andrei Liviu Nicolicioiu, Julius Von K&Auml;gelgen, Armin Keki&Auml;f, Frederik Tr&Auml;uble, Andrea Dittadi, Bernhard Sch&Auml;kopf</p>	20379
<p>Simplified State Space Layers for Sequence Modeling</p> <p>Jimmy T.H. Smith, Andrew Warrington, Scott Linderman</p>	20395
<p>Weakly Supervised Explainable Phrasal Reasoning with Neural Fuzzy Logic</p> <p>Zijun Wu, Zi Xuan Zhang, Atharva Naik, Zhijian Mei, Mauajama Firdaus, Lili Mou</p>	20430
<p>Scalable and Equivariant Spherical CNNs by Discrete-Continuous (DISCO) Convolutions</p> <p>Jeremy Ocampo, Matthew Alexander Price, Jason McEwen</p>	20449
<p>Diffusion Models for Causal Discovery via Topological Ordering</p> <p>Pedro Sanchez, Xiao Liu, Alison Q O'Neil, Sotirios A. Tsafaris</p>	20468
<p>Deconstructing Distributions: A Pointwise Framework of Learning</p> <p>Gal Kaplun, Nikhil Ghosh, Saurabh Garg, Boaz Barak, Preetum Nakkiran</p>	20488
<p>Graph Neural Network-Inspired Kernels for Gaussian Processes in Semi-Supervised Learning</p> <p>Zehao Niu, Mihai Anitescu, Jie Chen</p>	20508
<p>Meta Temporal Point Processes</p> <p>Wonho Bae, Mohamed Osama Ahmed, Frederick Tung, Gabriel L. Oliveira</p>	20527

<p>Efficiently Computing Nash Equilibria in Adversarial Team Markov Games</p> <p><i>Fivos Kalogiannis, Ioannis Anagnostides, Ioannis Panageas, Emmanouil-Vasileios Vataakis-Gkaragkounis, Vaggos Chatziafratis, Stelios Andrew Stavroulakis</i></p>	20538
<p>Neural DAG Scheduling via One-Shot Priority Sampling</p> <p><i>Wonseok Jeon, Mukul Gagrani, Burak Bartan, Weiliang Will Zeng, Harris Teague, Piero Zappi, Christopher Lott</i></p>	20553
<p>Direct Embedding of Temporal Network Edges via Time-Decayed Line Graphs</p> <p><i>Sudhanshu Chanpuriya, Ryan A. Rossi, Sungchul Kim, Tong Yu, Jane Hoffswell, Nedim Lipka, Shunan Guo, Cameron N Musco</i></p>	20574
<p>Learning Sparse Group Models Through Boolean Relaxation</p> <p><i>Yijie Wang, Yuan Zhou, Xiaoqing Huang, Kun Huang, Jie Zhang, Jianzhu Ma</i></p>	20587
<p>Planning Goals for Exploration</p> <p><i>Edward S. Hu, Richard Chang, Oleh Rybkin, Dinesh Jayaraman</i></p>	20613
<p>Provably Auditing Ordinary Least Squares in Low Dimensions</p> <p><i>Ankur Moitra, Dhruv Rohatgi</i></p>	20635
<p>Contrastive Learning Can Find An Optimal Basis For Approximately View-Invariant Functions</p> <p><i>Daniel D. Johnson, Ayoub El Hanchi, Chris J. Maddison</i></p>	20668
<p>A Theoretical Understanding of Shallow Vision Transformers: Learning, Generalization, and Sample Complexity</p> <p><i>Hongkang Li, Meng Wang, Sijia Liu, Pin-Yu Chen</i></p>	20707
<p>PiFold: Toward effective and efficient protein inverse folding</p> <p><i>Zhangyang Gao, Cheng Tan, Stan Z. Li</i></p>	20752
<p>Progress measures for grokking via mechanistic interpretability</p> <p><i>Neel Nanda, Lawrence Chan, Tom Lieberum, Jess Smith, Jacob Steinhardt</i></p>	20766
<p>Distributed Differential Privacy in Multi-Armed Bandits</p> <p><i>Sayak Ray Chowdhury, Xingyu Zhou</i></p>	20801
<p>Link Prediction with Non-Contrastive Learning</p> <p><i>William Shiao, Zhichun Guo, Tong Zhao, Evangelos E. Papalexakis, Yozen Liu, Neil Shah</i></p>	20835
<p>Provably Efficient Lifelong Reinforcement Learning with Linear Representation</p> <p><i>Sanae Amani, Lin Yang, Ching-An Cheng</i></p>	20854
<p>Re-Imagen: Retrieval-Augmented Text-to-Image Generator</p> <p><i>Wenhu Chen, Hexiang Hu, Chitwan Saharia, William W. Cohen</i></p>	20896
<p>Improved Sample Complexity for Reward-free Reinforcement Learning under Low-rank MDPs</p> <p><i>Yuan Cheng, Ruiquan Huang, Yingbin Liang, Jing Yang</i></p>	20921

<p>Sequential Attention for Feature Selection</p> <p><i>Taisuke Yasuda, Mohammadhossein Bateni, Lin Chen, Matthew Fahrbach, Gang Fu, Vahab Mirrokni</i></p>	20953
<p>Disentangling the Mechanisms Behind Implicit Regularization in SGD</p> <p><i>Zachary Novack, Simran Kaur, Tanya Marwah, Saurabh Garg, Zachary Chase Lipton</i></p>	20974
<p>Certiably Robust Policy Learning against Adversarial Multi-Agent Communication</p> <p><i>Yanchao Sun, Ruijie Zheng, Parisa Hassanzadeh, Yongyuan Liang, Soheil Feizi, Sumitra Ganesh, Furong Huang</i></p>	20989
<p>SoftMatch: Addressing the Quantity-Quality Tradeoff in Semi-supervised Learning</p> <p><i>Hao Chen, Ran Tao, Yue Fan, Yidong Wang, Jindong Wang, Bernt Schiele, Xing Xie, Bhiksha Raj, Marios Savvides</i></p>	21019
<p>On Accelerated Perceptrons and Beyond</p> <p><i>Guanghui Wang, Rafael Hanashiro, Etash Kumar Guha, Jacob Abernethy</i></p>	21040
<p>Spectral Decomposition Representation for Reinforcement Learning</p> <p><i>Tongzheng Ren, Tianjun Zhang, Lisa Lee, Joseph E. Gonzalez, Dale Schuurmans, Bo Dai</i></p>	21052
<p>Confidence Estimation Using Unlabeled Data</p> <p><i>Chen Li, Xiaoling Hu, Chao Chen</i></p>	21084
<p>FreeMatch: Self-adaptive Thresholding for Semi-supervised Learning</p> <p><i>Yidong Wang, Hao Chen, Qiang Heng, Wenxin Hou, Yue Fan, Zhen Wu, Jindong Wang, Marios Savvides, Takahiro Shinozaki, Bhiksha Raj, Bernt Schiele, Xing Xie</i></p>	21104
<p>Words are all you need? Language as an approximation for human similarity judgments</p> <p><i>Raja Marjeh, Pol Van Rijn, Ilya Sucholutsky, Theodore Summers, Harin Lee, Thomas L. Griffiths, Nori Jacoby</i></p>	21124
<p>ROCO: A General Framework for Evaluating Robustness of Combinatorial Optimization Solvers on Graphs</p> <p><i>Han Lu, Zenan Li, Runzhong Wang, Qibing Ren, Xijun Li, Mingxuan Yuan, Jia Zeng, Xiaokang Yang, Junchi Yan</i></p>	21152
<p>Latent Variable Representation for Reinforcement Learning</p> <p><i>Tongzheng Ren, Chenjun Xiao, Tianjun Zhang, Na Li, Zhaoran Wang, sujay sanghavi, Dale Schuurmans, Bo Dai</i></p>	21175
<p>Federated Nearest Neighbor Machine Translation</p> <p><i>Yichao Du, Zhirui Zhang, Bingzhe Wu, Lemao Liu, Tong Xu, Enhong Chen</i></p>	21215
<p>Extreme Q-Learning: MaxEnt RL without Entropy</p> <p><i>Divyansh Garg, Joey Hejna, Matthieu Geist, Stefano Ermon</i></p>	21235
<p>Moderate Coreset: A Universal Method of Data Selection for Real-world Data-efficient Deep Learning</p> <p><i>Xiaobo Xia, Jiale Liu, Jun Yu, Xu Shen, Bo Han, Tongliang Liu</i></p>	21260

<p>Learning on Large-scale Text-attributed Graphs via Variational Inference</p> <p><i>Jianan Zhao, Meng Qu, Chaozhuo Li, Hao Yan, Qian Liu, Rui Li, Xing Xie, Jian Tang</i></p>	21280
<p>SAM as an Optimal Relaxation of Bayes</p> <p><i>Thomas Mllenhoff, Mohammad Emtiyaz Khan</i></p>	21293
<p>A System for Morphology-Task Generalization via Unified Representation and Behavior Distillation</p> <p><i>Hiroki Furuta, Yusuke Iwasawa, Yutaka Matsuo, Shixiang Shane Gu</i></p>	21318
<p>Machine Unlearning of Federated Clusters</p> <p><i>Chao Pan, Jin Sima, Saurav Prakash, Vishal Rana, Olgica Milenkovic</i></p>	21352
<p>An efficient encoder-decoder architecture with top-down attention for speech separation</p> <p><i>Kai Li, Runxuan Yang, Xiaolin Hu</i></p>	21379
<p>Automatic Chain of Thought Prompting in Large Language Models</p> <p><i>Zhuosheng Zhang, Aston Zhang, Mu Li, Alex Smola</i></p>	21395
<p>LipsFormer: Introducing Lipschitz Continuity to Vision Transformers</p> <p><i>Xianbiao Qi, Jianan Wang, Yihao Chen, Yukai Shi, Lei Zhang</i></p>	21427
<p>DocPrompting: Generating Code by Retrieving the Docs</p> <p><i>Shuyan Zhou, Uri Alon, Frank F. Xu, Zhengbao Jiang, Graham Neubig</i></p>	21451
<p>Graph Signal Sampling for Inductive One-Bit Matrix Completion: a Closed-form Solution</p> <p><i>Chao Chen, Haoyu Geng, Gang Zeng, Zhaobing Han, Hua Chai, Xiaokang Yang, Junchi Yan</i></p>	21470
<p>Learning Kernelized Contextual Bandits in a Distributed and Asynchronous Environment</p> <p><i>Chuanhao Li, Huazheng Wang, Mengdi Wang, Hongning Wang</i></p>	21495
<p>Towards Better Selective Classification</p> <p><i>Leo Feng, Mohamed Osama Ahmed, Hossein Hajimirsadeghi, Amir H. Abdi</i></p>	21520
<p>Represent to Control Partially Observed Systems: Representation Learning with Provable Sample Efficiency</p> <p><i>Lingxiao Wang, Qi Cai, Zhuoran Yang, Zhaoran Wang</i></p>	21545
<p>Latent Bottlenecked Attentive Neural Processes</p> <p><i>Leo Feng, Hossein Hajimirsadeghi, Yoshua Bengio, Mohamed Osama Ahmed</i></p>	21557
<p>Language Models Can Teach Themselves to Program Better</p> <p><i>Patrick Haluptzok, Matthew Bowers, Adam Tauman Kalai</i></p>	21574
<p>Fairness and Accuracy under Domain Generalization</p> <p><i>Thai-Hoang Pham, Xueru Zhang, Ping Zhang</i></p>	21597

<p>Optimal Conservative Offline RL with General Function Approximation via Augmented Lagrangian</p> <p><i>Paria Rashidinejad, Hanlin Zhu, Kunhe Yang, Stuart Russell, Jiantao Jiao</i></p>	21638
<p>An Adaptive Policy to Employ Sharpness-Aware Minimization</p> <p><i>Weisen Jiang, Hansi Yang, Yu Zhang, James Kwok</i></p>	21681
<p>Efficient Attention via Control Variates</p> <p><i>Lin Zheng, Jianbo Yuan, Chong Wang, Lingpeng Kong</i></p>	21699
<p>On the Soft-Subnetwork for Few-Shot Class Incremental Learning</p> <p><i>Haeyong Kang, Jaehong Yoon, Sultan Rizky Hikmawan Madjid, Sung Ju Hwang, Chang D. Yoo</i></p>	21735
<p>CLIPSep: Learning Text-queried Sound Separation with Noisy Unlabeled Videos</p> <p><i>Hao-Wen Dong, Naoya Takahashi, Yuki Mitsufuji, Julian McAuley, Taylor Berg-Kirkpatrick</i></p>	21758
<p>Offline RL for Natural Language Generation with Implicit Language Q Learning</p> <p><i>Charlie Victor Snell, Ilya Kostrikov, Yi Su, Sherry Yang, Sergey Levine</i></p>	21780
<p>Soft Neighbors are Positive Supporters in Contrastive Visual Representation Learning</p> <p><i>Chongjian GE, Jiangliu Wang, Zhan Tong, Shoufa Chen, Yibing Song, Ping Luo</i></p>	21811
<p>Explaining Temporal Graph Models through an Explorer-Navigator Framework</p> <p><i>Wenwen Xia, Mincai Lai, Caihua Shan, Yao Zhang, Xinnan Dai, Xiang Li, Dongsheng Li</i></p>	21827
<p>Equivariant Descriptor Fields: SE(3)-Equivariant Energy-Based Models for End-to-End Visual Robotic Manipulation Learning</p> <p><i>Hyunwoo Ryu, Hong-in Lee, Jeong-Hoon Lee, Jongeun Choi</i></p>	21838
<p>Interpretability in the Wild: a Circuit for Indirect Object Identification in GPT-2 Small</p> <p><i>Kevin Ro Wang, Alexandre Variengien, Arthur Conmy, Buck Shlegeris, Jacob Steinhardt</i></p>	21870
<p>3D UX-Net: A Large Kernel Volumetric ConvNet Modernizing Hierarchical Transformer for Medical Image Segmentation</p> <p><i>Ho Hin Lee, Shunxing Bao, Yuankai Huo, Bennett A. Landman</i></p>	21891
<p>Adversarial Diversity in Hanabi</p> <p><i>Brandon Cui, Andrei Lupu, Samuel Sokota, Hengyuan Hu, David J Wu, Jakob Nicolaus Foerster</i></p>	21906
<p>Re-calibrating Feature Attributions for Model Interpretation</p> <p><i>Peiyu Yang, NAVEED AKHTAR, Zeyi Wen, Mubarak Shah, Ajmal Saeed Mian</i></p>	21922
<p>ROSCOE: A Suite of Metrics for Scoring Step-by-Step Reasoning</p> <p><i>Olga Golovneva, Moya Peng Chen, Spencer Poff, Martin Corredor, Luke Zettlemoyer, Maryam Fazel-Zarandi, Asli Celikyilmaz</i></p>	21946
<p>M-L2O: Towards Generalizable Learning-to-Optimize by Test-Time Fast Self-Adaptation</p> <p><i>Junjie Yang, Xuxi Chen, Tianlong Chen, Zhangyang Wang, Yingbin Liang</i></p>	22003

<p>CodeGen: An Open Large Language Model for Code with Multi-Turn Program Synthesis</p> <p>Erik Nijkamp, Bo Pang, Hiroaki Hayashi, Lifu Tu, Huan Wang, Yingbo Zhou, Silvio Savarese, Caiming Xiong</p>	22025
<p>Understanding the Role of Nonlinearity in Training Dynamics of Contrastive Learning</p> <p>Yuandong Tian</p>	22050
<p>PD-MORL: Preference-Driven Multi-Objective Reinforcement Learning Algorithm</p> <p>Toygun Basaklar, Suat Gumussoy, Umit Ogras</p>	22089
<p>MoDem: Accelerating Visual Model-Based Reinforcement Learning with Demonstrations</p> <p>Nicklas Hansen, Yixin Lin, Hao Su, Xiaolong Wang, Vikash Kumar, Aravind Rajeswaran</p>	22113
<p>Data Continuity Matters: Improving Sequence Modeling with Lipschitz Regularizer</p> <p>Eric Qu, Xufang Luo, Dongsheng Li</p>	22133
<p>Causal Estimation for Text Data with (Apparent) Overlap Violations</p> <p>Lin Gui, Victor Veitch</p>	22156
<p>GLM-130B: An Open Bilingual Pre-trained Model</p> <p>Aohan Zeng, Xiao Liu, Zhengxiao Du, Zihan Wang, Hanyu Lai, Ming Ding, Zhuoyi Yang, Yifan Xu, Wendi Zheng, Xiao Xia, Weng Lam Tam, Zixuan Ma, Yufei Xue, Jidong Zhai, Wenguang Chen, Zhiyuan Liu, Peng Zhang, Yuxiao Dong, Jie Tang</p>	22176
<p>CLIP-Dissect: Automatic Description of Neuron Representations in Deep Vision Networks</p> <p>Tuomas Oikarinen, Tsui-Wei Weng</p>	22232
<p>Multi-level Protein Structure Pre-training via Prompt Learning</p> <p>Zeyuan Wang, Qiang Zhang, Shuang-Wei HU, Haoran Yu, Xurui Jin, Zhichen Gong, Huajun Chen</p>	22264
<p>Label-free Concept Bottleneck Models</p> <p>Tuomas Oikarinen, Subhro Das, Lam M. Nguyen, Tsui-Wei Weng</p>	22283
<p>Learning to Extrapolate: A Transductive Approach</p> <p>Aviv Netanyahu, Abhishek Gupta, Max Simchowitz, Kaiqing Zhang, Pulkit Agrawal</p>	22311
<p>Spacetime Representation Learning</p> <p>Marc T. Law, James Lucas</p>	22345
<p>Memorization-Dilation: Modeling Neural Collapse Under Noise</p> <p>Duc Anh Nguyen, Ron Levie, Julian Lienen, Eyke HÅllermeier, Gitta Kutyniok</p>	22373
<p>Calibration Matters: Tackling Maximization Bias in Large-scale Advertising Recommendation Systems</p> <p>Yewen Fan, Nian Si, Kun Zhang</p>	22385
<p>Understanding Embodied Reference with Touch-Line Transformer</p> <p>Yang Li, Xiaoxue Chen, Hao Zhao, Jiangtao Gong, Guyue Zhou, Federico Rossano, Yixin Zhu</p>	22412

<p>Mastering the Game of No-Press Diplomacy via Human-Regularized Reinforcement Learning and Planning</p> <p>Anton Bakhtin, David J Wu, Adam Lerer, Jonathan Gray, Athul Paul Jacob, Gabriele Farina, Alexander H Miller, Noam Brown</p>	22428
<p>PAC-NeRF: Physics Augmented Continuum Neural Radiance Fields for Geometry-Agnostic System Identification</p> <p>Xuan Li, Yi-Ling Qiao, Peter Yichen Chen, Krishna Murthy Jatavallabhula, Ming Lin, Chenfanfu Jiang, Chuang Gan</p>	22457
<p>Scaling Forward Gradient With Local Losses</p> <p>Mengye Ren, Simon Kornblith, Renjie Liao, Geoffrey Hinton</p>	22471
<p>Free Lunch for Domain Adversarial Training: Environment Label Smoothing</p> <p>YiFan Zhang, xue wang, Jian Liang, Zhang Zhang, Liang Wang, Rong Jin, Tieniu Tan</p>	22502
<p>On Compositional Uncertainty Quantification for Seq2seq Graph Parsing</p> <p>Zi Lin, Du Phan, Panupong Pasupat, Jeremiah Zhe Liu, Jingbo Shang</p>	22540
<p>Imbalanced Semi-supervised Learning with Bias Adaptive Classifier</p> <p>Renzhen Wang, Xixi Jia, Quanzhang Wang, Yichen Wu, Deyu Meng</p>	22561
<p>Equivariant Shape-Conditioned Generation of 3D Molecules for Ligand-Based Drug Design</p> <p>Keir Adams, Connor W. Coley</p>	22583
<p>Any-scale Balanced Samplers for Discrete Space</p> <p>Haoran Sun, Bo Dai, Charles Sutton, Dale Schuurmans, Hanjun Dai</p>	22629
<p>Edgeformers: Graph-Empowered Transformers for Representation Learning on Textual-Edge Networks</p> <p>Bowen Jin, Yu Zhang, Yu Meng, Jiawei Han</p>	22646
<p>More Centralized Training, Still Decentralized Execution: Multi-Agent Conditional Policy Factorization</p> <p>Jiangxing Wang, Deheng Ye, Zongqing Lu</p>	22664
<p>Targeted Hyperparameter Optimization with Lexicographic Preferences Over Multiple Objectives</p> <p>Shaokun Zhang, Feiran Jia, Chi Wang, Qingyun Wu</p>	22682
<p>Flow Matching for Generative Modeling</p> <p>Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, Matthew Le</p>	22698
<p>Preference Transformer: Modeling Human Preferences using Transformers for RL</p> <p>Changyeon Kim, Jongjin Park, Jinwoo Shin, Honglak Lee, Pieter Abbeel, Kimin Lee</p>	22726

<p>Winning Both the Accuracy of Floating Point Activation and the Simplicity of Integer Arithmetic</p> <p>Yulhwa Kim, Jaeyong Jang, Jehun Lee, Jihoon Park, Jeonghoon Kim, Byeongwook Kim, Baeseong park, Se Jung Kwon, Dongsoo Lee, jae-joon kim</p>	22745
<p>Task Ambiguity in Humans and Language Models</p> <p>Alex Tamkin, Kunal Handa, Avash Shrestha, Noah Goodman</p>	22763
<p>Long-Tailed Partial Label Learning via Dynamic Rebalancing</p> <p>Feng Hong, Jiangchao Yao, Zhihan Zhou, Ya Zhang, Yanfeng Wang</p>	22795
<p>FluidLab: A Differentiable Environment for Benchmarking Complex Fluid Manipulation</p> <p>Zhou Xian, Bo Zhu, Zhenjia Xu, Hsiao-Yu Tung, Antonio Torralba, Katerina Fragkiadaki, Chuang Gan</p>	22819
<p>Mosaic Representation Learning for Self-supervised Visual Pre-training</p> <p>Zhaoqing Wang, Ziyu Chen, Yaqian Li, Yandong Guo, Jun Yu, Mingming Gong, Tongliang Liu</p>	22838
<p>The hidden uniform cluster prior in self-supervised learning</p> <p>Mido Assran, Randall Balestriero, Quentin Duval, Florian Bordes, Ishan Misra, Piotr Bojanowski, Pascal Vincent, Michael Rabbat, Nicolas Ballas</p>	22854
<p>An Equal-Size Hard EM Algorithm for Diverse Dialogue Generation</p> <p>Yujiao Wen, Yongchang Hao, Yanshuai Cao, Lili Mou</p>	22880
<p>Near-Optimal Deployment Efficiency in Reward-Free Reinforcement Learning with Linear Function Approximation</p> <p>Dan Qiao, Yu-Xiang Wang</p>	22898
<p>Asymptotic Instance-Optimal Algorithms for Interactive Decision Making</p> <p>Kefan Dong, Tengyu Ma</p>	22935
<p>Max-Margin Works while Large Margin Fails: Generalization without Uniform Convergence</p> <p>Margalit Glasgow, Colin Wei, Mary Wootters, Tengyu Ma</p>	22947
<p>Pushing the Accuracy-Group Robustness Frontier with Introspective Self-play</p> <p>Jeremiah Zhe Liu, Krishnamurthy Dj Dvijotham, Jihyeon Lee, Quan Yuan, Balaji Lakshminarayanan, Deepak Ramachandran</p>	22959
<p>Squeeze Training for Adversarial Robustness</p> <p>Qizhang Li, Yiwen Guo, Wangmeng Zuo, Hao Chen</p>	23001
<p>DeBERTaV3: Improving DeBERTa using ELECTRA-Style Pre-Training with Gradient-Disentangled Embedding Sharing</p> <p>Pengcheng He, Jianfeng Gao, Weizhu Chen</p>	23017
<p>On The Relative Error of Random Fourier Features for Preserving Kernel Distance</p> <p>Kuan Cheng, Shaofeng H.-C. Jiang, Luojuan Wei, Zhide Wei</p>	23033
<p>Population-size-Aware Policy Optimization for Mean-Field Games</p> <p>Pengdeng Li, Xinrun Wang, Shuxin Li, Hau Chan, Bo An</p>	23056

<p>What Do Self-Supervised Vision Transformers Learn?</p> <p><i>Namuk Park, Wonjae Kim, Byeongho Heo, Taekyung Kim, Sangdoon Yun</i></p>	23088
<p>Mind's Eye: Grounded Language Model Reasoning through Simulation</p> <p><i>Ruibo Liu, Jason Wei, Shixiang Shane Gu, Te-Yen Wu, Soroush Vosoughi, Claire Cui, Denny Zhou, Andrew M. Dai</i></p>	23107
<p>Knowledge-in-Context: Towards Knowledgeable Semi-Parametric Language Models</p> <p><i>Xiaoman Pan, Wenlin Yao, Hongming Zhang, Dian Yu, Dong Yu, Jianshu Chen</i></p>	23125
<p>Budgeted Training for Vision Transformer</p> <p><i>zhuofan xia, Xuran Pan, Xuan Jin, Yuan He, Hui Xue', Shiji Song, Gao Huang</i></p>	23154
<p>When Data Geometry Meets Deep Function: Generalizing Offline Reinforcement Learning</p> <p><i>Jianxiong Li, Xianyuan Zhan, Haoran Xu, Xiangyu Zhu, Jingjing Liu, Ya-Qin Zhang</i></p>	23170
<p>Incremental Learning of Structured Memory via Closed-Loop Transcription</p> <p><i>Shengbang Tong, Xili Dai, Ziyang Wu, Mingyang Li, Brent Yi, Yi Ma</i></p>	23204
<p>Minimum Variance Unbiased N:M Sparsity for the Neural Gradients</p> <p><i>Brian Chmiel, Itay Hubara, Ron Banner, Daniel Soudry</i></p>	23224
<p>Symmetric Pruning in Quantum Neural Networks</p> <p><i>Xinbiao Wang, Junyu Liu, Tongliang Liu, Yong Luo, Yuxuan Du, Dacheng Tao</i></p>	23236
<p>Solving Constrained Variational Inequalities via a First-order Interior Point-based Method</p> <p><i>Tong Yang, Michael Jordan, Tatjana Chavdarova</i></p>	23261
<p>Visually-Augmented Language Modeling</p> <p><i>Weizhi Wang, Li Dong, Hao Cheng, Haoyu Song, Xiaodong Liu, Xifeng Yan, Jianfeng Gao, Furu Wei</i></p>	23315
<p>Unsupervised 3D Object Learning through Neuron Activity aware Plasticity</p> <p><i>Beomseok Kang, Biswadeep Chakraborty, Saibal Mukhopadhyay</i></p>	23332
<p>Not All Tasks Are Born Equal: Understanding Zero-Shot Generalization</p> <p><i>Jing Zhou, Zongyu Lin, Yanan Zheng, Jian Li, Zhilin Yang</i></p>	23356
<p>Complexity-Based Prompting for Multi-step Reasoning</p> <p><i>Yao Fu, Hao Peng, Ashish Sabharwal, Peter Clark, Tushar Khot</i></p>	23377
<p>Building a Subspace of Policies for Scalable Continual Learning</p> <p><i>Jean-Baptiste Gaya, Thang Doan, Lucas Caccia, Laure Soulier, Ludovic Denoyer, Roberta Raileanu</i></p>	23392
<p>Decompositional Generation Process for Instance-Dependent Partial Label Learning</p> <p><i>Congyu Qiao, Ning Xu, Xin Geng</i></p>	23420
<p>Indiscriminate Poisoning Attacks on Unsupervised Contrastive Learning</p> <p><i>Hao He, Kaiwen Zha, Dina Katabi</i></p>	23436

Unsupervised Semantic Segmentation with Self-supervised Object-centric Representations <i>Andrii Zadaianchuk, Mattheaus Kleindessner, Yi Zhu, Francesco Locatello, Thomas Brox</i>	23452
Valid P-Value for Deep Learning-driven Salient Region <i>Miwa Daiki, Vo Nguyen Le Duy, Ichiro Takeuchi</i>	23475
Quantized Compressed Sensing with Score-Based Generative Models <i>Xiangming Meng, Yoshiyuki Kabashima</i>	23487
Near-optimal Coresets for Robust Clustering <i>Lingxiao Huang, Shaofeng H.-C. Jiang, Jianing Lou, Xuan Wu</i>	23517
CUTS: Neural Causal Discovery from Irregular Time-Series Data <i>Yuxiao Cheng, Runzhao Yang, Tingxiong Xiao, Zongren Li, Jinli Suo, Kunlun He, Qionghai Dai</i>	23538
Learning Diffusion Bridges on Constrained Domains <i>Xingchao Liu, Lemeng Wu, Mao Ye, qiang liu</i>	23562
Data augmentation alone can improve adversarial training <i>Lin Li, Michael W. Spratling</i>	23591
Language Models are Realistic Tabular Data Generators <i>Vadim Borisov, Kathrin Sessler, Tobias Leemann, Martin Pawelczyk, Gjergji Kasneci</i>	23614
Tensor-Based Sketching Method for the Low-Rank Approximation of Data Streams. <i>Cuiyu Liu, Xiao Chuanfu, Mingshuo Ding, Chao Yang</i>	23632
Inequality phenomenon in L_{∞}-adversarial training, and its unrealized threats <i>Ranjie Duan, YueFeng Chen, Yao Zhu, Xiaojun Jia, Rong Zhang, Hui Xue'</i>	23645
Flow Straight and Fast: Learning to Generate and Transfer Data with Rectified Flow <i>Xingchao Liu, Chengyue Gong, qiang liu</i>	23666
Better Teacher Better Student: Dynamic Prior Knowledge for Knowledge Distillation <i>Martin Zong, Zengyu Qiu, Xinzhu Ma, Kunlin Yang, Chunya Liu, Jun Hou, Shuai Yi, Wanli Ouyang</i>	23699
Quality-Similar Diversity via Population Based Reinforcement Learning <i>Shuang Wu, Jian Yao, Haobo Fu, Ye Tian, Chao Qian, Yaodong Yang, QIANG FU, Yang Wei</i>	23721
Efficient recurrent architectures through activity sparsity and sparse back-propagation through time <i>Anand Subramoney, Khaleelulla Khan Nazeer, Mark SchÃ¶ne, Christian Mayr, David Kappel</i>	23747
Discovering Generalizable Multi-agent Coordination Skills from Multi-task Offline Data <i>Fuxiang Zhang, Chengxing Jia, Yi-Chen Li, Lei Yuan, Yang Yu, Zongzhang Zhang</i>	23761
Learning Vortex Dynamics for Fluid Inference and Prediction <i>Yitong Deng, Hong-Xing Yu, Jiajun Wu, Bo Zhu</i>	23785

<p>λ-DARTS: Mitigating Performance Collapse by Harmonizing Operation Selection among Cells</p> <p>Sajad Movahedi, Melika Adabinejad, Ayyoob Imani, Arezou Keshavarz, Mostafa Dehghani, Azadeh Shakery, Babak N Araabi</p>	23807
<p>Bayesian Oracle for bounding information gain in neural encoding models</p> <p>Konstantin-Klemens Lurz, Mohammad Bashiri, Edgar Y. Walker, Fabian H. Sinz</p>	23834
<p>NAGphormer: A Tokenized Graph Transformer for Node Classification in Large Graphs</p> <p>Jinsong Chen, Kaiyuan Gao, Gaichao Li, Kun He</p>	23845
<p>Scaling Pareto-Efficient Decision Making via Offline Multi-Objective RL</p> <p>Baiting Zhu, Meihua Dang, Aditya Grover</p>	23863
<p>Out-of-Distribution Detection based on In-Distribution Data Patterns Memorization with Modern Hopfield Energy</p> <p>Jinsong Zhang, Qiang Fu, Xu Chen, Lun Du, Zelin Li, Gang Wang, xiaoguang Liu, Shi Han, Dongmei Zhang</p>	23881
<p>Policy-Based Self-Competition for Planning Problems</p> <p>Jonathan Pirnay, Quirin GÄttl, Jakob Burger, Dominik Gerhard Grimm</p>	23900
<p>Leveraging Unlabeled Data to Track Memorization</p> <p>Mahsa Forouzesh, Hanie Sedghi, Patrick Thiran</p>	23924
<p>Neural-based classification rule learning for sequential data</p> <p>Marine Collery, Philippe Bonnard, François Fages, Remy Kusters</p>	23974
<p>Almost Linear Constant-Factor Sketching for ℓ_1 and Logistic Regression</p> <p>Alexander Munteanu, Simon Omlor, David Woodruff</p>	23989
<p>Capturing the Motion of Every Joint: 3D Human Pose and Shape Estimation with Independent Tokens</p> <p>Sen Yang, Wen Heng, Gang Liu, GUOZHONG LUO, Wankou Yang, Gang YU</p>	24024
<p>MAESTRO: Open-Ended Environment Design for Multi-Agent Reinforcement Learning</p> <p>Mikayel Samvelyan, Akbir Khan, Michael D Dennis, Minqi Jiang, Jack Parker-Holder, Jakob Nicolaus Foerster, Roberta Raileanu, Tim Rocktäschel</p>	24041
<p>Gradient Gating for Deep Multi-Rate Learning on Graphs</p> <p>T. Konstantin Rusch, Benjamin Paul Chamberlain, Michael W. Mahoney, Michael M. Bronstein, Siddhartha Mishra</p>	24067
<p>Global Explainability of GNNs via Logic Combination of Learned Concepts</p> <p>Steve Azzolin, Antonio Longa, Pietro Barbiero, Pietro Lio, Andrea Passerini</p>	24088
<p>Wasserstein Auto-encoded MDPs: Formal Verification of Efficiently Distilled RL Policies with Many-sided Guarantees</p> <p>Florent Delgrange, Ann Nowe, Guillermo Perez</p>	24107

Lower Bounds on the Depth of Integral ReLU Neural Networks via Lattice Polytopes <i>Christian Alexander Haase, Christoph Hertrich, Georg Loho</i>	24137
Graph-based Deterministic Policy Gradient for Repetitive Combinatorial Optimization Problems <i>Zhongyuan Zhao, Ananthram Swami, Santiago Segarra</i>	24150
MAST: Masked Augmentation Subspace Training for Generalizable Self-Supervised Priors <i>Chen Huang, Hanlin Goh, Jiatao Gu, Joshua M. Susskind</i>	24171
The Implicit Bias of Minima Stability in Multivariate Shallow ReLU Networks <i>Mor Shpigel Nacson, Rotem Mulayoff, Greg Ongie, Tomer Michaeli, Daniel Soudry</i>	24189
How Sharpness-Aware Minimization Minimizes Sharpness? <i>Kaiyue Wen, Tengyu Ma, Zhiyuan Li</i>	24226
Active Learning for Object Detection with Evidential Deep Learning and Hierarchical Uncertainty Aggregation <i>Younghyun Park, Wonjeong Choi, Soyeong Kim, Dong-Jun Han, Jaekyun Moon</i>	24302
Learning ReLU networks to high uniform accuracy is intractable <i>Julius Berner, Philipp Grohs, Felix Voigtlaender</i>	24322
A law of adversarial risk, interpolation, and label noise <i>Daniel Paleka, Amartya Sanyal</i>	24344
Fisher-Legendre (FishLeg) optimization of deep neural networks <i>Jezabel R Garcia, Federica Freddi, Stathi Fotiadis, Maolin Li, Sattar Vakili, Alberto Bernacchia, Guillaume Hennequin</i>	24366
Understanding the Generalization of Adam in Learning Neural Networks with Proper Regularization <i>Difan Zou, Yuan Cao, Yuanzhi Li, Quanquan Gu</i>	24388
SLTUNET: A Simple Unified Model for Sign Language Translation <i>Biao Zhang, Mathias Müller, Rico Sennrich</i>	24427
User-Interactive Offline Reinforcement Learning <i>Phillip Swazinna, Steffen Udluft, Thomas Runkler</i>	24445
Scaffolding a Student to Instill Knowledge <i>Anil Kag, Durmus Alp Emre Acar, Aditya Gangrade, Venkatesh Saligrama</i>	24463
Learnable Graph Convolutional Attention Networks <i>Adrián Javaloy, Pablo Sanchez Martin, Amit Levi, Isabel Valera</i>	24480
FastFill: Efficient Compatible Model Update <i>Florian Jaeckle, Fartash Faghri, Ali Farhadi, Oncel Tuzel, Hadi Pouransari</i>	24515

<p>Explaining RL Decisions with Trajectories</p> <p><i>Shripad Vilasrao Deshmukh, Arpan Dasgupta, Balaji Krishnamurthy, Nan Jiang, Chirag Agarwal, Georgios Theodorou, Jayakumar Subramanian</i></p>	24539
<p>A Primal-Dual Framework for Transformers and Neural Networks</p> <p><i>Tan Minh Nguyen, Tam Minh Nguyen, Nhat Ho, Andrea L. Bertozzi, Richard Baraniuk, Stanley Osher</i></p>	24559
<p>Online Low Rank Matrix Completion</p> <p><i>Soumyabrata Pal, Prateek Jain</i></p>	24585
<p>Contrastive Learning for Unsupervised Domain Adaptation of Time Series</p> <p><i>Yilmazcan Ozyurt, Stefan Feuerriegel, Ce Zhang</i></p>	24597
<p>Anti-Symmetric DGN: a stable architecture for Deep Graph Networks</p> <p><i>Alessio Gravina, Davide Bacciu, Claudio Gallicchio</i></p>	24637
<p>Statistical Theory of Differentially Private Marginal-based Data Synthesis Algorithms</p> <p><i>Ximing Li, Chendi Wang, Guang Cheng</i></p>	24656
<p>Consolidator: Mergable Adapter with Group Connections for Visual Adaptation</p> <p><i>Tianxiang Hao, Hui Chen, Yuchen Guo, Guiguang Ding</i></p>	24668
<p>Confidential-PROFIT: Confidential PROOF of Fair Training of Trees</p> <p><i>Ali Shahin Shamsabadi, Sierra Calanda Wyllie, Nicholas Franzese, Natalie Dullerud, Sbastien Gams, Nicolas Papernot, Xiao Wang, Adrian Weller</i></p>	24685
<p>Domain Generalisation via Domain Adaptation: An Adversarial Fourier Amplitude Approach</p> <p><i>Minyoung Kim, Da Li, Timothy Hospedales</i></p>	24708
<p>Dual Algorithmic Reasoning</p> <p><i>Daniilo Numeroso, Davide Bacciu, Petar Velickovi</i></p>	24729
<p>TDR-CL: Targeted Doubly Robust Collaborative Learning for Debaised Recommendations</p> <p><i>Haoxuan Li, Yan Lyu, Chunyuan Zheng, Peng Wu</i></p>	24745
<p>DAVA: Disentangling Adversarial Variational Autoencoder</p> <p><i>Benjamin Estermann, Roger Wattenhofer</i></p>	24760
<p>Learning Input-agnostic Manipulation Directions in StyleGAN with Text Guidance</p> <p><i>Yoonjeon Kim, Hyunsu Kim, Junho Kim, Yunjey Choi, Eunho Yang</i></p>	24785
<p>\mathscr{N}-WL: A New Hierarchy of Expressivity for Graph Neural Networks</p> <p><i>Qing Wang, Dillon Ze Chen, Asiri Wijesinghe, Shouheng Li, Muhammad Farhan</i></p>	24814
<p>Sampling-free Inference for Ab-Initio Potential Energy Surface Networks</p> <p><i>Nicholas Gao, Stephan Gnnemann</i></p>	24854
<p>Value Memory Graph: A Graph-Structured World Model for Offline Reinforcement Learning</p> <p><i>Deyao Zhu, Li Erran Li, Mohamed Elhoseiny</i></p>	24876

<p>Optimizing Spca-based Continual Learning: A Theoretical Approach</p> <p>Chunchun Yang, Malik Tiomoko, Zengfu Wang</p>	24900
<p>Revisit Finetuning strategy for Few-Shot Learning to Transfer the Emdeddings</p> <p>Heng Wang, Tan Yue, Xiang Ye, Zihang He, Bohan Li, Yong Li</p>	24921
<p>CogVideo: Large-scale Pretraining for Text-to-Video Generation via Transformers</p> <p>Wenyi Hong, Ming Ding, Wendi Zheng, Xinghan Liu, Jie Tang</p>	24932
<p>Offline Reinforcement Learning via High-Fidelity Generative Behavior Modeling</p> <p>Huayu Chen, Cheng Lu, Chengyang Ying, Hang Su, Jun Zhu</p>	24956
<p>Robust Explanation Constraints for Neural Networks</p> <p>Matthew Robert Wicker, Juyeon Heo, Luca Costabello, Adrian Weller</p>	24977
<p>Training language models to summarize narratives improves brain alignment</p> <p>Khai Loong Aw, Mariya Toneva</p>	24990
<p>Theoretical Characterization of the Generalization Performance of Overfitted Meta-Learning</p> <p>Peizhong Ju, Yingbin Liang, Ness Shroff</p>	25027
<p>Defending against Adversarial Audio via Diffusion Model</p> <p>Shutong Wu, Jiong Xiao Wang, Wei Ping, Weili Nie, Chaowei Xiao</p>	25079
<p>Sampling-based inference for large linear models, with application to linearised Laplace</p> <p>Javier Antoran, Shreyas Padhy, Riccardo Barbano, Eric Nalisnick, David Janz, Jos�� Miguel Hern��ndez-Lobato</p>	25100
<p>StableDR: Stabilized Doubly Robust Learning for Recommendation on Data Missing Not at Random</p> <p>Haoxuan Li, Chunyuan Zheng, Peng Wu</p>	25133
<p>PASHA: Efficient HPO and NAS with Progressive Resource Allocation</p> <p>Ondrej Bohdal, Lukas Balles, Martin Wistuba, Beyza Ermis, Cedric Archambeau, Giovanni Zappella</p>	25151
<p>Holistic Adversarially Robust Pruning</p> <p>Qi Zhao, Christian Wressnegger</p>	25171
<p>That Label's got Style: Handling Label Style Bias for Uncertain Image Segmentation</p> <p>Kilian Zepf, Eike Petersen, Jes Frellsen, Aasa Feragen</p>	25193
<p>DeCap: Decoding CLIP Latents for Zero-Shot Captioning via Text-Only Training</p> <p>Wei Li, Linchao Zhu, Longyin Wen, Yi Yang</p>	25212
<p>Timing is Everything: Learning to Act Selectively with Costly Actions and Budgetary Constraints</p> <p>David Henry Mguni, Aivar Sootla, Juliusz Krzysztof Ziomek, Oliver Slumbers, Zipeng Dai, Kun Shao, Jun Wang</p>	25230

<p>Guess the Instruction! Flipped Learning Makes Language Models Stronger Zero-Shot Learners</p> <p>Seonghyeon Ye, Doyoung Kim, Joel Jang, Joongbo Shin, Minjoon Seo</p>	25241
<p>Compressing multidimensional weather and climate data into neural networks</p> <p>Langwen Huang, Torsten Hoefler</p>	25260
<p>Learning Fast and Slow for Online Time Series Forecasting</p> <p>Quang Pham, Chenghao Liu, Doyen Sahoo, Steven Hoi</p>	25277
<p>Understanding Neural Coding on Latent Manifolds by Sharing Features and Dividing Ensembles</p> <p>Martin Bjerke, Lukas Schott, Kristopher T Jensen, Claudia Battistin, David A. Klindt, Benjamin Adric Dunn</p>	25299
<p>Structure by Architecture: Structured Representations without Regularization</p> <p>Felix Leeb, Giulia Lanzillotta, Yashas Annadani, Michel Besserve, Stefan Bauer, Bernhard Schölkopf</p>	25325
<p>On Explaining Neural Network Robustness with Activation Path</p> <p>Ziping Jiang</p>	25360
<p>A GNN-Guided Predict-and-Search Framework for Mixed-Integer Linear Programming</p> <p>Qingyu Han, Linxin Yang, Qian Chen, Xiang Zhou, Dong Zhang, Akang Wang, Ruoyu Sun, Xiaodong Luo</p>	25380
<p>LexMAE: Lexicon-Bottlenecked Pretraining for Large-Scale Retrieval</p> <p>Tao Shen, Xiubo Geng, Chongyang Tao, Can Xu, Xiaolong Huang, Binxing Jiao, Linjun Yang, Daxin Jiang</p>	25395
<p>Short-Term Memory Convolutions</p> <p>Grzegorz Stefański, Krzysztof Arendt, Paweł Daniłuk, Bartłomiej Jasik, Artur Szumaczuł</p>	25414
<p>Fundamental Limits in Formal Verification of Message-Passing Neural Networks</p> <p>Marco Sälzler, Martin Lange</p>	25428
<p>Boosting the Cycle Counting Power of Graph Neural Networks with \mathbb{R}^2-GNNs</p> <p>Yinan Huang, Xingang Peng, Jianzhu Ma, Muhan Zhang</p>	25446
<p>Improving the imputation of missing data with Markov Blanket discovery</p> <p>Yang Liu, Anthony Constantinou</p>	25473
<p>Learning to Solve Constraint Satisfaction Problems with Recurrent Transformer</p> <p>Zhun Yang, Adam Ishay, Joohyung Lee</p>	25487
<p>Model-based Causal Bayesian Optimization</p> <p>Scott Sussex, Anastasia Makarova, Andreas Krause</p>	25509
<p>TANGOS: Regularizing Tabular Neural Networks through Gradient Orthogonalization and Specialization</p> <p>Alan Jeffares, Tennison Liu, Jonathan Crabb, Fergus Imrie, Mihaela van der Schaar</p>	25533

<p>DynaMS: Dyanmic Margin Selection for Efficient Deep Learning</p> <p><i>Jiaxing Wang, Yong Li, Jingwei Zhuo, Xupeng Shi, WEIZHONG ZHANG, Lixing Gong, Tong Tao, Pengzhang Liu, Yongjun Bao, Weipeng Yan</i></p>	25560
<p>DiffusER: Diffusion via Edit-based Reconstruction</p> <p><i>Machel Reid, Vincent Josua Hellendoorn, Graham Neubig</i></p>	25580
<p>Multi-objective optimization via equivariant deep hypervolume approximation</p> <p><i>Jim Boelrijk, Bernd Ensing, Patrick Forr</i></p>	25592
<p>Compositional Prompt Tuning with Motion Cues for Open-vocabulary Video Relation Detection</p> <p><i>Kaifeng Gao, Long Chen, Hanwang Zhang, Jun Xiao, Qianru Sun</i></p>	25613
<p>Unbiased Supervised Contrastive Learning</p> <p><i>Carlo Alberto Barbano, Benoit Dufumier, Enzo Tartaglione, Marco Grangetto, Pietro Gori</i></p>	25629
<p>The Curious Case of Benign Memorization</p> <p><i>Sotiris Anagnostidis, Gregor Bachmann, Lorenzo Noci, Thomas Hofmann</i></p>	25653
<p>Learning Probabilistic Topological Representations Using Discrete Morse Theory</p> <p><i>Xiaoling Hu, Dimitris Samaras, Chao Chen</i></p>	25675
<p>Learning topology-preserving data representations</p> <p><i>Ilya Trofimov, Daniil Cherniavskii, Eduard Tulchinskii, Nikita Balabin, Evgeny Burnaev, Serguei Barannikov</i></p>	25694
<p>Sequential Learning of Neural Networks for Prequential MDL</p> <p><i>Jorg Bornschein, Yazhe Li, Marcus Hutter</i></p>	25724
<p>Representational Dissimilarity Metric Spaces for Stochastic Neural Networks</p> <p><i>Lyndon Duong, Jingyang Zhou, Josue Nassar, Jules Berman, Jeroen Olieslagers, Alex H Williams</i></p>	25745
<p>Localized Randomized Smoothing for Collective Robustness Certification</p> <p><i>Jan Schuchardt, Tom Wollschlaeger, Aleksandar Bojchevski, Stephan Günemann</i></p>	25780
<p>ExpressivE: A Spatio-Functional Embedding For Knowledge Graph Completion</p> <p><i>Aleksandar Pavlović, Emanuel Sallinger</i></p>	25827
<p>Why adversarial training can hurt robust accuracy</p> <p><i>Jacob Clarysse, Julia Härrmann, Fanny Yang</i></p>	25872
<p>Copy is All You Need</p> <p><i>Tian Lan, Deng Cai, Yan Wang, Heyan Huang, Xian-Ling Mao</i></p>	25901
<p>Leveraging Importance Weights in Subset Selection</p> <p><i>Gui Citovsky, Giulia DeSalvo, Sanjiv Kumar, Sri Kumar Ramalingam, Afshin Rostamizadeh, Yunjuan Wang</i></p>	25917

<p>Correlative Information Maximization Based Biologically Plausible Neural Networks for Correlated Source Separation</p> <p><i>Bariscan Bozkurt, AteÅŸ ÅŸsferdiyaroÅŸlu, Cengiz Pehlevan, Alper Tunga Erdogan</i></p>	25935
<p>Efficient Planning in a Compact Latent Action Space</p> <p><i>zhengyao jiang, Tianjun Zhang, Michael Janner, Yueying Li, Tim RocktÄšschel, Edward Grefenstette, Yuandong Tian</i></p>	25969
<p>Enhancing the Inductive Biases of Graph Neural ODE for Modeling Physical Systems</p> <p><i>Suresh Bishnoi, Ravinder Bhattoo, Jayadeva Jayadeva, Sayan Ranu, N M Anoop Krishnan</i></p>	25992
<p>Contrastive Meta-Learning for Partially Observable Few-Shot Learning</p> <p><i>Adam Jelley, Amos Storkey, Antreas Antoniou, Sam Devlin</i></p>	26020
<p>Selection-Inference: Exploiting Large Language Models for Interpretable Logical Reasoning</p> <p><i>Antonia Creswell, Murray Shanahan, Irina Higgins</i></p>	26038
<p>Imitating Human Behaviour with Diffusion Models</p> <p><i>Tim Pearce, Tabish Rashid, Anssi Kanervisto, Dave Bignell, Mingfei Sun, Raluca Georgescu, Sergio Valcarcel Macua, Shan Zheng Tan, Ida Momennejad, Katja Hofmann, Sam Devlin</i></p>	26074
<p>Understanding weight-magnitude hyperparameters in training binary networks</p> <p><i>Joris Quist, Yunqiang Li, Jan van Gemert</i></p>	26098
<p>This Looks Like It Rather Than That: ProtoKNN For Similarity-Based Classifiers</p> <p><i>Yuki Ukai, Tsubasa Hirakawa, Takayoshi Yamashita, Hironobu Fujiyoshi</i></p>	26110
<p>UNICORN: A Unified Backdoor Trigger Inversion Framework</p> <p><i>Zhenting Wang, Kai Mei, Juan Zhai, Shiqing Ma</i></p>	26127
<p>Latent Graph Inference using Product Manifolds</p> <p><i>Haitz SÄŸez de OcÄŸriz Borde, Anees Kazi, Federico Barbero, Pietro Lio</i></p>	26147
<p>Federated Learning as Variational Inference: A Scalable Expectation Propagation Approach</p> <p><i>Han Guo, Philip Greengard, Hongyi Wang, Andrew Gelman, Yoon Kim, Eric Xing</i></p>	26183
<p>Identifiability Results for Multimodal Contrastive Learning</p> <p><i>Imant Daunhawer, Alice Bizeul, Emanuele Palumbo, Alexander Marx, Julia E Vogt</i></p>	26200
<p>A View From Somewhere: Human-Centric Face Representations</p> <p><i>Jerone Theodore Alexander Andrews, Przemyslaw Joniak, Alice Xiang</i></p>	26223
<p>CANIFE: Crafting Canaries for Empirical Privacy Measurement in Federated Learning</p> <p><i>Samuel Maddock, Alexandre Sablayrolles, Pierre Stock</i></p>	26253
<p>Edge Guided GANs with Contrastive Learning for Semantic Image Synthesis</p> <p><i>Hao Tang, XIAOJUAN QI, Guolei Sun, Dan Xu, Nicu Sebe, Radu Timofte, Luc Van Gool</i></p>	26273

<p>HyperDeepONet: learning operator with complex target function space using the limited resources via hypernetwork</p> <p><i>Jae Yong Lee, SungWoong CHO, Hyung Ju Hwang</i></p>	26290
<p>Pseudo-label Training and Model Inertia in Neural Machine Translation</p> <p><i>Benjamin Hsu, Anna Currey, Xing Niu, Maria Nadejde, Georgiana Dinu</i></p>	26316
<p>How robust is unsupervised representation learning to distribution shift?</p> <p><i>Yuge Shi, Imant Daunhawer, Julia E Vogt, Philip Torr, Amartya Sanyal</i></p>	26334
<p>Probabilistically Robust Recourse: Navigating the Trade-offs between Costs and Robustness in Algorithmic Recourse</p> <p><i>Martin Pawelczyk, Teresa Datta, Johan Van den Heuvel, Gjergji Kasneci, Himabindu Lakkaraju</i></p>	26354
<p>A Self-Attention Ansatz for Ab-initio Quantum Chemistry</p> <p><i>Ingrid von Glehn, James S Spencer, David Pfau</i></p>	26380
<p>AE-FLOW: Autoencoders with Normalizing Flows for Medical Images Anomaly Detection</p> <p><i>Yuzhong Zhao, Qiaoqiao Ding, Xiaoqun Zhang</i></p>	26401
<p>Computing all Optimal Partial Transports</p> <p><i>Abhijeet Phatak, Sharath Raghvendra, CHITTARANJAN TRIPATHY, Kaiyi Zhang</i></p>	26415
<p>In-context Reinforcement Learning with Algorithm Distillation</p> <p><i>Michael Laskin, Luyu Wang, Junhyuk Oh, Emilio Parisotto, Stephen Spencer, Richie Steigerwald, DJ Strouse, Steven Stenberg Hansen, Angelos Filos, Ethan Brooks, maxime gazeau, Himanshu Sahn, Satinder Singh, Volodymyr Mnih</i></p>	26427
<p>BALTO: fast tensor program optimization with diversity-based active learning</p> <p><i>Jun Bi, Xiaqing Li, Qi Guo, Rui Zhang, Yuanbo Wen, Xing Hu, Zidong Du, Xinkai Song, Yifan Hao, Yunji Chen</i></p>	26453
<p>Learning What and Where: Disentangling Location and Identity Tracking Without Supervision</p> <p><i>Manuel Traub, Sebastian Otte, Tobias Menge, Matthias Karlbauer, Jannik Thuemmel, Martin V. Butz</i></p>	26465
<p>On the Trade-Off between Actionable Explanations and the Right to be Forgotten</p> <p><i>Martin Pawelczyk, Tobias Leemann, Asia Biega, Gjergji Kasneci</i></p>	26496
<p>A Theory of Dynamic Benchmarks</p> <p><i>Ali Shirali, Rediet Abebe, Moritz Hardt</i></p>	26519
<p>EPISODE: Episodic Gradient Clipping with Periodic Resampled Corrections for Federated Learning with Heterogeneous Data</p> <p><i>Michael Crawshaw, Yajie Bao, Mingrui Liu</i></p>	26543
<p>Revisiting adapters with adversarial training</p> <p><i>Sylvestre-Alvise Rebuffi, Francesco Croce, Sven Gowal</i></p>	26578

<p>Bidirectional Language Models Are Also Few-shot Learners</p> <p>Ajay Patel, Bryan Li, Mohammad Sadegh Rasooli, Noah Constant, Colin Raffel, Chris Callison-Burch</p>	26598
<p>Energy-Based Test Sample Adaptation for Domain Generalization</p> <p>Zehao Xiao, Xiantong Zhen, Shengcai Liao, Cees G. M. Snoek</p>	26623
<p>A Higher Precision Algorithm for Computing the \mathbb{L}_1-Wasserstein Distance</p> <p>Pankaj K Agarwal, Sharath Raghvendra, Pouyan Shirzadian, Rachita Sowle</p>	26648
<p>Learning Rationalizable Equilibria in Multiplayer Games</p> <p>Yuanhao Wang, Dingwen Kong, Yu Bai, Chi Jin</p>	26660
<p>HypeR: Multitask Hyper-Prompted Training Enables Large-Scale Retrieval Generalization</p> <p>ZeFeng Cai, Chongyang Tao, Tao Shen, Can Xu, Xiubo Geng, Xin Alex Lin, Liang He, Daxin Jiang</p>	26685
<p>Online Boundary-Free Continual Learning by Scheduled Data Prior</p> <p>Hyunseo Koh, Minhyuk Seo, Jihwan Bang, Hwanjun Song, Deokki Hong, Seulki Park, Jung-Woo Ha, Jonghyun Choi</p>	26701
<p>Achieving Near-Optimal Individual Regret & Low Communications in Multi-Agent Bandits</p> <p>Xuchuang Wang, Lin Yang, Yu-Zhen Janice Chen, Xutong Liu, Mohammad Hajiesmaili, Don Towsley, John C.S. Lui</p>	26724
<p>Subquadratic Algorithms for Kernel Matrices via Kernel Density Estimation</p> <p>Ainesh Bakshi, Piotr Indyk, Praneeth Kacham, Sandeep Silwal, Samson Zhou</p>	26750
<p>Diffusion Policies as an Expressive Policy Class for Offline Reinforcement Learning</p> <p>Zhendong Wang, Jonathan J Hunt, Mingyuan Zhou</p>	26790
<p>Learning to Induce Causal Structure</p> <p>Nan Rosemary Ke, Silvia Chiappa, Jane X Wang, Jorg Bornschein, Anirudh Goyal, Melanie Rey, Theophane Weber, Matthew Botvinick, Michael Curtis Mozer, Danilo Jimenez Rezende</p>	26807
<p>Optimistic Exploration with Learned Features Provably Solves Markov Decision Processes with Neural Dynamics</p> <p>Sirui Zheng, Lingxiao Wang, Shuang Qiu, Zuyue Fu, Zhuoran Yang, Csaba Szepesvari, Zhaoran Wang</p>	26837
<p>Stay Moral and Explore: Learn to Behave Morally in Text-based Games</p> <p>Zijing Shi, Meng Fang, Yunqiu Xu, Ling Chen, Yali Du</p>	26891
<p>Towards a Unified Theoretical Understanding of Non-contrastive Learning via Rank Differential Mechanism</p> <p>Zhijian Zhuo, Yifei Wang, Jinwen Ma, Yisen Wang</p>	26909
<p>Bridging the Gap to Real-World Object-Centric Learning</p> <p>Maximilian Seitzer, Max Horn, Andrii Zadaianchuk, Dominik Zietlow, Tianjun Xiao, Carl-Johann Simon-Gabriel, Tong He, Zheng Zhang, Bernhard Schölkopf, Thomas Brox, Francesco Locatello</p>	26930

VIP: Towards Universal Visual Reward and Representation via Value-Implicit Pre-Training Yecheng Jason Ma, Shagun Sodhani, Dinesh Jayaraman, Osbert Bastani, Vikash Kumar, Amy Zhang	26973
Neuromechanical Autoencoders: Learning to Couple Elastic and Neural Network Nonlinearity Deniz Oktay, Mehran Mirramezani, Eder Medina, Ryan P Adams	27007
Strong inductive biases provably prevent harmless interpolation Michael Aerni, Marco Milanta, Konstantin Donhauser, Fanny Yang	27023
Self-supervised learning with rotation-invariant kernels LÃ©on Zheng, Gilles Puy, Elisa Riccietti, Patrick Perez, RÃ©mi Gribonval	27079
Exploring the Limits of Differentially Private Deep Learning with Group-wise Clipping Jiyan He, Xuechen Li, Da Yu, Huishuai Zhang, Janardhan Kulkarni, Yin Tat Lee, Arturs Backurs, Nenghai Yu, Jiang Bian	27106
VIPeR: Provably Efficient Algorithm for Offline RL with Neural Function Approximation Thanh Nguyen-Tang, Raman Arora	27131
Temporal Disentanglement of Representations for Improved Generalisation in Reinforcement Learning Mhairi Dunion, Trevor McInroe, Kevin Sebastian Luck, Josiah P. Hanna, Stefano V Albrecht	27175
Understanding The Robustness of Self-supervised Learning Through Topic Modeling Zeping Luo, Shiyu Wu, Cindy Weng, Mo Zhou, Rong Ge	27191
FIFA: Making Fairness More Generalizable in Classifiers Trained on Imbalanced Data Zhun Deng, Jiayao Zhang, Linjun Zhang, Ting Ye, Yates Coley, Weijie J Su, James Zou	27222
Transformers Learn Shortcuts to Automata Bingbin Liu, Jordan T. Ash, Surbhi Goel, Akshay Krishnamurthy, Cyril Zhang	27244
\$k\$NN Prompting: Beyond-Context Learning with Calibration-Free Nearest Neighbor Inference Benfeng Xu, Quan Wang, Zhendong Mao, Yajuan Lyu, Qiaoqiao She, Yongdong Zhang	27308
Autoregressive Conditional Neural Processes Wessel Bruinsma, Stratis Markou, James Requeima, Andrew Y. K. Foong, Tom Andersson, Anna Vaughan, Anthony Buonomo, Scott Hosking, Richard E Turner	27332
Robust Scheduling with GFlowNets David W Zhang, Corrado Rainone, Markus Peschl, Roberto Bondesan	27389
Decentralized Optimistic Hyperpolicy Mirror Descent: Provably No-Regret Learning in Markov Games Wenhao Zhan, Jason D. Lee, Zhuoran Yang	27406
PAC Reinforcement Learning for Predictive State Representations Wenhao Zhan, Masatoshi Uehara, Wen Sun, Jason D. Lee	27445

<p>Spectral Augmentation for Self-Supervised Learning on Graphs</p> <p><i>Lu Lin, Jinghui Chen, Hongning Wang</i></p>	27481
<p>MACTA: A Multi-agent Reinforcement Learning Approach for Cache Timing Attacks and Detection</p> <p><i>Jiaxun Cui, Xiaomeng Yang, Mulong Luo, Geunbae Lee, Peter Stone, Hsien-Hsin S. Lee, Benjamin Lee, G. Edward Suh, Wenjie Xiong, Yuandong Tian</i></p>	27508
<p>\mathcal{O}-GNN: incorporating ring priors into molecular modeling</p> <p><i>Jinhua Zhu, Kehan Wu, Bohan Wang, Yingce Xia, Shufang Xie, Qi Meng, Lijun Wu, Tao Qin, Wengang Zhou, Houqiang Li, Tie-Yan Liu</i></p>	27530
<p>Latent Neural ODEs with Sparse Bayesian Multiple Shooting</p> <p><i>Valerii Iakovlev, Cagatay Yildiz, Markus Heinonen, Harri Lähdesmäki</i></p>	27549
<p>Emergence of Maps in the Memories of Blind Navigation Agents</p> <p><i>Erik Wijmans, Manolis Savva, Irfan Essa, Stefan Lee, Ari S. Morcos, Dhruv Batra</i></p>	27580
<p>Learning Fair Graph Representations via Automated Data Augmentations</p> <p><i>Hongyi Ling, Zhimeng Jiang, Youzhi Luo, Shuiwang Ji, Na Zou</i></p>	27606
<p>Learning About Progress From Experts</p> <p><i>Jake Bruce, Ankit Anand, Bogdan Mazouze, Rob Fergus</i></p>	27624
<p>Understanding DDPM Latent Codes Through Optimal Transport</p> <p><i>Valentin Khrulkov, Gleb Ryzhakov, Andrei Chertkov, Ivan Oseledets</i></p>	27650
<p>Interneurons accelerate learning dynamics in recurrent neural networks for statistical adaptation</p> <p><i>David Lipshutz, Cengiz Pehlevan, Dmitri Chklovskii</i></p>	27665
<p>Generating Sequences by Learning to Self-Correct</p> <p><i>Sean Welleck, Ximing Lu, Peter West, Faeze Brahman, Tianxiao Shen, Daniel Khoshabi, Yejin Choi</i></p>	27681
<p>Disentanglement of Correlated Factors via Hausdorff Factorized Support</p> <p><i>Karsten Roth, Mark Ibrahim, Zeynep Akata, Pascal Vincent, Diane Bouchacourt</i></p>	27700
<p>Choreographer: Learning and Adapting Skills in Imagination</p> <p><i>Pietro Mazzaglia, Tim Verbelen, Bart Dhoedt, Alexandre Lacoste, Sai Rajeswar</i></p>	27727
<p>TTN: A Domain-Shift Aware Batch Normalization in Test-Time Adaptation</p> <p><i>Hyesu Lim, Byeonggeun Kim, Jaegul Choo, Sungha Choi</i></p>	27752
<p>Sequential Gradient Coding For Straggler Mitigation</p> <p><i>Nikhil Krishnan Muralee Krishnan, MohammadReza Ebrahimi, Ashish J Khisti</i></p>	27771
<p>Revisiting Populations in multi-agent Communication</p> <p><i>Paul Michel, Mathieu Rita, Kory Wallace Mathewson, Olivier Tieleman, Angeliki Lazaridou</i></p>	27800

<p>Amortised Invariance Learning for Contrastive Self-Supervision</p> <p><i>Ruchika Chavhan, Jan Stuehmer, Calum Heggan, Mehrdad Yaghoobi, Timothy Hospedales</i></p>	27816
<p>Neural Implicit Shape Editing using Boundary Sensitivity</p> <p><i>Arturs Berzins, Moritz Ibing, Leif Kobbelt</i></p>	27839
<p>Improved Learning-augmented Algorithms for k-means and k-medians Clustering</p> <p><i>Thy Dinh Nguyen, Anamay Chaturvedi, Huy Nguyen</i></p>	27854
<p>Canary in a Coalmine: Better Membership Inference with Ensembled Adversarial Queries</p> <p><i>Yuxin Wen, Arpit Bansal, Hamid Kazemi, Eitan Borgnia, Micah Goldblum, Jonas Geiping, Tom Goldstein</i></p>	27873
<p>SoftZoo: A Soft Robot Co-design Benchmark For Locomotion In Diverse Environments</p> <p><i>Tsun-Hsuan Wang, Pingchuan Ma, Andrew Everett Spielberg, Zhou Xian, Hao Zhang, Joshua B. Tenenbaum, Daniela Rus, Chuang Gan</i></p>	27887
<p>ImageNet-X: Understanding Model Mistakes with Factor of Variation Annotations</p> <p><i>Badr Youbi Idrissi, Diane Bouchacourt, Randall Balestriero, Ivan Evtimov, Caner Hazirbas, Nicolas Ballas, Pascal Vincent, Michal Drozdal, David Lopez-Paz, Mark Ibrahim</i></p>	27907
<p>Meta Learning to Bridge Vision and Language Models for Multimodal Few-Shot Learning</p> <p><i>Ivona Najdenkoska, Xiantong Zhen, Marcel Worring</i></p>	27931
<p>Scaling Laws For Deep Learning Based Image Reconstruction</p> <p><i>Tobit Klug, Reinhard Heckel</i></p>	27944
<p>Proposal-Contrastive Pretraining for Object Detection from Fewer Data</p> <p><i>Quentin Bouniot, Romaric Audigier, Angelique Loesch, Amaury Habrard</i></p>	27983
<p>In-Situ Text-Only Adaptation of Speech Models with Low-Overhead Speech Imputations</p> <p><i>Ashish Mittal, Sunita Sarawagi, Preethi Jyothi</i></p>	27996
<p>Aligning Model and Macaque Inferior Temporal Cortex Representations Improves Model-to-Human Behavioral Alignment and Adversarial Robustness</p> <p><i>Joel Dapello, Kohitij Kar, Martin Schrimpf, Robert Baldwin Geary, Michael Ferguson, David Daniel Cox, James J. DiCarlo</i></p>	28012
<p>Metadata Archaeology: Unearthing Data Subsets by Leveraging Training Dynamics</p> <p><i>Shoaib Ahmed Siddiqui, Nitarshan Rajkumar, Tegan Maharaj, David Krueger, Sara Hooker</i></p>	28025
<p>A probabilistic framework for task-aligned intra- and inter-area neural manifold estimation</p> <p><i>Edoardo Balzani, Jean-Paul G Noel, Pedro Herrero-Vidal, Dora E Angelaki, Cristina Savin</i></p>	28046
<p>MMVAE+: Enhancing the Generative Quality of Multimodal VAEs without Compromises</p> <p><i>Emanuele Palumbo, Imant Daunhawer, Julia E Vogt</i></p>	28068

<p>Heterogeneous Neuronal and Synaptic Dynamics for Spike-Efficient Unsupervised Learning: Theory and Design Principles</p> <p><i>Biswadeep Chakraborty, Saibal Mukhopadhyay</i></p>	28095
<p>Efficient Conditionally Invariant Representation Learning</p> <p><i>Roman Pogodin, Namrata Deka, Yazhe Li, Danica J. Sutherland, Victor Veitch, Arthur Gretton</i></p>	28131
<p>(Certified!!) Adversarial Robustness for Free!</p> <p><i>Nicholas Carlini, Florian Tramer, Krishnamurthy Dj Dvijotham, Leslie Rice, Mingjie Sun, J Zico Kolter</i></p>	28151
<p>HiT-MDP: Learning the SMDP option framework on MDPs with Hidden Temporal Embeddings</p> <p><i>Chang Li, Dongjin Song, Dacheng Tao</i></p>	28165
<p>Making Better Decision by Directly Planning in Continuous Control</p> <p><i>Jinhua Zhu, Yue Wang, Lijun Wu, Tao Qin, Wengang Zhou, Tie-Yan Liu, Houqiang Li</i></p>	28197
<p>SP2 : A Second Order Stochastic Polyak Method</p> <p><i>Shuang Li, William Joseph Swartworth, Martin Takáč, Deanna Needell, Robert M. Gower</i></p>	28222
<p>On the Convergence of AdaGrad(Norm) on \mathbb{R}^d: Beyond Convexity, Non-Asymptotic Rate and Acceleration</p> <p><i>Zijian Liu, Ta Duy Nguyen, Alina Ene, Huy Nguyen</i></p>	28257
<p>Order Matters: Agent-by-agent Policy Optimization</p> <p><i>Xihuai Wang, Zheng Tian, Ziyu Wan, Ying Wen, Jun Wang, Weinan Zhang</i></p>	28301
<p>Accelerating Hamiltonian Monte Carlo via Chebyshev Integration Time</p> <p><i>Jun-Kun Wang, Andre Wibisono</i></p>	28336
<p>Equivariance-aware Architectural Optimization of Neural Networks</p> <p><i>Kaitlin Maile, Dennis George Wilson, Patrick Forr</i></p>	28354
<p>Hebbian Deep Learning Without Feedback</p> <p><i>Adrien Journ, Hector Garcia Rodriguez, Qinghai Guo, Timoleon Moraitis</i></p>	28373
<p>Planning with Large Language Models for Code Generation</p> <p><i>Shun Zhang, Zhenfang Chen, Yikang Shen, Mingyu Ding, Joshua B. Tenenbaum, Chuang Gan</i></p>	28401
<p>Summarization Programs: Interpretable Abstractive Summarization with Neural Modular Trees</p> <p><i>Swarnadeep Saha, Shiyue Zhang, Peter Hase, Mohit Bansal</i></p>	28429
<p>Symmetries, Flat Minima, and the Conserved Quantities of Gradient Flow</p> <p><i>Bo Zhao, Jordan Ganev, Robin Walters, Rose Yu, Nima Dehmamy</i></p>	28452
<p>Hyperparameter Optimization through Neural Network Partitioning</p> <p><i>Bruno Kacper Mlodozeniec, Matthias Reisser, Christos Louizos</i></p>	28494
<p>Few-shot Backdoor Attacks via Neural Tangent Kernels</p> <p><i>Jonathan Hayase, Sewoong Oh</i></p>	28521

<p>Noise-Robust De-Duplication at Scale</p> <p><i>Emily Silcock, Luca D'Amico-Wong, Jinglin Yang, Melissa Dell</i></p>	28536
<p>Logical Message Passing Networks with One-hop Inference on Atomic Formulas</p> <p><i>Zihao Wang, Yangqiu Song, Ginny Wong, Simon See</i></p>	28563
<p>Lossless Adaptation of Pretrained Vision Models For Robotic Manipulation</p> <p><i>Mohit Sharma, Claudio Fantacci, Yuxiang Zhou, Skanda Koppula, Nicolas Heess, Jon Scholz, Yusuf Aytar</i></p>	28583
<p>Sparse Upcycling: Training Mixture-of-Experts from Dense Checkpoints</p> <p><i>Aran Komatsuzaki, Joan Puigcerver, James Lee-Thorp, Carlos Riquelme Ruiz, Basil Mustafa, Joshua Ainslie, Yi Tay, Mostafa Dehghani, Neil Houlsby</i></p>	28601
<p>Variational Information Pursuit for Interpretable Predictions</p> <p><i>Aditya Chattopadhyay, Kwan Ho Ryan Chan, Benjamin David Haeffele, Donald Geman, Rene Vidal</i></p>	28628
<p>Revisiting Robustness in Graph Machine Learning</p> <p><i>Lukas Gosch, Daniel Sturm, Simon Geisler, Stephan GÅ¼nnemann</i></p>	28661
<p>Understanding Train-Validation Split in Meta-Learning with Neural Networks</p> <p><i>Xinzhe Zuo, Zixiang Chen, Huaxiu Yao, Yuan Cao, Quanquan Gu</i></p>	28715
<p>Can We Find Nash Equilibria at a Linear Rate in Markov Games?</p> <p><i>Zhuoqing Song, Jason D. Lee, Zhuoran Yang</i></p>	28771
<p>Using Language to Extend to Unseen Domains</p> <p><i>Lisa Dunlap, Clara Mohri, Devin Guillory, Han Zhang, Trevor Darrell, Joseph E. Gonzalez, Aditi Raghunathan, Anna Rohrbach</i></p>	28833
<p>CrAM: A Compression-Aware Minimizer</p> <p><i>Alexandra Peste, Adrian Vladu, Eldar Kurtic, Christoph H Lampert, Dan Alistarh</i></p>	28845
<p>Transfer Learning with Deep Tabular Models</p> <p><i>Roman Levin, Valeriia Cherepanova, Avi Schwarzschild, Arpit Bansal, C. Bayan Bruss, Tom Goldstein, Andrew Gordon Wilson, Micah Goldblum</i></p>	28868
<p>FIT: A Metric for Model Sensitivity</p> <p><i>Ben Zandonati, Adrian Alan Pol, Maurizio Pierini, Olya Sirkin, Tal Kopetz</i></p>	28902
<p>Investigating Multi-task Pretraining and Generalization in Reinforcement Learning</p> <p><i>Adrien Ali Taiga, Rishabh Agarwal, Jesse Farebrother, Aaron Courville, Marc G Bellemare</i></p>	28924
<p>Bayes-MIL: A New Probabilistic Perspective on Attention-based Multiple Instance Learning for Whole Slide Images</p> <p><i>Yufei CUI, Ziquan Liu, Xiangyu Liu, Xue Liu, Cong Wang, Tei-Wei Kuo, Chun Jason Xue, Antoni B. Chan</i></p>	28947

<p>Modeling the Data-Generating Process is Necessary for Out-of-Distribution Generalization</p> <p><i>Jivat Neet Kaur, Emre Kiciman, Amit Sharma</i></p>	28958
<p>Task-customized Masked Autoencoder via Mixture of Cluster-conditional Experts</p> <p><i>Zhili LIU, Kai Chen, Jianhua Han, Lanqing HONG, Hang Xu, Zhenguo Li, James Kwok</i></p>	28985
<p>Long Range Language Modeling via Gated State Spaces</p> <p><i>Harsh Mehta, Ankit Gupta, Ashok Cutkosky, Behnam Neyshabur</i></p>	29000
<p>Phenaki: Variable Length Video Generation from Open Domain Textual Descriptions</p> <p><i>Ruben Villegas, Mohammad Babaeizadeh, Pieter-Jan Kindermans, Hernan Moraldo, Han Zhang, Mohammad Taghi Saffar, Santiago Castro, Julius Kunze, Dumitru Erhan</i></p>	29020
<p>Differentially Private Adaptive Optimization with Delayed Preconditioners</p> <p><i>Tian Li, Manzil Zaheer, Ken Liu, Sashank J. Reddi, Hugh Brendan McMahan, Virginia Smith</i></p>	29034
<p>SemPPL: Predicting Pseudo-Labels for Better Contrastive Representations</p> <p><i>Matko BoÅ¡njak, Pierre Harvey Richemond, Nenad Tomasev, Florian Strub, Jacob C Walker, Felix Hill, Lars Holger Buesing, Razvan Pascanu, Charles Blundell, Jovana Mitrovic</i></p>	29060
<p>Taking a Step Back with KCal: Multi-Class Kernel-Based Calibration for Deep Neural Networks</p> <p><i>Zhen Lin, Shubhendu Trivedi, Jimeng Sun</i></p>	29085
<p>Agree to Disagree: Diversity through Disagreement for Better Transferability</p> <p><i>Matteo Pagliardini, Martin Jaggi, François Fleuret, Sai Praneeth Karimireddy</i></p>	29099
<p>Optimizing Bi-Encoder for Named Entity Recognition via Contrastive Learning</p> <p><i>Sheng Zhang, Hao Cheng, Jianfeng Gao, Hoifung Poon</i></p>	29123
<p>Effective passive membership inference attacks in federated learning against overparameterized models</p> <p><i>Jiacheng Li, Ninghui Li, Bruno Ribeiro</i></p>	29143
<p>The Lie Derivative for Measuring Learned Equivariance</p> <p><i>Nate Gruver, Marc Anton Finzi, Micah Goldblum, Andrew Gordon Wilson</i></p>	29171
<p>DexDeform: Dexterous Deformable Object Manipulation with Human Demonstrations and Differentiable Physics</p> <p><i>Sizhe Li, Zhiao Huang, Tao Chen, Tao Du, Hao Su, Joshua B. Tenenbaum, Chuang Gan</i></p>	29180
<p>Composing Task Knowledge With Modular Successor Feature Approximators</p> <p><i>Wilka Torrico Carvalho, Angelos Filos, Richard Lewis, Honglak Lee, Satinder Singh</i></p>	29201
<p>Performance Bounds for Model and Policy Transfer in Hidden-parameter MDPs</p> <p><i>Haotian Fu, Jiayu Yao, Omer Gottesman, Finale Doshi-Velez, George Konidaris</i></p>	29213
<p>The Surprising Effectiveness of Equivariant Models in Domains with Latent Symmetry</p> <p><i>Dian Wang, Jung Yeon Park, Neel Sortur, Lawson L.S. Wong, Robin Walters, Robert Platt</i></p>	29242

<p>Distributed Extra-gradient with Optimal Complexity and Communication Guarantees</p> <p><i>Ali Ramezani-Kebrya, Kimon Antonakopoulos, Igor Krawczuk, Justin Deschenaux, Volkan Cevher</i></p>	29269
<p>Towards Interpretable Deep Reinforcement Learning with Human-Friendly Prototypes</p> <p><i>Eoin M. Kenny, Mycal Tucker, Julie Shah</i></p>	29302
<p>Accelerated Single-Call Methods for Constrained Min-Max Optimization</p> <p><i>Yang Cai, Weiqiang Zheng</i></p>	29321
<p>ContraNorm: A Contrastive Learning Perspective on Oversmoothing and Beyond</p> <p><i>Xiaojun Guo, Yifei Wang, Tianqi Du, Yisen Wang</i></p>	29349
<p>Don't forget the nullspace! Nullspace occupancy as a mechanism for out of distribution failure</p> <p><i>Daksh Idrani, Vivek Madan, Naman Goyal, David J. Schwab, Shanmukha Ramakrishna Vedantam</i></p>	29368
<p>Error Sensitivity Modulation based Experience Replay: Mitigating Abrupt Representation Drift in Continual Learning</p> <p><i>Fahad Sarfraz, Elahe Arani, Bahram Zonooz</i></p>	29393
<p>Last Layer Re-Training is Sufficient for Robustness to Spurious Correlations</p> <p><i>Polina Kirichenko, Pavel Izmailov, Andrew Gordon Wilson</i></p>	29411
<p>Verifying the Union of Manifolds Hypothesis for Image Data</p> <p><i>Bradley CA Brown, Anthony L. Caterini, Brendan Leigh Ross, Jesse C Cresswell, Gabriel Loaiza-Ganem</i></p>	29448
<p>Planning with Sequence Models through Iterative Energy Minimization</p> <p><i>Hongyi Chen, Yilun Du, Yiye Chen, Joshua B. Tenenbaum, Patricio A. Vela</i></p>	29472
<p>Pseudoinverse-Guided Diffusion Models for Inverse Problems</p> <p><i>Jiaming Song, Arash Vahdat, Morteza Mardani, Jan Kautz</i></p>	29490
<p>Augmentation with Projection: Towards an Effective and Efficient Data Augmentation Paradigm for Distillation</p> <p><i>Ziqi Wang, Yuexin Wu, Frederick Liu, Daogao Liu, Le Hou, Hongkun Yu, Jing Li, Heng Ji</i></p>	29520
<p>Provable Defense Against Geometric Transformations</p> <p><i>Rem Yang, Jacob Laurel, Sasa Misailovic, Gagandeep Singh</i></p>	29540
<p>Evaluating Representations with Readout Model Switching</p> <p><i>Yazhe Li, Jorg Bornschein, Marcus Hutter</i></p>	29559
<p>TEMPERA: Test-Time Prompt Editing via Reinforcement Learning</p> <p><i>Tianjun Zhang, Xuezhi Wang, Denny Zhou, Dale Schuurmans, Joseph E. Gonzalez</i></p>	29582
<p>Is Conditional Generative Modeling all you need for Decision Making?</p> <p><i>Anurag Ajay, Yilun Du, Abhi Gupta, Joshua B. Tenenbaum, Tommi S. Jaakkola, Pulkit Agrawal</i></p>	29598

<p>Do We Really Need Complicated Model Architectures For Temporal Networks?</p> <p>Weilin Cong, Si Zhang, Jian Kang, Baichuan Yuan, Hao Wu, Xin Zhou, Hanghang Tong, Mehrdad Mahdavi</p>	29620
<p>How gradient estimator variance and bias impact learning in neural networks</p> <p>Arna Ghosh, Yuhan Helena Liu, Guillaume Lajoie, Konrad Kording, Blake Aaron Richards</p>	29647
<p>3D Equivariant Diffusion for Target-Aware Molecule Generation and Affinity Prediction</p> <p>Jiaqi Guan, Wesley Wei Qian, Xingang Peng, Yufeng Su, Jian Peng, Jianzhu Ma</p>	29660
<p>Explicitly Minimizing the Blur Error of Variational Autoencoders</p> <p>Gustav Bredell, Kyriakos Flouris, Krishna Chaitanya, Ertunc Erdil, Ender Konukoglu</p>	29682
<p>DAG Learning on the Permutahedron</p> <p>Valentina Zantedeschi, Luca Franceschi, Jean Kaddour, Matt Kusner, Vlad Niculae</p>	29698
<p>SMART: Self-supervised Multi-task pretraining with control Transformers</p> <p>Yanchao Sun, Shuang Ma, Ratnesh Madaan, Rogerio Bonatti, Furong Huang, Ashish Kapoor</p>	29723
<p>Semi-Parametric Inducing Point Networks and Neural Processes</p> <p>Richa Rastogi, Yair Schiff, Alon Hachohen, Zhaozhi Li, Ian Lee, Yuntian Deng, Mert R. Sabuncu, Volodymyr Kuleshov</p>	29744
<p>Associative Memory Augmented Asynchronous Spatiotemporal Representation Learning for Event-based Perception</p> <p>Uday Kamal, Saurabh Dash, Saibal Mukhopadhyay</p>	29771
<p>Scalable Batch-Mode Deep Bayesian Active Learning via Equivalence Class Annealing</p> <p>Renyu Zhang, Aly A Khan, Robert L. Grossman, Yuxin Chen</p>	29791
<p>DINO as a von Mises-Fisher mixture model</p> <p>Hariprasath Govindarajan, Per SidaŃn, Jacob Roll, Fredrik Lindsten</p>	29819
<p>Relative Behavioral Attributes: Filling the Gap between Symbolic Goal Specification and Reward Learning from Human Preferences</p> <p>Lin Guan, Karthik Valmeekam, Subbarao Kambhampati</p>	29838
<p>Semantic Uncertainty: Linguistic Invariances for Uncertainty Estimation in Natural Language Generation</p> <p>Lorenz Kuhn, Yarin Gal, Sebastian Farquhar</p>	29857
<p>Federated Learning from Small Datasets</p> <p>Michael Kamp, Jonas Fischer, Jilles Vreeken</p>	29876
<p>Spatio-temporal point processes with deep non-stationary kernels</p> <p>Zheng Dong, Xiuyuan Cheng, Yao Xie</p>	29889
<p>Contrastive Corpus Attribution for Explaining Representations</p> <p>Chris Lin, Hugh Chen, Chanwoo Kim, Su-In Lee</p>	29907

<p>Where to Diffuse, How to Diffuse, and How to Get Back: Automated Learning for Multivariate Diffusions</p> <p><i>Raghav Singhal, Mark Goldstein, Rajesh Ranganath</i></p>	29937
<p>Grounding Graph Network Simulators using Physical Sensor Observations</p> <p><i>Jonas LinkerhÄgner, Niklas Freymuth, Paul Maria Scheikl, Franziska Mathis-Ullrich, Gerhard Neumann</i></p>	29965
<p>DensePure: Understanding Diffusion Models for Adversarial Robustness</p> <p><i>Chaowei Xiao, Zhongzhu Chen, Kun Jin, Jiong Xiao Wang, Weili Nie, Mingyan Liu, Anima Anandkumar, Bo Li, Dawn Song</i></p>	29992
<p>Rhino: Deep Causal Temporal Relationship Learning with History-dependent Noise</p> <p><i>Wenbo Gong, Joel Jennings, Cheng Zhang, Nick Pawlowski</i></p>	30015
<p>Discovering Evolution Strategies via Meta-Black-Box Optimization</p> <p><i>Robert Tjarko Lange, Tom Schaul, Yutian Chen, Tom Zahavy, Valentin Dalibard, Chris Lu, Satinder Singh, Sebastian Flennerhag</i></p>	30045
<p>Towards convergence to Nash equilibria in two-team zero-sum games</p> <p><i>Fivos Kalogiannis, Ioannis Panageas, Emmanouil-Vasileios Vlatakis-Gkaragkounis</i></p>	30070
<p>ReAct: Synergizing Reasoning and Acting in Language Models</p> <p><i>Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik R Narasimhan, Yuan Cao</i></p>	30084
<p>Personalized Reward Learning with Interaction-Grounded Learning (IGL)</p> <p><i>Jessica Maghakian, Paul Mineiro, Kishan Panaganti, Mark Rucker, Akanksha Saran, Cheng Tan</i></p>	30117
<p>Make-A-Video: Text-to-Video Generation without Text-Video Data</p> <p><i>Uriel Singer, Adam Polyak, Thomas Hayes, Xi Yin, Jie An, Songyang Zhang, Qiyuan Hu, Harry Yang, Oron Ashual, Oran Gafni, Devi Parikh, Sonal Gupta, Yaniv Taigman</i></p>	30133
<p>Solving Continuous Control via Q-learning</p> <p><i>Tim Seyde, Peter Werner, Wilko Schwarting, Igor Gilitschenski, Martin Riedmiller, Daniela Rus, Markus Wulfmeier</i></p>	30149
<p>Hyper-Decision Transformer for Efficient Online Policy Adaptation</p> <p><i>Mengdi Xu, Yuchen Lu, Yikang Shen, Shun Zhang, Ding Zhao, Chuang Gan</i></p>	30172
<p>Beyond Lipschitz: Sharp Generalization and Excess Risk Bounds for Full-Batch GD</p> <p><i>Konstantinos Nikolakakis, Farzin Haddadpour, Amin Karbasi, Dionysios Kalogerias</i></p>	30191
<p>Bispectral Neural Networks</p> <p><i>Sophia Sanborn, Christian A Shewmake, Bruno Olshausen, Christopher J. Hillar</i></p>	30222
<p>$O(T^{-1})$ Convergence of Optimistic-Follow-the-Regularized-Leader in Two-Player Zero-Sum Markov Games</p> <p><i>Yuepeng Yang, Cong Ma</i></p>	30241

<p>Mitigating Gradient Bias in Multi-objective Learning: A Provably Convergent Approach</p> <p>Heshan Devaka Fernando, Han Shen, Miao Liu, Subhajit Chaudhury, Keerthiram Murugesan, Tianyi Chen</p>	30260
<p>Causal Reasoning in the Presence of Latent Confounders via Neural ADMG Learning</p> <p>Matthew Ashman, Chao Ma, Agrin Hilmkil, Joel Jennings, Cheng Zhang</p>	30292
<p>Integrating Symmetry into Differentiable Planning with Steerable Convolutions</p> <p>Linfeng Zhao, Xupeng Zhu, Lingzhi Kong, Robin Walters, Lawson L.S. Wong</p>	30311
<p>DySR: Adaptive Super-Resolution via Algorithm and System Co-design</p> <p>Syed Zawad, Cheng Li, Zhewei Yao, Elton Zheng, Yuxiong He, Feng Yan</p>	30352
<p>Fake It Until You Make It : Towards Accurate Near-Distribution Novelty Detection</p> <p>Hossein Mirzaei, Mohammadreza Salehi, Sajjad Shahabi, Efstratios Gavves, Cees G. M. Snoek, Mohammad Sabokrou, Mohammad Hossein Rohban</p>	30366
<p>Analogy-Forming Transformers for Few-Shot 3D Parsing</p> <p>Nikolaos Gkanatsios, Mayank Singh, Zhaoyuan Fang, Shubham Tulsiani, Katerina Fragkiadaki</p>	30393
<p>Depth Separation with Multilayer Mean-Field Networks</p> <p>Yunwei Ren, Mo Zhou, Rong Ge</p>	30421
<p>Implicit Bias of Large Depth Networks: a Notion of Rank for Nonlinear Functions</p> <p>Arthur Jacot</p>	30465
<p>Function-space regularized R�nyi divergences</p> <p>Jeremiah Birrell, Yannis Pantazis, Paul Dupuis, Luc Rey-Bellet, Markos Katsoulakis</p>	30476
<p>Why (and When) does Local SGD Generalize Better than SGD?</p> <p>Xinran Gu, Kaifeng Lyu, Longbo Huang, Sanjeev Arora</p>	30504
<p>Images as Weight Matrices: Sequential Image Generation Through Synaptic Learning Rules</p> <p>Kazuki Irie, J�rgen Schmidhuber</p>	30580
<p>Learning Low Dimensional State Spaces with Overparameterized Recurrent Neural Nets</p> <p>Edo Cohen-Karlik, Itamar Menuhin-Gruman, Raja Giryes, Nadav Cohen, Amir Globerson</p>	30597
<p>Learning rigid dynamics with face interaction graph networks</p> <p>Kelsey R Allen, Yulia Rubanova, Tatiana Lopez-Guevara, William F Whitney, Alvaro Sanchez-Gonzalez, Peter Battaglia, Tobias Pfaff</p>	30629
<p>Synthetic Data Generation of Many-to-Many Datasets via Random Graph Generation</p> <p>Kai Xu, Georgi Ganev, Emile Joubert, Rees Davison, Olivier Van Acker, Luke Robinson</p>	30655
<p>Disentanglement with Biological Constraints: A Theory of Functional Cell Types</p> <p>James C. R. Whittington, Will Dorrell, Surya Ganguli, Timothy Behrens</p>	30672

<p><i>Is Model Ensemble Necessary? Model-based RL via a Single Model with Lipschitz Regularized Value Function</i></p> <p><i>Ruijie Zheng, Xiyao Wang, Huazhe Xu, Furong Huang</i></p>	30709
<p><i>Adversarial Imitation Learning with Preferences</i></p> <p><i>Aleksandar Taranovic, Andras Gabor Kupcsik, Niklas Freymuth, Gerhard Neumann</i></p>	30730
<p><i>Decision Transformer under Random Frame Dropping</i></p> <p><i>Kaizhe Hu, Ray Chen Zheng, Yang Gao, Huazhe Xu</i></p>	30747
<p><i>Score-based Continuous-time Discrete Diffusion Models</i></p> <p><i>Haoran Sun, Lijun Yu, Bo Dai, Dale Schuurmans, Hanjun Dai</i></p>	30767
<p><i>Scaling up and Stabilizing Differentiable Planning with Implicit Differentiation</i></p> <p><i>Linfeng Zhao, Huazhe Xu, Lawson L.S. Wong</i></p>	30783
<p><i>Conservative Bayesian Model-Based Value Expansion for Offline Policy Optimization</i></p> <p><i>Jihwan Jeong, Xiaoyu Wang, Michael Gimelfarb, Hyunwoo Kim, Baher abdulhai, Scott Sanner</i></p>	30804
<p><i>Universal Few-shot Learning of Dense Prediction Tasks with Visual Token Matching</i></p> <p><i>Donggyun Kim, Jinwoo Kim, Seongwoong Cho, Chong Luo, Seunghoon Hong</i></p>	30816
<p><i>Subsampling in Large Graphs Using Ricci Curvature</i></p> <p><i>Shushan Wu, Huimin Cheng, Jiazhong Cai, Ping Ma, Wenxuan Zhong</i></p>	30842
<p><i>Sampling is as easy as learning the score: theory for diffusion models with minimal data assumptions</i></p> <p><i>Sitan Chen, Sinho Chewi, Jerry Li, Yuanzhi Li, Adil Salim, Anru Zhang</i></p>	30868
<p><i>Learning multi-scale local conditional probability models of images</i></p> <p><i>Zahra Kadkhodaie, Florentin Guth, Stéphanie Mallat, Eero P Simoncelli</i></p>	30896
<p><i>Is Reinforcement Learning (Not) for Natural Language Processing: Benchmarks, Baselines, and Building Blocks for Natural Language Policy Optimization</i></p> <p><i>Rajkumar Ramamurthy, Prithviraj Ammanabrolu, Kiant Brantley, Jack Hessel, Rafet Sifa, Christian Bauckhage, Hannaneh Hajishirzi, Yejin Choi</i></p>	30912
<p><i>Optimal Transport for Offline Imitation Learning</i></p> <p><i>Yicheng Luo, zhengyao jiang, Samuel Cohen, Edward Grefenstette, Marc Peter Deisenroth</i></p>	30973
<p><i>A Statistical Framework for Personalized Federated Learning and Estimation: Theory, Algorithms, and Privacy</i></p> <p><i>Kaan Ozkara, Antonious M. Girgis, Deepesh Data, Suhas Diggavi</i></p>	30990
<p><i>The Asymmetric Maximum Margin Bias of Quasi-Homogeneous Neural Networks</i></p> <p><i>Daniel Kunin, Atsushi Yamamura, Chao Ma, Surya Ganguli</i></p>	31027

<p>Contrastive Audio-Visual Masked Autoencoder</p> <p><i>Yuan Gong, Andrew Rouditchenko, Alexander H. Liu, David Harwath, Leonid Karlinsky, Hilde Kuehne, James R. Glass</i></p>	31068
<p>Calibrating the Rigged Lottery: Making All Tickets Reliable</p> <p><i>Bowen Lei, Ruqi Zhang, Dongkuan Xu, Bani Mallick</i></p>	31097
<p>Reward Design with Language Models</p> <p><i>Minae Kwon, Sang Michael Xie, Kalesha Bullard, Dorsa Sadigh</i></p>	31120
<p>KwikBucks: Correlation Clustering with Cheap-Weak and Expensive-Strong Signals</p> <p><i>Sandeep Silwal, Sara Ahmadian, Andrew Nystrom, Andrew McCallum, Deepak Ramachandran, Seyed Mehran Kazemi</i></p>	31138
<p>Recitation-Augmented Language Models</p> <p><i>Zhiqing Sun, Xuezhi Wang, Yi Tay, Yiming Yang, Denny Zhou</i></p>	31171
<p>DreamFusion: Text-to-3D using 2D Diffusion</p> <p><i>Ben Poole, Ajay Jain, Jonathan T. Barron, Ben Mildenhall</i></p>	31194
<p>Language models are multilingual chain-of-thought reasoners</p> <p><i>Freda Shi, Mirac Suzgun, Markus Freitag, Xuezhi Wang, Suraj Srivats, Soroush Vosoughi, Hyung Won Chung, Yi Tay, Sebastian Ruder, Denny Zhou, Dipanjan Das, Jason Wei</i></p>	31214
<p>FIGARO: Controllable Music Generation using Learned and Expert Features</p> <p><i>Dimitri von R��tte, Luca Biggio, Yannic Kilcher, Thomas Hofmann</i></p>	31234
<p>On the duality between contrastive and non-contrastive self-supervised learning</p> <p><i>Quentin Garrido, Yubei Chen, Adrien Bardes, Laurent Najman, Yann LeCun</i></p>	31252
<p>Using Both Demonstrations and Language Instructions to Efficiently Learn Robotic Tasks</p> <p><i>Albert Yu, Ray Mooney</i></p>	31280
<p>HomoDistil: Homotopic Task-Agnostic Distillation of Pre-trained Transformers</p> <p><i>Chen Liang, Haoming Jiang, Zheng Li, Xianfeng Tang, Bing Yin, Tuo Zhao</i></p>	31302
<p>On The Specialization of Neural Modules</p> <p><i>Devon Jarvis, Richard Klein, Benjamin Rosman, Andrew M Saxe</i></p>	31321
<p>Learning MLPs on Graphs: A Unified View of Effectiveness, Robustness, and Efficiency</p> <p><i>Yijun Tian, Chuxu Zhang, Zhichun Guo, Xiangliang Zhang, Nitesh Chawla</i></p>	31352
<p>A VAE for Transformers with Nonparametric Variational Information Bottleneck</p> <p><i>James Henderson, Fabio James Fehr</i></p>	31364
<p>Learning where and when to reason in neuro-symbolic inference</p> <p><i>Cristina Cornelio, Jan Stuehmer, Shell Xu Hu, Timothy Hospedales</i></p>	31392

<p>Towards Understanding GD with Hard and Conjugate Pseudo-labels for Test-Time Adaptation</p> <p><i>Jun-Kun Wang, Andre Wibisono</i></p>	31412
<p>Serving Graph Compression for Graph Neural Networks</p> <p><i>Si Si, Felix Yu, Ankit Singh Rawat, Cho-Jui Hsieh, Sanjiv Kumar</i></p>	31427
<p>On Achieving Optimal Adversarial Test Error</p> <p><i>Justin D. Li, Matus Telgarsky</i></p>	31439
<p>ESCHER: Eschewing Importance Sampling in Games by Computing a History Value Function to Estimate Regret</p> <p><i>Stephen Marcus McAleer, Gabriele Farina, Marc Lanctot, Tuomas Sandholm</i></p>	31472
<p>Can Neural Networks Learn Implicit Logic from Physical Reasoning?</p> <p><i>Aaron Traylor, Roman Feiman, Ellie Pavlick</i></p>	31494
<p>Transformer-based World Models Are Happy With 100k Interactions</p> <p><i>Jan Robine, Marc Härtel, Tobias Uelwer, Stefan Harmeling</i></p>	31505
<p>Revisiting Intrinsic Reward for Exploration in Procedurally Generated Environments</p> <p><i>Kaixin Wang, Kuangqi Zhou, Bingyi Kang, Jiashi Feng, Shuicheng YAN</i></p>	31526
<p>LogicDP: Creating Labels for Graph Data via Inductive Logic Programming</p> <p><i>Yuan Yang, Faramarz Fekri, James Clayton Kerce, Ali Payani</i></p>	31546
<p>Dichotomy of Control: Separating What You Can Control from What You Cannot</p> <p><i>Sherry Yang, Dale Schuurmans, Pieter Abbeel, Ofir Nachum</i></p>	31561
<p>AGRO: Adversarial discovery of error-prone Groups for Robust Optimization</p> <p><i>Bhargavi Paranjape, Pradeep Dasigi, Vivek Srikumar, Luke Zettlemoyer, Hannaneh Hajishirzi</i></p>	31582
<p>TypeT5: Seq2seq Type Inference using Static Analysis</p> <p><i>Jiayi Wei, Greg Durrett, Isil Dillig</i></p>	31610
<p>Label Propagation with Weak Supervision</p> <p><i>Rattana Pukdee, Dylan Sam, Pradeep Kumar Ravikumar, Nina Balcan</i></p>	31629
<p>FunkNN: Neural Interpolation for Functional Generation</p> <p><i>AmirEhsan Khorashadizadeh, Anadi Chaman, Valentin Debarnot, Ivan Dokmanić</i></p>	31656
<p>Brain-like representational straightening of natural movies in robust feedforward neural networks</p> <p><i>Tahereh Toosi, Elias Issa</i></p>	31673
<p>Promptagator: Few-shot Dense Retrieval From 8 Examples</p> <p><i>Zhuyun Dai, Vincent Y Zhao, Ji Ma, Yi Luan, Jianmo Ni, Jing Lu, Anton Bakalov, Kelvin Guu, Keith Hall, Ming-Wei Chang</i></p>	31694

<p>Learning Structured Representations by Embedding Class Hierarchy</p> <p><i>Siqi Zeng, Remi Tachet des Combes, Han Zhao</i></p>	31716
<p>Is Forgetting Less a Good Inductive Bias for Forward Transfer?</p> <p><i>Jiefeng Chen, Timothy Nguyen, Dilan Gorur, Arslan Chaudhry</i></p>	31728
<p>Exploring perceptual straightness in learned visual representations</p> <p><i>Anne Harrington, Vasha DuTell, Ayush Tewari, Mark Hamilton, Simon Stent, Ruth Rosenholtz, William T. Freeman</i></p>	31753
<p>Efficient approximation of neural population structure and correlations with probabilistic circuits</p> <p><i>Koosha Khalvati, Samantha Johnson, Stefan Mihalas, Michael A Buice</i></p>	31782
<p>Neural Compositional Rule Learning for Knowledge Graph Reasoning</p> <p><i>Kewei Cheng, Nesreen Ahmed, Yizhou Sun</i></p>	31796
<p>Progressive Mix-Up for Few-Shot Supervised Multi-Source Domain Transfer</p> <p><i>Ronghang Zhu, Ronghang Zhu, Xiang Yu, Sheng Li</i></p>	31808
<p>A Simple Approach for Visual Room Rearrangement: 3D Mapping and Semantic Search</p> <p><i>Brandon Trabucco, Gunnar A Sigurdsson, Robinson Piramuthu, Gaurav S. Sukhatme, Ruslan Salakhutdinov</i></p>	31825
<p>The Onset of Variance-Limited Behavior for Networks in the Lazy and Rich Regimes</p> <p><i>Alexander Atanasov, Blake Bordelon, Sabarish Sainathan, Cengiz Pehlevan</i></p>	31843
<p>A Theoretical Framework for Inference and Learning in Predictive Coding Networks</p> <p><i>Beren Millidge, Yuhang Song, Tommaso Salvatori, Thomas Lukasiewicz, Rafal Bogacz</i></p>	31877
<p>Backpropagation at the Infinitesimal Inference Limit of Energy-Based Models: Unifying Predictive Coding, Equilibrium Propagation, and Contrastive Hebbian Learning</p> <p><i>Beren Millidge, Yuhang Song, Tommaso Salvatori, Thomas Lukasiewicz, Rafal Bogacz</i></p>	31901
<p>f-DM: A Multi-stage Diffusion Model via Progressive Signal Transformation</p> <p><i>Jiatao Gu, Shuangfei Zhai, Yizhe Zhang, Miguel Ángel Bautista, Joshua M. Susskind</i></p>	31915
<p>A Mixture-of-Expert Approach to RL-based Dialogue Management</p> <p><i>Yinlam Chow, Azamat Tulepbergenov, Ofir Nachum, Dhawal Gupta, Moonkyung Ryu, Mohammad Ghavamzadeh, Craig Boutilier</i></p>	31942
<p>FedExp: Speeding Up Federated Averaging via Extrapolation</p> <p><i>Divyansh Jhunjhunwala, Shiqiang Wang, Gauri Joshi</i></p>	31955
<p>Multiple sequence alignment as a sequence-to-sequence learning problem</p> <p><i>Edo Dotan, Yonatan Belinkov, Oren Avram, Elya Wygoda, Noa Ecker, Michael Alburquerque, Omri Keren, Gil Loewenthal, Tal Pupko</i></p>	31990

<p>Systematic Rectification of Language Models via Dead-end Analysis</p> <p><i>Meng Cao, Mehdi Fatemi, Jackie CK Cheung, Samira Shabanian</i></p>	32011
<p>Compositionality with Variation Reliably Emerges in Neural Networks</p> <p><i>Henry Conklin, Kenny Smith</i></p>	32032
<p>First Steps Toward Understanding the Extrapolation of Nonlinear Models to Unseen Domains</p> <p><i>Kefan Dong, Tengyu Ma</i></p>	32052
<p>Projective Proximal Gradient Descent for Nonconvex Nonsmooth Optimization: Fast Convergence Without Kurdyka-Lojasiewicz (KL) Property</p> <p><i>Yingzhen Yang, Ping Li</i></p>	32064
<p>Deep Learning From Crowdsourced Labels: Coupled Cross-Entropy Minimization, Identifiability, and Regularization</p> <p><i>Shahana Ibrahim, Tri Nguyen, Xiao Fu</i></p>	32086
<p>Progressive Prompts: Continual Learning for Language Models</p> <p><i>Anastasia Razdaibiedina, Yuning Mao, Rui Hou, Madian Khabisa, Mike Lewis, Amjad Almahairi</i></p>	32125
<p>GOGGLE: Generative Modelling for Tabular Data by Learning Relational Structure</p> <p><i>Tennison Liu, Zhaozhi Qian, Jeroen Berrevoets, Mihaela van der Schaar</i></p>	32148
<p>Interpretable Geometric Deep Learning via Learnable Randomness Injection</p> <p><i>Siqi Miao, Yunan Luo, Mia Liu, Pan Li</i></p>	32170
<p>How I Learned to Stop Worrying and Love Retraining</p> <p><i>Max Zimmer, Christoph Spiegel, Sebastian Pokutta</i></p>	32193
<p>Denoising Diffusion Samplers</p> <p><i>Francisco Vargas, Will Sussman Grathwohl, Arnaud Doucet</i></p>	32242
<p>Interpretations of Domain Adaptations via Layer Variational Analysis</p> <p><i>Huan-Hsin Tseng, Hsin-Yi Lin, Kuo-Hsuan Hung, Yu Tsao</i></p>	32272
<p>Provable Robustness against Wasserstein Distribution Shifts via Input Randomization</p> <p><i>Aounon Kumar, Alexander Levine, Tom Goldstein, Soheil Feizi</i></p>	32300
<p>Empowering Graph Representation Learning with Test-Time Graph Transformation</p> <p><i>Wei Jin, Tong Zhao, Jiayuan Ding, Yozen Liu, Jiliang Tang, Neil Shah</i></p>	32327
<p>How Does Semi-supervised Learning with Pseudo-labelers Work? A Case Study</p> <p><i>Yiwen Kou, Zixiang Chen, Yuan Cao, Quanquan Gu</i></p>	32354
<p>Learning Language Representations with Logical Inductive Bias</p> <p><i>Jianshu Chen</i></p>	32402
<p>Artificial Neuronal Ensembles with Learned Context Dependent Gating</p> <p><i>Matthew James Tilley, Michelle Miller, David Freedman</i></p>	32430

<p>Panning for Gold in Federated Learning: Targeted Text Extraction under Arbitrarily Large-Scale Aggregation</p> <p>Hong-Min Chu, Jonas Geiping, Liam H Fowl, Micah Goldblum, Tom Goldstein</p>	32443
<p>gDDIM: Generalized denoising diffusion implicit models</p> <p>Qinsheng Zhang, Molei Tao, Yongxin Chen</p>	32459
<p>Fast Sampling of Diffusion Models with Exponential Integrator</p> <p>Qinsheng Zhang, Yongxin Chen</p>	32490
<p>Characteristic Neural Ordinary Differential Equation</p> <p>Xingzi Xu, Ali Hasan, Khalil Elkhail, Jie Ding, Vahid Tarokh</p>	32523
<p>Efficiently Controlling Multiple Risks with Pareto Testing</p> <p>Bracha Laufer-Goldshtein, Adam Fisch, Regina Barzilay, Tommi S. Jaakkola</p>	32558
<p>Greedy Actor-Critic: A New Conditional Cross-Entropy Method for Policy Improvement</p> <p>Samuel Neumann, Sungsu Lim, Ajin George Joseph, Yangchen Pan, Adam White, Martha White</p>	32584
<p>Learning Harmonic Molecular Representations on Riemannian Manifold</p> <p>Yiqun Wang, Yuning Shen, Shi Chen, Lihao Wang, Fei YE, Hao Zhou</p>	32611
<p>Phase2vec: dynamical systems embedding with a physics-informed convolutional network</p> <p>Matt Ricci, Noa Moriel, Zoe Piran, Mor Nitzan</p>	32636
<p>Information Plane Analysis for Dropout Neural Networks</p> <p>Linara Adilova, Bernhard C Geiger, Asja Fischer</p>	32657
<p>TextGrad: Advancing Robustness Evaluation in NLP by Gradient-Driven Optimization</p> <p>Bairu Hou, Jinghan Jia, Yihua Zhang, Guanhua Zhang, Yang Zhang, Sijia Liu, Shiyu Chang</p>	32676
<p>BC-IRL: Learning Generalizable Reward Functions from Demonstrations</p> <p>Andrew Szot, Amy Zhang, Dhruv Batra, Zsolt Kira, Franziska Meier</p>	32693
<p>Understanding Influence Functions and Datamodels via Harmonic Analysis</p> <p>Nikunj Saunshi, Arushi Gupta, Mark Braverman, Sanjeev Arora</p>	32712
<p>Evolve Smoothly, Fit Consistently: Learning Smooth Latent Dynamics For Advection-Dominated Systems</p> <p>Zhong Yi Wan, Leonardo Zepeda-Nunez, Anudhyan Boral, Fei Sha</p>	32734
<p>Fundamental limits on the robustness of image classifiers</p> <p>Zheng Dai, David Gifford</p>	32759
<p>Robust Graph Dictionary Learning</p> <p>Weijie Liu, Jiahao Xie, Chao Zhang, Makoto Yamada, Nenggan Zheng, Hui Qian</p>	32788

<p>How Much Data Are Augmentations Worth? An Investigation into Scaling Laws, Invariance, and Implicit Regularization</p> <p><i>Jonas Geiping, Micah Goldblum, Gowthami Somepalli, Ravid Shwartz-Ziv, Tom Goldstein, Andrew Gordon Wilson</i></p>	32810
<p>Time Will Tell: New Outlooks and A Baseline for Temporal Multi-View 3D Object Detection</p> <p><i>Jinhyung Park, Chenfeng Xu, Shijia Yang, Kurt Keutzer, Kris M. Kitani, Masayoshi Tomizuka, Wei Zhan</i></p>	32841
<p>Binding Language Models in Symbolic Languages</p> <p><i>Zhoujun Cheng, Tianbao Xie, Peng Shi, Chengzu Li, Rahul Nadkarni, Yushi Hu, Caiming Xiong, Dragomir Radev, Mari Ostendorf, Luke Zettlemoyer, Noah A. Smith, Tao Yu</i></p>	32866
<p>On the Sensitivity of Reward Inference to Misspecified Human Models</p> <p><i>Joey Hong, Kush Bhatia, Anca Dragan</i></p>	32893
<p>Confidence-Conditioned Value Functions for Offline Reinforcement Learning</p> <p><i>Joey Hong, Aviral Kumar, Sergey Levine</i></p>	32910
<p>Learning Iterative Neural Optimizers for Image Steganography</p> <p><i>Xiangyu Chen, Varsha Kishore, Kilian Q Weinberger</i></p>	32926
<p>SIMPLE: A Gradient Estimator for k-Subset Sampling</p> <p><i>Kareem Ahmed, Zhe Zeng, Mathias Niepert, Guy Van den Broeck</i></p>	32947
<p>Information-Theoretic Diffusion</p> <p><i>Xianghao Kong, Rob Brekelmans, Greg Ver Steeg</i></p>	32963
<p>Momentum Stiefel Optimizer, with Applications to Suitably-Orthogonal Attention, and Optimal Transport</p> <p><i>Lingkai Kong, Yuqing Wang, Molei Tao</i></p>	32975
<p>Neural Causal Models for Counterfactual Identification and Estimation</p> <p><i>Kevin MUYUAN XIA, YUSHU PAN, ELIAS BAREINBOIM</i></p>	33005
<p>Replay Memory as An Empirical MDP: Combining Conservative Estimation with Experience Replay</p> <p><i>Hongming Zhang, Chenjun Xiao, Han Wang, Jun Jin, bo xu, Martin MÅ¼ller</i></p>	33019
<p>Random Laplacian Features for Learning with Hyperbolic Space</p> <p><i>Tao Yu, Christopher De Sa</i></p>	33046
<p>Transferable Unlearnable Examples</p> <p><i>Jie Ren, Han Xu, Yuxuan Wan, Xingjun Ma, Lichao Sun, Jiliang Tang</i></p>	33069
<p>Supervision Complexity and its Role in Knowledge Distillation</p> <p><i>Hrayr Harutyunyan, Ankit Singh Rawat, Aditya Krishna Menon, Seungyeon Kim, Sanjiv Kumar</i></p>	33085

<p>Fantastic Rewards and How to Tame Them: A Case Study on Reward Learning for Task-oriented Dialogue Systems</p> <p><i>Yihao Feng, Shentao Yang, Shujian Zhang, Jianguo Zhang, Caiming Xiong, Mingyuan Zhou, Huan Wang</i></p>	33110
<p>A Time Series is Worth 64 Words: Long-term Forecasting with Transformers</p> <p><i>Yuqi Nie, Nam H Nguyen, Phanwadee Sinthong, Jayant Kalagnanam</i></p>	33132
<p>When and Why Vision-Language Models Behave like Bags-Of-Words, and What to Do About It?</p> <p><i>Mert Yuksekgonul, Federico Bianchi, Pratyusha Kalluri, Dan Jurafsky, James Zou</i></p>	33156
<p>Effectively Modeling Time Series with Simple Discrete State Spaces</p> <p><i>Michael Zhang, Khaled Kamal Saab, Michael Poli, Tri Dao, Karan Goel, Christopher Re</i></p>	33176
<p>Energy-Inspired Self-Supervised Pretraining for Vision Models</p> <p><i>Ze Wang, Jiang Wang, Zicheng Liu, Qiang Qiu</i></p>	33214
<p>Self-Ensemble Protection: Training Checkpoints Are Good Data Protectors</p> <p><i>Sizhe Chen, Geng Yuan, Xinwen Cheng, Yifan Gong, Minghai Qin, Yanzhi Wang, Xiaolin Huang</i></p>	33231
<p>Decomposed Prompting: A Modular Approach for Solving Complex Tasks</p> <p><i>Tushar Khot, Harsh Trivedi, Matthew Finlayson, Yao Fu, Kyle Richardson, Peter Clark, Ashish Sabharwal</i></p>	33243
<p>Competitive Physics Informed Networks</p> <p><i>Qi Zeng, Yash Kothari, Spencer H Bryngelson, Florian Tobias Schaefer</i></p>	33312
<p>Equal Improbability: A New Fairness Notion Considering the Long-term Impact</p> <p><i>Ozgur Guldogan, Yuchen Zeng, Jy-yong Sohn, Ramtin Pedarsani, Kangwook Lee</i></p>	33329
<p>Composite Slice Transformer: An Efficient Transformer with Composition of Multi-Scale Multi-Range Attentions</p> <p><i>Mingu Lee, Saurabh Pitre, Tianyu Jiang, Pierre-David Letourneau, Matthew J Morse, Kanghwan Jang, Joseph Soriaga, Parham Noorzad, Hsin-Pai Cheng, Christopher Lott</i></p>	33360
<p>Dual Student Networks for Data-Free Model Stealing</p> <p><i>James Beetham, Navid Kardan, Ajmal Saeed Mian, Mubarak Shah</i></p>	33385
<p>Building Normalizing Flows with Stochastic Interpolants</p> <p><i>Michael Samuel Albergo, Eric Vanden-Eijnden</i></p>	33397
<p>Noise Is Not the Main Factor Behind the Gap Between Sgd and Adam on Transformers, But Sign Descent Might Be</p> <p><i>Frederik Kunstner, Jacques Chen, Jonathan Wilder Lavington, Mark Schmidt</i></p>	33426
<p>A Stable and Scalable Method for Solving Initial Value PDEs with Neural Networks</p> <p><i>Marc Anton Finzi, Andres Potapczynski, Matthew Choptuik, Andrew Gordon Wilson</i></p>	33454

A Control-Centric Benchmark for Video Prediction Stephen Tian, Chelsea Finn, Jiajun Wu	33468
Spotlight: Mobile UI Understanding using Vision-Language Models with a Focus Gang Li, Yang Li	33491
Forward Super-Resolution: How Can GANs Learn Hierarchical Generative Models for Real-World Distributions Zeyuan Allen-Zhu, Yuanzhi Li	33507
Towards Understanding Ensemble, Knowledge Distillation and Self-Distillation in Deep Learning Zeyuan Allen-Zhu, Yuanzhi Li	33520
Min-Max Multi-objective Bilevel Optimization with Applications in Robust Machine Learning Alex Gu, Songtao Lu, Parikshit Ram, Tsui-Wei Weng	33532
Continual Pre-training of Language Models Zixuan Ke, Yijia Shao, Haowei Lin, Tatsuya Konishi, Gyuhak Kim, Bing Liu	33568
$\mathcal{SE}(3)$-Equivariant Attention Networks for Shape Reconstruction in Function Space Evangelos Chatzipantazis, Stefanos Pertigkiozoglou, Edgar Dobriban, Kostas Daniilidis	33584
A new characterization of the edge of stability based on a sharpness measure aware of batch gradient distribution Sungyoon Lee, Cheongjae Jang	33613
OPTQ: Accurate Quantization for Generative Pre-trained Transformers Elias Frantar, Saleh Ashkboos, Torsten Hoefer, Dan Alistarh	33657
Discrete Predictor-Corrector Diffusion Models for Image Synthesis Jose Lezama, Tim Salimans, Lu Jiang, Huiwen Chang, Jonathan Ho, Irfan Essa	33673
Enhancing Meta Learning via Multi-Objective Soft Improvement Functions Runsheng Yu, Weiyu Chen, Xinrun Wang, James Kwok	33692
What learning algorithm is in-context learning? Investigations with linear models Ekin Akyüz, Dale Schuurmans, Jacob Andreas, Tengyu Ma, Denny Zhou	33712
Pruning Deep Neural Networks from a Sparsity Perspective Enmao Diao, Ganghua Wang, Jiawei Zhang, Yuhong Yang, Jie Ding, Vahid Tarokh	33741
Large Language Models are Human-Level Prompt Engineers Yongchao Zhou, Andrei Ioan Muresanu, Ziyen Han, Keiran Paster, Silviu Pitis, Harris Chan, Jimmy Ba	33753
Real-time variational method for learning neural trajectory and its dynamics Matthew Dowling, Yuan Zhao, Il Memming Park	33796

<p>Implicit regularization in Heavy-ball momentum accelerated stochastic gradient descent</p> <p>Avrajit Ghosh, He Lyu, Xitong Zhang, Rongrong Wang</p>	33814
<p>Implicit Regularization for Group Sparsity</p> <p>Jiangyuan Li, Thanh V Nguyen, Chinmay Hegde, Raymond K. W. Wong</p>	33825
<p>Diffusion-based Image Translation using disentangled style and content representation</p> <p>Gihyun Kwon, Jong Chul Ye</p>	33837
<p>Offline Q-learning on Diverse Multi-Task Data Both Scales And Generalizes</p> <p>Aviral Kumar, Rishabh Agarwal, Xinyang Geng, George Tucker, Sergey Levine</p>	33859
<p>Safe Exploration Incurs Nearly No Additional Sample Complexity for Reward-Free RL</p> <p>Ruiquan Huang, Jing Yang, Yingbin Liang</p>	33879
<p>Avoiding spurious correlations via logit correction</p> <p>Sheng Liu, Xu Zhang, Nitesh Sekhar, Yue Wu, Prateek Singhal, Carlos Fernandez-Granda</p>	33916
<p>Broken Neural Scaling Laws</p> <p>Ethan Caballero, Kshitij Gupta, Irina Rish, David Krueger</p>	33933
<p>Guiding Energy-based Models via Contrastive Latent Variables</p> <p>Hankook Lee, Jongheon Jeong, Sejun Park, Jinwoo Shin</p>	33965
<p>Adaptive Optimization in the ∞-Width Limit</p> <p>Etai Littwin, Greg Yang</p>	33983
<p>Decepticons: Corrupted Transformers Breach Privacy in Federated Learning for Language Models</p> <p>Liam H Fowl, Jonas Geiping, Steven Reich, Yuxin Wen, Wojciech Czaja, Micah Goldblum, Tom Goldstein</p>	34016
<p>Unsupervised Meta-learning via Few-shot Pseudo-supervised Contrastive Learning</p> <p>Huiwon Jang, Hankook Lee, Jinwoo Shin</p>	34039
<p>Unsupervised Learning for Combinatorial Optimization Needs Meta Learning</p> <p>Haoyu Peter Wang, Pan Li</p>	34057
<p>Reparameterization through Spatial Gradient Scaling</p> <p>Alexander Detkov, Mohammad Salameh, Muhammad Fetrat, Jialin Zhang, Robin Luwei, SHANGLING JUI, Di Niu</p>	34075
<p>Mind the Pool: Convolutional Neural Networks Can Overfit Input Size</p> <p>Bilal Alsallakh, David Yan, Narine Kokhlikyan, Vivek Miglani, Orion Reblitz-Richardson, Pamela Bhattacharya</p>	34094
<p>Diffusion-GAN: Training GANs with Diffusion</p> <p>Zhendong Wang, Huangjie Zheng, Pengcheng He, Weizhu Chen, Mingyuan Zhou</p>	34111

<p>REPAIR: REnormalizing Permuted Activations for Interpolation Repair</p> <p><i>Keller Jordan, Hanie Sedghi, Olga Saukh, Rahim Entezari, Behnam Neyshabur</i></p>	34137
<p>Compositional Task Representations for Large Language Models</p> <p><i>NAN SHAO, Zefan Cai, Hanwei xu, Chonghua Liao, Yanan Zheng, Zhilin Yang</i></p>	34154
<p>Federated Neural Bandits</p> <p><i>Zhongxiang Dai, Yao Shu, Arun Verma, Flint Xiaofeng Fan, Bryan Kian Hsiang Low, Patrick Jaillet</i></p>	34172
<p>Neural Image-based Avatars: Generalizable Radiance Fields for Human Avatar Modeling</p> <p><i>YoungJoong Kwon, Dahun Kim, Duygu Ceylan, Henry Fuchs</i></p>	34207
<p>Interaction-Based Disentanglement of Entities for Object-Centric World Models</p> <p><i>Akihiro Nakano, Masahiro Suzuki, Yutaka Matsuo</i></p>	34220
<p>Simple Emergent Action Representations from Multi-Task Policy Training</p> <p><i>Pu Hua, Yubei Chen, Huazhe Xu</i></p>	34241
<p>Weighted Clock Logic Point Process</p> <p><i>Ruixuan Yan, Yunshi Wen, Debarun Bhattacharjya, Ronny Luss, Tengfei Ma, Achille Fokoue, Anak Agung Julius</i></p>	34263
<p>Memorization Capacity of Neural Networks with Conditional Computation</p> <p><i>Erdem Koyuncu</i></p>	34287
<p>Faster Last-iterate Convergence of Policy Optimization in Zero-Sum Markov Games</p> <p><i>Shicong Cen, Yuejie Chi, Simon Shaolei Du, Lin Xiao</i></p>	34298
<p>Is the Performance of My Deep Network Too Good to Be True? A Direct Approach to Estimating the Bayes Error in Binary Classification</p> <p><i>Takashi Ishida, Ikko Yamane, Nontawat Charoenphakdee, Gang Niu, Masashi Sugiyama</i></p>	34340
<p>UniKGQA: Unified Retrieval and Reasoning for Solving Multi-hop Question Answering Over Knowledge Graph</p> <p><i>Jinhao Jiang, Kun Zhou, Xin Zhao, Ji-Rong Wen</i></p>	34362
<p>Causal Confusion and Reward Misidentification in Preference-Based Reward Learning</p> <p><i>Jeremy Tien, Jerry Zhi-Yang He, Zackory Erickson, Anca Dragan, Daniel S. Brown</i></p>	34378
<p>Bias Propagation in Federated Learning</p> <p><i>Hongyan Chang, Reza Shokri</i></p>	34405
<p>The In-Sample Softmax for Offline Reinforcement Learning</p> <p><i>Chenjun Xiao, Han Wang, Yangchen Pan, Adam White, Martha White</i></p>	34431
<p>Neural Bregman Divergences for Distance Learning</p> <p><i>Fred Lu, Edward Raff, Francis Ferraro</i></p>	34455

<p>Replicable Bandits</p> <p><i>Hossein Esfandiari, Alkis Kalavasis, Amin Karbasi, Andreas Krause, Vahab Mirrokni, Grigoris Velegkas</i></p>	34473
<p>Augmentation Component Analysis: Modeling Similarity via the Augmentation Overlaps</p> <p><i>Lu Han, Han-Jia Ye, De-Chuan Zhan</i></p>	34496
<p>SmartFRZ: An Efficient Training Framework using Attention-Based Layer Freezing</p> <p><i>Sheng Li, Geng Yuan, Yue Dai, Youtao Zhang, Yanzhi Wang, Xulong Tang</i></p>	34519
<p>Characterizing intrinsic compositionality in transformers with Tree Projections</p> <p><i>Shikhar Murty, Pratyusha Sharma, Jacob Andreas, Christopher D Manning</i></p>	34534
<p>Linearly Mapping from Image to Text Space</p> <p><i>Jack Merullo, Louis Castricato, Carsten Eickhoff, Ellie Pavlick</i></p>	34549
<p>GEASS: Neural causal feature selection for high-dimensional biological data</p> <p><i>Mingze Dong, Yuval Kluger</i></p>	34584
<p>Excess Risk of Two-Layer ReLU Neural Networks in Teacher-Student Settings and its Superiority to Kernel Methods</p> <p><i>Shunta Akiyama, Taiji Suzuki</i></p>	34603
<p>PINTO: Faithful Language Reasoning Using Prompt-Generated Rationales</p> <p><i>PeiFeng Wang, Aaron Chan, Filip Ilievski, Muhao Chen, Xiang Ren</i></p>	34632
<p>Learning Achievement Structure for Structured Exploration in Domains with Sparse Reward</p> <p><i>Zihan Zhou, Animesh Garg</i></p>	34649
<p>Stateful Active Facilitator: Coordination and Environmental Heterogeneity in Cooperative Multi-Agent Reinforcement Learning</p> <p><i>Dianbo Liu, Vedant Shah, Oussama Boussif, Cristian Meo, Anirudh Goyal, Tianmin Shu, Michael Curtis Mozer, Nicolas Heess, Yoshua Bengio</i></p>	34664
<p>The Role of Coverage in Online Reinforcement Learning</p> <p><i>Tengyang Xie, Dylan J Foster, Yu Bai, Nan Jiang, Sham M. Kakade</i></p>	34687
<p>The Tilted Variational Autoencoder: Improving Out-of-Distribution Detection</p> <p><i>Griffin Floto, Stefan Kremer, Mihai Nica</i></p>	34728
<p>Git Re-Basin: Merging Models modulo Permutation Symmetries</p> <p><i>Samuel Ainsworth, Jonathan Hayase, Siddhartha Srinivasa</i></p>	34746
<p>Quantile Risk Control: A Flexible Framework for Bounding the Probability of High-Loss Predictions</p> <p><i>Jake Snell, Thomas P Zollo, Zhun Deng, Toniann Pitassi, Richard Zemel</i></p>	34775
<p>HiCLIP: Contrastive Language-Image Pretraining with Hierarchy-aware Attention</p> <p><i>Shijie Geng, Jianbo Yuan, Yu Tian, Yuxiao Chen, Yongfeng Zhang</i></p>	34794

<p>Over-Training with Mixup May Hurt Generalization</p> <p><i>Zixuan Liu, Ziqiao Wang, Hongyu Guo, Yongyi Mao</i></p>	34815
<p>Minimalistic Unsupervised Representation Learning with the Sparse Manifold Transform</p> <p><i>Yubei Chen, Zeyu Yun, Yi Ma, Bruno Olshausen, Yann LeCun</i></p>	34835
<p>The Best of Both Worlds: Accurate Global and Personalized Models through Federated Learning with Data-Free Hyper-Knowledge Distillation</p> <p><i>Huancheng Chen, Chaining Wang, Haris Vikalo</i></p>	34859
<p>Simple initialization and parametrization of sinusoidal networks via their kernel bandwidth</p> <p><i>Filipe de Avila Belbute-Peres, J Zico Kolter</i></p>	34883
<p>Scale-invariant Bayesian Neural Networks with Connectivity Tangent Kernel</p> <p><i>SungYub Kim, Sihwan Park, Kyung-Su Kim, Eunho Yang</i></p>	34917
<p>On the Importance and Applicability of Pre-Training for Federated Learning</p> <p><i>Hong-You Chen, Cheng-Hao Tu, Ziwei Li, Han Wei Shen, Wei-Lun Chao</i></p>	34947
<p>On Representing Linear Programs by Graph Neural Networks</p> <p><i>Ziang Chen, Jialin Liu, Xinshang Wang, Wotao Yin</i></p>	34972
<p>On Representing Mixed-Integer Linear Programs by Graph Neural Networks</p> <p><i>Ziang Chen, Jialin Liu, Xinshang Wang, Wotao Yin</i></p>	35001
<p>Chasing All-Round Graph Representation Robustness: Model, Training, and Optimization</p> <p><i>Chunhui Zhang, Yijun Tian, Mingxuan Ju, Zheyuan Liu, Yanfang Ye, Nitesh Chawla, Chuxu Zhang</i></p>	35026
<p>Robust Algorithms on Adaptive Inputs from Bounded Adversaries</p> <p><i>Yeshwanth Cherapanamjeri, Sandeep Silwal, David Woodruff, Fred Zhang, Qiuyi Zhang, Samson Zhou</i></p>	35042
<p>Proto-Value Networks: Scaling Representation Learning with Auxiliary Tasks</p> <p><i>Jesse Farebrother, Joshua Greaves, Rishabh Agarwal, Charline Le Lan, Ross Goroshin, Pablo Samuel Castro, Marc G Bellemare</i></p>	35070
<p>Offline Reinforcement Learning with Differentiable Function Approximation is Provably Efficient</p> <p><i>Ming Yin, Mengdi Wang, Yu-Xiang Wang</i></p>	35092
<p>WikiWhy: Answering and Explaining Cause-and-Effect Questions</p> <p><i>Matthew Ho, Aditya Sharma, Justin Chang, Michael Saxon, Sharon Levy, Yujie Lu, William Yang Wang</i></p>	35143
<p>Ask Me Anything: A simple strategy for prompting language models</p> <p><i>Simran Arora, Avaniika Narayan, Mayee F Chen, Laurel Orr, Neel Guha, Kush Bhatia, Ines Chami, Christopher Re</i></p>	35162

<p>Efficient Deep Reinforcement Learning Requires Regulating Overfitting</p> <p>Qiyang Li, Aviral Kumar, Ilya Kostrikov, Sergey Levine</p>	35234
<p>TextShield: Beyond Successfully Detecting Adversarial Sentences in text classification</p> <p>Lingfeng Shen, Ze Zhang, Haiyun Jiang, Ying Chen</p>	35261
<p>Robust Multivariate Time-Series Forecasting: Adversarial Attacks and Defense Mechanisms</p> <p>Linbo Liu, Youngsuk Park, Trong Nghia Hoang, Hilaf Hasson, Luke Huan</p>	35279
<p>Masked Distillation with Receptive Tokens</p> <p>Tao Huang, Yuan Zhang, Shan You, Fei Wang, Chen Qian, Jian Cao, Chang Xu</p>	35297
<p>Understanding the Covariance Structure of Convolutional Filters</p> <p>Asher Trockman, Devin Willmott, J Zico Kolter</p>	35312
<p>Auto-Encoding Goodness of Fit</p> <p>Aaron Palmer, Zhiyi Chi, Derek Aguiar, Jinbo Bi</p>	35341
<p>Recursive Time Series Data Augmentation</p> <p>Amine Mohamed Aboussalah, Minjae Kwon, Raj G Patel, Cheng Chi, Chi-Guhn Lee</p>	35384
<p>Language Models Are Greedy Reasoners: A Systematic Formal Analysis of Chain-of-Thought</p> <p>Abulhair Saparov, He He</p>	35413
<p>Specformer: Spectral Graph Neural Networks Meet Transformers</p> <p>Deyu Bo, Chuan Shi, Lele Wang, Renjie Liao</p>	35433
<p>STUNT: Few-shot Tabular Learning with Self-generated Tasks from Unlabeled Tables</p> <p>Jaehyun Nam, Jihoon Tack, Kyungmin Lee, Hankook Lee, Jinwoo Shin</p>	35451
<p>CktGNN: Circuit Graph Neural Network for Electronic Design Automation</p> <p>Zehao Dong, Weidong Cao, Muhan Zhang, Dacheng Tao, Yixin Chen, Xuan Zhang</p>	35470
<p>Can Agents Run Relay Race with Strangers? Generalization of RL to Out-of-Distribution Trajectories</p> <p>Li-Cheng Lan, Huan Zhang, Cho-Jui Hsieh</p>	35489
<p>Training-Free Structured Diffusion Guidance for Compositional Text-to-Image Synthesis</p> <p>Weixi Feng, Xuehai He, Tsu-Jui Fu, Varun Jampani, Arjun Reddy Akula, Pradyumna Narayana, Sugato Basu, Xin Eric Wang, William Yang Wang</p>	35506
<p>Neural Systematic Binder</p> <p>Gautam Singh, Yeongbin Kim, Sungjin Ahn</p>	35527
<p>Editing models with task arithmetic</p> <p>Gabriel Ilharco, Marco Tulio Ribeiro, Mitchell Wortsman, Ludwig Schmidt, Hannaneh Hajishirzi, Ali Farhadi</p>	35554

<p>What Makes Convolutional Models Great on Long Sequence Modeling?</p> <p><i>Yuhong Li, Tianle Cai, Yi Zhang, Deming Chen, Debadeepta Dey</i></p>	35585
<p>Understanding and Adopting Rational Behavior by Bellman Score Estimation</p> <p><i>Kuno Kim, Stefano Ermon</i></p>	35604
<p>Pareto Invariant Risk Minimization: Towards Mitigating the Optimization Dilemma in Out-of-Distribution Generalization</p> <p><i>Yongqiang Chen, Kaiwen Zhou, Yatao Bian, Binghui Xie, Bingzhe Wu, Yonggang Zhang, MA KAILI, Han Yang, Peilin Zhao, Bo Han, James Cheng</i></p>	35616
<p>A Closer Look at Model Adaptation using Feature Distortion and Simplicity Bias</p> <p><i>Puja Trivedi, Danai Koutra, Jayaraman J. Thiagarajan</i></p>	35666
<p>Computational Language Acquisition with Theory of Mind</p> <p><i>Andy Liu, Hao Zhu, Emmy Liu, Yonatan Bisk, Graham Neubig</i></p>	35683
<p>Combinatorial Pure Exploration of Causal Bandits</p> <p><i>Nuoya Xiong, Wei Chen</i></p>	35695
<p>A theoretical study of inductive biases in contrastive learning</p> <p><i>Jeff Z. HaoChen, Tengyu Ma</i></p>	35706
<p>Seeing Differently, Acting Similarly: Heterogeneously Observable Imitation Learning</p> <p><i>Xin-Qiang Cai, Yao-Xiang Ding, Zixuan Chen, Yuan Jiang, Masashi Sugiyama, Zhi-Hua Zhou</i></p>	35724
<p>Imitating Graph-Based Planning with Goal-Conditioned Policies</p> <p><i>Junsu Kim, Younggyo Seo, Sungsoo Ahn, Kyunghwan Son, Jinwoo Shin</i></p>	35745
<p>Learning to reason over visual objects</p> <p><i>Shanka Subhra Mondal, Taylor Whittington Webb, Jonathan Cohen</i></p>	35764
<p>On Pre-training Language Model for Antibody</p> <p><i>Danqing Wang, Fei YE, Hao Zhou</i></p>	35786
<p>Rethinking Symbolic Regression: Morphology and Adaptability in the Context of Evolutionary Algorithms</p> <p><i>Kei Sen Fong, Shelvia Wongso, Mehul Motani</i></p>	35811
<p>GNNInterpreter: A Probabilistic Generative Model-Level Explanation for Graph Neural Networks</p> <p><i>Xiaoqi Wang, Han Wei Shen</i></p>	35827
<p>UniMax: Fairer and More Effective Language Sampling for Large-Scale Multilingual Pretraining</p> <p><i>Hyung Won Chung, Xavier Garcia, Adam Roberts, Yi Tay, Orhan Firat, Sharan Narang, Noah Constant</i></p>	35852
<p>FLIP: A Provable Defense Framework for Backdoor Mitigation in Federated Learning</p> <p><i>Kaiyuan Zhang, Guan hong Tao, Qiuling Xu, Siyuan Cheng, Shengwei An, Yingqi Liu, Shiwei Feng, Guangyu Shen, Pin-Yu Chen, Shiqing Ma, Xiangyu Zhang</i></p>	35871

<p>Sparse Distributed Memory is a Continual Learner</p> <p>Trenton Bricken, Xander Davies, Deepak Singh, Dmitry Krotov, Gabriel Kreiman</p>	35897
<p>Ordered GNN: Ordering Message Passing to Deal with Heterophily and Over-smoothing</p> <p>Yunchong Song, Chenghu Zhou, Xinbing Wang, Zhouhan Lin</p>	35954
<p>A Kernel Perspective of Skip Connections in Convolutional Networks</p> <p>Daniel Barzilai, Amnon Geifman, Meirav Galun, Ronen Basri</p>	35972
<p>Improved Training of Physics-Informed Neural Networks Using Energy-Based Priors: a Study on Electrical Impedance Tomography</p> <p>Akarsh Pokkunuru, Pedram Rooshenas, Thilo Strauss, Anuj Abhishek, Taufiqar Khan</p>	36010
<p>Multi-Objective Online Learning</p> <p>Jiyan Jiang, Wenpeng Zhang, Shiji Zhou, Lihong Gu, Xiaodong Zeng, Wenwu Zhu</p>	36030
<p>TrojText: Test-time Invisible Textual Trojan Insertion</p> <p>Qian Lou, Yepeng Liu, Bo Feng</p>	36060
<p>Safe Reinforcement Learning From Pixels Using a Stochastic Latent Representation</p> <p>Yannick Hogewind, Thiago D. Simão, Tal Kachman, Nils Jansen</p>	36072
<p>Mid-Vision Feedback</p> <p>Michael Maynard, Eadom T Dessalene, Cornelia Fermuller, Yiannis Aloimonos</p>	36086
<p>Coupled Multiwavelet Operator Learning for Coupled Differential Equations</p> <p>Xiongye Xiao, Defu Cao, Ruochen Yang, Gaurav Gupta, Gengshuo Liu, Chenzhong Yin, Radu Balan, Paul Bogdan</p>	36104
<p>Backpropagation through Combinatorial Algorithms: Identity with Projection Works</p> <p>Subham Sekhar Sahoo, Anselm Paulus, Marin Vlastelica, Václav Musil, Volodymyr Kuleshov, Georg Martius</p>	36125
<p>Feature selection and low test error in shallow low-rotation ReLU networks</p> <p>Matus Telgarsky</p>	36152
<p>Bitrate-Constrained DRO: Beyond Worst Case Robustness To Unknown Group Shifts</p> <p>Amrith Setlur, Don Dennis, Benjamin Eysenbach, Aditi Raghunathan, Chelsea Finn, Virginia Smith, Sergey Levine</p>	36190
<p>CASR: Generating Complex Sequences with Autoregressive Self-Boost Refinement</p> <p>Hongwei Han, Mengyu Zhou, Shi Han, Xiu Li, Dongmei Zhang</p>	36203
<p>UL2: Unifying Language Learning Paradigms</p> <p>Yi Tay, Mostafa Dehghani, Vinh Q. Tran, Xavier Garcia, Jason Wei, Xuezhi Wang, Hyung Won Chung, Dara Bahri, Tal Schuster, Steven Zheng, Denny Zhou, Neil Houlsby, Donald Metzler</p>	36221

<p>Arbitrary Virtual Try-on Network: Characteristics Representation and Trade-off between Body and Clothing</p> <p><i>Yu Liu, Mingbo Zhao, Zhao Zhang, Jicong Fan, Yang Lou, Shuicheng Yan</i></p>	36254
<p>Efficient Certified Training and Robustness Verification of Neural ODEs</p> <p><i>Mustafa Zeqiri, Mark Niklas Mueller, Marc Fischer, Martin Vechev</i></p>	36277
<p>MA-BERT: Towards Matrix Arithmetic-only BERT Inference by Eliminating Complex Non-Linear Functions</p> <p><i>Neo Wei Ming, Zhehui Wang, Cheng Liu, Rick Siow Mong Goh, Tao Luo</i></p>	36291
<p>Improving Differentiable Neural Architecture Search by Encouraging Transferability</p> <p><i>Parth Sheth, Pengtao Xie</i></p>	36307
<p>Scaling Up Probabilistic Circuits by Latent Variable Distillation</p> <p><i>Anji Liu, Honghua Zhang, Guy Van den Broeck</i></p>	36347
<p>Selective Frequency Network for Image Restoration</p> <p><i>Yuning Cui, Yi Tao, Zhenshan Bing, Wenqi Ren, Xinwei Gao, Xiaochun Cao, Kai Huang, Alois Knoll</i></p>	36361
<p>Transfer NAS with Meta-learned Bayesian Surrogates</p> <p><i>Gresa Shala, Thomas Elsken, Frank Hutter, Josif Grabocka</i></p>	36374
<p>Hierarchical Abstraction for Combinatorial Generalization in Object Rearrangement</p> <p><i>Michael Chang, Alyssa Li Dayan, Franziska Meier, Thomas L. Griffiths, Sergey Levine, Amy Zhang</i></p>	36391
<p>Regression with Label Differential Privacy</p> <p><i>Badih Ghazi, Pritish Kamath, Ravi Kumar, Ethan Leeman, Pasin Manurangsi, Avinash Varadarajan, Chiyuan Zhang</i></p>	36404
<p>Certified Training: Small Boxes are All You Need</p> <p><i>Mark Niklas Mueller, Franziska Eckert, Marc Fischer, Martin Vechev</i></p>	36432
<p>Learning with Logical Constraints but without Shortcut Satisfaction</p> <p><i>Zenan Li, Zehua Liu, Yuan Yao, Jingwei Xu, Taolue Chen, Xiaoxing Ma, Jian Li</i></p>	36446
<p>Leveraging Large Language Models for Multiple Choice Question Answering</p> <p><i>Joshua Robinson, David Wingate</i></p>	36470
<p>Fast Nonlinear Vector Quantile Regression</p> <p><i>Aviv A. Rosenberg, Sanketh Vedula, Yaniv Romano, Alexander Bronstein</i></p>	36498
<p>Guarded Policy Optimization with Imperfect Online Demonstrations</p> <p><i>Zhenghai Xue, Zhenghao Peng, Quanyi Li, Zhihan Liu, Bolei Zhou</i></p>	36533
<p>Preserving Pre-trained Features Helps Calibrate Fine-tuned Language Models</p> <p><i>Guande He, Jianfei Chen, Jun Zhu</i></p>	36553

<p>RandProx: Primal-Dual Optimization Algorithms with Randomized Proximal Updates</p> <p><i>Laurent Condat, Peter Richtárik</i></p>	36574
<p>Learning in temporally structured environments</p> <p><i>Matt Jones, Tyler R. Scott, Mengye Ren, Gamaleldin Fathy Elsayed, Katherine Hermann, David Mayo, Michael Curtis Mozer</i></p>	36601
<p>Protein Sequence and Structure Co-Design with Equivariant Translation</p> <p><i>Chence Shi, Chuanrui Wang, Jiarui Lu, Bozitao Zhong, Jian Tang</i></p>	36620
<p>Gray-Box Gaussian Processes for Automated Reinforcement Learning</p> <p><i>Gresa Shala, André Biedenkapp, Frank Hutter, Josif Grabocka</i></p>	36640
<p>Implicit Bias in Leaky ReLU Networks Trained on High-Dimensional Data</p> <p><i>Spencer Frei, Gal Vardi, Peter Bartlett, Nathan Srebro, Wei Hu</i></p>	36677
<p>Evolving Populations of Diverse RL Agents with MAP-Elites</p> <p><i>Thomas PIERROT, Arthur Flajolet</i></p>	36721
<p>Individual Privacy Accounting with Gaussian Differential Privacy</p> <p><i>Antti Koskela, Marlon Tobaben, Antti Honkela</i></p>	36741
<p>LiftedCL: Lifting Contrastive Learning for Human-Centric Perception</p> <p><i>Ziwei Chen, Qiang Li, Xiaofeng Wang, Wankou Yang</i></p>	36766
<p>Confidence-Based Feature Imputation for Graphs with Partially Known Features</p> <p><i>Daeho Um, Jiwoong Park, Seulki Park, Jin young Choi</i></p>	36783
<p>Asynchronous Distributed Bilevel Optimization</p> <p><i>Yang Jiao, Kai Yang, Tiancheng Wu, Dongjin Song, Chengtao Jian</i></p>	36811
<p>Uniform-in-time propagation of chaos for the mean-field gradient Langevin dynamics</p> <p><i>Taiji Suzuki, Atsushi Nitanda, Denny Wu</i></p>	36848
<p>Neuroevolution is a Competitive Alternative to Reinforcement Learning for Skill Discovery</p> <p><i>Felix Chalumeau, Raphael Boige, Bryan Lim, Valentin Macé, Maxime Allard, Arthur Flajolet, Antoine Cully, Thomas PIERROT</i></p>	36885
<p>Zeroth-Order Optimization with Trajectory-Informed Derivative Estimation</p> <p><i>Yao Shu, Zhongxiang Dai, Weicong Sng, Arun Verma, Patrick Jaillet, Bryan Kian Hsiang Low</i></p>	36920
<p>A Message Passing Perspective on Learning Dynamics of Contrastive Learning</p> <p><i>Yifei Wang, Qi Zhang, Tianqi Du, Jiansheng Yang, Zhouchen Lin, Yisen Wang</i></p>	36947
<p>Geometrically regularized autoencoders for non-Euclidean data</p> <p><i>Cheongjae Jang, Yonghyeon Lee, Yung-Kyun Noh, Frank C. Park</i></p>	36963
<p>Concept-level Debugging of Part-Prototype Networks</p> <p><i>Andrea Bontempelli, Stefano Teso, Katya Tentori, Fausto Giunchiglia, Andrea Passerini</i></p>	36998

<p>Mole-BERT: Rethinking Pre-training Graph Neural Networks for Molecules</p> <p><i>Jun Xia, Chengshuai Zhao, Bozhen Hu, Zhangyang Gao, Cheng Tan, Yue Liu, Siyuan Li, Stan Z. Li</i></p>	37019
<p>Modeling content creator incentives on algorithm-curated platforms</p> <p><i>Jiri Hron, Karl Krauth, Michael Jordan, Niki Kilbertus, Sarah Dean</i></p>	37037
<p>Actionable Neural Representations: Grid Cells from Minimal Constraints</p> <p><i>Will Dorrell, Peter E. Latham, Timothy E. J. Behrens, James C. R. Whittington</i></p>	37059
<p>Behavior Proximal Policy Optimization</p> <p><i>Zifeng Zhuang, Kun LEI, Jinxin Liu, Donglin Wang, Yilang Guo</i></p>	37106
<p>Can BERT Refrain from Forgetting on Sequential Tasks? A Probing Study</p> <p><i>Mingxu Tao, Yansong Feng, Dongyan Zhao</i></p>	37126
<p>Iterative Circuit Repair Against Formal Specifications</p> <p><i>Matthias Cosler, Frederik Schmitt, Christopher Hahn, Bernd Finkbeiner</i></p>	37145
<p>Boosting Causal Discovery via Adaptive Sample Reweighting</p> <p><i>An Zhang, Fangfu Liu, Wenchang Ma, Zhibo Cai, Xiang Wang, Tat-Seng Chua</i></p>	37168
<p>Unveiling the sampling density in non-uniform geometric graphs</p> <p><i>Raffaele Paolino, Aleksandar Bojchevski, Stephan Gnnemann, Gitta Kutyniok, Ron Levie</i></p>	37191
<p>Decision S4: Efficient Sequence-Based RL via State Spaces Layers</p> <p><i>Shmuel Bar David, Itamar Zimmerman, Eliya Nachmani, Lior Wolf</i></p>	37219
<p>Cross-Layer Retrospective Retrieving via Layer Attention</p> <p><i>Yanwen Fang, Yuxi CAI, Jintai Chen, Jingyu Zhao, Guangjian Tian, Guodong Li</i></p>	37240
<p>Learning Group Importance using the Differentiable Hypergeometric Distribution</p> <p><i>Thomas M. Sutter, Laura Manduchi, Alain Ryser, Julia E Vogt</i></p>	37266
<p>Making Substitute Models More Bayesian Can Enhance Transferability of Adversarial Examples</p> <p><i>Qizhang Li, Yiwen Guo, Wangmeng Zuo, Hao Chen</i></p>	37295
<p>Learning a Data-Driven Policy Network for Pre-Training Automated Feature Engineering</p> <p><i>Liyao Li, Haobo Wang, Liangyu Zha, Qingyi Huang, Sai Wu, Gang Chen, Junbo Zhao</i></p>	37311
<p>Don't fear the unlabelled: safe semi-supervised learning via debiasing</p> <p><i>Hugo Schmutz, Olivier HUMBERT, Pierre-Alexandre Mattei</i></p>	37330
<p>Online Bias Correction for Task-Free Continual Learning</p> <p><i>Aristotelis Chrysakis, Marie-Francine Moens</i></p>	37372
<p>FairGBM: Gradient Boosting with Fairness Constraints</p> <p><i>Andr Cruz, Catarina G Belm, Joo Bravo, Pedro Saleiro, Pedro Bizarro</i></p>	37386

<p>Few-Shot Domain Adaptation For End-to-End Communication</p> <p>Jayaram Raghuram, Yijing Zeng, Dolores Garcia, Rafael Ruiz, Somesh Jha, Joerg Widmer, Suman Banerjee</p>	37409
<p>CodeBPE: Investigating Subtokenization Options for Large Language Model Pretraining on Source Code</p> <p>Nadezhda Chirkova, Sergey Troshin</p>	37444
<p>Task-Aware Information Routing from Common Representation Space in Lifelong Learning</p> <p>Prashant Shivaram Bhat, Bahram Zonooz, Elahe Arani</p>	37462
<p>EquiMod: An Equivariance Module to Improve Visual Instance Discrimination</p> <p>Alexandre DEVILLERS, Mathieu Lefort</p>	37482
<p>Share Your Representation Only: Guaranteed Improvement of the Privacy-Utility Tradeoff in Federated Learning</p> <p>Zebang Shen, Jiayuan Ye, Anmin Kang, Hamed Hassani, Reza Shokri</p>	37496
<p>Humanly Certifying Superhuman Classifiers</p> <p>Qiongkai Xu, Christian Walder, Chenchen Xu</p>	37531
<p>LS-IQ: Implicit Reward Regularization for Inverse Reinforcement Learning</p> <p>Firas Al-Hafez, Davide Tateo, Oleg Arenz, Guoping Zhao, Jan Peters</p>	37548
<p>WiNeRT: Towards Neural Ray Tracing for Wireless Channel Modelling and Differentiable Simulations</p> <p>Tribhuvanesh Orekondy, Pratik Kumar, Shreya Kadambi, Hao Ye, Joseph Soriaga, Arash Behboodi</p>	37574
<p>Curriculum-based Co-design of Morphology and Control of Voxel-based Soft Robots</p> <p>Yuxing Wang, Shuang Wu, Haobo Fu, QIANG FU, Tiantian Zhang, Yongzhe Chang, Xueqian Wang</p>	37593
<p>Solving stochastic weak Minty variational inequalities without increasing batch size</p> <p>Thomas Pethick, Olivier Fercoq, Puya Latafat, Panagiotis Patrinos, Volkan Cevher</p>	37613
<p>QAID: Question Answering Inspired Few-shot Intent Detection</p> <p>Asaf Yehudai, Matan Vetzler, Yosi Mass, Koren Lazar, Doron Cohen, Boaz Carmeli</p>	37647
<p>Transformer-based model for symbolic regression via joint supervised learning</p> <p>Wenqiang Li, Weijun Li, Linjun Sun, Min Wu, Lina Yu, Jingyi Liu, Yanjie Li, Songsong Tian</p>	37660
<p>Learning Uncertainty for Unknown Domains with Zero-Target-Assumption</p> <p>Yu Yu, Hassan Sajjad, Jia Xu</p>	37675
<p>Mini-batch k-means terminates within $O(d/\epsilon)$ iterations</p> <p>Gregory Schwartzman</p>	37691
<p>Human-Guided Fair Classification for Natural Language Processing</p> <p>Florian E. Dörner, Momchil Peychev, Nikola Konstantinov, Naman Goel, Elliott Ash, Martin Vechev</p>	37701

<p>Encoding Recurrence into Transformers</p> <p><i>Feiqing Huang, Kexin Lu, Yuxi CAI, Zhen Qin, Yanwen Fang, Guangjian Tian, Guodong Li</i></p>	37731
<p>Softened Symbol Grounding for Neuro-symbolic Systems</p> <p><i>Zenan Li, Yuan Yao, Taolue Chen, Jingwei Xu, Chun Cao, Xiaoxing Ma, Jian L^u</i></p>	37754
<p>Learning Continuous Normalizing Flows For Faster Convergence To Target Distribution via Ascent Regularizations</p> <p><i>Shuangshuang Chen, Sihao Ding, Yiannis Karayiannidis, Mårten Björkman</i></p>	37772
<p>Finding Actual Descent Directions for Adversarial Training</p> <p><i>Fabian Latorre, Igor Krawczuk, Leello Tadesse Dadi, Thomas Pethick, Volkan Cevher</i></p>	37797
<p>TVSPrupe - Pruning Non-discriminative filters via Total Variation separability of intermediate representations without fine tuning</p> <p><i>Chaitanya Murti, Tanay Narshana, Chiranjib Bhattacharyya</i></p>	37814
<p>DFPC: Data flow driven pruning of coupled channels without data.</p> <p><i>Tanay Narshana, Chaitanya Murti, Chiranjib Bhattacharyya</i></p>	37840
<p>Factorized Fourier Neural Operators</p> <p><i>Alasdair Tran, Alexander Mathews, Lexing Xie, Cheng Soon Ong</i></p>	37865
<p>Quantifying and Mitigating the Impact of Label Errors on Model Disparity Metrics</p> <p><i>Julius Adebayo, Melissa Hall, Bowen Yu, Bobbie Chern</i></p>	37882