

8th International Conference on Learning Representations (ICLR 2020)

Addis Ababa, Ethiopia
30 April 2020

Volume 1 of 17

ISBN: 978-1-7138-7274-0

Printed from e-media with permission by:

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



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

Copyright© (2020) by International Conference on Learning Representations
All rights reserved.

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

For permission requests, please contact International Conference on Learning Representations
at the address below.

International Conference on Learning Representations
2710 E Corridor Drive
Appleton, WI 54913

www.iclr.cc

Additional copies of this publication are available from:

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

TABLE OF CONTENTS

VOLUME 1

POSTERS PRESENTATIONS

PAY ATTENTION TO FEATURES, TRANSFER LEARN FASTER CNNs.....	1
<i>Kafeng Wang, Xitong Gao, Yiren Zhao, Xingjian Li, Dejing Dou, Cheng-Zhong Xu</i>	
GRADIENTS AS FEATURES FOR DEEP REPRESENTATION LEARNING	15
<i>Fangzhou Mu, Yingyu Liang, Yin Li</i>	
MONOTONIC MULTIHEAD ATTENTION.....	27
<i>Xutai Ma, Juan Miguel Pino, James Cross, Liezl Puzon, Jiatao Gu</i>	
MASSIVELY MULTILINGUAL SPARSE WORD REPRESENTATIONS.....	42
<i>Gábor Berend</i>	
QUERY-EFFICIENT META ATTACK TO DEEP NEURAL NETWORKS	58
<i>Jiawei Du, Hu Zhang, Joey Tianyi Zhou, Yi Yang, Jiashi Feng</i>	
BREAKING CERTIFIED DEFENSES: SEMANTIC ADVERSARIAL EXAMPLES WITH SPOOFED ROBUSTNESS CERTIFICATES	73
<i>Amin Ghiasi, Ali Shafahi, Tom Goldstein</i>	
ENABLING DEEP SPIKING NEURAL NETWORKS WITH HYBRID CONVERSION AND SPIKE TIMING DEPENDENT BACKPROPAGATION	89
<i>Nitin Rathi, Gopalakrishnan Srinivasan, Priyadarshini Panda, Kaushik Roy</i>	
HOW TO OWN THE NAS IN YOUR SPARE TIME	103
<i>Sanghyun Hong, Michael Davinroy, Yigitcan Kaya, Dana Dachman-Soled, Tudor Dumitras</i>	
THE SHAPE OF DATA: INTRINSIC DISTANCE FOR DATA DISTRIBUTIONS	119
<i>Anton Tsitsulin, Marina Munkhoeva, Davide Mottin, Panagiotis Karras, Alex Bronstein, Ivan Oseledets, Emmanuel Mueller</i>	
UNDERSTANDING GENERALIZATION IN RECURRENT NEURAL NETWORKS	139
<i>Zhuozhuo Tu, Fengxiang He, Dacheng Tao</i>	
CONSERVATIVE UNCERTAINTY ESTIMATION BY FITTING PRIOR NETWORKS.....	156
<i>Kamil Ciosek, Vincent Fortuin, Ryota Tomioka, Katja Hofmann, Richard Turner</i>	
NAS-BENCH-1SHOT1: BENCHMARKING AND DISSECTING ONE-SHOT NEURAL ARCHITECTURE SEARCH.....	177
<i>Arber Zela, Julien Siems, Frank Hutter</i>	
LEARNING TO COORDINATE MANIPULATION SKILLS VIA SKILL BEHAVIOR DIVERSIFICATION	197
<i>Youngwoon Lee, Jingyun Yang, Joseph J. Lim</i>	
ROBUST SUBSPACE RECOVERY LAYER FOR UNSUPERVISED ANOMALY DETECTION.....	212
<i>Chieh-Hsin Lai, Dongmian Zou, Gilad Lerman</i>	
LEARNING NEARLY DECOMPOSABLE VALUE FUNCTIONS VIA COMMUNICATION MINIMIZATION.....	240
<i>Tonghan Wang*, Jianhao Wang*, Chongyi Zheng, Chongjie Zhang</i>	
EXTREME CLASSIFICATION VIA ADVERSARIAL SOFTMAX APPROXIMATION.....	255
<i>Robert Bamler, Stephan Mandt</i>	
INFORMATION GEOMETRY OF ORTHOGONAL INITIALIZATIONS AND TRAINING.....	269
<i>Piotr Aleksander Sokól, Il Memming Park</i>	
MIXED PRECISION DNNs: ALL YOU NEED IS A GOOD PARAMETRIZATION.....	286
<i>Stefan Uhlich, Lukas Mauch, Fabien Cardinaux, Kazuki Yoshiyama, Javier Alonso Garcia, Stephen Tiedemann, Thomas Kemp, Akira Nakamura</i>	

CO-ATTENTIVE EQUIVARIANT NEURAL NETWORKS: FOCUSING EQUIVARIANCE ON TRANSFORMATIONS CO-OCCURRING IN DATA.....	307
<i>David W. Romero, Mark Hoogendoorn</i>	
DEEP ORIENTATION UNCERTAINTY LEARNING BASED ON A BINGHAM LOSS	320
<i>Igor Gilitschenski, Roshni Sahoo, Wilko Schwarting, Alexander Amini, Sertac Karaman, Daniela Rus</i>	
CRITICAL INITIALISATION IN CONTINUOUS APPROXIMATIONS OF BINARY NEURAL NETWORKS.....	332
<i>George Stamatescu, Federica Gerace, Carlo Lucibello, Ian Fuss, Langford White</i>	
SUB-POLICY ADAPTATION FOR HIERARCHICAL REINFORCEMENT LEARNING	355
<i>Alexander Li, Carlos Florensa, Ignasi Clavera, Pieter Abbeel</i>	
EPISODIC REINFORCEMENT LEARNING WITH ASSOCIATIVE MEMORY	370
<i>Guangxiang Zhu*, Zichuan Lin*, Guangwen Yang, Chongjie Zhang</i>	
PITFALLS OF IN-DOMAIN UNCERTAINTY ESTIMATION AND ENSEMBLING IN DEEP LEARNING.....	385
<i>Arsenii Ashukha, Alexander Lyzhov, Dmitry Molchanov, Dmitry Vetrov</i>	
DIFFTAICHI: DIFFERENTIABLE PROGRAMMING FOR PHYSICAL SIMULATION.....	415
<i>Yuanming Hu, Luke Anderson, Tzu-Mao Li, Qi Sun, Nathan Carr, Jonathan Ragan-Kelley, Fredo Durand</i>	
DOMAIN ADAPTIVE MULTIBRANCH NETWORKS	435
<i>Róger Bermúdez-Chacón, Mathieu Salzmann, Pascal Fua</i>	
PREDICTION POISONING: TOWARDS DEFENSES AGAINST DNN MODEL STEALING ATTACKS	454
<i>Tribhuvanesh Orekondy, Bernt Schiele, Mario Fritz</i>	
SCALOR: GENERATIVE WORLD MODELS WITH SCALABLE OBJECT REPRESENTATIONS.....	471
<i>Jindong Jiang*, Sepehr Janghorbani*, Gerard De Melo, Sungjin Ahn</i>	
NEURAL TANGENT KERNELS, TRANSPORTATION MAPPINGS, AND UNIVERSAL APPROXIMATION	493
<i>Ziwei Ji, Matus Telgarsky, Ruicheng Xian</i>	
LEARNING TO MOVE WITH AFFORDANCE MAPS	521
<i>William Qi, Ravi Teja Mullapudi, Saurabh Gupta, Deva Ramanan</i>	
DIFFERENTIABLE LEARNING OF NUMERICAL RULES IN KNOWLEDGE GRAPHS	536
<i>Po-Wei Wang, Daria Stepanova, Csaba Domokos, J. Zico Kolter</i>	
CONSISTENCY REGULARIZATION FOR GENERATIVE ADVERSARIAL NETWORKS	548
<i>Han Zhang, Zizhao Zhang, Augustus Odena, Honglak Lee</i>	
ON GENERALIZATION ERROR BOUNDS OF NOISY GRADIENT METHODS FOR NON-CONVEX LEARNING	567
<i>Jian Li, Xuanyuan Luo, Mingda Qiao</i>	
SCALE-EQUIVARIANT STEERABLE NETWORKS.....	598
<i>Ivan Sosnovik, Michal Szmaja, Arnold Smeulders</i>	
CLASSIFICATION-BASED ANOMALY DETECTION FOR GENERAL DATA.....	612
<i>Liron Bergman, Yedid Hoshen</i>	
UNRESTRICTED ADVERSARIAL EXAMPLES VIA SEMANTIC MANIPULATION	624
<i>Anand Bhattad, Min Jin Chong, Kaizhao Liang, Bo Li, D. A. Forsyth</i>	
DISCRIMINATIVE PARTICLE FILTER REINFORCEMENT LEARNING FOR COMPLEX PARTIAL OBSERVATIONS	643
<i>Xiao Ma, Peter Karkus, David Hsu, Wee Sun Lee, Nan Ye</i>	
LEARNING TO GROUP: A BOTTOM-UP FRAMEWORK FOR 3D PART DISCOVERY IN UNSEEN CATEGORIES	661
<i>Tiange Luo, Kaichun Mo, Zhiao Huang, Jiarui Xu, Siyu Hu, Liwei Wang, Hao Su</i>	

STATE ALIGNMENT-BASED IMITATION LEARNING	684
<i>Fangchen Liu, Zhan Ling, Tongzhou Mu, Hao Su</i>	
LIPSCHITZ CONSTANT ESTIMATION OF NEURAL NETWORKS VIA SPARSE POLYNOMIAL OPTIMIZATION	699
<i>Fabian Latorre, Paul Rolland, Volkan Cevher</i>	

VOLUME 2

EFFECT OF ACTIVATION FUNCTIONS ON THE TRAINING OF OVERPARAMETRIZED NEURAL NETS	715
<i>Abhishek Panigrahi, Abhishek Shetty, Navin Goyal</i>	
PROVABLE FILTER PRUNING FOR EFFICIENT NEURAL NETWORKS	789
<i>Lucas Liebenwein, Cenk Baykal, Harry Lang, Dan Feldman, Daniela Rus</i>	
END TO END TRAINABLE ACTIVE CONTOURS VIA DIFFERENTIABLE RENDERING.....	818
<i>Shir Gur, Tal Shaharabany, Lior Wolf</i>	
COMPOSITIONAL LANGUAGE CONTINUAL LEARNING	831
<i>Yuanpeng Li, Liang Zhao, Kenneth Church, Mohamed Elhoseiny</i>	
ADVERSARIAL LIPSCHITZ REGULARIZATION.....	846
<i>Dávid Terjék</i>	
A FUNCTION SPACE VIEW OF BOUNDED NORM INFINITE WIDTH RELU NETS: THE MULTIVARIATE CASE	863
<i>Greg Ongie, Rebecca Willett, Daniel Soudry, Nathan Srebro</i>	
LITE TRANSFORMER WITH LONG-SHORT RANGE ATTENTION	891
<i>Zhanghao Wu*, Zhijian Liu*, Ji Lin, Yujun Lin, Song Han</i>	
MUTUAL INFORMATION GRADIENT ESTIMATION FOR REPRESENTATION LEARNING.....	904
<i>Liangjian Wen, Yiji Zhou, Lirong He, Mingyuan Zhou, Zenglin Xu</i>	
REGULARIZING ACTIVATIONS IN NEURAL NETWORKS VIA DISTRIBUTION MATCHING WITH THE WASSERSTEIN METRIC	918
<i>Taejong Joo, Donggu Kang, Byunghoon Kim</i>	
TRANSFERRING OPTIMALITY ACROSS DATA DISTRIBUTIONS VIA HOMOTOPY METHODS.....	931
<i>Matilde Gargiani, Andrea Zanelli, Quoc Tran Dinh, Moritz Diehl, Frank Hutter</i>	
LATENT NORMALIZING FLOWS FOR MANY-TO-MANY CROSS-DOMAIN MAPPINGS	953
<i>Shweta Mahajan, Iryna Gurevych, Stefan Roth</i>	
DYNAMIC MODEL PRUNING WITH FEEDBACK.....	968
<i>Tao Lin, Sebastian U. Stich, Luis Barba, Daniil Dmitriev, Martin Jaggi</i>	
ON THE INTERACTION BETWEEN SUPERVISION AND SELF-PLAY IN EMERGENT COMMUNICATION.....	990
<i>Ryan Lowe*, Abhinav Gupta*, Jakob Foerster, Douwe Kiela, Joelle Pineau</i>	
A META-TRANSFER OBJECTIVE FOR LEARNING TO DISENTANGLE CAUSAL MECHANISMS	1005
<i>Yoshua Bengio, Tristan Deleu, Nasim Rahaman, Nan Rosemary Ke, Sebastien Lachapelle, Olexa Bilaniuk, Anirudh Goyal, Christopher Pal</i>	
EXPECTED INFORMATION MAXIMIZATION: USING THE I-PROJECTION FOR MIXTURE DENSITY ESTIMATION	1032
<i>Philipp Becker, Oleg Arenz, Gerhard Neumann</i>	
DEEP AUDIO PRIORS EMERGE FROM HARMONIC CONVOLUTIONAL NETWORKS	1048
<i>Zhoutong Zhang, Yunyun Wang, Chuang Gan, Jiajun Wu, Joshua B. Tenenbaum, Antonio Torralba, William T. Freeman</i>	
A CLOSER LOOK AT THE APPROXIMATION CAPABILITIES OF NEURAL NETWORKS	1060
<i>Kai Fong Ernest Chong</i>	
RESIDUAL ENERGY-BASED MODELS FOR TEXT GENERATION	1076
<i>Yuntian Deng, Anton Bakhtin, Myle Ott, Arthur Szlam, Marc'Aurelio Ranzato</i>	

ATOMNAS: FINE-GRAINED END-TO-END NEURAL ARCHITECTURE SEARCH	1094
<i>Jieru Mei, Yingwei Li, Xiaochen Lian, Xiaojie Jin, Linjie Yang, Alan Yuille, Jianchao Yang</i>	
AUGMIX: A SIMPLE DATA PROCESSING METHOD TO IMPROVE ROBUSTNESS AND UNCERTAINTY	1108
<i>Dan Hendrycks*, Norman Mu*, Ekin Dogus Cubuk, Barret Zoph, Justin Gilmer, Balaji Lakshminarayanan</i>	
MEMORY-BASED GRAPH NETWORKS	1123
<i>Amir Hosein Khasahmadi, Kaveh Hassani, Parsa Moradi, Leo Lee, Quaid Morris</i>	
VARIATIONAL TEMPLATE MACHINE FOR DATA-TO-TEXT GENERATION	1139
<i>Rong Ye, Wenxian Shi, Hao Zhou, Zhongyu Wei, Lei Li</i>	
PHASE TRANSITIONS FOR THE INFORMATION BOTTLENECK IN REPRESENTATION LEARNING	1153
<i>Tailin Wu, Ian Fischer</i>	
PERMUTATION EQUIVARIANT MODELS FOR COMPOSITIONAL GENERALIZATION IN LANGUAGE	1180
<i>Jonathan Gordon, David Lopez-Paz, Marco Baroni, Diane Bouchacourt</i>	
TRAINING BINARY NEURAL NETWORKS WITH REAL-TO-BINARY CONVOLUTIONS	1192
<i>Brais Martinez, Jing Yang, Adrian Bulat, Georgios Tzimiropoulos</i>	
STRUCTBERT: INCORPORATING LANGUAGE STRUCTURES INTO PRE-TRAINING FOR DEEP LANGUAGE UNDERSTANDING	1203
<i>Wei Wang, Bin Bi, Ming Yan, Chen Wu, Jiangnan Xia, Zuyi Bao, Liwei Peng, Luo Si</i>	
SMOOTH MARKETS: A BASIC MECHANISM FOR ORGANIZING GRADIENT-BASED LEARNERS	1213
<i>David Balduzzi, Wojciech M. Czarnecki, Tom Anthony, Ian Gemp, Edward Hughes, Joel Leibo, Georgios Piliouras, Thore Graepel</i>	
FAIR RESOURCE ALLOCATION IN FEDERATED LEARNING	1231
<i>Tian Li, Maziar Sanjabi, Ahmad Beirami, Virginia Smith</i>	
NEVER GIVE UP: LEARNING DIRECTED EXPLORATION STRATEGIES	1258
<i>Adrià Puigdomènech Badia, Pablo Sprechmann, Alex Vitvitskiy, Daniel Guo, Bilal Piot, Steven Kapturowski, Olivier Tieleman, Martin Arjovsky, Alexander Pritzel, Andrew Bolt, Charles Blundell</i>	
ADVECTIVENET: AN EULERIAN-LAGRANGIAN FLUIDIC RESERVOIR FOR POINT CLOUD PROCESSING	1286
<i>Xingzhe He, Helen Lu Cao, Bo Zhu</i>	
YOU CAN TEACH AN OLD DOG NEW TRICKS! ON TRAINING KNOWLEDGE GRAPH EMBEDDINGS	1299
<i>Daniel Ruffinelli, Samuel Broscheit, Rainer Gemulla</i>	
FUNCTIONAL REGULARISATION FOR CONTINUAL LEARNING WITH GAUSSIAN PROCESSES	1319
<i>Michalis K. Titsias, Jonathan Schwarz, Alexander G. De G. Matthews, Razvan Pascanu, Yee Whye Teh</i>	
DYNAMICS-AWARE EMBEDDINGS	1336
<i>William Whitney, Rajat Agarwal, Kyunghyun Cho, Abhinav Gupta</i>	
RAPP: NOVELTY DETECTION WITH RECONSTRUCTION ALONG PROJECTION PATHWAY	1352
<i>Ki Hyun Kim, Sangwoo Shim, Yongsub Lim, Jongseob Jeon, Jeongwoo Choi, Byungchan Kim, Andre S. Yoon</i>	
HYPERMODELS FOR EXPLORATION	1366
<i>Vikranth Dwaracherla, Xiuyuan Lu, Morteza Ibrahimi, Ian Osband, Zheng Wen, Benjamin Van Roy</i>	
META REINFORCEMENT LEARNING WITH AUTONOMOUS INFERENCE OF SUBTASK DEPENDENCIES	1382
<i>Sungryull Sohn, Hyunjae Woo, Jongwook Choi, Honglak Lee</i>	

BAYESOPT ADVERSARIAL ATTACK	1410
<i>Binxin Ru, Adam Cobb, Arno Blaas, Yarin Gal</i>	
MODEL-BASED REINFORCEMENT LEARNING FOR BIOLOGICAL SEQUENCE DESIGN	1426
<i>Christof Angermueller, David Dohan, David Belanger, Ramya Deshpande, Kevin Murphy, Lucy Colwell</i>	
BINARYDUO: REDUCING GRADIENT MISMATCH IN BINARY ACTIVATION NETWORK BY COUPLING BINARY ACTIVATIONS	1449
<i>Hyungjun Kim, Kyungsu Kim, Jinseok Kim, Jae-Joon Kim</i>	
MIXED-CURVATURE VARIATIONAL AUTOENCODERS.....	1468
<i>Ondrej Skopec, Octavian-Eugen Ganeva, Gary Bécigneul</i>	

VOLUME 3

DEMYSTIFYING INTER-CLASS DISENTANGLEMENT	1512
<i>Aviv Gabbay, Yedid Hoshen</i>	
UNDERSTANDING THE LIMITATIONS OF CONDITIONAL GENERATIVE MODELS.....	1534
<i>Ethan Fetaya, Joern-Henrik Jacobsen, Will Grathwohl, Richard Zemel</i>	
KEEP DOING WHAT WORKED: BEHAVIOR MODELLING PRIORS FOR OFFLINE REINFORCEMENT LEARNING	1550
<i>Noah Siegel, Jost Tobias Springenberg, Felix Berkenkamp, Abbas Abdolmaleki, Michael Neunert, Thomas Lampe, Roland Hafner, Nicolas Heess, Martin Riedmiller</i>	
EMPIRICAL BAYES TRANSDUCTIVE META-LEARNING WITH SYNTHETIC GRADIENTS	1571
<i>Shell Xu Hu, Pablo Garcia Moreno, Yang Xiao, Xi Shen, Guillaume Obozinski, Neil Lawrence, Andreas Damianou</i>	
SPIKE-BASED CAUSAL INFERENCE FOR WEIGHT ALIGNMENT	1587
<i>Jordan Guerguiev, Konrad Kording, Blake Richards</i>	
LOOKAHEAD: A FAR-SIGHTED ALTERNATIVE OF MAGNITUDE-BASED PRUNING	1599
<i>Sejun Park*, Jaeho Lee*, Sangwoo Mo, Jinwoo Shin</i>	
VARIABAD: A VERY GOOD METHOD FOR BAYES-ADAPTIVE DEEP RL VIA META-LEARNING.....	1619
<i>Luisa Zintgraf, Kyriacos Shiarlis, Maximilian Igl, Sebastian Schulze, Yarin Gal, Katja Hofmann, Shimon Whiteson</i>	
MAKING EFFICIENT USE OF DEMONSTRATIONS TO SOLVE HARD EXPLORATION PROBLEMS.....	1639
<i>Caglar Gulcehre, Tom Le Paine, Bobak Shahriari, Misha Denil, Matt Hoffman, Hubert Soyer, Richard Tanburn, Steven Kapturowski, Neil Rabinowitz, Duncan Williams, Gabriel Barth-Maron, Ziyu Wang, Nando De Freitas, Worlds Team</i>	
META-LEARNING CURIOSITY ALGORITHMS	1659
<i>Ferran Alet*, Martin F. Schneider*, Tomas Lozano-Perez, Leslie Pack Kaelbling</i>	
VQ-WAV2VEC: SELF-SUPERVISED LEARNING OF DISCRETE SPEECH REPRESENTATIONS.....	1681
<i>Alexei Baevski, Steffen Schneider, Michael Auli</i>	
INFINITE-HORIZON OFF-POLICY POLICY EVALUATION WITH MULTIPLE BEHAVIOR POLICIES	1693
<i>Xinyun Chen, Lu Wang, Yizhe Hang, Heng Ge, Hongyuan Zha</i>	
MMA TRAINING: DIRECT INPUT SPACE MARGIN MAXIMIZATION THROUGH ADVERSARIAL TRAINING.....	1709
<i>Gavin Weiguang Ding, Yash Sharma, Kry Yik Chau Lui, Ruitong Huang</i>	
INCORPORATING BERT INTO NEURAL MACHINE TRANSLATION.....	1737
<i>Jinhua Zhu, Yingce Xia, Lijun Wu, Di He, Tao Qin, Wengang Zhou, Houqiang Li, Tiejun Liu</i>	
STRUCTURED OBJECT-AWARE PHYSICS PREDICTION FOR VIDEO MODELING AND PLANNING.....	1755
<i>Jannik Kossen, Karl Stelzner, Marcel Hussing, Claas Voelcker, Kristian Kersting</i>	

LEARNING-AUGMENTED DATA STREAM ALGORITHMS	1774
<i>Tanqiu Jiang, Yi Li, Honghao Lin, Yisong Ruan, David P. Woodruff</i>	
ON THE RELATIONSHIP BETWEEN SELF-ATTENTION AND CONVOLUTIONAL LAYERS.....	1791
<i>Jean-Baptiste Cordonnier, Andreas Loukas, Martin Jaggi</i>	
SPIKEGRAD: AN ANN-EQUIVALENT COMPUTATION MODEL FOR IMPLEMENTING BACKPROPAGATION WITH SPIKES	1809
<i>Johannes C. Thiele, Olivier Bichler, Antoine Dupret</i>	
GRADIENT L1 REGULARIZATION FOR QUANTIZATION ROBUSTNESS	1822
<i>Milad Alizadeh, Arash Behboodi, Mart Van Baalen, Christos Louizos, Tijmen Blankevoort, Max Welling</i>	
TOWARD EVALUATING ROBUSTNESS OF DEEP REINFORCEMENT LEARNING WITH CONTINUOUS CONTROL	1837
<i>Tsui-Wei Weng, Krishnamurthy (Dj) Dvijotham*, Jonathan Uesato*, Kai Xiao*, Sven Gowal*, Robert Stanforth*, Pushmeet Kohli</i>	
DECENTRALIZED DEEP LEARNING WITH ARBITRARY COMMUNICATION COMPRESSION	1850
<i>Anastasia Koloskova*, Tao Lin*, Sebastian U Stich, Martin Jaggi</i>	
TRAINING GENERATIVE ADVERSARIAL NETWORKS FROM INCOMPLETE OBSERVATIONS USING FACTORISED DISCRIMINATORS	1872
<i>Daniel Stoller, Sebastian Ewert, Simon Dixon</i>	
COMBINING Q-LEARNING AND SEARCH WITH AMORTIZED VALUE ESTIMATES	1896
<i>Jessica B. Hamrick, Victor Bapst, Alvaro Sanchez-Gonzalez, Tobias Pfaff, Theophane Weber, Lars Buesing, Peter W. Battaglia</i>	
INFINITE-HORIZON DIFFERENTIABLE MODEL PREDICTIVE CONTROL	1923
<i>Sebastian East, Marco Gallieri, Jonathan Masci, Jan Koutnik, Mark Cannon</i>	
PROJECTION-BASED CONSTRAINED POLICY OPTIMIZATION.....	1938
<i>Tsung-Yen Yang, Justinian Rosca, Karthik Narasimhan, Peter J. Ramadge</i>	
YOU ONLY TRAIN ONCE: LOSS-CONDITIONAL TRAINING OF DEEP NETWORKS.....	1962
<i>Alexey Dosovitskiy, Josip Djolonga</i>	
GRAPHSAINTE: GRAPH SAMPLING BASED INDUCTIVE LEARNING METHOD	1979
<i>Hanqing Zeng, Hongkuan Zhou, Ajitesh Srivastava, Rajgopal Kannan, Viktor Prasanna</i>	
EFFICIENT PROBABILISTIC LOGIC REASONING WITH GRAPH NEURAL NETWORKS.....	1998
<i>Yuyu Zhang, Xinshi Chen, Yuan Yang, Arun Ramamurthy, Bo Li, Yuan Qi, Le Song</i>	
LOW-DIMENSIONAL STATISTICAL MANIFOLD EMBEDDING OF DIRECTED GRAPHS	2018
<i>Thorben Funke, Tian Guo, Alen Lancic, Nino Antulov-Fantulin</i>	
RIDE: REWARDING IMPACT-DRIVEN EXPLORATION FOR PROCEDURALLY- GENERATED ENVIRONMENTS	2036
<i>Roberta Raileanu, Tim Rocktäschel</i>	
SPACE: UNSUPERVISED OBJECT-ORIENTED SCENE REPRESENTATION VIA SPATIAL ATTENTION AND DECOMPOSITION	2057
<i>Zhixuan Lin, Yi-Fu Wu, Skand Vishwanath Peri, Weihao Sun, Gautam Singh, Fei Deng, Jindong Jiang, Sungjin Ahn</i>	
CROSS-LINGUAL ABILITY OF MULTILINGUAL BERT: AN EMPIRICAL STUDY	2079
<i>Karthikeyan K, Zihan Wang, Stephen Mayhew, Dan Roth</i>	
REDUCING TRANSFORMER DEPTH ON DEMAND WITH STRUCTURED DROPOUT	2091
<i>Angela Fan, Edouard Grave, Armand Joulin</i>	
NEURAL OUTLIER REJECTION FOR SELF-SUPERVISED KEYPOINT LEARNING.....	2107
<i>Jiexiong Tang, Hanme Kim, Vitor Guizilini, Sudeep Pillai, Rares Ambrus</i>	
B-SPLINE CNNs ON LIE GROUPS	2121
<i>Erik J Bekkers</i>	

QUANTIFYING POINT-PREDICTION UNCERTAINTY IN NEURAL NETWORKS VIA RESIDUAL ESTIMATION WITH AN I/O KERNEL	2146
<i>Xin Qiu, Elliot Meyerson, Risto Miikkulainen</i>	
EMPIR: ENSEMBLES OF MIXED PRECISION DEEP NETWORKS FOR INCREASED ROBUSTNESS AGAINST ADVERSARIAL ATTACKS	2181
<i>Sanchari Sen, Balaraman Ravindran, Anand Raghunathan</i>	
LEARNING TO EXPLORE USING ACTIVE NEURAL SLAM	2193
<i>Devendra Singh Chaplot, Dhiraj Gandhi, Saurabh Gupta, Abhinav Gupta, Ruslan Salakhutdinov</i>	
UNDERSTANDING AND IMPROVING INFORMATION TRANSFER IN MULTI-TASK LEARNING.....	2211
<i>Sen Wu, Hongyang R. Zhang, Christopher Ré</i>	
A STOCHASTIC DERIVATIVE FREE OPTIMIZATION METHOD WITH MOMENTUM.....	2239
<i>Eduard Gorbunov, Adel Bibi, Ozan Sener, El Houcine Bergou, Peter Richtarik</i>	
COMPRESSIVE TRANSFORMERS FOR LONG-RANGE SEQUENCE MODELLING	2267
<i>Jack W. Rae, Anna Potapenko, Siddhant M. Jayakumar, Chloe Hillier, Timothy P. Lillicrap</i>	
REINFORCED GENETIC ALGORITHM LEARNING FOR OPTIMIZING COMPUTATION GRAPHS	2286
<i>Aditya Paliwal, Felix Gimeno, Vinod Nair, Yujia Li, Miles Lubin, Pushmeet Kohli, Oriol Vinyals</i>	

VOLUME 4

LAGRANGIAN FLUID SIMULATION WITH CONTINUOUS CONVOLUTIONS	2310
<i>Benjamin Ummerhofer, Lukas Prantl, Nils Thuerey, Vladlen Koltun</i>	
LEARNING TO GUIDE RANDOM SEARCH.....	2325
<i>Ozan Sener, Vladlen Koltun</i>	
ROBUST ANOMALY DETECTION AND BACKDOOR ATTACK DETECTION VIA DIFFERENTIAL PRIVACY	2348
<i>Min Du, Ruoxi Jia, Dawn Song</i>	
DEEP PROBABILISTIC SUBSAMPLING FOR TASK-ADAPTIVE COMPRESSED SENSING.....	2367
<i>Iris A. M. Huijben, Bastiaan S. Veeling, Ruud J. G. Van Sloun</i>	
LEARNING ROBUST REPRESENTATIONS VIA MULTI-VIEW INFORMATION BOTTLENECK.....	2383
<i>Marco Federici, Anjan Dutta, Patrick Forré, Nate Kushman, Zeynep Akata</i>	
BATCH-SHAPING FOR LEARNING CONDITIONAL CHANNEL GATED NETWORKS	2409
<i>Babak Ehteshami Bejnordi, Tijmen Blankevoort, Max Welling</i>	
INDUCTIVE AND UNSUPERVISED REPRESENTATION LEARNING ON GRAPH STRUCTURED OBJECTS	2423
<i>Lichen Wang, Bo Zong, Qianqian Ma, Wei Cheng, Jingchao Ni, Wenchao Yu, Yanchi Liu, Dongjin Song, Haifeng Chen, Yun Fu</i>	
U-GAT-IT: UNSUPERVISED GENERATIVE ATTENTIONAL NETWORKS WITH ADAPTIVE LAYER-INSTANCE NORMALIZATION FOR IMAGE-TO-IMAGE TRANSLATION.....	2443
<i>Junho Kim, Minjae Kim, Hyeonwoo Kang, Kwang Hee Lee</i>	
MASKED BASED UNSUPERVISED CONTENT TRANSFER.....	2462
<i>Ron Mokady, Sagie Benaim, Lior Wolf, Amit Bermano</i>	
DROPEdge: TOWARDS DEEP GRAPH CONVOLUTIONAL NETWORKS ON NODE CLASSIFICATION	2502
<i>Yu Rong, Wenbing Huang, Tingyang Xu, Junzhou Huang</i>	
WHY NOT TO USE ZERO IMPUTATION? CORRECTING SPARSITY BIAS IN TRAINING NEURAL NETWORKS.....	2519
<i>Joonyoung Yi, Juhyuk Lee, Kwang Joon Kim, Sung Ju Hwang, Eunho Yang</i>	

PROBABILITY CALIBRATION FOR KNOWLEDGE GRAPH EMBEDDING MODELS.....	2546
<i>Pedro Tabacof, Luca Costabello</i>	
ON THE EQUIVALENCE BETWEEN POSITIONAL NODE EMBEDDINGS AND STRUCTURAL GRAPH REPRESENTATIONS	2561
<i>Balasubramaniam Srinivasan, Bruno Ribeiro</i>	
NEURAL EPITOME SEARCH FOR ARCHITECTURE-AGNOSTIC NETWORK COMPRESSION	2585
<i>Daquan Zhou, Xiaojie Jin, Qibin Hou, Kaixin Wang, Jianchao Yang, Jiashi Feng</i>	
HYPER-SAGNN: A SELF-ATTENTION BASED GRAPH NEURAL NETWORK FOR HYPERGRAPHS	2601
<i>Ruochi Zhang, Yuesong Zou, Jian Ma</i>	
A NEURAL DIRICHLET PROCESS MIXTURE MODEL FOR TASK-FREE CONTINUAL LEARNING.....	2619
<i>Soochan Lee, Junsoo Ha, Dongsu Zhang, Gunhee Kim</i>	
ON SOLVING MINIMAX OPTIMIZATION LOCALLY: A FOLLOW-THE-RIDGE APPROACH.....	2641
<i>Yuanhao Wang*, Guodong Zhang*, Jimmy Ba</i>	
DISTRIBUTIONALLY ROBUST NEURAL NETWORKS	2661
<i>Shiori Sagawa*, Pang Wei Koh*, Tatsunori B. Hashimoto, Percy Liang</i>	
KERNEL OF CYCLEGAN AS A PRINCIPAL HOMOGENEOUS SPACE.....	2680
<i>Nikita Moriakov, Jonas Adler, Jonas Teuwen</i>	
DON'T USE LARGE MINI-BATCHES, USE LOCAL SGD	2694
<i>Tao Lin, Sebastian U. Stich, Kumar Kshitij Patel, Martin Jaggi</i>	
PROVABLE ROBUSTNESS AGAINST ALL ADVERSARIAL LP-PERTURBATIONS FOR $P \geq 1$	2734
<i>Francesco Croce, Matthias Hein</i>	
ON UNIVERSAL EQUIVARIANT SET NETWORKS	2754
<i>Nimrod Segol, Yaron Lipman</i>	
TENSOR DECOMPOSITIONS FOR TEMPORAL KNOWLEDGE BASE COMPLETION	2765
<i>Timothée Lacroix, Guillaume Obozinski, Nicolas Usunier</i>	
NETWORK RANDOMIZATION: A SIMPLE TECHNIQUE FOR GENERALIZATION IN DEEP REINFORCEMENT LEARNING	2777
<i>Kimin Lee, Kibok Lee, Jinwoo Shin, Honglak Lee</i>	
ROBUSTNESS VERIFICATION FOR TRANSFORMERS	2799
<i>Zhouxing Shi, Huan Zhang, Kai-Wei Chang, Minlie Huang, Cho-Jui Hsieh</i>	
FANTASTIC GENERALIZATION MEASURES AND WHERE TO FIND THEM	2821
<i>Yiding Jiang*, Behnam Neyshabur*, Hossein Mobahi, Dilip Krishnan, Samy Bengio</i>	
NESTEROV ACCELERATED GRADIENT AND SCALE INVARIANCE FOR ADVERSARIAL ATTACKS	2854
<i>Jiadong Lin, Chuanbiao Song, Kun He, Liwei Wang, John E. Hopcroft</i>	
WEAKLY SUPERVISED DISENTANGLEMENT WITH GUARANTEES	2866
<i>Rui Shu, Yining Chen, Abhishek Kumar, Stefano Ermon, Ben Poole</i>	
DISCREPANCY RATIO: EVALUATING MODEL PERFORMANCE WHEN EVEN EXPERTS DISAGREE ON THE TRUTH.....	2902
<i>Igor Lovchinsky, Alon Daks, Israel Malkin, Pouya Samangouei, Ardavan Saeedi, Yang Liu, Swami Sankaranarayanan, Tomer Gafner, Ben Sternlieb, Patrick Maher, Nathan Silberman</i>	
ABDUCTIVE COMMONSENSE REASONING.....	2921
<i>Chandra Bhagavatula, Ronan Le Bras, Chaitanya Malaviya, Keisuke Sakaguchi, Ari Holtzman, Hannah Rashkin, Doug Downey, Wen-Tau Yih, Yejin Choi</i>	
VARIANCE REDUCTION WITH SPARSE GRADIENTS	2939
<i>Melih Elibol, Lihua Lei, Michael I. Jordan</i>	

BLOCKSWAP: FISHER-GUIDED BLOCK SUBSTITUTION FOR NETWORK COMPRESSION ON A BUDGET	2957
<i>Jack Turner, Elliot J. Crowley, Michael O'Boyle, Amos Storkey, Gavin Gray</i>	
LEARNING TRANSPORT COST FROM SUBSET CORRESPONDENCE	2972
<i>Ruishan Liu, Akshay Balsubramani, James Zou</i>	
RÉNYI FAIR INFERENCE	2987
<i>Sina Baharlouei, Maher Nouiehed, Ahmad Beirami, Meisam Razaviyayn</i>	
META DROPOUT: LEARNING TO PERTURB LATENT FEATURES FOR GENERALIZATION	3006
<i>Hae Beom Lee, Taewook Nam, Eunho Yang, Sung Ju Hwang</i>	
ADVERSARIAL AUTOAUGMENT	3021
<i>Xinyu Zhang, Qiang Wang, Jian Zhang, Zhao Zhong</i>	
STATE-ONLY IMITATION WITH TRANSITION DYNAMICS MISMATCH	3034
<i>Tanmay Gangwani, Jian Peng</i>	
MEASURING AND IMPROVING THE USE OF GRAPH INFORMATION IN GRAPH NEURAL NETWORKS	3051
<i>Yifan Hou, Jian Zhang, James Cheng, Kaili Ma, Richard T. B. Ma, Hongzhi Chen, Ming-Chang Yang</i>	
UNIVERSAL APPROXIMATION WITH CERTIFIED NETWORKS	3067
<i>Maximilian Baader, Matthew Mirman, Martin Vechev</i>	
EXPLAIN YOUR MOVE: UNDERSTANDING AGENT ACTIONS USING SPECIFIC AND RELEVANT FEATURE ATTRIBUTION	3085
<i>Nikaash Puri, Sukriti Verma, Piyush Gupta, Dhruv Kayastha, Shripad Deshmukh, Balaji Krishnamurthy, Sameer Singh</i>	
DEEP SYMBOLIC SUPEROPTIMIZATION WITHOUT HUMAN KNOWLEDGE.....	3099
<i>Hui Shi, Yang Zhang, Xinyun Chen, Yuandong Tian, Jishen Zhao</i>	

VOLUME 5

SAMPLE EFFICIENT POLICY GRADIENT METHODS WITH RECURSIVE VARIANCE REDUCTION	3117
<i>Pan Xu, Felicia Gao, Quanquan Gu</i>	
CERTIFIED DEFENSES FOR ADVERSARIAL PATCHES	3140
<i>Ping-Yeh Chiang*, Renkun Ni*, Ahmed Abdelkader, Chen Zhu, Christoph Studor, Tom Goldstein</i>	
CONTRASTIVE REPRESENTATION DISTILLATION.....	3156
<i>Yonglong Tian, Dilip Krishnan, Phillip Isola</i>	
A FRAMEWORK FOR ROBUSTNESS CERTIFICATION OF SMOOTHED CLASSIFIERS USING F-DIVERGENCES.....	3175
<i>Krishnamurthy (Dj) Dvijotham, Jamie Hayes, Borja Balle, Zico Kolter, Chongli Qin, Andras Gyorgy, Kai Xiao, Sven Gowal, Pushmeet Kohli</i>	
PADÉ ACTIVATION UNITS: END-TO-END LEARNING OF FLEXIBLE ACTIVATION FUNCTIONS IN DEEP NETWORKS.....	3201
<i>Alejandro Molina, Patrick Schramowski, Kristian Kersting</i>	
LEARNING TO RETRIEVE REASONING PATHS OVER WIKIPEDIA GRAPH FOR QUESTION ANSWERING	3218
<i>Akari Asai, Kazuma Hashimoto, Hannaneh Hajishirzi, Richard Socher, Caiming Xiong</i>	
A BASELINE FOR FEW-SHOT IMAGE CLASSIFICATION.....	3240
<i>Guneet Singh Dhillon, Pratik Chaudhari, Avinash Ravichandran, Stefano Soatto</i>	
ABSTRACT DIAGRAMMATIC REASONING WITH MULTIPLEX GRAPH NETWORKS	3260
<i>Duo Wang, Mateja Jamnik, Pietro Lio</i>	

ENVIRONMENTAL DRIVERS OF SYSTEMATICITY AND GENERALIZATION IN A SITUATED AGENT.....	3280
<i>Felix Hill, Andrew Lampinen, Rosalia Schneider, Stephen Clark, Matthew Botvinick, James L. McClelland, Adam Santoro</i>	
ELECTRA: PRE-TRAINING TEXT ENCODERS AS DISCRIMINATORS RATHER THAN GENERATORS.....	3295
<i>Kevin Clark, Minh-Thang Luong, Quoc V. Le, Christopher D. Manning</i>	
EVOLUTIONARY POPULATION CURRICULUM FOR SCALING MULTI-AGENT REINFORCEMENT LEARNING.....	3313
<i>Qian Long*, Zihan Zhou*, Abhinav Gupta, Fei Fang, Yi Wu†, Xiaolong Wang†</i>	
THINKING WHILE MOVING: DEEP REINFORCEMENT LEARNING WITH CONCURRENT CONTROL.....	3330
<i>Ted Xiao, Eric Jang, Dmitry Kalashnikov, Sergey Levine, Julian Ibarz, Karol Hausman, Alexander Herzog</i>	
JACOBIAN ADVERSARIALLY REGULARIZED NETWORKS FOR ROBUSTNESS.....	3347
<i>Alvin Chan, Yi Tay, Yew Soon Ong, Jie Fu</i>	
TOWARDS VERIFIED ROBUSTNESS UNDER TEXT DELETION INTERVENTIONS.....	3360
<i>Johannes Welbl, Po-Sen Huang, Robert Stanforth, Sven Gowal, Krishnamurthy (Dj) Dvijotham, Martin Szummer, Pushmeet Kohli</i>	
RGBD-GAN: UNSUPERVISED 3D REPRESENTATION LEARNING FROM NATURAL IMAGE DATASETS VIA RGBD IMAGE SYNTHESIS.....	3374
<i>Atsuhiko Noguchi, Tatsuya Harada</i>	
PROVABLE BENEFIT OF ORTHOGONAL INITIALIZATION IN OPTIMIZING DEEP LINEAR NETWORKS.....	3395
<i>Wei Hu, Lechao Xiao, Jeffrey Pennington</i>	
PLUG AND PLAY LANGUAGE MODELS: A SIMPLE APPROACH TO CONTROLLED TEXT GENERATION.....	3411
<i>Sumanth Dathathri, Andrea Madotto, Janice Lan, Jane Hung, Eric Frank, Piero Molino, Jason Yosinski, Rosanne Liu</i>	
RETHINKING THE HYPERPARAMETERS FOR FINE-TUNING.....	3445
<i>Hao Li, Pratik Chaudhari, Hao Yang, Michael Lam, Avinash Ravichandran, Rahul Bhotika, Stefano Soatto</i>	
QUERY2BOX: REASONING OVER KNOWLEDGE GRAPHS IN VECTOR SPACE USING BOX EMBEDDINGS.....	3465
<i>Hongyu Ren*, Weihua Hu*, Jure Leskovec</i>	
IMPLEMENTING INDUCTIVE BIAS FOR DIFFERENT NAVIGATION TASKS THROUGH DIVERSE RNN ATTRACTORS.....	3482
<i>Tie Xu, Omri Barak</i>	
PCMC-NET: FEATURE-BASED PAIRWISE CHOICE MARKOV CHAINS.....	3509
<i>Alix Lhéritier</i>	
MULTI-AGENT INTERACTIONS MODELING WITH CORRELATED POLICIES.....	3520
<i>Minghuan Liu, Ming Zhou, Weinan Zhang, Yuzheng Zhuang, Jun Wang, Wulong Liu, Yong Yu</i>	
ONCE-FOR-ALL: TRAIN ONE NETWORK AND SPECIALIZE IT FOR EFFICIENT DEPLOYMENT.....	3540
<i>Han Cai, Chuang Gan, Tianzhe Wang, Zhekai Zhang, Song Han</i>	
GENERALIZED CONVOLUTIONAL FOREST NETWORKS FOR DOMAIN GENERALIZATION AND VISUAL RECOGNITION.....	3555
<i>Jongbin Ryu, Gitaek Kwon, Ming-Hsuan Yang, Jongwoo Lim</i>	
SNODE: SPECTRAL DISCRETIZATION OF NEURAL ODES FOR SYSTEM IDENTIFICATION.....	3569
<i>Alessio Quaglino, Marco Gallieri, Jonathan Masci, Jan Koutnik</i>	
GUIDING PROGRAM SYNTHESIS BY LEARNING TO GENERATE EXAMPLES.....	3584
<i>Larissa Laich, Pavol Bielik, Martin Vechev</i>	

FAST NEURAL NETWORK ADAPTATION VIA PARAMETER REMAPPING AND ARCHITECTURE SEARCH.....	3606
<i>Jiemin Fang*, Yuzhu Sun*, Kangjian Peng*, Qian Zhang, Yuan Li, Wenyu Liu, Xinggong Wang</i>	
EXPLORATORY NOT EXPLANATORY: COUNTERFACTUAL ANALYSIS OF SALIENCY MAPS FOR DEEP REINFORCEMENT LEARNING.....	3620
<i>Akanksha Atrey, Kaleigh Clary, David Jensen</i>	
META-LEARNING DEEP ENERGY-BASED MEMORY MODELS.....	3643
<i>Sergey Bartunov, Jack Rae, Simon Osindero, Timothy Lillicrap</i>	
GRAPH CONVOLUTIONAL REINFORCEMENT LEARNING.....	3666
<i>Jiechuan Jiang, Chen Dun, Tiejun Huang, Zongqing Lu</i>	
THE CURIOUS CASE OF NEURAL TEXT DEGENERATION.....	3679
<i>Ari Holtzman, Jan Buys, Li Du, Maxwell Forbes, Yejin Choi</i>	
MULTILINGUAL ALIGNMENT OF CONTEXTUAL WORD REPRESENTATIONS.....	3695
<i>Steven Cao, Nikita Kitaev, Dan Klein</i>	
THE GAMBLER'S PROBLEM AND BEYOND.....	3710
<i>Baoxiang Wang, Shuai Li, Jiajin Li, Siu On Chan</i>	
GRAPHAF: A FLOW-BASED AUTOREGRESSIVE MODEL FOR MOLECULAR GRAPH GENERATION.....	3733
<i>Chence Shi*, Minkai Xu*, Zhaocheng Zhu, Weinan Zhang, Ming Zhang, Jian Tang</i>	
DOUBLE NEURAL COUNTERFACTUAL REGRET MINIMIZATION.....	3751
<i>Hui Li, Kailiang Hu, Shaohua Zhang, Yuan Qi, Le Song</i>	
NEURAL POLICY GRADIENT METHODS: GLOBAL OPTIMALITY AND RATES OF CONVERGENCE.....	3779
<i>Lingxiao Wang, Qi Cai, Zhuoran Yang, Zhaoran Wang</i>	
TRIPLE WINS: BOOSTING ACCURACY, ROBUSTNESS AND EFFICIENCY TOGETHER BY ENABLING INPUT-ADAPTIVE INFERENCE.....	3825
<i>Ting-Kuei Hu, Tianlong Chen, Haotao Wang, Zhangyang Wang</i>	
DYNAMIC SPARSE TRAINING: FIND EFFICIENT SPARSE NETWORK FROM SCRATCH WITH TRAINABLE MASKED LAYERS.....	3839
<i>Junjie Liu, Zhe Xu, Runbin Shi, Ray C. C. Cheung, Hayden K. H. So</i>	
BEYOND LINEARIZATION: ON QUADRATIC AND HIGHER-ORDER APPROXIMATION OF WIDE NEURAL NETWORKS.....	3853
<i>Yu Bai, Jason D. Lee</i>	
DEEP GRAPH MATCHING CONSENSUS.....	3888
<i>Matthias Fey, Jan E. Lenssen, Christopher Morris, Jonathan Masci, Nils M. Kriege</i>	

VOLUME 6

SELF-SUPERVISED LEARNING OF APPLIANCE USAGE.....	3911
<i>Chen-Yu Hsu, Abbas Zeitoun, Guang-He Lee, Dina Katabi, Tommi Jaakkola</i>	
QUANTUM ALGORITHMS FOR DEEP CONVOLUTIONAL NEURAL NETWORKS.....	3925
<i>Iordanis Kerenidis, Jonas Landman, Anupam Prakash</i>	
UNDERSTANDING L4-BASED DICTIONARY LEARNING: INTERPRETATION, STABILITY, AND ROBUSTNESS.....	3961
<i>Yuexiang Zhai, Hermish Mehta, Zhengyuan Zhou, Yi Ma</i>	
SYMPLECTIC ODE-NET: LEARNING HAMILTONIAN DYNAMICS WITH CONTROL.....	3989
<i>Yaofeng Desmond Zhong, Biswadip Dey, Amit Chakraborty</i>	
CONTROLLING GENERATIVE MODELS WITH CONTINUOUS FACTORS OF VARIATIONS.....	4006
<i>Antoine Plummerault, Hervé Le Borgne, Céline Hudelot</i>	
SIMPLE AND EFFECTIVE REGULARIZATION METHODS FOR TRAINING ON NOISILY LABELED DATA WITH GENERALIZATION GUARANTEE.....	4023
<i>Wei Hu, Zhiyuan Li, Dingli Yu</i>	

UNSUPERVISED CLUSTERING USING PSEUDO-SEMI-SUPERVISED LEARNING	4041
<i>Divam Gupta, Ramachandran Ramjee, Nipun Kwatra, Muthian Sivathanu</i>	
PAIRNORM: TACKLING OVERSMOOTHING IN GNNS.....	4056
<i>Lingxiao Zhao, Leman Akoglu</i>	
BLACK-BOX OFF-POLICY ESTIMATION FOR INFINITE-HORIZON REINFORCEMENT LEARNING.....	4073
<i>Ali Mousavi, Lihong Li, Qiang Liu, Denny Zhou</i>	
EMPIRICAL STUDIES ON THE PROPERTIES OF LINEAR REGIONS IN DEEP NEURAL NETWORKS.....	4090
<i>Xiao Zhang, Dongrui Wu</i>	
SNOW: SUBSCRIBING TO KNOWLEDGE VIA CHANNEL POOLING FOR TRANSFER & LIFELONG LEARNING OF CONVOLUTIONAL NEURAL NETWORKS	4107
<i>Chungkuk Yoo, Bumsoo Kang, Minsik Cho</i>	
SMOOTHNESS AND STABILITY IN GANS	4121
<i>Casey Chu, Kentaro Minami, Kenji Fukumizu</i>	
ADAPTIVE CORRELATED MONTE CARLO FOR CONTEXTUAL CATEGORICAL SEQUENCE GENERATION	4152
<i>Xinjie Fan, Yizhe Zhang, Zhendong Wang, Mingyuan Zhou</i>	
ON BONUS BASED EXPLORATION METHODS IN THE ARCADE LEARNING ENVIRONMENT.....	4172
<i>Adrien Ali Taiga, William Fedus, Marlos C. Machado, Aaron Courville, Marc G. Bellemare</i>	
IMPACT: IMPORTANCE WEIGHTED ASYNCHRONOUS ARCHITECTURES WITH CLIPPED TARGET NETWORKS	4196
<i>Michael Luo, Jiahao Yao, Richard Liaw, Eric Liang, Ion Stoica</i>	
HILLOC: LOSSLESS IMAGE COMPRESSION WITH HIERARCHICAL LATENT VARIABLE MODELS	4210
<i>James Townsend, Thomas Bird, Julius Kunze, David Barber</i>	
PHYSICS-AWARE DIFFERENCE GRAPH NETWORKS FOR SPARSELY-OBSERVED DYNAMICS.....	4224
<i>Sungyong Seo*, Chuizheng Meng*, Yan Liu</i>	
POLYLOGARITHMIC WIDTH SUFFICES FOR GRADIENT DESCENT TO ACHIEVE ARBITRARILY SMALL TEST ERROR WITH SHALLOW RELU NETWORKS.....	4239
<i>Ziwei Ji, Matus Telgarsky</i>	
LEARNING REPRESENTATIONS FOR BINARY-CLASSIFICATION WITHOUT BACKPROPAGATION.....	4262
<i>Mathias Lechner</i>	
FREQUENCY-BASED SEARCH-CONTROL IN DYNA.....	4279
<i>Yangchen Pan, Jincheng Mei, Amir-Massoud Farahmand</i>	
TOWARDS STABLE AND EFFICIENT TRAINING OF VERIFIABLY ROBUST NEURAL NETWORKS.....	4300
<i>Huan Zhang, Hongge Chen, Chaowei Xiao, Sven Gowal, Robert Stanforth, Bo Li, Duane Boning, Cho-Jui Hsieh</i>	
ITERATIVE ENERGY-BASED PROJECTION ON A NORMAL DATA MANIFOLD FOR ANOMALY LOCALIZATION	4325
<i>David Dehaene, Oriel Frigo, Sébastien Combrexelle, Pierre Eline</i>	
TOWARDS NEURAL NETWORKS THAT PROVABLY KNOW WHEN THEY DON'T KNOW	4342
<i>Alexander Meinke, Matthias Hein</i>	
BATCHENSEMBLE: AN ALTERNATIVE APPROACH TO EFFICIENT ENSEMBLE AND LIFELONG LEARNING	4360
<i>Yeming Wen, Dustin Tran, Jimmy Ba</i>	
INDUCTIVE REPRESENTATION LEARNING ON TEMPORAL GRAPHS.....	4379
<i>Da Xu, Chuanwei Ruan, Evren Korpeoglu, Sushant Kumar, Kannan Achan</i>	

GENERATIVE MODELS FOR EFFECTIVE ML ON PRIVATE, DECENTRALIZED DATASETS	4398
<i>Sean Augenstein, H. Brendan McMahan, Daniel Ramage, Swaroop Ramaswamy, Peter Kairouz, Mingqing Chen, Rajiv Mathews, Blaise Aguera Y Arcas</i>	
PICKING WINNING TICKETS BEFORE TRAINING BY PRESERVING GRADIENT FLOW	4424
<i>Chaoqi Wang, Guodong Zhang, Roger Grosse</i>	
CURRICULUM LOSS: ROBUST LEARNING AND GENERALIZATION AGAINST LABEL CORRUPTION	4435
<i>Yueming Lyu, Ivor W. Tsang</i>	
UNCERTAINTY-GUIDED CONTINUAL LEARNING WITH BAYESIAN NEURAL NETWORKS.....	4457
<i>Sayna Ebrahimi, Mohamed Elhoseiny, Trevor Darrell, Marcus Rohrbach</i>	
TRAINING RECURRENT NEURAL NETWORKS ONLINE BY LEARNING EXPLICIT STATE VARIABLES	4473
<i>Somjit Nath, Vincent Liu, Alan Chan, Xin Li, Adam White, Martha White</i>	
CROSS-LINGUAL ALIGNMENT VS JOINT TRAINING: A COMPARATIVE STUDY AND A SIMPLE UNIFIED FRAMEWORK	4494
<i>Zirui Wang*, Jiateng Xie*, Ruochen Xu, Yiming Yang, Graham Neubig, Jaime G. Carbonell</i>	
ROBUST REINFORCEMENT LEARNING FOR CONTINUOUS CONTROL WITH MODEL MISSPECIFICATION.....	4509
<i>Daniel J. Mankowitz, Nir Levine, Rae Jeong, Abbas Abdolmaleki, Jost Tobias Springenberg, Yuanyuan Shi, Jackie Kay, Todd Hester, Timothy Mann, Martin Riedmiller</i>	
DECOUPLING REPRESENTATION AND CLASSIFIER FOR LONG-TAILED RECOGNITION.....	4537
<i>Bingyi Kang, Saining Xie, Marcus Rohrbach, Zhicheng Yan, Albert Gordo, Jiashi Feng, Yannis Kalantidis</i>	
LEARNING FROM UNLABELLED VIDEOS USING CONTRASTIVE PREDICTIVE NEURAL 3D MAPPING	4553
<i>Adam W. Harley, Shrinidhi K. Lakshmikanth, Fangyu Li, Xian Zhou, Hsiao-Yu Fish Tung, Katerina Fragkiadaki</i>	
FROM INFERENCE TO GENERATION: END-TO-END FULLY SELF-SUPERVISED GENERATION OF HUMAN FACE FROM SPEECH.....	4572
<i>Hyeong-Seok Choi, Changdae Park, Kyogu Lee</i>	
LAMBDANET: PROBABILISTIC TYPE INFERENCE USING GRAPH NEURAL NETWORKS	4590
<i>Jiayi Wei, Maruth Goyal, Greg Durrett, Isil Dillig</i>	
MODEL-AUGMENTED ACTOR-CRITIC: BACKPROPAGATING THROUGH PATHS	4601
<i>Ignasi Clavera, Yao Fu, Pieter Abbeel</i>	
VARIATIONAL AUTOENCODERS FOR HIGHLY MULTIVARIATE SPATIAL POINT PROCESSES INTENSITIES	4615
<i>Baichuan Yuan, Xiaowei Wang, Jianxin Ma, Chang Zhou, Andrea L. Bertozzi, Hongxia Yang</i>	
DENOISING AND REGULARIZATION VIA EXPLOITING THE STRUCTURAL BIAS OF CONVOLUTIONAL GENERATORS	4629
<i>Reinhard Heckel And Mahdi Soltanolkotabi</i>	
REVISITING SELF-TRAINING FOR NEURAL SEQUENCE GENERATION	4660
<i>Junxian He, Jiatao Gu, Jiajun Shen, Marc'Aurelio Ranzato</i>	
TOWARDS A DEEP NETWORK ARCHITECTURE FOR STRUCTURED SMOOTHNESS.....	4675
<i>Haroun Habeeb, Oluwasanmi Koyejo</i>	
ON THE GLOBAL CONVERGENCE OF TRAINING DEEP LINEAR RESNETS	4689
<i>Difan Zou, Philip M. Long, Quanquan Gu</i>	

VOLUME 7

A CLOSER LOOK AT THE OPTIMIZATION LANDSCAPES OF GENERATIVE ADVERSARIAL NETWORKS	4715
<i>Hugo Berard, Gauthier Gidel, Amjad Almahairi, Pascal Vincent, Simon Lacoste-Julien</i>	

BIOLOGICALLY INSPIRED SLEEP ALGORITHM FOR INCREASED GENERALIZATION AND ADVERSARIAL ROBUSTNESS IN DEEP NEURAL NETWORKS	4733
<i>Timothy Tadros, Giri Krishnan, Ramyaa Ramyaa, Maxim Bazhenov</i>	
DISTRIBUTED BANDIT LEARNING: NEAR-OPTIMAL REGRET WITH EFFICIENT COMMUNICATION.....	4756
<i>Yuanhao Wang, Jiachen Hu, Xiaoyu Chen, Liwei Wang</i>	
SHIFTED AND SQUEEZED 8-BIT FLOATING POINT FORMAT FOR LOW-PRECISION TRAINING OF DEEP NEURAL NETWORKS.....	4787
<i>Leopold Cambier, Anahita Bhiwandiwalla, Ting Gong, Oguz H. Elibol, Mehran Nekuii, Hanlin Tang</i>	
INTRIGUING PROPERTIES OF ADVERSARIAL TRAINING AT SCALE	4799
<i>Cihang Xie, Alan Yuille</i>	
DEEP DOUBLE DESCENT: WHERE BIGGER MODELS AND MORE DATA HURT.....	4813
<i>Preetum Nakkiran, Gal Kaplun, Yamini Bansal, Tristan Yang, Boaz Barak, Ilya Sutskever</i>	
DECODING AS DYNAMIC PROGRAMMING FOR RECURRENT AUTOREGRESSIVE MODELS	4837
<i>Najam Zaidi, Trevor Cohn, Gholamreza Haffari</i>	
SYNTHESIZING PROGRAMMATIC POLICIES THAT INDUCTIVELY GENERALIZE.....	4848
<i>Jeevana Priya Inala, Osbert Bastani, Zenna Tavares, Armando Solar-Lezama</i>	
TRANSFORMER-XH: MULTI-EVIDENCE REASONING WITH EXTRA HOP ATTENTION	4869
<i>Chen Zhao, Chenyan Xiong, Corby Rosset, Xia Song, Paul Bennett, Saurabh Tiwary</i>	
GENERALIZATION THROUGH MEMORIZATION: NEAREST NEIGHBOR LANGUAGE MODELS	4885
<i>Urvashi Khandelwal, Omer Levy, Dan Jurafsky, Luke Zettlemoyer, Mike Lewis</i>	
SINGLE EPISODE POLICY TRANSFER IN REINFORCEMENT LEARNING	4898
<i>Jiachen Yang, Brenden Petersen, Hongyuan Zha, Daniel Faissol</i>	
TOWARDS STABILIZING BATCH STATISTICS IN BACKWARD PROPAGATION OF BATCH NORMALIZATION.....	4915
<i>Junjie Yan, Ruosi Wan, Xiangyu Zhang, Wei Zhang, Yichen Wei, Jian Sun</i>	
NEURQURI: NEURAL QUESTION REQUIREMENT INSPECTOR FOR ANSWERABILITY PREDICTION IN MACHINE READING COMPREHENSION.....	4932
<i>Seohyun Back, Sai Chetan Chinthakindi, Akhil Kedia, Haejun Lee, Jaegul Choo</i>	
THE EARLY PHASE OF NEURAL NETWORK TRAINING.....	4946
<i>Jonathan Frankle, David J. Schwab, Ari S. Morcos</i>	
RNNS INCREMENTALLY EVOLVING ON AN EQUILIBRIUM MANIFOLD: A PANACEA FOR VANISHING AND EXPLODING GRADIENTS?	4966
<i>Anil Kag, Ziming Zhang, Venkatesh Saligrama</i>	
EXTREME TENSORING FOR LOW-MEMORY PRECONDITIONING	4990
<i>Xinyi Chen, Naman Agarwal, Elad Hazan, Cyril Zhang, Yi Zhang</i>	
NON-AUTOREGRESSIVE DIALOG STATE TRACKING.....	5004
<i>Hung Le, Richard Socher, Steven C. H. Hoi</i>	
BAYESIAN META SAMPLING FOR FAST UNCERTAINTY ADAPTATION	5025
<i>Zhenyi Wang, Yang Zhao, Ping Yu, Ruiyi Zhang, Changyou Chen</i>	
ECONOMY STATISTICAL RECURRENT UNITS FOR INFERRING NONLINEAR GRANGER CAUSALITY	5047
<i>Saurabh Khanna, Vincent Y. F. Tan</i>	
MEMO: A DEEP NETWORK FOR FLEXIBLE COMBINATION OF EPISODIC MEMORIES	5066
<i>Andrea Banino, Adrià Puigdomènech Badia, Raphael Köster, Martin J. Chadwick, Vinicius Zambaldi, Demis Hassabis, Caswell Barry, Matthew Botvinick, Dharshan Kumaran, Charles Blundell</i>	

PROBABILISTIC CONNECTION IMPORTANCE INFERENCE AND LOSSLESS COMPRESSION OF DEEP NEURAL NETWORKS	5091
<i>Xin Xing, Long Sha, Pengyu Hong, Zuofeng Shang, Jun S. Liu</i>	
COHERENT GRADIENTS: AN APPROACH TO UNDERSTANDING GENERALIZATION IN GRADIENT DESCENT-BASED OPTIMIZATION	5103
<i>Satrajit Chatterjee</i>	
JELLY BEAN WORLD: A TESTBED FOR NEVER-ENDING LEARNING	5116
<i>Emmanouil Antonios Platanios, Abulhair Saparov, Tom Mitchell</i>	
LEARNING FROM EXPLANATIONS WITH NEURAL EXECUTION TREE	5133
<i>Ziqi Wang*, Yujia Qin*, Wenxuan Zhou, Jun Yan, Qinyuan Ye, Leonardo Neves, Zhiyuan Liu, Xiang Ren</i>	
DISCOVERING MOTOR PROGRAMS BY RECOMPOSING DEMONSTRATIONS	5151
<i>Tanmay Shankar, Shubham Tulsiani, Lerrel Pinto, Abhinav Gupta</i>	
CONVERGENCE OF GRADIENT METHODS ON BILINEAR ZERO-SUM GAMES	5163
<i>Guojun Zhang, Yaoliang Yu</i>	
COMPOSING TASK-AGNOSTIC POLICIES WITH DEEP REINFORCEMENT LEARNING	5192
<i>Ahmed H. Qureshi, Jacob J. Johnson, Yuzhe Qin, Taylor Henderson, Byron Boots, Michael C. Yip</i>	
THE LOCAL ELASTICITY OF NEURAL NETWORKS	5208
<i>Hangfeng He, Weijie Su</i>	
GRADIENT-BASED NEURAL DAG LEARNING	5227
<i>Sébastien Lachapelle, Philippe Brouillard, Tristan Deleu, Simon Lacoste-Julien</i>	
COMPOSITION-BASED MULTI-RELATIONAL GRAPH CONVOLUTIONAL NETWORKS	5250
<i>Shikhar Vashishth, Soumya Sanyal, Vikram Nitin, Partha Talukdar</i>	
CAPSULES WITH INVERTED DOT-PRODUCT ATTENTION ROUTING	5266
<i>Yao-Hung Hubert Tsai, Nitish Srivastava, Hanlin Goh, Ruslan Salakhutdinov</i>	
FSNET: COMPRESSION OF DEEP CONVOLUTIONAL NEURAL NETWORKS BY FILTER SUMMARY	5281
<i>Yingzhen Yang, Jiahui Yu, Nebojsa Jojic, Jun Huan, Thomas S. Huang</i>	
ON THE NEED FOR TOPOLOGY-AWARE GENERATIVE MODELS FOR MANIFOLD- BASED DEFENSES	5295
<i>Uyeong Jang, Susmit Jha, Somesh Jha</i>	
NEURAL EXECUTION OF GRAPH ALGORITHMS	5319
<i>Petar Velickovic, Rex Ying, Matilde Padovano, Raia Hadsell, Charles Blundell</i>	
BERTSCORE: EVALUATING TEXT GENERATION WITH BERT	5333
<i>Tianyi Zhang*, Varsha Kishore*, Felix Wu*, Kilian Q. Weinberger, Yoav Artzi</i>	
AUGMENTING NON-COLLABORATIVE DIALOG SYSTEMS WITH EXPLICIT SEMANTIC AND STRATEGIC DIALOG HISTORY	5376
<i>Yiheng Zhou, Yulia Tsvetkov, Alan W Black, Zhou Yu</i>	
PREDICTION, CONSISTENCY, CURVATURE: REPRESENTATION LEARNING FOR LOCALLY-LINEAR CONTROL	5388
<i>Nir Levine, Yinlam Chow, Rui Shu, Ang Li, Mohammad Ghavamzadeh, Hung Bui</i>	
GRAPH CONSTRAINED REINFORCEMENT LEARNING FOR NATURAL LANGUAGE ACTION SPACES	5416
<i>Prithviraj Ammanabrolu, Matthew Hausknecht</i>	
TOWARDS FAST ADAPTATION OF NEURAL ARCHITECTURES WITH META LEARNING	5440
<i>Dongze Lian, Yin Zheng, Yintao Xu, Yanxiong Lu, Leyu Lin, Peilin Zhao, Junzhou Huang, Shenghua Gao</i>	
VARIATIONAL HETERO-ENCODER RANDOMIZED GANS FOR JOINT IMAGE-TEXT MODELING	5456
<i>Hao Zhang, Bo Chen, Long Tian, Zhengjue Wang, Mingyuan Zhou</i>	

HIGHER-ORDER FUNCTION NETWORKS FOR LEARNING COMPOSABLE 3D OBJECT REPRESENTATIONS.....	5489
<i>Eric Mitchell, Selim Engin, Volkan Isler, Daniel D Lee</i>	
NEURAL MODULE NETWORKS FOR REASONING OVER TEXT	5508
<i>Nitish Gupta, Kevin Lin, Dan Roth, Sameer Singh, Matt Gardner</i>	

VOLUME 8

IMPROVED MEMORY IN RECURRENT NEURAL NETWORKS WITH SEQUENTIAL NON-NORMAL DYNAMICS.....	5525
<i>Emin Orhan, Xaq Pitkow</i>	
LEARN TO EXPLAIN EFFICIENTLY VIA NEURAL LOGIC INDUCTIVE LEARNING	5540
<i>Yuan Yang, Le Song</i>	
IMPROVING NEURAL LANGUAGE GENERATION WITH SPECTRUM CONTROL	5555
<i>Lingxiao Wang, Jing Huang, Kevin Huang, Ziniu Hu, Guangtao Wang, Quanquan Gu</i>	
SPAN RECOVERY FOR DEEP NEURAL NETWORKS WITH APPLICATIONS TO INPUT OBFUSCATION	5570
<i>Rajesh Jayaram, David P. Woodruff, Qiuyi Zhang</i>	
OBLIQUE DECISION TREES FROM DERIVATIVES OF RELU NETWORKS	5588
<i>Guang-He Lee, Tommi S. Jaakkola</i>	
PRECISION GATING: IMPROVING NEURAL NETWORK EFFICIENCY WITH DYNAMIC DUAL-PRECISION ACTIVATIONS	5606
<i>Yichi Zhang, Ritchie Zhao, Weizhe Hua, Nayun Xu, G. Edward Suh, Zhiru Zhang</i>	
PAC CONFIDENCE SETS FOR DEEP NEURAL NETWORKS VIA CALIBRATED PREDICTION	5619
<i>Sangdon Park, Osbert Bastani, Nikolai Matni, Insup Lee</i>	
DD-PPO: LEARNING NEAR-PERFECT POINTGOAL NAVIGATORS FROM 2.5 BILLION FRAMES.....	5648
<i>Erik Wijmans, Abhishek Kadian, Ari Morcos, Stefan Lee, Irfan Essa, Devi Parikh, Manolis Savva, Dhruv Batra</i>	
LEARNING TO LEARN BY ZERO-TH-ORDER ORACLE.....	5669
<i>Yangjun Ruan, Yuanhao Xiong, Sashank Reddi, Sanjiv Kumar, Cho-Jui Hsieh</i>	
METAPIX: FEW-SHOT VIDEO RETARGETING	5686
<i>Jessica Lee, Deva Ramanan, Rohit Girdhar</i>	
SLOWMO: IMPROVING COMMUNICATION-EFFICIENT DISTRIBUTED SGD WITH SLOW MOMENTUM.....	5699
<i>Jiayu Wang, Vinayak Tantia, Nicolas Ballas, Michael Rabbat</i>	
PSEUDO-LIDAR++: ACCURATE DEPTH FOR 3D OBJECT DETECTION IN AUTONOMOUS DRIVING	5726
<i>Yurong You, Yan Wang, Wei-Lun Chao, Divyansh Garg, Geoff Pleiss, Bharath Hariharan, Mark Campbell, Kilian Q. Weinberger</i>	
FOUR THINGS EVERYONE SHOULD KNOW TO IMPROVE BATCH NORMALIZATION	5748
<i>Cecilia Summers, Michael J. Dinneen</i>	
LEARNING TO SOLVE THE CREDIT ASSIGNMENT PROBLEM	5766
<i>Benjamin James Lansdell, Prashanth Ravi Prakash, Konrad Paul Kording</i>	
SAMPLING-FREE LEARNING OF BAYESIAN QUANTIZED NEURAL NETWORKS.....	5785
<i>Jiahao Su, Milan Cvitkovic, Furong Huang</i>	
DEFINE: DEEP FACTORIZED INPUT TOKEN EMBEDDINGS FOR NEURAL SEQUENCE MODELING.....	5811
<i>Sachin Mehta, Rik Koncel-Kedziorski, Mohammad Rastegari, Hannaneh Hajishirzi</i>	
DBA: DISTRIBUTED BACKDOOR ATTACKS AGAINST FEDERATED LEARNING.....	5830
<i>Chulin Xie, Keli Huang, Pin-Yu Chen, Bo Li</i>	

FAST IS BETTER THAN FREE: REVISITING ADVERSARIAL TRAINING	5849
<i>Eric Wong, Leslie Rice, J. Zico Kolter</i>	
THIEVES ON SESAME STREET! MODEL EXTRACTION OF BERT-BASED APIS	5866
<i>Kalpesh Krishna, Gaurav Singh Tomar, Ankur P. Parikh, Nicolas Papernot, Mohit Iyyer</i>	
UNDERSTANDING KNOWLEDGE DISTILLATION IN NON-AUTOREGRESSIVE MACHINE TRANSLATION	5885
<i>Chunting Zhou, Jiatao Gu, Graham Neubig</i>	
LOCALITY AND COMPOSITIONALITY IN ZERO-SHOT LEARNING	5902
<i>Tristan Sylvain, Linda Petrini, Devon Hjelm</i>	
RECURRENT NEURAL CIRCUITS FOR CONTOUR DETECTION	5923
<i>Drew Linsley*, Junkyung Kim*, Alekh Ashok, Thomas Serre</i>	
CHAMELEON: ADAPTIVE CODE OPTIMIZATION FOR EXPEDITED DEEP NEURAL NETWORK COMPILATION	5946
<i>Byoung Hoon Ahn, Prannoy Pilligundla, Amir Yazdanbakhsh, Hadi Esmaeilzadeh</i>	
INTRINSIC MOTIVATION FOR ENCOURAGING SYNERGISTIC BEHAVIOR	5963
<i>Rohan Chitnis, Shubham Tulsiani, Saurabh Gupta, Abhinav Gupta</i>	
RACT: TOWARD AMORTIZED RANKING-CRITICAL TRAINING FOR COLLABORATIVE FILTERING.....	5978
<i>Sam Lobel*, Chunyuan Li*, Jianfeng Gao, Lawrence Carin</i>	
SIGN-OPT: A QUERY-EFFICIENT HARD-LABEL ADVERSARIAL ATTACK	5997
<i>Minhao Cheng, Simranjit Singh, Patrick H. Chen, Pin-Yu Chen, Sijia Liu, Cho-Jui Hsieh</i>	
PLAYING THE LOTTERY WITH REWARDS AND MULTIPLE LANGUAGES: LOTTERY TICKETS IN RL AND NLP.....	6013
<i>Haonan Yu, Sergey Edunov, Yuandong Tian, Ari S. Morcos</i>	
LEARNING SPACE PARTITIONS FOR NEAREST NEIGHBOR SEARCH	6025
<i>Yihe Dong, Piotr Indyk, Ilya Razenshteyn, Tal Wagner</i>	
DEEPPV2D: VIDEO TO DEPTH WITH DIFFERENTIABLE STRUCTURE FROM MOTION	6041
<i>Zachary Teed, Jia Deng</i>	
TOWARDS BETTER UNDERSTANDING OF ADAPTIVE GRADIENT ALGORITHMS IN GENERATIVE ADVERSARIAL NETS.....	6060
<i>Mingrui Liu, Youssef Mroueh, Jerret Ross, Wei Zhang, Xiaodong Cui, Payel Das, Tianbao Yang</i>	
ROBUST AND INTERPRETABLE BLIND IMAGE DENOISING VIA BIAS-FREE CONVOLUTIONAL NEURAL NETWORKS.....	6083
<i>Sreyas Mohan, Zahra Kadkhodaie, Eero P. Simoncelli, Carlos Fernandez-Granda</i>	
CM3: COOPERATIVE MULTI-GOAL MULTI-STAGE MULTI-AGENT REINFORCEMENT LEARNING.....	6105
<i>Jiachen Yang, Alireza Nakhaei, David Isele, Kikuo Fujimura, Hongyuan Zha</i>	
DEEP IMITATIVE MODELS FOR FLEXIBLE INFERENCE, PLANNING, AND CONTROL	6129
<i>Nicholas Rhinehart, Rowan McAllister, Sergey Levine</i>	
PRE-TRAINING TASKS FOR EMBEDDING-BASED LARGE-SCALE RETRIEVAL	6149
<i>Wei-Cheng Chang, Felix X. Yu, Yin-Wen Chang, Yiming Yang, Sanjiv Kumar</i>	
ARE TRANSFORMERS UNIVERSAL APPROXIMATORS OF SEQUENCE-TO-SEQUENCE FUNCTIONS?.....	6161
<i>Chulhee Yun, Srinadh Bhojanapalli, Ankit Singh Rawat, Sashank Reddi, Sanjiv Kumar</i>	
META-DATASET: A DATASET OF DATASETS FOR LEARNING TO LEARN FROM FEW EXAMPLES	6184
<i>Eleni Triantafillou, Tyler Zhu, Vincent Dumoulin, Pascal Lamblin, Utku Evci, Kelvin Xu, Ross Goroshin, Carles Gelada, Kevin Swersky, Pierre-Antoine Manzagol, Hugo Larochelle</i>	
ONE-SHOT PRUNING OF RECURRENT NEURAL NETWORKS BY JACOBIAN SPECTRUM EVALUATION	6208
<i>Shunshi Zhang, Bradly C. Stadie</i>	

DIFFERENTIALLY PRIVATE META-LEARNING	6220
<i>Jeffrey Li, Mikhail Khodak, Sebastian Caldas, Ameet Talwalkar</i>	
BRIDGING MODE CONNECTIVITY IN LOSS LANDSCAPES AND ADVERSARIAL ROBUSTNESS	6238
<i>Pu Zhao, Pin-Yu Chen, Payel Das, Karthikeyan Natesan Ramamurthy, Xue Lin</i>	
OVERLEARNING REVEALS SENSITIVE ATTRIBUTES	6266
<i>Congzheng Song, Vitaly Shmatikov</i>	
ADVERSARIALLY ROBUST TRANSFER LEARNING	6278
<i>Ali Shafahi, Parsa Saadatpanah, Chen Zhu, Amin Ghiasi, Christoph Studer, David Jacobs, Tom Goldstein</i>	
LEARNING TO LINK	6292
<i>Maria-Florina Balcan, Travis Dick, Manuel Lang</i>	
DETECTING EXTRAPOLATION WITH LOCAL ENSEMBLES	6312
<i>David Madras, James Atwood, Alexander D'Amour</i>	

VOLUME 9

GLOBAL RELATIONAL MODELS OF SOURCE CODE	6331
<i>Vincent J. Hellendoorn, Charles Sutton, Rishabh Singh, Petros Maniatis, David Bieber</i>	
SELECTION VIA PROXY: EFFICIENT DATA SELECTION FOR DEEP LEARNING	6343
<i>Cody Coleman, Christopher Yeh, Stephen Mussmann, Baharan Mirzasoleiman, Peter Bailis, Percy Liang, Jure Leskovec, Matei Zaharia</i>	
SHORT AND SPARSE DECONVOLUTION --- A GEOMETRIC APPROACH	6368
<i>Yenson Lau, Qing Qu, Han-Wen Kuo, Pengcheng Zhou, Yuqian Zhang, John Wright</i>	
STOCHASTIC WEIGHT AVERAGING IN PARALLEL: LARGE-BATCH TRAINING THAT GENERALIZES WELL	6384
<i>Vipul Gupta, Santiago Aklé Serrano, Dennis Decoste</i>	
ADJUSTABLE REAL-TIME STYLE TRANSFER	6396
<i>Mohammad Babaeizadeh, Golnaz Ghiasi</i>	
UNPAIRED POINT CLOUD COMPLETION ON REAL SCANS USING ADVERSARIAL TRAINING	6413
<i>Xuelin Chen, Baoquan Chen, Niloy J. Mitra</i>	
EFFICIENT RIEMANNIAN OPTIMIZATION ON THE STIEFEL MANIFOLD VIA THE CAYLEY TRANSFORM	6430
<i>Jun Li, Fuxin Li, Sinisa Todorovic</i>	
AMRL: AGGREGATED MEMORY FOR REINFORCEMENT LEARNING	6446
<i>Jacob Beck, Kamil Ciosek, Sam Devlin, Sebastian Tschiatschek, Cheng Zhang, Katja Hofmann</i>	
SCALABLE MODEL COMPRESSION BY ENTROPY PENALIZED REPARAMETERIZATION	6466
<i>Deniz Oktay, Johannes Ballé, Saurabh Singh, Abhinav Shrivastava</i>	
DYNAMIC TIME LAG REGRESSION: PREDICTING WHAT & WHEN	6477
<i>Mandar Chandorkar, Cyril Furtlehner, Bala Poduval, Enrico Camporeale, Michele Sebag</i>	
SEMI-SUPERVISED GENERATIVE MODELING FOR CONTROLLABLE SPEECH SYNTHESIS	6498
<i>Raza Habib, Soroosh Mariooryad, Matt Shannon, Eric Battenberg, Rj Skerry-Ryan, Daisy Stanton, David Kao, Tom Bagby</i>	
NEURAL TEXT GENERATION WITH UNLIKELIHOOD TRAINING	6516
<i>Sean Welleck, Iliia Kulikov, Stephen Roller, Emily Dinan, Kyunghyun Cho, Jason Weston</i>	
PURE AND SPURIOUS CRITICAL POINTS: A GEOMETRIC STUDY OF LINEAR NETWORKS	6534
<i>Matthew Trager, Kathlén Kohn, Joan Bruna</i>	

LEARNING HEURISTICS FOR QUANTIFIED BOOLEAN FORMULAS THROUGH REINFORCEMENT LEARNING	6559
<i>Gil Lederman, Markus Rabe, Sanjit Seshia, Edward A. Lee</i>	
CAQL: CONTINUOUS ACTION Q-LEARNING	6577
<i>Moonkyung Ryu, Yinlam Chow, Ross Anderson, Christian Tjandraatmadja, Craig Boutilier</i>	
ADAPTIVE STRUCTURAL FINGERPRINTS FOR GRAPH ATTENTION NETWORKS	6602
<i>Kai Zhang, Yaokang Zhu, Jun Wang, Jie Zhang</i>	
REMIXMATCH: SEMI-SUPERVISED LEARNING WITH DISTRIBUTION MATCHING AND AUGMENTATION ANCHORING	6614
<i>David Berthelot, Nicholas Carlini, Ekin D. Cubuk, Alex Kurakin, Kihyuk Sohn, Han Zhang, Colin Raffel</i>	
IDENTITY CRISIS: MEMORIZATION AND GENERALIZATION UNDER EXTREME OVERPARAMETERIZATION	6627
<i>Chiyuan Zhang, Samy Bengio, Moritz Hardt, Michael C. Mozer, Yoram Singer</i>	
IMPROVED SAMPLE COMPLEXITIES FOR DEEP NEURAL NETWORKS AND ROBUST CLASSIFICATION VIA AN ALL-LAYER MARGIN	6666
<i>Colin Wei, Tengyu Ma</i>	
AUGMENTING GENETIC ALGORITHMS WITH DEEP NEURAL NETWORKS FOR EXPLORING THE CHEMICAL SPACE	6701
<i>Akshatkumar Nigam, Pascal Friederich, Mario Krenn, Alan Aspuru-Guzik</i>	
RAPID LEARNING OR FEATURE REUSE? TOWARDS UNDERSTANDING THE EFFECTIVENESS OF MAML	6715
<i>Aniruddh Raghu, Maithra Raghu, Samy Bengio, Oriol Vinyals</i>	
IMITATION LEARNING VIA OFF-POLICY DISTRIBUTION MATCHING	6736
<i>Ilya Kostrikov, Ofir Nachum, Jonathan Tompson</i>	
REANALYSIS OF VARIANCE REDUCED TEMPORAL DIFFERENCE LEARNING	6747
<i>Tengyu Xu, Zhe Wang, Yi Zhou, Yingbin Liang</i>	
MINIMIZING FLOPS TO LEARN EFFICIENT SPARSE REPRESENTATIONS	6774
<i>Biswajit Paria, Chih-Kuan Yeh, Ian E. H. Yen, Ning Xu, Pradeep Ravikumar, Barnabás Póczos</i>	
BUDGETED TRAINING: RETHINKING DEEP NEURAL NETWORK TRAINING UNDER RESOURCE CONSTRAINTS	6794
<i>Mengtian Li, Ersin Yumer, Deva Ramanan</i>	
DEEP SEMI-SUPERVISED ANOMALY DETECTION	6814
<i>Lukas Ruff, Robert A. Vandermeulen, Nico Görnitz, Alexander Binder, Emmanuel Müller, Klaus-Robert Müller, Marius Kloft</i>	
SIGN BITS ARE ALL YOU NEED FOR BLACK-BOX ATTACKS	6837
<i>Abdullah Al-Dujaili, Una-May O'Reilly</i>	
REINFORCED ACTIVE LEARNING FOR IMAGE SEGMENTATION	6862
<i>Arantxa Casanova, Pedro O. Pinheiro, Negar Rostamzadeh, Christopher J. Pal</i>	
ON THE "STEERABILITY" OF GENERATIVE ADVERSARIAL NETWORKS	6879
<i>Ali Jahanian*, Lucy Chai*, Phillip Isola</i>	
LEARNED STEP SIZE QUANTIZATION	6910
<i>Steven K. Esser, Jeffrey L. McKinstry, Deepika Bablani, Rathinakumar Appuswamy, Dharmendra S. Modha</i>	
STOCHASTIC CONDITIONAL GENERATIVE NETWORKS WITH BASIS DECOMPOSITION	6922
<i>Ze Wang, Xiuyuan Cheng, Guillermo Sapiro, Qiang Qiu</i>	
LANGUAGE GANS FALLING SHORT	6939
<i>Massimo Caccia, Lucas Caccia, William Fedus, Hugo Larochelle, Joelle Pineau, Laurent Charlin</i>	

GENESIS: GENERATIVE SCENE INFERENCE AND SAMPLING WITH OBJECT-CENTRIC LATENT REPRESENTATIONS	6957
<i>Martin Engelcke, Adam R. Kosiosek, Oiwi Parker Jones, Ingmar Posner</i>	
UNDERSTANDING THE LIMITATIONS OF VARIATIONAL MUTUAL INFORMATION ESTIMATORS	6974
<i>Jiaming Song, Stefano Ermon</i>	
FEATURE INTERACTION INTERPRETABILITY: A CASE FOR EXPLAINING AD-RECOMMENDATION SYSTEMS VIA NEURAL INTERACTION DETECTION.....	6992
<i>Michael Tsang, Dehua Cheng, Hanpeng Liu, Xue Feng, Eric Zhou, Yan Liu</i>	
UNSUPERVISED MODEL SELECTION FOR VARIATIONAL DISENTANGLED REPRESENTATION LEARNING.....	7011
<i>Sunny Duan, Loic Matthey, Andre Saraiva, Nick Watters, Chris Burgess, Alexander Lerchner, Irina Higgins</i>	
A THEORETICAL ANALYSIS OF THE NUMBER OF SHOTS IN FEW-SHOT LEARNING.....	7040
<i>Tianshi Cao, Marc T Law, Sanja Fidler</i>	
DYNAMICAL DISTANCE LEARNING FOR SEMI-SUPERVISED AND UNSUPERVISED SKILL DISCOVERY	7057
<i>Kristian Hartikainen, Xinyang Geng, Tuomas Haarnoja, Sergey Levine</i>	
ON THE VARIANCE OF THE ADAPTIVE LEARNING RATE AND BEYOND	7074
<i>Liyuan Liu, Haoming Jiang, Pengcheng He, Weizhu Chen, Xiaodong Liu, Jianfeng Gao, Jiawei Han</i>	
QUANTIFYING THE COST OF RELIABLE PHOTO AUTHENTICATION VIA HIGH-PERFORMANCE LEARNED LOSSY REPRESENTATIONS.....	7087
<i>Pawel Korus, Nasir Memon</i>	
OPTION DISCOVERY USING DEEP SKILL CHAINING	7114
<i>Akhil Bagaria, George Konidaris</i>	

VOLUME 10

V4D: 4D CONVOLUTIONAL NEURAL NETWORKS FOR VIDEO-LEVEL REPRESENTATION LEARNING.....	7135
<i>Shiwen Zhang, Sheng Guo, Weilin Huang, Matthew R. Scott, Limin Wang</i>	
LEARNING TO REPRESENT PROGRAMS WITH PROPERTY SIGNATURES	7151
<i>Augustus Odena, Charles Sutton</i>	
GENERATIVE RATIO MATCHING NETWORKS	7164
<i>Akash Srivastava, Kai Xu, Michael U. Gutmann, Charles Sutton</i>	
IN SEARCH FOR A SAT-FRIENDLY BINARIZED NEURAL NETWORK ARCHITECTURE.....	7182
<i>Nina Narodytska, Hongce Zhang, Aarti Gupta, Toby Walsh</i>	
ACTOR-CRITIC PROVABLY FINDS NASH EQUILIBRIA OF LINEAR-QUADRATIC MEAN-FIELD GAMES.....	7199
<i>Zuyue Fu, Zhuoran Yang, Yongxin Chen, Zhaoran Wang</i>	
THE ASYMPTOTIC SPECTRUM OF THE HESSIAN OF DNN THROUGHOUT TRAINING	7253
<i>Arthur Jacot, Franck Gabriel, Clement Hongler</i>	
TREE-STRUCTURED ATTENTION WITH HIERARCHICAL ACCUMULATION.....	7279
<i>Xuan-Phi Nguyen, Shafiq Joty, Steven Hoi, Richard Socher</i>	
DEEP 3D PAN VIA LOCAL ADAPTIVE "T-SHAPED" CONVOLUTIONS WITH GLOBAL AND LOCAL ADAPTIVE DILATIONS	7294
<i>Juan Luis Gonzalez Bello, Munchurl Kim</i>	
LOW-RESOURCE KNOWLEDGE-GROUNDED DIALOGUE GENERATION	7314
<i>Xueliang Zhao, Wei Wu, Chongyang Tao, Can Xu, Dongyan Zhao, Rui Yan</i>	
A TARGET-AGNOSTIC ATTACK ON DEEP MODELS: EXPLOITING SECURITY VULNERABILITIES OF TRANSFER LEARNING	7330
<i>Shahbaz Rezaei, Xin Liu</i>	

ON COMPUTATION AND GENERALIZATION OF GENERATIVE ADVERSARIAL IMITATION LEARNING	7344
<i>Minshuo Chen, Yizhou Wang, Tianyi Liu, Zhuoran Yang, Xingguo Li, Zhaoran Wang, Tuo Zhao</i>	
FEW-SHOT LEARNING ON GRAPHS VIA SUPER-CLASSES BASED ON GRAPH SPECTRAL MEASURES	7375
<i>Jatin Chauhan, Deepak Nathani, Manohar Kaul</i>	
MULTIPLICATIVE INTERACTIONS AND WHERE TO FIND THEM	7394
<i>Siddhant M. Jayakumar, Wojciech M. Czarnecki, Jacob Menick, Jonathan Schwarz, Jack Rae, Simon Osindero, Yee Whye Teh, Tim Harley, Razvan Pascanu</i>	
CONTINUAL LEARNING WITH BAYESIAN NEURAL NETWORKS FOR NON- STATIONARY DATA	7410
<i>Richard Kurle, Botond Cseke, Alexej Klushyn, Patrick Van Der Smagt, Stephan Günnemann</i>	
SADAM: A VARIANT OF ADAM FOR STRONGLY CONVEX FUNCTIONS	7433
<i>Guanghui Wang, Shiyin Lu, Quan Cheng, Wei-Wei Tu, Lijun Zhang</i>	
GENERALIZATION BOUNDS FOR DEEP CONVOLUTIONAL NEURAL NETWORKS	7454
<i>Philip M. Long, Hanie Sedghi</i>	
A FAIR COMPARISON OF GRAPH NEURAL NETWORKS FOR GRAPH CLASSIFICATION	7469
<i>Federico Errica, Marco Podda, Davide Bacciu, Alessio Micheli</i>	
FINDING AND VISUALIZING WEAKNESSES OF DEEP REINFORCEMENT LEARNING AGENTS	7483
<i>Christian Rupprecht, Cyril Ibrahim, Christopher J. Pal</i>	
COMPUTATION REALLOCATION FOR OBJECT DETECTION	7505
<i>Feng Liang, Chen Lin, Ronghao Guo, Ming Sun, Wei Wu, Junjie Yan, Wanli Ouyang</i>	
FROM VARIATIONAL TO DETERMINISTIC AUTOENCODERS	7517
<i>Partha Ghosh, Mehdi S. M. Sajjadi, Antonio Vergari, Michael Black, Bernhard Scholkopf</i>	
ADVERSARIAL ROBUST REPRESENTATIONS WITH SMOOTH ENCODERS	7542
<i>Taylan Cemgil, Sumedh Ghaisas, Krishnamurthy (Dj) Dvijotham, Pushmeet Kohli</i>	
ASSEMBLENET: SEARCHING FOR MULTI-STREAM NEURAL CONNECTIVITY IN VIDEO ARCHITECTURES	7562
<i>Michael S. Ryoo, Aj Piergiovanni, Mingxing Tan, Anelia Angelova</i>	
RECLOR: A READING COMPREHENSION DATASET REQUIRING LOGICAL REASONING	7577
<i>Weihao Yu, Zihang Jiang, Yanfei Dong, Jiashi Feng</i>	
ORDER LEARNING AND ITS APPLICATION TO AGE ESTIMATION	7603
<i>Kyungsun Lim, Nyeong-Ho Shin, Young-Yoon Lee, Chang-Su Kim</i>	
EFFICIENT AND INFORMATION-PRESERVING FUTURE FRAME PREDICTION AND BEYOND	7623
<i>Wei Yu, Yichao Lu, Steve Easterbrook, Sanja Fidler</i>	
NAS EVALUATION IS FRUSTRATINGLY HARD	7637
<i>Antoine Yang, Pedro M. Esperança, Fabio M. Carlucci</i>	
CLN2INV: LEARNING LOOP INVARIANTS WITH CONTINUOUS LOGIC NETWORKS	7650
<i>Gabriel Ryan, Justin Wong, Jianan Yao, Ronghui Gu, Suman Jana</i>	
SCALABLE NEURAL METHODS FOR REASONING WITH A SYMBOLIC KNOWLEDGE BASE	7671
<i>William W. Cohen, Haitian Sun, R. Alex Hofer, Matthew Siegler</i>	
A CONSTRUCTIVE PREDICTION OF THE GENERALIZATION ERROR ACROSS SCALES	7688
<i>Jonathan S. Rosenfeld, Amir Rosenfeld, Yonatan Belinkov, Nir Shavit</i>	
AN INDUCTIVE BIAS FOR DISTANCES: NEURAL NETS THAT RESPECT THE TRIANGLE INEQUALITY	7718
<i>Silviu Pitis, Harris Chan, Kiarash Jamali, Jimmy Ba</i>	

PHYSICS-AS-INVERSE-GRAPHICS: UNSUPERVISED PHYSICAL PARAMETER ESTIMATION FROM VIDEO	7747
<i>Miguel Jaques, Michael Burke, Timothy Hospedales</i>	
COUNTERFACTUALS UNCOVER THE MODULAR STRUCTURE OF DEEP GENERATIVE MODELS	7763
<i>Michel Besserve, Arash Mehrjou, Rémy Sun, Bernhard Schölkopf</i>	
GAP-AWARE MITIGATION OF GRADIENT STALENESS	7789
<i>Saar Barkai, Ido Hakimi, Assaf Schuster</i>	
ENSEMBLE DISTRIBUTION DISTILLATION	7819
<i>Andrey Malinin, Bruno Mlodozienec, Mark Gales</i>	
DEFORMABLE KERNELS: ADAPTING EFFECTIVE RECEPTIVE FIELDS FOR OBJECT DEFORMATION	7841
<i>Hang Gao, Xizhou Zhu, Stephen Lin, Jifeng Dai</i>	
VL-BERT: PRE-TRAINING OF GENERIC VISUAL-LINGUISTIC REPRESENTATIONS	7856
<i>Weijie Su, Xizhou Zhu, Yue Cao, Bin Li, Lewei Lu, Furu Wei, Jifeng Dai</i>	
OPTIMISTIC EXPLORATION EVEN WITH A PESSIMISTIC INITIALISATION	7872
<i>Tabish Rashid, Bei Peng, Wendelin Boehmer, Shimon Whiteson</i>	
CERTIFIED ROBUSTNESS FOR TOP-K PREDICTIONS AGAINST ADVERSARIAL PERTURBATIONS VIA RANDOMIZED SMOOTHING.....	7901
<i>Jinyuan Jia, Xiaoyu Cao, Binghui Wang, Neil Zhenqiang Gong</i>	

VOLUME 11

IDENTIFYING THROUGH FLOWS FOR RECOVERING LATENT REPRESENTATIONS	7928
<i>Shen Li, Bryan Hooi, Gim Hee Lee</i>	
ROBUST TRAINING WITH ENSEMBLE CONSENSUS.....	7943
<i>Jisoo Lee, Sae-Young Chung</i>	
SELF-ADVERSARIAL LEARNING WITH COMPARATIVE DISCRIMINATION FOR TEXT GENERATION.....	7958
<i>Wangchunshu Zhou, Tao Ge, Ke Xu, Furu Wei, Ming Zhou</i>	
VID2GAME: CONTROLLABLE CHARACTERS EXTRACTED FROM REAL-WORLD VIDEOS	7975
<i>Oran Gafni, Lior Wolf, Yaniv Taigman</i>	
ACTION SEMANTICS NETWORK: CONSIDERING THE EFFECTS OF ACTIONS IN MULTIAGENT SYSTEMS	7991
<i>Weixun Wang, Tianpei Yang, Yong Liu, Jianye Hao, Xiaotian Hao, Yujing Hu, Yingfeng Chen, Changjie Fan, Yang Gao</i>	
LEARNING EFFICIENT PARAMETER SERVER SYNCHRONIZATION POLICIES FOR DISTRIBUTED SGD.....	8009
<i>Rong Zhu, Sheng Yang, Andreas Pfadler, Zhengping Qian, Jingren Zhou</i>	
RELATIONAL STATE-SPACE MODEL FOR STOCHASTIC MULTI-OBJECT SYSTEMS	8019
<i>Fan Yang, Ling Chen, Fan Zhou, Yusong Gao, Wei Cao</i>	
PIECEWISE LINEAR ACTIVATIONS SUBSTANTIALLY SHAPE THE LOSS SURFACES OF NEURAL NETWORKS.....	8037
<i>Fengxiang He, Bohan Wang, Dacheng Tao</i>	
NOVELTY DETECTION VIA BLURRING	8082
<i>Sungik Choi, Sae-Young Chung</i>	
BOUNDS ON OVER-PARAMETERIZATION FOR GUARANTEED EXISTENCE OF DESCENT PATHS IN SHALLOW RELU NETWORKS	8096
<i>Arsalan Sharifnassab, Saber Salehkaleybar, S. Jamaloddin Golestani</i>	
DATA-INDEPENDENT NEURAL PRUNING VIA CORESETS.....	8121
<i>Ben Mussay, Margarita Osadchy, Vladimir Braverman, Samson Zhou, Dan Feldman</i>	

DEEP NETWORK CLASSIFICATION BY SCATTERING AND HOMOTOPY DICTIONARY LEARNING.....	8135
<i>John Zarka, Louis Thiry, Tomas Angles, Stephane Mallat</i>	
Q-LEARNING WITH UCB EXPLORATION IS SAMPLE EFFICIENT FOR INFINITE- HORIZON MDP	8147
<i>Yuanhao Wang, Kefan Dong, Xiaoyu Chen, Liwei Wang</i>	
MIXOUT: EFFECTIVE REGULARIZATION TO FINETUNE LARGE-SCALE PRETRAINED LANGUAGE MODELS	8165
<i>Cheolhyoung Lee, Kyunghyun Cho, Wanmo Kang</i>	
I AM GOING MAD: MAXIMUM DISCREPANCY COMPETITION FOR COMPARING CLASSIFIERS ADAPTIVELY.....	8182
<i>Haotao Wang, Tianlong Chen, Zhangyang Wang, Kede Ma</i>	
BLACK-BOX ADVERSARIAL ATTACK WITH TRANSFERABLE MODEL-BASED EMBEDDING.....	8195
<i>Zhichao Huang, Tong Zhang</i>	
COMPOSITIONAL LANGUAGES EMERGE IN A NEURAL ITERATED LEARNING MODEL	8215
<i>Yi Ren, Shangmin Guo, Matthieu Labeau, Shay B. Cohen, Simon Kirby</i>	
POPULATION-GUIDED PARALLEL POLICY SEARCH FOR REINFORCEMENT LEARNING	8237
<i>Whiyoung Jung, Giseung Park, Youngchul Sung</i>	
VARIATIONAL RECURRENT MODELS FOR SOLVING PARTIALLY OBSERVABLE CONTROL TASKS	8264
<i>Dongqi Han, Kenji Doya, Jun Tani</i>	
GAT: GENERATIVE ADVERSARIAL TRAINING FOR ADVERSARIAL EXAMPLE DETECTION AND ROBUST CLASSIFICATION.....	8283
<i>Xuwan Yin, Soheil Kolouri, Gustavo K Rohde</i>	
DETECTING AND DIAGNOSING ADVERSARIAL IMAGES WITH CLASS-CONDITIONAL CAPSULE RECONSTRUCTIONS	8310
<i>Yao Qin, Nicholas Frosst, Sara Sabour, Colin Raffel, Garrison Cottrell, Geoffrey Hinton</i>	
MACER: ATTACK-FREE AND SCALABLE ROBUST TRAINING VIA MAXIMIZING CERTIFIED RADIUS.....	8328
<i>Runtian Zhai, Chen Dan, Di He, Huan Zhang, Boqing Gong, Pradeep Ravikumar, Cho-Jui Hsieh, Liwei Wang</i>	
SEMANTICALLY-GUIDED REPRESENTATION LEARNING FOR SELF-SUPERVISED MONOCULAR DEPTH	8348
<i>Vitor Guizilini, Rui Hou, Jie Li, Rares Ambrus, Adrien Gaidon</i>	
STOCHASTIC AUC MAXIMIZATION WITH DEEP NEURAL NETWORKS	8362
<i>Mingrui Liu, Zhuoning Yuan, Yiming Ying, Tianbao Yang</i>	
DIFFERENCE-SEEKING GENERATIVE ADVERSARIAL NETWORK--UNSEEN SAMPLE GENERATION.....	8386
<i>Yi Lin Sung, Sung-Hsien Hsieh, Soo-Chang Pei, Chun-Shien Lu</i>	
FASTERSEG: SEARCHING FOR FASTER REAL-TIME SEMANTIC SEGMENTATION.....	8407
<i>Wuyang Chen, Xinyu Gong, Xianming Liu, Qian Zhang, Yuan Li, Zhangyang Wang</i>	
LEARNING EXECUTION THROUGH NEURAL CODE FUSION	8421
<i>Zhan Shi, Kevin Swersky, Daniel Tarlow, Parthasarathy Ranganathan, Milad Hashemi</i>	
EDITABLE NEURAL NETWORKS.....	8434
<i>Anton Sinitsin, Vsevolod Plokhotnyuk, Dmitry Pyrkin, Sergei Popov, Artem Babenko</i>	
CAN GRADIENT CLIPPING MITIGATE LABEL NOISE?.....	8446
<i>Aditya Krishna Menon, Ankit Singh Rawat, Sashank J. Reddi, Sanjiv Kumar</i>	
PRETRAINED ENCYCLOPEDIA: WEAKLY SUPERVISED KNOWLEDGE-PRETRAINED LANGUAGE MODEL.....	8472
<i>Wenhan Xiong, Jingfei Du, William Yang Wang, Veselin Stoyanov</i>	

PRUNED GRAPH SCATTERING TRANSFORMS	8485
<i>Vassilis N. Ioannidis, Siheng Chen, Georgios B. Giannakis</i>	
GLAD: LEARNING SPARSE GRAPH RECOVERY	8499
<i>Harsh Shrivastava, Xinshi Chen, Binghong Chen, Guanghui Lan, Srinivas Aluru, Han Liu, Le Song</i>	
VIDEOFLOW: A CONDITIONAL FLOW-BASED MODEL FOR STOCHASTIC VIDEO GENERATION.....	8521
<i>Manoj Kumar, Mohammad Babaeizadeh, Dumitru Erhan, Chelsea Finn, Sergey Levine, Laurent Dinh, Durk Kingma</i>	
ADVERSARIAL POLICIES: ATTACKING DEEP REINFORCEMENT LEARNING	8539
<i>Adam Gleave, Michael Dennis, Cody Wild, Neel Kant, Sergey Levine, Stuart Russell</i>	
ESCAPING SADDLE POINTS FASTER WITH STOCHASTIC MOMENTUM.....	8555
<i>Jun-Kun Wang, Chi-Heng Lin, Jacob Abernethy</i>	
FEW-SHOT TEXT CLASSIFICATION WITH DISTRIBUTIONAL SIGNATURES	8588
<i>Yujia Bao, Menghua Wu, Shiyu Chang, Regina Barzilay</i>	
GEOMETRIC INSIGHTS INTO THE CONVERGENCE OF NONLINEAR TD LEARNING	8612
<i>David Brandfonbrener, Joan Bruna</i>	
LEARNING SELF-CORRECTABLE POLICIES AND VALUE FUNCTIONS FROM DEMONSTRATIONS WITH NEGATIVE SAMPLING	8629
<i>Yuping Luo, Huazhe Xu, Tengyu Ma</i>	
EXPLORING MODEL-BASED PLANNING WITH POLICY NETWORKS	8647
<i>Tingwu Wang, Jimmy Ba</i>	
ON IDENTIFIABILITY IN TRANSFORMERS	8668
<i>Gino Brunner, Yang Liu, Damian Pascual, Oliver Richter, Massimiliano Ciaramita, Roger Wattenhofer</i>	
AUTOMATED CURRICULUM GENERATION THROUGH SETTER-SOLVER INTERACTIONS	8703
<i>Sebastien Racaniere, Andrew Lampinen, Adam Santoro, David Reichert, Vlad Firoiu, Timothy Lillicrap</i>	
PROGRESSIVE MEMORY BANKS FOR INCREMENTAL DOMAIN ADAPTATION.....	8721
<i>Nabiha Asghar, Lili Mou, Kira A. Selby, Kevin D. Pantasdo, Pascal Poupart, Xin Jiang</i>	

VOLUME 12

WHAT GRAPH NEURAL NETWORKS CANNOT LEARN: DEPTH VS WIDTH	8737
<i>Andreas Loukas</i>	
RTFM: GENERALISING TO NEW ENVIRONMENT DYNAMICS VIA READING	8754
<i>Victor Zhong, Tim Rocktäschel, Edward Grefenstette</i>	
INPUT COMPLEXITY AND OUT-OF-DISTRIBUTION DETECTION WITH LIKELIHOOD-BASED GENERATIVE MODELS.....	8771
<i>Joan Serra, David Álvarez, Vicenç Gómez, Olga Slizovskaia, José F. Núñez, Jordi Luque</i>	
FUNCTIONAL VS. PARAMETRIC EQUIVALENCE OF RELU NETWORKS	8786
<i>Mary Phuong, Christoph H. Lampert</i>	
DISENTANGLING FACTORS OF VARIATIONS USING FEW LABELS.....	8809
<i>Francesco Locatello, Michael Tschannen, Stefan Bauer, Gunnar Rätsch, Bernhard Schölkopf, Olivier Bachem</i>	
A CRITICAL ANALYSIS OF SELF-SUPERVISION, OR WHAT WE CAN LEARN FROM A SINGLE IMAGE.....	8833
<i>Asano Ym., Rupprecht C., Vedaldi A.</i>	
ACCELERATING SGD WITH MOMENTUM FOR OVER-PARAMETERIZED LEARNING	8849
<i>Chaoyue Liu, Mikhail Belkin</i>	

INTERPRETABLE COMPLEX-VALUED NEURAL NETWORKS FOR PRIVACY PROTECTION	8874
<i>Liyao Xiang, Hao Zhang, Haotian Ma, Yifan Zhang, Jie Ren, Quanshi Zhang</i>	
V-MPO: ON-POLICY MAXIMUM A POSTERIORI POLICY OPTIMIZATION FOR DISCRETE AND CONTINUOUS CONTROL.....	8887
<i>H. Francis Song, Abbas Abdolmaleki, Jost Tobias Springenberg, Aidan Clark, Hubert Soyer, Jack W. Rae, Seb Noury, Arun Ahuja, Siqi Liu, Dhruva Tirumala, Nicolas Heess, Dan Belov, Martin Riedmiller, Matthew M. Botvinick</i>	
IMPROVING ADVERSARIAL ROBUSTNESS REQUIRES REVISITING MISCLASSIFIED EXAMPLES.....	8909
<i>Yisen Wang, Difan Zou, Jinfeng Yi, James Bailey, Xingjun Ma, Quanquan Gu</i>	
DIVIDEMIX: LEARNING WITH NOISY LABELS AS SEMI-SUPERVISED LEARNING	8923
<i>Junnan Li, Richard Socher, Steven C. H. Hoi</i>	
FOOLING DETECTION ALONE IS NOT ENOUGH: ADVERSARIAL ATTACK AGAINST MULTIPLE OBJECT TRACKING.....	8937
<i>Yunhan Jia, Yantao Lu, Junjie Shen, Qi Alfred Chen, Hao Chen, Zhenyu Zhong, Tao Wei</i>	
WATCH, TRY, LEARN: META-LEARNING FROM DEMONSTRATIONS AND REWARDS.....	8952
<i>Allan Zhou, Eric Jang, Daniel Kappler, Alex Herzog, Mohi Khansari, Paul Wohlhart, Yunfei Bai, Mrinal Kalakrishnan, Sergey Levine, Chelsea Finn</i>	
LOGIC AND THE 2-SIMPLICIAL TRANSFORMER	8969
<i>James Clift, Dmitry Doryn, Daniel Mufet, James Wallbridge</i>	
AE-OT: A NEW GENERATIVE MODEL BASED ON EXTENDED SEMI-DISCRETE OPTIMAL TRANSPORT	8996
<i>Dongsheng An, Yang Guo, Na Lei, Zhongxuan Luo, Shing-Tung Yau, Xianfeng Gu</i>	
EXPLORATION IN REINFORCEMENT LEARNING WITH DEEP COVERING OPTIONS.....	9015
<i>Yuu Jinnai, Jee Won Park, Marlos C. Machado, George Konidaris</i>	
LEARNING DISENTANGLED REPRESENTATIONS FOR COUNTERFACTUAL REGRESSION	9028
<i>Negar Hassanpour, Russell Greiner</i>	
ANALYSIS OF VIDEO FEATURE LEARNING IN TWO-STREAM CNNs ON THE EXAMPLE OF ZEBRAFISH SWIM BOUT CLASSIFICATION	9039
<i>Bennet Breier, Arno Onken</i>	
ROBUST LOCAL FEATURES FOR IMPROVING THE GENERALIZATION OF ADVERSARIAL TRAINING.....	9057
<i>Chuanbiao Song, Kun He, Jiadong Lin, Liwei Wang, John E. Hopcroft</i>	
REINFORCEMENT LEARNING WITH COMPETITIVE ENSEMBLES OF INFORMATION-CONSTRAINED PRIMITIVES.....	9069
<i>Anirudh Goyal, Shagun Sodhani, Jonathan Binas, Xue Bin Peng, Sergey Levine, Yoshua Bengio</i>	
LEARNING THE ARROW OF TIME FOR PROBLEMS IN REINFORCEMENT LEARNING	9090
<i>Nasim Rahaman, Steffen Wolf, Anirudh Goyal, Roman Remme, Yoshua Bengio</i>	
THE VARIATIONAL BANDWIDTH BOTTLENECK: STOCHASTIC EVALUATION ON AN INFORMATION BUDGET	9116
<i>Anirudh Goyal, Yoshua Bengio, Matthew Botvinick, Sergey Levine</i>	
THE IMPLICIT BIAS OF DEPTH: HOW INCREMENTAL LEARNING DRIVES GENERALIZATION.....	9133
<i>Daniel Gissin, Shai Shalev-Shwartz, Amit Daniely</i>	
RETHINKING SOFTMAX CROSS-ENTROPY LOSS FOR ADVERSARIAL ROBUSTNESS.....	9158
<i>Tianyu Pang, Kun Xu, Yinpeng Dong, Chao Du, Ning Chen, Jun Zhu</i>	

MEASURING COMPOSITIONAL GENERALIZATION: A COMPREHENSIVE METHOD ON REALISTIC DATA	9177
<i>Daniel Keysers, Nathanael Schärli, Nathan Scales, Hylke Buisman, Daniel Furrer, Sergii Kashubin, Nikola Momchev, Danila Sinopalnikov, Lukasz Stafiniak, Tibor Tihon, Dmitry Tsarkov, Xiao Wang, Marc Van Zee, Olivier Bousquet</i>	
THEORY AND EVALUATION METRICS FOR LEARNING DISENTANGLED REPRESENTATIONS.....	9215
<i>Kien Do, Truyen Tran</i>	
MIXUP INFERENCE: BETTER EXPLOITING MIXUP TO DEFEND ADVERSARIAL ATTACKS	9245
<i>Tianyu Pang*, Kun Xu*, Jun Zhu</i>	
DYNAMICALLY PRUNED MESSAGE PASSING NETWORKS FOR LARGE-SCALE KNOWLEDGE GRAPH REASONING	9259
<i>Xiaoran Xu, Wei Feng, Yunsheng Jiang, Xiaohui Xie, Zhiqing Sun, Zhi-Hong Deng</i>	
ARE PRE-TRAINED LANGUAGE MODELS AWARE OF PHRASES? SIMPLE BUT STRONG BASELINES FOR GRAMMAR INDUCTION.....	9289
<i>Taeuk Kim, Jihun Choi, Daniel Edmiston, Sang-Goo Lee</i>	
FSPool: LEARNING SET REPRESENTATIONS WITH FEATUREWISE SORT POOLING	9311
<i>Yan Zhang, Jonathon Hare, Adam Prügel-Bennett</i>	
MULTI-AGENT REINFORCEMENT LEARNING FOR NETWORKED SYSTEM CONTROL	9331
<i>Tianshu Chu, Sandeep Chinchali, Sachin Katti</i>	
HIERARCHICAL FORESIGHT: SELF-SUPERVISED LEARNING OF LONG-HORIZON TASKS VIA VISUAL SUBGOAL GENERATION.....	9348
<i>Suraj Nair, Chelsea Finn</i>	
NEURAL STORED-PROGRAM MEMORY	9365
<i>Hung Le, Truyen Tran, Svetha Venkatesh</i>	
ES-MAML: SIMPLE HESSIAN-FREE META LEARNING.....	9392
<i>Xingyou Song, Wenbo Gao, Yuxiang Yang, Krzysztof Choromanski, Aldo Pacchiano, Yunhao Tang</i>	
TABFACT: A LARGE-SCALE DATASET FOR TABLE-BASED FACT VERIFICATION	9411
<i>Wenhu Chen, Hongmin Wang, Jianshu Chen, Yunkai Zhang, Hong Wang, Shiyang Li, Xiyou Zhou, William Yang Wang</i>	
IMPLICIT BIAS OF GRADIENT DESCENT BASED ADVERSARIAL TRAINING ON SEPARABLE DATA	9436
<i>Yan Li, Ethan X. Fang, Huan Xu, Tuo Zhao</i>	
IMAGE-GUIDED NEURAL OBJECT RENDERING.....	9461
<i>Justus Thies, Michael Zollhöfer, Christian Theobalt, Marc Stamminger, Matthias Nießner</i>	
KNOWLEDGE CONSISTENCY BETWEEN NEURAL NETWORKS AND BEYOND	9477
<i>Ruofan Liang, Tianlin Li, Longfei Li, Jing Wang, Quanshi Zhang</i>	
LAZY-CFR: FAST AND NEAR-OPTIMAL REGRET MINIMIZATION FOR EXTENSIVE GAMES WITH IMPERFECT INFORMATION.....	9492
<i>Yichi Zhou, Tongzheng Ren, Jialian Li, Dong Yan, Jun Zhu</i>	
ADDITIVE POWERS-OF-TWO QUANTIZATION: AN EFFICIENT NON-UNIFORM DISCRETIZATION FOR NEURAL NETWORKS.....	9512
<i>Yuhang Li, Xin Dong, Wei Wang</i>	

VOLUME 13

ENHANCING TRANSFORMATION-BASED DEFENSES AGAINST ADVERSARIAL ATTACKS WITH A DISTRIBUTION CLASSIFIER.....	9527
<i>Connie Kou, Hwee Kuan Lee, Ee-Chien Chang, Teck Khim Ng</i>	
OBSERVATIONAL OVERFITTING IN REINFORCEMENT LEARNING	9546
<i>Xingyou Song, Yiding Jiang, Stephen Tu, Yilun Du, Behnam Neyshabur</i>	

ON MUTUAL INFORMATION MAXIMIZATION FOR REPRESENTATION LEARNING	9575
<i>Michael Tschannen, Josip Djolonga, Paul K. Rubenstein, Sylvain Gelly, Mario Lucic</i>	
RANKING POLICY GRADIENT	9591
<i>Kaixiang Lin, Jiayu Zhou</i>	
SVQN: SEQUENTIAL VARIATIONAL SOFT Q-LEARNING NETWORKS	9617
<i>Shiyu Huang, Hang Su, Jun Zhu, Ting Chen</i>	
UNDERSTANDING ARCHITECTURES LEARNT BY CELL-BASED NEURAL ARCHITECTURE SEARCH	9634
<i>Yao Shu, Wei Wang, Shaofeng Cai</i>	
AUTOQ: AUTOMATED KERNEL-WISE NEURAL NETWORK QUANTIZATION.....	9655
<i>Qian Lou, Feng Guo, Minje Kim, Lantao Liu, Lei Jiang.</i>	
POLY-ENCODERS: ARCHITECTURES AND PRE-TRAINING STRATEGIES FOR FAST AND ACCURATE MULTI-SENTENCE SCORING.....	9666
<i>Samuel Humeau, Kurt Shuster, Marie-Anne Lachaux, Jason Weston</i>	
A LEARNING-BASED ITERATIVE METHOD FOR SOLVING VEHICLE ROUTING PROBLEMS.....	9680
<i>Hao Lu, Xingwen Zhang, Shuang Yang</i>	
TRANSFERABLE PERTURBATIONS OF DEEP FEATURE DISTRIBUTIONS	9695
<i>Nathan Inkawhich, Kevin Liang, Lawrence Carin, Yiran Chen</i>	
CONTINUAL LEARNING WITH ADAPTIVE WEIGHTS (CLAW).....	9709
<i>Tameem Adel, Han Zhao, Richard E. Turner</i>	
SCALABLE AND ORDER-ROBUST CONTINUAL LEARNING WITH ADDITIVE PARAMETER DECOMPOSITION.....	9728
<i>Jaehong Yoon, Saehoon Kim, Eunho Yang, Sung Ju Hwang</i>	
WEAKLY SUPERVISED CLUSTERING BY EXPLOITING UNIQUE CLASS COUNT.....	9743
<i>Mustafa Umit Oner, Hwee Kuan Lee, Wing-Kin Sung</i>	
LINEAR SYMMETRIC QUANTIZATION OF NEURAL NETWORKS FOR LOW-PRECISION INTEGER HARDWARE	9768
<i>Xiandong Zhao, Ying Wang, Xuyi Cai, Cheng Liu, Lei Zhang</i>	
TO RELIEVE YOUR HEADACHE OF TRAINING AN MRF, TAKE ADVIL	9784
<i>Chongxuan Li, Chao Du, Kun Xu, Max Welling, Jun Zhu, Bo Zhang</i>	
AUTOMATED RELATIONAL META-LEARNING	9804
<i>Huaxiu Yao, Xian Wu, Zhiqiang Tao, Yaliang Li, Bolin Ding, Ruirui Li, Zhenhui Li</i>	
N-BEATS: NEURAL BASIS EXPANSION ANALYSIS FOR INTERPRETABLE TIME SERIES FORECASTING	9823
<i>Boris N. Oreshkin, Dmitri Carпов, Nicolas Chapados, Yoshua Bengio</i>	
DEEP LEARNING OF DETERMINANTAL POINT PROCESSES VIA PROPER SPECTRAL SUB-GRADIENT	9854
<i>Tianshu Yu, Yikang Li, Baoxin Li</i>	
DISTANCE-BASED LEARNING FROM ERRORS FOR CONFIDENCE CALIBRATION	9866
<i>Chen Xing, Sercan Arik, Zizhao Zhang, Tomas Pfister</i>	
CURVATURE GRAPH NETWORK.....	9878
<i>Ze Ye, Kin Sum Liu, Tengfei Ma, Jie Gao, Chao Chen</i>	
LEARNING EXPENSIVE COORDINATION: AN EVENT-BASED DEEP RL APPROACH.....	9893
<i>Zhenyu Shi*, Runsheng Yu*, Xinrun Wang*, Rundong Wang, Youzhi Zhang, Hanjiang Lai, Bo An</i>	
LAMOL: LANGUAGE MODELING FOR LIFELONG LANGUAGE LEARNING	9919
<i>Fan-Keng Sun*, Cheng-Hao Ho*, Hung-Yi Lee</i>	
PROXSGD: TRAINING STRUCTURED NEURAL NETWORKS UNDER REGULARIZATION AND CONSTRAINTS	9934
<i>Yang Yang, Yaxiong Yuan, Avraam Chatzimichailidis, Ruud Jg Van Sloun, Lei Lei, Symeon Chatzinotas</i>	

DIVERSE TRAJECTORY FORECASTING WITH DETERMINANTAL POINT PROCESSES.....	9949
<i>Ye Yuan, Kris M. Kitani</i>	
EVALUATING THE SEARCH PHASE OF NEURAL ARCHITECTURE SEARCH	9964
<i>Kaicheng Yu, Christian Sciuto, Martin Jaggi, Claudiu Musat, Mathieu Salzmann</i>	
DEEPOYER: LEARNING SPARSER NEURAL NETWORK WITH DIFFERENTIABLE SCALE-INVARIANT SPARSITY MEASURES	9980
<i>Huanrui Yang, Wei Wen, Hai Li</i>	
DEPTH-ADAPTIVE TRANSFORMER.....	9998
<i>Maha Elbayad, Jiatao Gu, Edouard Grave, Michael Auli</i>	
FEDERATED ADVERSARIAL DOMAIN ADAPTATION.....	10013
<i>Xingchao Peng, Zijun Huang, Yizhe Zhu, Kate Saenko</i>	
MAXMIN Q-LEARNING: CONTROLLING THE ESTIMATION BIAS OF Q-LEARNING	10032
<i>Qingfeng Lan, Yangchen Pan, Alona Fyshe, Martha White</i>	
AUTOMATICALLY DISCOVERING AND LEARNING NEW VISUAL CATEGORIES WITH RANKING STATISTICS	10052
<i>Kai Han, Sylvestre-Alvise Rebuffi, Sebastien Ehrhardt, Andrea Vedaldi, Andrew Zisserman</i>	
MUTUAL MEAN-TEACHING: PSEUDO LABEL REFINERY FOR UNSUPERVISED DOMAIN ADAPTATION ON PERSON RE-IDENTIFICATION	10065
<i>Yixiao Ge, Dapeng Chen, Hongsheng Li</i>	
NEURAL OBLIVIOUS DECISION ENSEMBLES FOR DEEP LEARNING ON TABULAR DATA.....	10080
<i>Sergei Popov, Stanislav Morozov, Artem Babenko</i>	
SQIL: IMITATION LEARNING VIA REINFORCEMENT LEARNING WITH SPARSE REWARDS.....	10092
<i>Siddharth Reddy, Anca D. Dragan, Sergey Levine</i>	
GRAPH INFERENCE LEARNING FOR SEMI-SUPERVISED CLASSIFICATION	10106
<i>Chunyan Xu, Zhen Cui, Xiaobin Hong, Tong Zhang, Jian Yang, Wei Liu</i>	
LEARNING DEEP GRAPH MATCHING WITH CHANNEL-INDEPENDENT EMBEDDING AND HUNGARIAN ATTENTION	10117
<i>Tianshu Yu, Runzhong Wang, Junchi Yan, Baoxin Li</i>	
STRUCTPOOL: STRUCTURED GRAPH POOLING VIA CONDITIONAL RANDOM FIELDS	10132
<i>Hao Yuan, Shuiwang Ji</i>	
ON THE WEAKNESSES OF REINFORCEMENT LEARNING FOR NEURAL MACHINE TRANSLATION	10144
<i>Leshem Choshen, Lior Fox, Zohar Aizenbud, Omri Abend</i>	
SHARING KNOWLEDGE IN MULTI-TASK DEEP REINFORCEMENT LEARNING.....	10158
<i>Carlo D'Eramo, Davide Tateo, Andrea Bonarini, Marcello Restelli, Jan Peters</i>	
REINFORCEMENT LEARNING BASED GRAPH-TO-SEQUENCE MODEL FOR NATURAL QUESTION GENERATION.....	10176
<i>Yu Chen, Lingfei Wu, Mohammed J. Zaki</i>	
SELF: LEARNING TO FILTER NOISY LABELS WITH SELF-ENSEMBLING.....	10193
<i>Duc Tam Nguyen, Chaithanya Kumar Mummadi, Thi Phuong Nhung Ngo, Thi Hoai Phuong Nguyen, Laura Beggel, Thomas Brox</i>	
LARGE BATCH OPTIMIZATION FOR DEEP LEARNING: TRAINING BERT IN 76 MINUTES	10209
<i>Yang You, Jing Li, Sashank Reddi, Jonathan Hseu, Sanjiv Kumar, Srinadh Bhojanapalli, Xiaodan Song, James Demmel, Kurt Keutzer, Cho-Jui Hsieh</i>	

SPOTLIGHT PRESENTATIONS

GEOM-GCN: GEOMETRIC GRAPH CONVOLUTIONAL NETWORKS.....	10247
<i>Hongbin Pei, Bingzhe Wei, Kevin Chen-Chuan Chang, Yu Lei, Bo Yang</i>	
AN EXPONENTIAL LEARNING RATE SCHEDULE FOR DEEP LEARNING	10259
<i>Zhiyuan Li, Sanjeev Arora</i>	

PROGRESSIVE LEARNING AND DISENTANGLEMENT OF HIERARCHICAL REPRESENTATIONS.....	10288
<i>Zhiyuan Li, Jaideep Vithal Murkute, Prashnna Kumar Gyawali, Linwei Wang</i>	
RECONSTRUCTING CONTINUOUS DISTRIBUTIONS OF 3D PROTEIN STRUCTURE FROM CRYO-EM IMAGES	10302
<i>Ellen D. Zhong, Tristan Bepler, Joseph H. Davis, Bonnie Berger</i>	
GENERALIZATION OF TWO-LAYER NEURAL NETWORKS: AN ASYMPTOTIC VIEWPOINT.....	10322
<i>Jimmy Ba, Murat Erdogdu, Taiji Suzuki, Denny Wu, Tianzong Zhang</i>	

VOLUME 14

DEPTH-WIDTH TRADE-OFFS FOR RELU NETWORKS VIA SHARPOVSKY'S THEOREM.....	10362
<i>Vaggos Chatziafratis, Sai Ganesh Nagarajan, Ioannis Panageas, Xiao Wang</i>	
ENERGY-BASED MODELS FOR ATOMIC-RESOLUTION PROTEIN CONFORMATIONS	10381
<i>Yilun Du, Joshua Meier, Jerry Ma, Rob Fergus, Alexander Rives</i>	
A MUTUAL INFORMATION MAXIMIZATION PERSPECTIVE OF LANGUAGE REPRESENTATION LEARNING.....	10397
<i>Lingpeng Kong, Cyprien De Masson D'Autume, Lei Yu, Wang Ling, Zihang Dai, Dani Yogatama</i>	
UNBIASED CONTRASTIVE DIVERGENCE ALGORITHM FOR TRAINING ENERGY-BASED LATENT VARIABLE MODELS	10409
<i>Yixuan Qiu, Lingsong Zhang, Xiao Wang</i>	
MAKING SENSE OF REINFORCEMENT LEARNING AND PROBABILISTIC INFERENCE	10426
<i>Brendan O'Donoghue, Ian Osband, Catalin Ionescu</i>	
DEEP LEARNING FOR SYMBOLIC MATHEMATICS	10442
<i>Guillaume Lample, François Charton</i>	
SUMO: UNBIASED ESTIMATION OF LOG MARGINAL PROBABILITY FOR LATENT VARIABLE MODELS.....	10466
<i>Yucen Luo, Alex Beatson, Mohammad Norouzi, Jun Zhu, David Duvenaud, Ryan P. Adams, Ricky T. Q. Chen</i>	
DEEPSPHERE: A GRAPH-BASED SPHERICAL CNN.....	10486
<i>Michaël Defferrard, Martino Milani, Frédéric Gusset, Nathanaël Perraudin</i>	
NEURAL ARITHMETIC UNITS.....	10504
<i>Andreas Madsen, Alexander Rosenberg Johansen</i>	
TRUTH OR BACKPROPAGANDA? AN EMPIRICAL INVESTIGATION OF DEEP LEARNING THEORY	10535
<i>Micah Goldblum, Jonas Geiping, Avi Schwarzschild, Michael Moeller, Tom Goldstein</i>	
DRAWING EARLY-BIRD TICKETS: TOWARD MORE EFFICIENT TRAINING OF DEEP NETWORKS.....	10553
<i>Haoran You, Chaojian Li, Pengfei Xu, Yonggan Fu, Yue Wang, Xiaohan Chen, Richard G. Baraniuk, Zhangyang Wang, Yingyan Lin</i>	
IMPROVING GENERALIZATION IN META REINFORCEMENT LEARNING USING LEARNED OBJECTIVES.....	10566
<i>Louis Kirsch, Sjoerd Van Steenkiste, Juergen Schmidhuber</i>	
KALEIDOSCOPE: AN EFFICIENT, LEARNABLE REPRESENTATION FOR ALL STRUCTURED LINEAR MAPS	10587
<i>Tri Dao, Nimit Sohoni, Albert Gu, Matthew Eichhorn, Amit Blonder, Megan Leszczynski, Atri Rudra, Christopher Ré</i>	
DISENTANGLEMENT BY NONLINEAR ICA WITH GENERAL INCOMPRESSIBLE-FLOW NETWORKS (GIN).....	10632
<i>Peter Sorrenson, Carsten Rother, Ullrich Köthe</i>	

CONTINUAL LEARNING WITH HYPERNETWORKS	10655
<i>Johannes Von Oswald, Christian Henning, Benjamin F. Grewe, João Sacramento</i>	
AT STABILITY'S EDGE: HOW TO ADJUST HYPERPARAMETERS TO PRESERVE MINIMA SELECTION IN ASYNCHRONOUS TRAINING OF NEURAL NETWORKS?.....	10683
<i>Niv Giladi, Mor Shpigel Nacson, Elad Hoffer, Daniel Soudry</i>	
COMPRESSION BASED BOUND FOR NON-COMPRESSED NETWORK: UNIFIED GENERALIZATION ERROR ANALYSIS OF LARGE COMPRESSIBLE DEEP NEURAL NETWORK	10704
<i>Taiji Suzuki, Hiroshi Abe, Tomoaki Nishimura</i>	
EXPLANATION BY PROGRESSIVE EXAGGERATION	10738
<i>Sumedha Singla, Brian Pollack, Junxiang Chen, Kayhan Batmanghelich</i>	
DIRECTIONAL MESSAGE PASSING FOR MOLECULAR GRAPHS.....	10758
<i>Johannes Gasteiger, Janek Groß, Stephan Günnemann</i>	
LEARNING FROM RULES GENERALIZING LABELED EXEMPLARS	10771
<i>Abhijeet Awasthi, Sabyasachi Ghosh, Rasna Goyal, Sunita Sarawagi</i>	
TRAINING INDIVIDUALLY FAIR ML MODELS WITH SENSITIVE SUBSPACE ROBUSTNESS	10788
<i>Mikhail Yurochkin, Amanda Bower, Yuekai Sun</i>	
WHAT CAN NEURAL NETWORKS REASON ABOUT?	10806
<i>Keyulu Xu, Jingling Li, Mozhi Zhang, Simon S. Du, Ken-Ichi Kawarabayashi, Stefanie Jegelka</i>	
WORD2KET: SPACE-EFFICIENT WORD EMBEDDINGS INSPIRED BY QUANTUM ENTANGLEMENT	10824
<i>Aliakbar Panahi, Seyran Saeedi, Tom Arodz</i>	
TOWARDS HIERARCHICAL IMPORTANCE ATTRIBUTION: EXPLAINING COMPOSITIONAL SEMANTICS FOR NEURAL SEQUENCE MODELS	10834
<i>Xisen Jin, Zhongyu Wei, Junyi Du, Xiangyang Xue, Xiang Ren</i>	
SPECTRAL EMBEDDING OF REGULARIZED BLOCK MODELS.....	10849
<i>Nathan De Lara, Thomas Bonald</i>	
MAXIMUM LIKELIHOOD CONSTRAINT INFERENCE FOR INVERSE REINFORCEMENT LEARNING.....	10862
<i>Dexter R. R. Scobee, S. Shankar Sastry</i>	
META-LEARNING ACQUISITION FUNCTIONS FOR TRANSFER LEARNING IN BAYESIAN OPTIMIZATION	10874
<i>Michael Volpp, Lukas P. Fröhlich, Kirsten Fischer, Andreas Doerr, Stefan Falkner, Frank Hutter, Christian Daniel</i>	
THE INGREDIENTS OF REAL WORLD ROBOTIC REINFORCEMENT LEARNING	10896
<i>Henry Zhu, Justin Yu, Abhishek Gupta, Dhruv Shah, Kristian Hartikainen, Avi Singh, Vikash Kumar, Sergey Levine</i>	
SCALING AUTOREGRESSIVE VIDEO MODELS	10917
<i>Dirk Weissenborn, Oscar Täckström, Jakob Uszkoreit</i>	
DIFFERENTIATION OF BLACKBOX COMBINATORIAL SOLVERS	10941
<i>Marin Vlastelica Pogancic, Anselm Paulus, Vit Musil, Georg Martius, Michal Rolinek</i>	
HARNESSING THE POWER OF INFINITELY WIDE DEEP NETS ON SMALL-DATA TASKS	10960
<i>Sanjeev Arora, Simon S. Du, Zhiyuan Li, Ruslan Salakhutdinov, Ruosong Wang, Dingli Yu</i>	
THE INTRIGUING ROLE OF MODULE CRITICALITY IN THE GENERALIZATION OF DEEP NETWORKS	10973
<i>Niladri Chatterji, Behnam Neyshabur, Hanie Sedghi</i>	
SELF-LABELLING VIA SIMULTANEOUS CLUSTERING AND REPRESENTATION LEARNING.....	10992
<i>Asano Ym., Rupprecht C., Vedaldi A.</i>	

NEURAL TANGENTS: FAST AND EASY INFINITE NEURAL NETWORKS IN PYTHON	11014
<i>Roman Novak, Lechao Xiao, Jiri Hron, Jaehoon Lee, Alexander A. Alemi, Jascha Sohl-Dickstein, Samuel S. Schoenholz</i>	
SEQUENTIAL LATENT KNOWLEDGE SELECTION FOR KNOWLEDGE-GROUNDED DIALOGUE	11033
<i>Byeongchang Kim, Jaewoo Ahn, Gunhee Kim</i>	
MEASURING THE RELIABILITY OF REINFORCEMENT LEARNING ALGORITHMS.....	11047
<i>Stephanie C. Y. Chan, Samuel Fishman, Anoop Korattikara, John Canny, Sergio Guadarrama</i>	
STABLE RANK NORMALIZATION FOR IMPROVED GENERALIZATION IN NEURAL NETWORKS AND GANS.....	11083
<i>Amartya Sanyal, Philip H. Torr, Puneet K. Dokania</i>	
DISAGREEMENT-REGULARIZED IMITATION LEARNING.....	11113
<i>Kiante Brantley, Wen Sun, Mikael Henaff</i>	

VOLUME 15

MODEL BASED REINFORCEMENT LEARNING FOR ATARI.....	11132
<i>Lukasz Kaiser, Mohammad Babaeizadeh, Piotr Milos, Blazej Osinski, Roy H Campbell, Konrad Czechowski, Dumitru Erhan, Chelsea Finn, Piotr Kozakowski, Sergey Levine, Afroz Mohiuddin, Ryan Sepassi, George Tucker, Henryk Michalewski</i>	
UNDERSTANDING WHY NEURAL NETWORKS GENERALIZE WELL THROUGH GSNR OF PARAMETERS.....	11160
<i>Jinlong Liu, Yunzhi Bai, Guoqing Jiang, Ting Chen, Huayan Wang</i>	
A LATENT MORPHOLOGY MODEL FOR OPEN-VOCABULARY NEURAL MACHINE TRANSLATION	11175
<i>Duygu Ataman, Wilker Aziz, Alexandra Birch</i>	
AND THE BIT GOES DOWN: REVISITING THE QUANTIZATION OF NEURAL NETWORKS.....	11190
<i>Pierre Stock, Armand Joulin, Rémi Gribonval, Benjamin Graham, Hervé Jégou</i>	
KERNELIZED WASSERSTEIN NATURAL GRADIENT.....	11201
<i>M Arbel, A Gretton, W Li, G Montufar</i>	
FREELB: ENHANCED ADVERSARIAL TRAINING FOR NATURAL LANGUAGE UNDERSTANDING	11232
<i>Chen Zhu, Yu Cheng, Zhe Gan, Siqi Sun, Tom Goldstein, Jingjing Liu</i>	
BEHAVIOUR SUITE FOR REINFORCEMENT LEARNING	11246
<i>Ian Osband, Yotam Doron, Matteo Hessel, John Aslanides, Eren Sezener, Andre Saraiva, Katrina McKinney, Tor Lattimore, Csaba Szepesvari, Satinder Singh, Benjamin Van Roy, Richard Sutton, David Silver, Hado Van Hasselt</i>	
STRATEGIES FOR PRE-TRAINING GRAPH NEURAL NETWORKS.....	11265
<i>Weihua Hu*, Bowen Liu*, Joseph Gomes, Marinka Zitnik, Percy Liang, Vijay Pande, Jure Leskovec</i>	
NAS-BENCH-201: EXTENDING THE SCOPE OF REPRODUCIBLE NEURAL ARCHITECTURE SEARCH.....	11287
<i>Xuanyi Dong, Yi Yang</i>	
EMERGENT TOOL USE FROM MULTI-AGENT AUTOCURRICULA.....	11303
<i>Bowen Baker, Ingmar Kanitscheider, Todor Markov, Yi Wu, Glenn Powell, Bob McGrew, Igor Mordatch</i>	
A PROBABILISTIC FORMULATION OF UNSUPERVISED TEXT STYLE TRANSFER	11331
<i>Junxian He, Xinyi Wang, Graham Neubig, Taylor Berg-Kirkpatrick</i>	
DREAM TO CONTROL: LEARNING BEHAVIORS BY LATENT IMAGINATION	11345
<i>Danijar Hafner, Timothy Lillicrap, Jimmy Ba, Mohammad Norouzi</i>	
REAL OR NOT REAL, THAT IS THE QUESTION.....	11365
<i>Yuanbo Xiangli*, Yubin Deng*, Bo Dai*, Chen Change Loy, Dahua Lin</i>	

NEURAL SYMBOLIC READER: SCALABLE INTEGRATION OF DISTRIBUTED AND SYMBOLIC REPRESENTATIONS FOR READING COMPREHENSION	11377
<i>Xinyun Chen, Chen Liang, Adams Wei Yu, Denny Zhou, Dawn Song, Quoc V. Le</i>	
NETWORK DECONVOLUTION.....	11393
<i>Chengxi Ye, Matthew Evanusa, Hua He, Anton Mitrokhin, Tom Goldstein, James A. Yorke, Cornelia Fermuller, Yiannis Aloimonos</i>	
SIMPLIFIED ACTION DECODER FOR DEEP MULTI-AGENT REINFORCEMENT LEARNING.....	11413
<i>Hengyuan Hu, Jakob N Foerster</i>	
IS A GOOD REPRESENTATION SUFFICIENT FOR SAMPLE EFFICIENT REINFORCEMENT LEARNING?.....	11427
<i>Simon S. Du, Sham M. Kakade, Ruosong Wang, Lin F. Yang</i>	
LEARNING THE DIFFERENCE THAT MAKES A DIFFERENCE WITH COUNTERFACTUALLY-AUGMENTED DATA	11447
<i>Divyansh Kaushik, Eduard Hovy, Zachary Lipton</i>	
ASYMPTOTICS OF WIDE NETWORKS FROM FEYNMAN DIAGRAMS.....	11464
<i>Ethan Dyer, Guy Gur-Ari</i>	
SYMPLECTIC RECURRENT NEURAL NETWORKS.....	11499
<i>Zhengdao Chen, Jianyu Zhang, Martin Arjovsky, Léon Bottou</i>	
LEARNING TO PLAN IN HIGH DIMENSIONS VIA NEURAL EXPLORATION-EXPLOITATION TREES.....	11522
<i>Binghong Chen, Bo Dai, Qinjie Lin, Guo Ye, Han Liu, Le Song</i>	
DISENTANGLING NEURAL MECHANISMS FOR PERCEPTUAL GROUPING	11548
<i>Junhyung Kim*, Drew Linsley*, Kalpit Thakkar, Thomas Serre</i>	
ALBERT: A LITE BERT FOR SELF-SUPERVISED LEARNING OF LANGUAGE REPRESENTATIONS.....	11574
<i>Zhenzhong Lan, Mingda Chen, Sebastian Goodman, Kevin Gimpel, Piyush Sharma, Radu Soricut</i>	
THE BREAK-EVEN POINT ON OPTIMIZATION TRAJECTORIES OF DEEP NEURAL NETWORKS.....	11591
<i>Stanislaw Jastrzebski, Maciej Szymczak, Stanislav Fort, Devansh Arpit, Jacek Tabor, Kyunghyun Cho*, Krzysztof Geras*</i>	
THE LOGICAL EXPRESSIVENESS OF GRAPH NEURAL NETWORKS	11612
<i>Pablo Barceló, Egor V. Kostylev, Mikael Monet, Jorge Pérez, Juan Reutter, Juan Pablo Silva</i>	
CLEVRER: COLLISION EVENTS FOR VIDEO REPRESENTATION AND REASONING	11633
<i>Kexin Yi*, Chuang Gan*, Yunzhu Li, Pushmeet Kohli, Jiajun Wu, Antonio Torralba, Joshua B. Tenenbaum</i>	
LEARNING COMPOSITIONAL KOOPMAN OPERATORS FOR MODEL-BASED CONTROL.....	11652
<i>Yunzhu Li, Hao He, Jiajun Wu, Dina Katabi, Antonio Torralba</i>	
DOUBLY ROBUST BIAS REDUCTION IN INFINITE HORIZON OFF-POLICY ESTIMATION.....	11666
<i>Ziyang Tang*, Yihao Feng*, Lihong Li, Dengyong Zhou, Qiang Liu</i>	
DEEP NEUROETHOLOGY OF A VIRTUAL RODENT	11686
<i>Josh Merel, Diego Aldarondo, Jesse Marshall, Yival Tassa, Greg Wayne, Bence Olveczky</i>	
EMERGENCE OF FUNCTIONAL AND STRUCTURAL PROPERTIES OF THE HEAD DIRECTION SYSTEM BY OPTIMIZATION OF RECURRENT NEURAL NETWORKS	11706
<i>Christopher J. Cueva, Peter Y. Wang, Matthew Chin, Xue-Xin Wei</i>	
DURATION-OF-STAY STORAGE ASSIGNMENT UNDER UNCERTAINTY	11723
<i>Michael Lingzhi Li, Elliott Wolf, Daniel Wintz</i>	
INDUCTIVE MATRIX COMPLETION BASED ON GRAPH NEURAL NETWORKS	11738
<i>Muhan Zhang, Yixin Chen</i>	
CONDITIONAL LEARNING OF FAIR REPRESENTATIONS	11752
<i>Han Zhao, Amanda Coston, Tameem Adel, Geoffrey J. Gordon</i>	

GRADIENTLESS DESCENT: HIGH-DIMENSIONAL ZERO-ORDER OPTIMIZATION	11769
<i>Daniel Golovin, John Karro, Greg Kochanski, Chansoo Lee, Xingyou Song, Qiuyi Zhang</i>	
ESTIMATING COUNTERFACTUAL TREATMENT OUTCOMES OVER TIME THROUGH ADVERSARIAL BALANCED REPRESENTATIONS	11790
<i>Ioana Bica, Ahmed M Alaa, James Jordon, Mihaela Van Der Schaar</i>	
COPHY: COUNTERFACTUAL LEARNING OF PHYSICAL DYNAMICS	11818
<i>Fabien Baradel, Natalia Neverova, Julien Mille, Greg Mori, Christian Wolf</i>	
HAMILTONIAN GENERATIVE NETWORKS	11832
<i>Peter Toth, Danilo J. Rezende, Andrew Jaegle, Sébastien Racanière, Aleksandar Botev, Irina Higgins</i>	
HOW MUCH POSITION INFORMATION DO CONVOLUTIONAL NEURAL NETWORKS ENCODE?	11851
<i>Md Amirul Islam*, Sen Jia*, Neil D. B. Bruce</i>	
SLICED CRAMER SYNAPTIC CONSOLIDATION FOR PRESERVING DEEPLY LEARNED REPRESENTATIONS	11862
<i>Soheil Kolouri, Nicholas A. Ketz, Andrea Soltoggio, Praveen K. Pilly</i>	
HOPPITY: LEARNING GRAPH TRANSFORMATIONS TO DETECT AND FIX BUGS IN PROGRAMS	11879
<i>Elizabeth Dinella, Hanjun Dai, Ziyang Li, Mayur Naik, Le Song, Ke Wang</i>	
INFLUENCE-BASED MULTI-AGENT EXPLORATION	11896
<i>Tonghan Wang*, Jianhao Wang*, Yi Wu, Chongjie Zhang</i>	
META-LEARNING WITHOUT MEMORIZATION	11919
<i>Mingzhang Yin, George Tucker, Mingyuan Zhou, Sergey Levine, Chelsea Finn</i>	

VOLUME 16

FINITE DEPTH AND WIDTH CORRECTIONS TO THE NEURAL TANGENT KERNEL	11940
<i>Boris Hanin, Mihai Nica</i>	
RIDGE REGRESSION: STRUCTURE, CROSS-VALIDATION, AND SKETCHING	11965
<i>Sifan Liu, Edgar Dobriban</i>	
CROSS-DOMAIN FEW-SHOT CLASSIFICATION VIA LEARNED FEATURE-WISE TRANSFORMATION	11992
<i>Hung-Yu Tseng, Hsin-Ying Lee, Jia-Bin Huang, Ming-Hsuan Yang</i>	
DDSP: DIFFERENTIABLE DIGITAL SIGNAL PROCESSING	12010
<i>Jesse Engel, Lamtharn (Hanoi) Hantrakul, Chenjie Gu, Adam Roberts</i>	
ENCODING WORD ORDER IN COMPLEX EMBEDDINGS	12029
<i>Benyou Wang, Donghao Zhao, Christina Lioma, Qiuchi Li, Peng Zhang, Jakob Grue Simonsen</i>	
ENHANCING ADVERSARIAL DEFENSE BY K-WINNERS-TAKE-ALL	12044
<i>Chang Xiao, Peilin Zhong, Changxi Zheng</i>	
ONLINE AND STOCHASTIC OPTIMIZATION BEYOND LIPSCHITZ CONTINUITY: A RIEMANNIAN APPROACH	12074
<i>Kimon Antonakopoulos, E. Veronica Belmega, Panayotis Mertikopoulos</i>	
PC-DARTS: PARTIAL CHANNEL CONNECTIONS FOR MEMORY-EFFICIENT ARCHITECTURE SEARCH	12093
<i>Yuhui Xu, Lingxi Xie, Xiaopeng Zhang, Xin Chen, Guo-Jun Qi, Qi Tian, Hongkai Xiong</i>	
TRANQUIL CLOUDS: NEURAL NETWORKS FOR LEARNING TEMPORALLY COHERENT FEATURES IN POINT CLOUDS	12106
<i>Lukas Prantl, Nuttapon Chentanez, Stefan Jeschke, Nils Thuerey</i>	
NEURAL MACHINE TRANSLATION WITH UNIVERSAL VISUAL REPRESENTATION	12122
<i>Zhuosheng Zhang, Kehai Chen, Rui Wang, Masao Utiyama, Eiichiro Sumita, Zuchao Li, Hai Zhao</i>	

WHITE NOISE ANALYSIS OF NEURAL NETWORKS	12136
<i>Ali Borji, Sikun Lin</i>	
SKIP CONNECTIONS MATTER: ON THE TRANSFERABILITY OF ADVERSARIAL EXAMPLES GENERATED WITH RESNETS	12164
<i>Dongxian Wu, Yisen Wang, Shu-Tao Xia, James Bailey, Xingjun Ma</i>	
A SIGNAL PROPAGATION PERSPECTIVE FOR PRUNING NEURAL NETWORKS AT INITIALIZATION.....	12179
<i>Namhoon Lee, Thalaiyasingam Ajanthan, Stephen Gould, Philip H. S. Torr</i>	
INTENSITY-FREE LEARNING OF TEMPORAL POINT PROCESSES.....	12195
<i>Oleksandr Shchur, Marin Bilos, Stephan Günnemann</i>	
LEARNING TO CONTROL PDES WITH DIFFERENTIABLE PHYSICS	12216
<i>Philipp Holl, Nils Thuerey, Vladlen Koltun</i>	
ESTIMATING GRADIENTS FOR DISCRETE RANDOM VARIABLES BY SAMPLING WITHOUT REPLACEMENT	12244
<i>Wouter Kool, Herke Van Hoof, Max Welling</i>	
DEFENDING AGAINST PHYSICALLY REALIZABLE ATTACKS ON IMAGE CLASSIFICATION	12272
<i>Tong Wu, Liang Tong, Yevgeniy Vorobeychik</i>	
ON ROBUSTNESS OF NEURAL ORDINARY DIFFERENTIAL EQUATIONS	12298
<i>Hanshu Yan, Jiawei Du, Vincent Tan, Jiashi Feng</i>	
INFOGRAPH: UNSUPERVISED AND SEMI-SUPERVISED GRAPH-LEVEL REPRESENTATION LEARNING VIA MUTUAL INFORMATION MAXIMIZATION	12313
<i>Fan-Yun Sun, Jordan Hoffman, Vikas Verma, Jian Tang</i>	
MULTI-SCALE REPRESENTATION LEARNING FOR SPATIAL FEATURE DISTRIBUTIONS USING GRID CELLS.....	12329
<i>Gengchen Mai, Krzysztof Janowicz, Bo Yan, Rui Zhu, Ling Cai, Ni Lao</i>	
GRAPH NEURAL NETWORKS EXPONENTIALLY LOSE EXPRESSIVE POWER FOR NODE CLASSIFICATION	12344
<i>Kenta Oono, Taiji Suzuki</i>	
SPARSE CODING WITH GATED LEARNED ISTA.....	12381
<i>Kailun Wu, Yiwon Guo, Ziang Li, Changshui Zhang</i>	
PROGRAM GUIDED AGENT.....	12408
<i>Shao-Hua Sun, Te-Lin Wu, Joseph J. Lim</i>	
 <u>ORAL PRESENTATIONS</u>	
FEDERATED LEARNING WITH MATCHED AVERAGING.....	12432
<i>Hongyi Wang, Mikhail Yurochkin, Yuekai Sun, Dimitris Papailiopoulos, Yasaman Khazaeni</i>	
DIFFERENTIABLE REASONING OVER A VIRTUAL KNOWLEDGE BASE	12448
<i>Bhuwan Dhingra, Manzil Zaheer, Vidhisha Balachandran, Graham Neubig, Ruslan Salakhutdinov, William W. Cohen</i>	
ADVERSARIAL TRAINING AND PROVABLE DEFENSES: BRIDGING THE GAP	12464
<i>Mislav Balunovic, Martin Vechev</i>	
GRADIENT DESCENT MAXIMIZES THE MARGIN OF HOMOGENEOUS NEURAL NETWORKS.....	12482
<i>Kaifeng Lyu, Jian Li</i>	
CONVOLUTIONAL CONDITIONAL NEURAL PROCESSES.....	12534
<i>Jonathan Gordon, Wessel P. Bruinsma, Andrew Y. K. Foong, James Requeima, Yann Dubois, Richard E. Turner</i>	
META-LEARNING WITH WARPED GRADIENT DESCENT	12566
<i>Sebastian Flennerhag, Andrei A. Rusu, Razvan Pascanu, Francesco Visin, Hujun Yin, Raia Hadsell</i>	

SEED RL: SCALABLE AND EFFICIENT DEEP-RL WITH ACCELERATED CENTRAL INFERENCE.....	12594
<i>Lasse Espeholt, Raphaël Marinier, Piotr Stanczyk, Ke Wang, Marcin Michalski?</i>	
HIGH FIDELITY SPEECH SYNTHESIS WITH ADVERSARIAL NETWORKS	12614
<i>Mikolaj Binkowski, Jeff Donahue, Sander Dieleman, Aidan Clark, Erich Elsen, Norman Casagrande, Luis C. Cobo, Karen Simonyan</i>	
A GENERALIZED TRAINING APPROACH FOR MULTIAGENT LEARNING	12631
<i>Paul Muller, Shayegan Omidshafiei, Mark Rowland, Karl Tuyls, Julien Perolat, Siqi Liu, Daniel Hennes, Luke Marris, Marc Lanctot, Edward Hughes, Zhe Wang, Guy Lever, Nicolas Heess, Thore Graepel, Remi Munos</i>	
BUILDING DEEP EQUIVARIANT CAPSULE NETWORKS	12666
<i>Sai Raam Venkataraman, S. Balasubramanian, R. Raghunatha Sarma</i>	
RESTRICTING THE FLOW: INFORMATION BOTTLENECKS FOR ATTRIBUTION.....	12685
<i>Karl Schulz, Leon Sixt, Federico Tombari, Tim Landgraf</i>	
INTRINSICALLY MOTIVATED DISCOVERY OF DIVERSE PATTERNS IN SELF-ORGANIZING SYSTEMS	12703
<i>Chris Reinke, Mayalen Etcheverry, Pierre-Yves Oudeyer</i>	
CAUSAL DISCOVERY WITH REINFORCEMENT LEARNING	12732
<i>Shengyu Zhu, Ignavier Ng, Zhitang Chen</i>	

VOLUME 17

ROTATION-INVARIANT CLUSTERING OF NEURONAL RESPONSES IN PRIMARY VISUAL CORTEX.....	12749
<i>Ivan Ustyuzhaninov, Santiago A. Cadena, Emmanouil Froudarakis, Paul G. Fahey, Edgar Y. Walker, Erick Cobos, Jacob Reimer, Fabian H. Sinz, Andreas S. Tolias, Matthias Bethge, Alexander S. Ecker</i>	
REFORMER: THE EFFICIENT TRANSFORMER.....	12764
<i>Nikita Kitaev, Lukasz Kaiser, Anselm Levskaya</i>	
TARGET-EMBEDDING AUTOENCODERS FOR SUPERVISED REPRESENTATION LEARNING.....	12776
<i>Daniel Jarrett, Mihaela Van Der Schaar</i>	
WATCH THE UNOBSERVED: A SIMPLE APPROACH TO PARALLELIZING MONTE CARLO TREE SEARCH.....	12810
<i>Anji Liu, Jianshu Chen, Mingze Yu, Yu Zhai, Xuwen Zhou, Ji Liu</i>	
RNA SECONDARY STRUCTURE PREDICTION BY LEARNING UNROLLED ALGORITHMS	12831
<i>Xinshi Chen, Yu Li, Ramzan Umarov, Xin Gao, Le Song</i>	
LEARNING TO BALANCE: BAYESIAN META-LEARNING FOR IMBALANCED AND OUT-OF-DISTRIBUTION TASKS	12850
<i>Hae Beom Lee, Hayeon Lee, Donghyun Na, Saehoon Kim, Minseop Park, Eunho Yang, Sung Ju Hwang</i>	
FAST TASK INFERENCE WITH VARIATIONAL INTRINSIC SUCCESSOR FEATURES	12865
<i>Steven Hansen, Will Dabney, Andre Barreto, David Warde-Farley, Tom Van De Wiele, Volodymyr Mnih</i>	
IMPLEMENTATION MATTERS IN DEEP RL: A CASE STUDY ON PPO AND TRPO.....	12883
<i>Logan Engstrom, Andrew Ilyas, Shibani Santurkar, Dimitris Tsipras, Firdaus Janoos, Larry Rudolph, Aleksander Madry</i>	
A CLOSER LOOK AT DEEP POLICY GRADIENTS	12897
<i>Andrew Ilyas, Logan Engstrom, Shibani Santurkar, Dimitris Tsipras, Firdaus Janoos, Larry Rudolph, Aleksander Madry</i>	
UNDERSTANDING AND ROBUSTIFYING DIFFERENTIABLE ARCHITECTURE SEARCH.....	12924
<i>Arber Zela, Thomas Elsken, Tonmoy Saikia, Yassine MARRAKCHI, Thomas Brox, Frank Hutter</i>	

DEEP BATCH ACTIVE LEARNING BY DIVERSE, UNCERTAIN GRADIENT LOWER BOUNDS	12952
<i>Jordan T. Ash, Chicheng Zhang, Akshay Krishnamurthy, John Langford, Alekh Agarwal</i>	
GEOMETRIC ANALYSIS OF NONCONVEX OPTIMIZATION LANDSCAPES FOR OVERCOMPLETE LEARNING	12978
<i>Qing Qu, Yuexiang Zhai, Xiao Li, Yuqian Zhang, Zhihui Zhu</i>	
A THEORY OF USABLE INFORMATION UNDER COMPUTATIONAL CONSTRAINTS	13039
<i>Yilun Xu, Shengjia Zhao, Jiaming Song, Russell Stewart, Stefano Ermon</i>	
MATHEMATICAL REASONING IN LATENT SPACE.....	13063
<i>Dennis Lee, Christian Szegedy, Markus Rabe, Sarah Loos, Kshitij Bansal</i>	
META-Q-LEARNING	13075
<i>Rasool Fakoor, Pratik Chaudhari, Stefano Soatto, Alexander J. Smola</i>	
COMPARING REWINDING AND FINE-TUNING IN NEURAL NETWORK PRUNING	13092
<i>Alex Renda, Jonathan Frankle, Michael Carbin</i>	
HARNESSING STRUCTURES FOR VALUE-BASED PLANNING AND REINFORCEMENT LEARNING.....	13123
<i>Yuzhe Yang, Guo Zhang, Zhi Xu, Dina Katabi</i>	
GRAPHZOOM: A MULTI-LEVEL SPECTRAL APPROACH FOR ACCURATE AND SCALABLE GRAPH EMBEDDING	13151
<i>Chenhui Deng, Zhiqiang Zhao, Yongyu Wang, Zhiru Zhang, Zhuo Feng</i>	
OPTIMAL STRATEGIES AGAINST GENERATIVE ATTACKS.....	13171
<i>Roy Mor, Erez Peterfreund, Matan Gavish, Amir Globerson</i>	
DYNAMICS-AWARE UNSUPERVISED DISCOVERY OF SKILLS	13207
<i>Archit Sharma, Shixiang Gu, Sergey Levine, Vikash Kumar, Karol Hausman</i>	
YOUR CLASSIFIER IS SECRETLY AN ENERGY BASED MODEL AND YOU SHOULD TREAT IT LIKE ONE.....	13228
<i>Will Grathwohl, Kuan-Chieh Wang, Joern-Henrik Jacobsen, David Duvenaud, Mohammad Norouzi, Kevin Swersky</i>	
CYCLICAL STOCHASTIC GRADIENT MCMC FOR BAYESIAN DEEP LEARNING.....	13251
<i>Ruqi Zhang, Chunyuan Li, Jianyi Zhang, Changyou Chen, Andrew Gordon Wilson</i>	
MIRROR-GENERATIVE NEURAL MACHINE TRANSLATION	13278
<i>Zaixiang Zheng, Hao Zhou, Shujian Huang, Lei Li, Xin-Yu Dai, Jiajun Chen</i>	
LEARNING HIERARCHICAL DISCRETE LINGUISTIC UNITS FROM VISUALLY-GROUNDED SPEECH.....	13294
<i>David Harwath*, Wei-Ning Hsu*, James Glass</i>	
MOGRIFIER LSTM	13316
<i>Gábor Melis, Tomáš Kociský, Phil Blunsom</i>	
POSTERIOR SAMPLING FOR MULTI-AGENT REINFORCEMENT LEARNING: SOLVING EXTENSIVE GAMES WITH IMPERFECT INFORMATION.....	13329
<i>Yichi Zhou, Jialian Li, Jun Zhu</i>	
WHY GRADIENT CLIPPING ACCELERATES TRAINING: A THEORETICAL JUSTIFICATION FOR ADAPTIVITY	13347
<i>Jingzhao Zhang, Tianxing He, Suvrit Sra, Ali Jadbabaie</i>	
NEURAL NETWORK BRANCHING FOR NEURAL NETWORK VERIFICATION.....	13368
<i>Jingyue Lu, M. Pawan Kumar</i>	
CONTRASTIVE LEARNING OF STRUCTURED WORLD MODELS	13392
<i>Thomas Kipf, Elise Van Der Pol, Max Welling</i>	
DATA-DEPENDENT GAUSSIAN PRIOR OBJECTIVE FOR LANGUAGE GENERATION	13413
<i>Zuchao Li, Rui Wang, Kehai Chen, Masso Utiyama, Eiichiro Sumita, Zhuosheng Zhang, Hai Zhao</i>	
ON THE CONVERGENCE OF FEDAVG ON NON-IID DATA.....	13431
<i>Xiang Li, Kaixuan Huang, Wenhao Yang, Shusen Wang, Zhihua Zhang</i>	

PRINCIPLED WEIGHT INITIALIZATION FOR HYPERNETWORKS	13457
<i>Oscar Chang, Lampros Flokas, Hod Lipson</i>	
GENDICE: GENERALIZED OFFLINE ESTIMATION OF STATIONARY VALUES	13478
<i>Ruiyi Zhang*, Bo Dai*, Lihong Li, Dale Schuurmans</i>	
BACKPACK: PACKING MORE INTO BACKPROP	13503
<i>Felix Dangel, Frederik Kunstner, Philipp Hennig</i>	
CATER: A DIAGNOSTIC DATASET FOR COMPOSITIONAL ACTIONS & TEMPORAL REASONING.....	13523
<i>Rohit Girdhar, Deva Ramanan</i>	
Author Index	