

36th International Conference on Machine Learning (ICML 2019)

Proceedings of Machine Learning Research Volume 97

Long Beach, California, USA
9 – 15 June 2019

Part 1 of 19

ISBN: 978-1-5108-8698-8

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© (2019) by the Authors
All rights reserved.

Printed by Curran Associates, Inc. (2019)

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

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

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

rasmussen@ptd.net

Additional copies of this publication are available from:

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

TABLE OF CONTENTS

PART 1

| | |
|--|-----|
| ARES AND MARS ADVERSARIAL AND MMD-MINIMIZING REGRESSION FOR SDES | 1 |
| <i>Gabriele Abbati, Philippe Wenk, Michael A. Osborne, Andreas Krause, Bernhard Schölkopf, Stefan Bauer</i> | |
| DYNAMIC WEIGHTS IN MULTI-OBJECTIVE DEEP REINFORCEMENT LEARNING | 13 |
| <i>Axel Abels, Diederik Roijers, Tom Lenaerts, Ann Nowé, Denis Steckelmacher</i> | |
| MIXHOP: HIGHER-ORDER GRAPH CONVOLUTIONAL ARCHITECTURES VIA SPARSIFIED NEIGHBORHOOD MIXING | 32 |
| <i>Sami Abu-El-Haija, Bryan Perozzi, Amol Kapoor, Nazanin Alipourfard, Kristina Lerman, Hrayr Harutyunyan, Greg Ver Steeg, Aram Galstyan</i> | |
| COMMUNICATION-CONSTRAINED INFERENCE AND THE ROLE OF SHARED RANDOMNESS | 42 |
| <i>Jayadev Acharya, Clement Canonne, Himanshu Tyagi</i> | |
| DISTRIBUTED LEARNING WITH SUBLINEAR COMMUNICATION | 60 |
| <i>Jayadev Acharya, Chris De Sa, Dylan Foster, Karthik Sridharan</i> | |
| COMMUNICATION COMPLEXITY IN LOCALLY PRIVATE DISTRIBUTION ESTIMATION AND HEAVY HITTERS | 85 |
| <i>Jayadev Acharya, Ziteng Sun</i> | |
| LEARNING MODELS FROM DATA WITH MEASUREMENT ERROR: TACKLING UNDERREPORTING | 95 |
| <i>Roy Adams, Yuelong Ji, Xiaobin Wang, Suchi Saria</i> | |
| TIBGM: A TRANSFERABLE AND INFORMATION-BASED GRAPHICAL MODEL APPROACH FOR REINFORCEMENT LEARNING | 106 |
| <i>Tameem Adel, Adrian Weller</i> | |
| PAC LEARNABILITY OF NODE FUNCTIONS IN NETWORKED DYNAMICAL SYSTEMS | 119 |
| <i>Abhijin Adiga, Chris J Kuhlman, Madhav Marathe, S Ravi, Anil Vullikanti</i> | |
| STATIC AUTOMATIC BATCHING IN TENSORFLOW | 129 |
| <i>Ashish Agarwal</i> | |
| EFFICIENT FULL-MATRIX ADAPTIVE REGULARIZATION | 139 |
| <i>Naman Agarwal, Brian Bullins, Xinyi Chen, Elad Hazan, Karan Singh, Cyril Zhang, Yi Zhang</i> | |
| ONLINE CONTROL WITH ADVERSARIAL DISTURBANCES | 154 |
| <i>Naman Agarwal, Brian Bullins, Elad Hazan, Sham Kakade, Karan Singh</i> | |
| FAIR REGRESSION: QUANTITATIVE DEFINITIONS AND REDUCTION-BASED ALGORITHMS | 166 |
| <i>Alekh Agarwal, Miroslav Dudík, Zhiwei Steven Wu</i> | |
| LEARNING TO GENERALIZE FROM SPARSE AND UNDERSPECIFIED REWARDS | 184 |
| <i>Rishabh Agarwal, Chen Liang, Dale Schuurmans, Mohammad Norouzi</i> | |
| THE KERNEL INTERACTION TRICK: FAST BAYESIAN DISCOVERY OF PAIRWISE INTERACTIONS IN HIGH DIMENSIONS | 199 |
| <i>Raj Agrawal, Brian Trippe, Jonathan Huggins, Tamara Broderick</i> | |
| UNDERSTANDING THE IMPACT OF ENTROPY ON POLICY OPTIMIZATION | 215 |
| <i>Zafarali Ahmed, Nicolas Le Roux, Mohammad Norouzi, Dale Schuurmans</i> | |
| FAIRWASHING: THE RISK OF RATIONALIZATION | 240 |
| <i>Ulrich Aivodji, Hiromi Arai, Olivier Fortineau, Sébastien Gams, Satoshi Hara, Alain Tapp</i> | |
| ADAPTIVE STOCHASTIC NATURAL GRADIENT METHOD FOR ONE-SHOT NEURAL ARCHITECTURE SEARCH | 253 |
| <i>Youhei Akimoto, Shinichi Shirakawa, Nozomu Yoshinari, Kento Uchida, Shota Saito, Kouhei Nishida</i> | |
| PROJECTIONS FOR APPROXIMATE POLICY ITERATION ALGORITHMS | 267 |
| <i>Riad Akrouf, Joni Pajarinen, Jan Peters, Gerhard Neumann</i> | |
| VALIDATING CAUSAL INFERENCE MODELS VIA INFLUENCE FUNCTIONS | 281 |
| <i>Ahmed Alaa, Mihaela Van Der Schaar</i> | |
| MULTI-OBJECTIVE TRAINING OF GENERATIVE ADVERSARIAL NETWORKS WITH MULTIPLE DISCRIMINATORS | 292 |
| <i>Isabela Albuquerque, Joao Monteiro, Thang Doan, Breandan Considine, Tiago Falk, Ioannis Mitliagkas</i> | |
| GRAPH ELEMENT NETWORKS: ADAPTIVE, STRUCTURED COMPUTATION AND MEMORY | 321 |
| <i>Ferran Alet, Adarsh Keshav Jeewajee, Maria Bauza Villalonga, Alberto Rodriguez, Tomas Lozano-Perez, Leslie Kaelbling</i> | |
| ANALOGIES EXPLAINED: TOWARDS UNDERSTANDING WORD EMBEDDINGS | 337 |
| <i>Carl Allen, Timothy Hospedales</i> | |
| INFINITE MIXTURE PROTOTYPES FOR FEW-SHOT LEARNING | 348 |
| <i>Kelsey Allen, Evan Shelhamer, Hanul Shin, Joshua Tenenbaum</i> | |
| A CONVERGENCE THEORY FOR DEEP LEARNING VIA OVER-PARAMETERIZATION | 362 |
| <i>Zeyuan Allen-Zhu, Yuanzhi Li, Zhao Song</i> | |
| ASYNCHRONOUS BATCH BAYESIAN OPTIMISATION WITH IMPROVED LOCAL PENALISATION | 373 |
| <i>Ahsan Alvi, Binxin Ru, Jan-Peter Calliess, Stephen Roberts, Michael A. Osborne</i> | |
| BOUNDING USER CONTRIBUTIONS: A BIAS-VARIANCE TRADE-OFF IN DIFFERENTIAL PRIVACY | 388 |
| <i>Kareem Amin, Alex Kulesza, Andres Munoz, Sergei Vassilvitskii</i> | |

| | |
|--|-----|
| EXPLAINING DEEP NEURAL NETWORKS WITH A POLYNOMIAL TIME ALGORITHM FOR SHAPLEY VALUE APPROXIMATION | 400 |
| <i>Marco Ancona, Cengiz Oztireli, Markus Gross</i> | |
| SCALING UP ORDINAL EMBEDDING: A LANDMARK APPROACH | 416 |
| <i>Jesse Anderton, Javed Aslam</i> | |
| SORTING OUT LIPSCHITZ FUNCTION APPROXIMATION | 432 |
| <i>Cem Anil, James Lucas, Roger Grosse</i> | |
| SPARSE MULTI-CHANNEL VARIATIONAL AUTOENCODER FOR THE JOINT ANALYSIS OF HETEROGENEOUS DATA | 453 |
| <i>Luigi Antelmi, Nicholas Ayache, Philippe Robert, Marco Lorenzi</i> | |
| UNSUPERVISED LABEL NOISE MODELING AND LOSS CORRECTION | 465 |
| <i>Eric Arazo, Diego Ortego, Paul Albert, Noel O'Connor, Kevin McGuinness</i> | |
| FINE-GRAINED ANALYSIS OF OPTIMIZATION AND GENERALIZATION FOR OVERPARAMETERIZED TWO-LAYER NEURAL NETWORKS | 477 |
| <i>Sanjeev Arora, Simon Du, Wei Hu, Zhiyuan Li, Ruosong Wang</i> | |
| DISTRIBUTED WEIGHTED MATCHING VIA RANDOMIZED COMPOSABLE CORESETS | 503 |
| <i>Sepehr Assadi, Mohammadhossein Bateni, Vahab Mirrokni</i> | |
| STOCHASTIC GRADIENT PUSH FOR DISTRIBUTED DEEP LEARNING | 514 |
| <i>Mahmoud Assran, Nicolas Loizou, Nicolas Ballas, Mike Rabbat</i> | |
| BAYESIAN OPTIMIZATION OF COMPOSITE FUNCTIONS | 547 |
| <i>Raul Astudillo, Peter Frazier</i> | |
| LINEAR-COMPLEXITY DATA-PARALLEL EARTH MOVER'S DISTANCE APPROXIMATIONS | 565 |
| <i>Kubilay Atas, Thomas Mittelholzer</i> | |
| BENEFITS AND PITFALLS OF THE EXPONENTIAL MECHANISM WITH APPLICATIONS TO HILBERT SPACES AND FUNCTIONAL PCA | 577 |
| <i>Jordan Awan, Ana Kenney, Matthew Reimherr, Aleksandra Slavkovic</i> | |
| FEATURE GROUPING AS A STOCHASTIC REGULARIZER FOR HIGH-DIMENSIONAL STRUCTURED DATA | 592 |
| <i>Sergul Aydore, Bertrand Thirion, Gael Varoquaux</i> | |
| BEYOND THE CHINESE RESTAURANT AND PITMAN-YOR PROCESSES: STATISTICAL MODELS WITH DOUBLE POWER-LAW BEHAVIOR | 604 |
| <i>Fadhel Ayed, Juho Lee, Francois Caron</i> | |
| SCALABLE FAIR CLUSTERING | 623 |
| <i>Arturs Backurs, Piotr Indyk, Krzysztof Onak, Baruch Schieber, Ali Vakilian, Tal Wagner</i> | |
| ENTROPIC GANS MEET VAES: A STATISTICAL APPROACH TO COMPUTE SAMPLE LIKELIHOODS IN GANS | 635 |
| <i>Yogesh Balaji, Hamed Hassani, Rama Chellappa, Soheil Feizi</i> | |
| PROVABLE GUARANTEES FOR GRADIENT-BASED META-LEARNING | 651 |
| <i>Maria-Florina Balcan, Mikhail Khodak, Ameet Talwalkar</i> | |
| OPEN-ENDED LEARNING IN SYMMETRIC ZERO-SUM GAMES | 676 |
| <i>David Balduzzi, Marta Garnelo, Yoram Bachrach, Wojciech Czarnecki, Julien Perolat, Max Jaderberg, Thore Graepel</i> | |
| CONCRETE AUTOENCODERS: DIFFERENTIABLE FEATURE SELECTION AND RECONSTRUCTION | 694 |
| <i>Muhammed Fatih Balin, Abubakar Abid, James Zou</i> | |
| HOLIST: AN ENVIRONMENT FOR MACHINE LEARNING OF HIGHER ORDER LOGIC THEOREM PROVING | 712 |
| <i>Kshitij Bansal, Sarah Loos, Markus Rabe, Christian Szegedy, Stewart Wilcox</i> | |

PART 2

| | |
|--|-----|
| STRUCTURED AGENTS FOR PHYSICAL CONSTRUCTION | 722 |
| <i>Victor Bapst, Alvaro Sanchez-Gonzalez, Carl Doersch, Kimberly Stachenfeld, Pushmeet Kohli, Peter Battaglia, Jessica Hamrick</i> | |
| LEARNING TO ROUTE IN SIMILARITY GRAPHS | 748 |
| <i>Dmitry Baranchuk, Dmitry Persiyanov, Anton Sinitin, Artem Babenko</i> | |
| A PERSONALIZED AFFECTIVE MEMORY MODEL FOR IMPROVING EMOTION RECOGNITION | 758 |
| <i>Pablo Barros, German Parisi, Stefan Wermter</i> | |
| SCALE-FREE ADAPTIVE PLANNING FOR DETERMINISTIC DYNAMICS & DISCOUNTED REWARDS | 772 |
| <i>Peter Bartlett, Victor Gabillon, Jennifer Healey, Michal Valko</i> | |
| PARETO OPTIMAL STREAMING UNSUPERVISED CLASSIFICATION | 792 |
| <i>Soumya Basu, Steven Gutstein, Brent Lance, Sanjay Shakkottai</i> | |
| CATEGORICAL FEATURE COMPRESSION VIA SUBMODULAR OPTIMIZATION | 814 |
| <i>Mohammadhossein Bateni, Lin Chen, Hossein Esfandiari, Thomas Fu, Vahab Mirrokni, Afshin Rostamizadeh</i> | |
| NOISE2SELF: BLIND DENOISING BY SELF-SUPERVISION | 826 |
| <i>Joshua Batson, Loic Royer</i> | |
| EFFICIENT OPTIMIZATION OF LOOPS AND LIMITS WITH RANDOMIZED TELESCOPING SUMS | 836 |
| <i>Alex Beatson, Ryan P Adams</i> | |
| RECURRENT KALMAN NETWORKS: FACTORIZED INFERENCE IN HIGH-DIMENSIONAL DEEP FEATURE SPACES | 855 |
| <i>Philipp Becker, Harit Pandya, Gregor Gebhardt, Cheng Zhao, C. James Taylor, Gerhard Neumann</i> | |

| | |
|---|------|
| SWITCHING LINEAR DYNAMICS FOR VARIATIONAL BAYES FILTERING | 870 |
| <i>Philip Becker-Ehmck, Jan Peters, Patrick Van Der Smagt</i> | |
| ACTIVE LEARNING FOR PROBABILISTIC STRUCTURED PREDICTION OF CUTS AND MATCHINGS | 884 |
| <i>Sima Behpour, Anqi Liu, Brian Ziebart</i> | |
| INVERTIBLE RESIDUAL NETWORKS | 894 |
| <i>Jens Behrmann, Will Grathwohl, Ricky T. Q. Chen, David Duvenaud, Joern-Henrik Jacobsen</i> | |
| GREEDY LAYERWISE LEARNING CAN SCALE TO IMAGENET | 911 |
| <i>Eugene Belilovsky, Michael Eickenberg, Edouard Oyallon</i> | |
| OVERCOMING MULTI-MODEL FORGETTING | 926 |
| <i>Yassine Benyahia, Kaicheng Yu, Kamil Bennani Smires, Martin Jaggi, Anthony C. Davison, Mathieu Salzmann, Claudiu Musat</i> | |
| OPTIMAL KRONECKER-SUM APPROXIMATION OF REAL TIME RECURRENT LEARNING | 939 |
| <i>Frederik Benzing, Marcelo Matheus Gaudy, Asier Mujika, Anders Martinsson, Angelika Steger</i> | |
| ADVERSARIALLY LEARNED REPRESENTATIONS FOR INFORMATION OBFUSCATION AND INFERENCE | 960 |
| <i>Martin Bertran, Natalia Martinez, Afroditi Papadaki, Qiang Qiu, Miguel Rodrigues, Galen Reeves, Guillermo Sapiro</i> | |
| BANDIT MULTICLASS LINEAR CLASSIFICATION: EFFICIENT ALGORITHMS FOR THE SEPARABLE CASE | 975 |
| <i>Alina Beygelzimer, David Pal, Balazs Szorenyi, Devanathan Thiruvkatachari, Chen-Yu Wei, Chicheng Zhang</i> | |
| ANALYZING FEDERATED LEARNING THROUGH AN ADVERSARIAL LENS | 1012 |
| <i>Arjun Nitin Bhagoji, Supriyo Chakraborty, Prateek Mittal, Seraphin Calo</i> | |
| OPTIMAL CONTINUOUS DR-SUBMODULAR MAXIMIZATION AND APPLICATIONS TO PROVABLE MEAN FIELD INFERENCE | 1029 |
| <i>Yatao Bian, Joachim Buhmann, Andreas Krause</i> | |
| MORE EFFICIENT OFF-POLICY EVALUATION THROUGH REGULARIZED TARGETED LEARNING | 1051 |
| <i>Aurelien Bibaut, Ivana Malenica, Nikos Vlassis, Mark Van Der Laan</i> | |
| A KERNEL PERSPECTIVE FOR REGULARIZING DEEP NEURAL NETWORKS | 1061 |
| <i>Alberto Bietti, Grégoire Mialon, Dexiong Chen, Julien Mairal</i> | |
| RETHINKING LOSSY COMPRESSION: THE RATE-DISTORTION-PERCEPTION TRADEOFF | 1081 |
| <i>Yochai Blau, Tomer Michaeli</i> | |
| CORRELATED BANDITS OR: HOW TO MINIMIZE MEAN-SQUARED ERROR ONLINE | 1103 |
| <i>Vinay Praneeth Boda, L. A. Prashanth</i> | |
| ADVERSARIAL ATTACKS ON NODE EMBEDDINGS VIA GRAPH POISONING | 1112 |
| <i>Aleksandar Bojchevski, Stephan Günnemann</i> | |
| ONLINE VARIANCE REDUCTION WITH MIXTURES | 1124 |
| <i>Zalán Borsos, Sebastian Curi, Kfir Yehuda Levy, Andreas Krause</i> | |
| COMPOSITIONAL FAIRNESS CONSTRAINTS FOR GRAPH EMBEDDINGS | 1139 |
| <i>Avishek Bose, William Hamilton</i> | |
| UNREPRODUCIBLE RESEARCH IS REPRODUCIBLE | 1150 |
| <i>Xavier Bouthillier, César Laurent, Pascal Vincent</i> | |
| BLENDED CONDITIONAL GRADIENTS | 1160 |
| <i>Gábor Braun, Sebastian Pokutta, Dan Tu, Stephen Wright</i> | |
| CORESETS FOR ORDERED WEIGHTED CLUSTERING | N/A |
| <i>Vladimir Braverman, Shaofeng H.-C. Jiang, Robert Krauthgamer, Xuan Wu</i> | |
| TARGET TRACKING FOR CONTEXTUAL BANDITS: APPLICATION TO DEMAND SIDE MANAGEMENT | 1184 |
| <i>Margaux Brégère, Pierre Gaillard, Yannig Goude, Gilles Stoltz</i> | |
| ACTIVE MANIFOLDS: A NON-LINEAR ANALOGUE TO ACTIVE SUBSPACES | 1204 |
| <i>Robert Bridges, Anthony Gruber, Christopher Felder, Miki Verma, Chelsey Hoff</i> | |
| CONDITIONING BY ADAPTIVE SAMPLING FOR ROBUST DESIGN | 1214 |
| <i>David Brookes, Hahnbeom Park, Jennifer Listgarten</i> | |
| EXTRAPOLATING BEYOND SUBOPTIMAL DEMONSTRATIONS VIA INVERSE REINFORCEMENT LEARNING FROM OBSERVATIONS | 1232 |
| <i>Daniel Brown, Wonjoon Goo, Prabhat Nagarajan, Scott Niekum</i> | |
| DEEP COUNTERFACTUAL REGRET MINIMIZATION | 1256 |
| <i>Noam Brown, Adam Lerer, Sam Gross, Tuomas Sandholm</i> | |
| UNDERSTANDING THE ORIGINS OF BIAS IN WORD EMBEDDINGS | 1275 |
| <i>Marc-Etienne Brunet, Colleen Alkalay-Houlihan, Ashton Anderson, Richard Zemel</i> | |
| LOW LATENCY PRIVACY PRESERVING INFERENCE | 1295 |
| <i>Alon Brutzkus, Ran Gilad-Bachrach, Oren Elisha</i> | |
| WHY DO LARGER MODELS GENERALIZE BETTER? A THEORETICAL PERSPECTIVE VIA THE XOR PROBLEM | 1310 |
| <i>Alon Brutzkus, Amir Globerson</i> | |
| ADVERSARIAL EXAMPLES FROM COMPUTATIONAL CONSTRAINTS | 1345 |
| <i>Sebastien Bubeck, Yin Tat Lee, Eric Price, Ilya Razenshteyn</i> | |
| SELF-SIMILAR EPOCHS: VALUE IN ARRANGEMENT | 1361 |
| <i>Eliav Buchnik, Edith Cohen, Avinatan Hasidim, Yossi Matias</i> | |
| LEARNING GENERATIVE MODELS ACROSS INCOMPARABLE SPACES | 1374 |
| <i>Charlotte Bunne, David Alvarez-Melis, Andreas Krause, Stefanie Jegelka</i> | |
| RATES OF CONVERGENCE FOR SPARSE VARIATIONAL GAUSSIAN PROCESS REGRESSION | 1390 |
| <i>David Burt, Carl Edward Rasmussen, Mark Van Der Wilk</i> | |

| | |
|---|------|
| WHAT IS THE EFFECT OF IMPORTANCE WEIGHTING IN DEEP LEARNING? | 1405 |
| <i>Jonathon Byrd, Zachary Lipton</i> | |

PART 3

| | |
|--|------|
| A QUANTITATIVE ANALYSIS OF THE EFFECT OF BATCH NORMALIZATION ON GRADIENT DESCENT | 1420 |
| <i>Yongqiang Cai, Qianxiao Li, Zuowei Shen</i> | |
| ACCELERATED LINEAR CONVERGENCE OF STOCHASTIC MOMENTUM METHODS IN WASSERSTEIN DISTANCES | 1445 |
| <i>Bugra Can, Mert Gurbuzbalaban, Lingjiong Zhu</i> | |
| ACTIVE EMBEDDING SEARCH VIA NOISY PAIRED COMPARISONS | 1493 |
| <i>Gregory Canal, Andy Massimino, Mark Davenport, Christopher Rozell</i> | |
| DYNAMIC LEARNING WITH FREQUENT NEW PRODUCT LAUNCHES: A SEQUENTIAL MULTINOMIAL LOGIT BANDIT PROBLEM | 1513 |
| <i>Junyu Cao, Wei Sun</i> | |
| COMPETING AGAINST NASH EQUILIBRIA IN ADVERSARIALY CHANGING ZERO-SUM GAMES | 1536 |
| <i>Adrian Rivera Cardoso, Jacob Abernethy, He Wang, Huan Xu</i> | |
| AUTOMATED MODEL SELECTION WITH BAYESIAN QUADRATURE | 1555 |
| <i>Henry Chai, Jean-Francois Ton, Michael A. Osborne, Roman Garnett</i> | |
| LEARNING ACTION REPRESENTATIONS FOR REINFORCEMENT LEARNING | 1565 |
| <i>Yash Chandak, Georgios Theodorou, James Kostas, Scott Jordan, Philip Thomas</i> | |
| DYNAMIC MEASUREMENT SCHEDULING FOR EVENT FORECASTING USING DEEP RL | 1583 |
| <i>Chun-Hao Chang, Mingjie Mai, Anna Goldenberg</i> | |
| ON SYMMETRIC LOSSES FOR LEARNING FROM CORRUPTED LABELS | 1597 |
| <i>Nontawat Charoenphakdee, Jongyeong Lee, Masashi Sugiyama</i> | |
| ONLINE LEARNING WITH KERNEL LOSSES | 1625 |
| <i>Niladri Chatterji, Aldo Pacchiano, Peter Bartlett</i> | |
| NEURAL NETWORK ATTRIBUTIONS: A CAUSAL PERSPECTIVE | 1660 |
| <i>Aditya Chattopadhyay, Piyushi Manupriya, Anirban Sarkar, Vineeth N Balasubramanian</i> | |
| PAC IDENTIFICATION OF MANY GOOD ARMS IN STOCHASTIC MULTI-ARMED BANDITS | 1677 |
| <i>Arghya Roy Chaudhuri, Shivaram Kalyanakrishnan</i> | |
| NEAREST NEIGHBOR AND KERNEL SURVIVAL ANALYSIS: NONASYMPTOTIC ERROR BOUNDS AND STRONG CONSISTENCY RATES | 1699 |
| <i>George Chen</i> | |
| STEIN POINT MARKOV CHAIN MONTE CARLO | 1737 |
| <i>Wilson Ye Chen, Alessandro Barp, Francois-Xavier Briol, Jackson Gorham, Mark Girolami, Lester Mackey, Chris Oates</i> | |
| PARTICLE FLOW BAYES' RULE | 1768 |
| <i>Xinshi Chen, Hanjun Dai, Le Song</i> | |
| PROPORTIONALLY FAIR CLUSTERING | 1782 |
| <i>Xingyu Chen, Brandon Fain, Liang Lyu, Kamesh Munagala</i> | |
| INFORMATION-THEORETIC CONSIDERATIONS IN BATCH REINFORCEMENT LEARNING | 1792 |
| <i>Jinglin Chen, Nan Jiang</i> | |
| GENERATIVE ADVERSARIAL USER MODEL FOR REINFORCEMENT LEARNING BASED RECOMMENDATION SYSTEM | 1818 |
| <i>Xinshi Chen, Shuang Li, Hui Li, Shaohua Jiang, Yuan Qi, Le Song</i> | |
| UNDERSTANDING AND UTILIZING DEEP NEURAL NETWORKS TRAINED WITH NOISY LABELS | 1833 |
| <i>Pengfei Chen, Ben Ben Liao, Guangyong Chen, Shengyu Zhang</i> | |
| A GRADUAL, SEMI-DISCRETE APPROACH TO GENERATIVE NETWORK TRAINING VIA EXPLICIT WASSERSTEIN MINIMIZATION | 1845 |
| <i>Yucheng Chen, Matus Telgarsky, Chao Zhang, Bolton Bailey, Daniel Hsu, Jian Peng</i> | |
| TRANSFERABILITY VS. DISCRIMINABILITY: BATCH SPECTRAL PENALIZATION FOR ADVERSARIAL DOMAIN ADAPTATION | 1859 |
| <i>Xinyang Chen, Sinan Wang, Mingsheng Long, Jianmin Wang</i> | |
| FAST INCREMENTAL VON NEUMANN GRAPH ENTROPY COMPUTATION: THEORY, ALGORITHM, AND APPLICATIONS | 1869 |
| <i>Pin-Yu Chen, Lingfei Wu, Sijia Liu, Indika Rajapakse</i> | |
| KATALYST: BOOSTING CONVEX KATAYUSHA FOR NON-CONVEX PROBLEMS WITH A LARGE CONDITION NUMBER | 1887 |
| <i>Zaiyi Chen, Yi Xu, Haoyuan Hu, Tianbao Yang</i> | |
| MULTIVARIATE-INFORMATION ADVERSARIAL ENSEMBLE FOR SCALABLE JOINT DISTRIBUTION MATCHING | 1901 |
| <i>Ziliang Chen, Zhanfu Yang, Xiaoxi Wang, Xiaodan Liang, Xiaopeng Yan, Guanbin Li, Liang Lin</i> | |
| ROBUST DECISION TREES AGAINST ADVERSARIAL EXAMPLES | 1911 |
| <i>Hongge Chen, Huan Zhang, Duane Boning, Cho-Jui Hsieh</i> | |
| RAFM: RANK-AWARE FACTORIZATION MACHINES | 1927 |
| <i>Xiaoshuang Chen, Yin Zheng, Jiaying Wang, Wenye Ma, Junzhou Huang</i> | |
| CONTROL REGULARIZATION FOR REDUCED VARIANCE REINFORCEMENT LEARNING | 1940 |
| <i>Richard Cheng, Abhinav Verma, Gabor Orosz, Swarat Chaudhuri, Yisong Yue, Joel Burdick</i> | |

| | |
|---|------|
| PREDICTOR-CORRECTOR POLICY OPTIMIZATION | 1957 |
| <i>Ching-An Cheng, Xinyan Yan, Nathan Ratliff, Byron Boots</i> | |
| VARIATIONAL INFERENCE FOR SPARSE NETWORK RECONSTRUCTION FROM COUNT DATA | 1988 |
| <i>Julien Chiquet, Stephane Robin, Mahendra Mariadassou</i> | |
| RANDOM WALKS ON HYPERGRAPHS WITH EDGE-DEPENDENT VERTEX WEIGHTS | 2002 |
| <i>Uthsav Chitra, Benjamin Raphael</i> | |
| NEURAL JOINT SOURCE-CHANNEL CODING | 2026 |
| <i>Kristy Choi, Kedar Tatwawadi, Aditya Grover, Tsachy Weissman, Stefano Ermon</i> | |
| BEYOND BACKPROP: ONLINE ALTERNATING MINIMIZATION WITH AUXILIARY VARIABLES | 2041 |
| <i>Anna Choromanska, Benjamin Cowen, Sadhana Kumaravel, Ronny Luss, Mattia Rigotti, Irina Rish, Paolo Diachille, Viatcheslav Gurev, Brian Kingsbury, Ravi Tejwani, Djallel Bouneffouf</i> | |
| UNIFYING ORTHOGONAL MONTE CARLO METHODS | 2061 |
| <i>Krzysztof Choromanski, Mark Rowland, Wenyu Chen, Adrian Weller</i> | |
| PROBABILITY FUNCTIONAL DESCENT: A UNIFYING PERSPECTIVE ON GANS, VARIATIONAL INFERENCE, AND REINFORCEMENT LEARNING | 2071 |
| <i>Casey Chu, Jose Blanchet, Peter Glynn</i> | |
| MEANSUM: A NEURAL MODEL FOR UNSUPERVISED MULTI-DOCUMENT ABSTRACTIVE SUMMARIZATION | 2088 |
| <i>Eric Chu, Peter Liu</i> | |
| WEAK DETECTION OF SIGNAL IN THE SPIKED WIGNER MODEL | 2111 |
| <i>Hye Won Chung, Ji Oon Lee</i> | |
| NEW RESULTS ON INFORMATION THEORETIC CLUSTERING | 2141 |
| <i>Ferdinando Cicalese, Eduardo Laber, Lucas Murtinho</i> | |

PART 4

| | |
|---|------|
| SENSITIVITY ANALYSIS OF LINEAR STRUCTURAL CAUSAL MODELS | 2202 |
| <i>Carlos Cinelli, Daniel Kumor, Bryant Chen, Judea Pearl, Elias Bareinboim</i> | |
| DIMENSIONALITY REDUCTION FOR TUKEY REGRESSION | 2223 |
| <i>Kenneth Clarkson, Ruosong Wang, David Woodruff</i> | |
| ON MEDIANS OF (RANDOMIZED) PAIRWISE MEANS | 2258 |
| <i>Pierre Laforgue, Stephan Clemencon, Patrice Bertail</i> | |
| QUANTIFYING GENERALIZATION IN REINFORCEMENT LEARNING | 2280 |
| <i>Karl Cobbe, Oleg Klimov, Chris Hesse, Taehoon Kim, John Schulman</i> | |
| EMPIRICAL ANALYSIS OF BEAM SEARCH PERFORMANCE DEGRADATION IN NEURAL SEQUENCE MODELS | 2294 |
| <i>Eldan Cohen, Christopher Beck</i> | |
| LEARNING LINEAR-QUADRATIC REGULATORS EFFICIENTLY WITH ONLY REGRET | 2313 |
| <i>Alon Cohen, Tomer Koren, Yishay Mansour</i> | |
| CERTIFIED ADVERSARIAL ROBUSTNESS VIA RANDOMIZED SMOOTHING | 2323 |
| <i>Jeremy Cohen, Elan Rosenfeld, Zico Kolter</i> | |
| GAUGE EQUIVARIANT CONVOLUTIONAL NETWORKS AND THE ICOSAHEDRAL CNN | 2357 |
| <i>Taco Cohen, Maurice Weiler, Berkay Kicanaoglu, Max Welling</i> | |
| CURIOUS: INTRINSICALLY MOTIVATED MODULAR MULTI-GOAL REINFORCEMENT LEARNING | 2372 |
| <i>Cédric Colas, Pierre Fournier, Mohamed Chetouani, Olivier Sigaud, Pierre-Yves Oudeyer</i> | |
| A FULLY DIFFERENTIABLE BEAM SEARCH DECODER | 2388 |
| <i>Ronan Collobert, Awni Hannun, Gabriel Synnaeve</i> | |
| SCALABLE METROPOLIS-HASTINGS FOR EXACT BAYESIAN INFERENCE WITH LARGE DATASETS | 2398 |
| <i>Rob Cornish, Paul Vanetti, Alexandre Bouchard-Cote, George Deligiannidis, Arnaud Doucet</i> | |
| ADJUSTMENT CRITERIA FOR GENERALIZING EXPERIMENTAL FINDINGS | 2430 |
| <i>Juan Correa, Jin Tian, Elias Bareinboim</i> | |
| ONLINE LEARNING WITH SLEEPING EXPERTS AND FEEDBACK GRAPHS | 2439 |
| <i>Corinna Cortes, Giulia Desalvo, Claudio Gentile, Mehryar Mohri, Scott Yang</i> | |
| ACTIVE LEARNING WITH DISAGREEMENT GRAPHS | 2468 |
| <i>Corinna Cortes, Giulia Desalvo, Mehryar Mohri, Ningshan Zhang, Claudio Gentile</i> | |
| SHAPE CONSTRAINTS FOR SET FUNCTIONS | 2484 |
| <i>Andrew Cotter, Maya Gupta, Heinrich Jiang, Erez Loidor, James Muller, Tamann Narayan, Serena Wang, Tao Zhu</i> | |
| TRAINING WELL-GENERALIZING CLASSIFIERS FOR FAIRNESS METRICS AND OTHER DATA-DEPENDENT CONSTRAINTS | 2499 |
| <i>Andrew Cotter, Maya Gupta, Heinrich Jiang, Nathan Srebro, Karthik Sridharan, Serena Wang, Blake Woodworth, Seungil You</i> | |
| MONGE BLUNTS BAYES: HARDNESS RESULTS FOR ADVERSARIAL TRAINING | 2523 |
| <i>Zac Cranko, Aditya Menon, Richard Nock, Cheng Soon Ong, Zhan Shi, Christian Walder</i> | |
| BOOSTED DENSITY ESTIMATION REMASTERED | 2544 |
| <i>Zac Cranko, Richard Nock</i> | |
| SUBMODULAR COST SUBMODULAR COVER WITH AN APPROXIMATE ORACLE | 2582 |
| <i>Victoria Crawford, Alan Kuhnle, My Thai</i> | |
| FLEXIBLY FAIR REPRESENTATION LEARNING BY DISENTANGLEMENT | 2600 |
| <i>Elliot Creager, David Madras, Joern-Henrik Jacobsen, Marissa Weis, Kevin Swersky, Toniann Pitassi, Richard Zemel</i> | |

| | |
|--|------|
| ANYTIME ONLINE-TO-BATCH, OPTIMISM AND ACCELERATION | 2613 |
| <i>Ashok Cutkosky</i> | |
| MATRIX-FREE PRECONDITIONING IN ONLINE LEARNING | 2622 |
| <i>Ashok Cutkosky, Tamas Sarlos</i> | |
| MINIMAL ACHIEVABLE SUFFICIENT STATISTIC LEARNING | 2645 |
| <i>Milan Cvitkovic, Günther Koliander</i> | |
| OPEN VOCABULARY LEARNING ON SOURCE CODE WITH A GRAPH-STRUCTURED CACHE | 2662 |
| <i>Milan Cvitkovic, Badal Singh, Animashree Anandkumar</i> | |
| THE VALUE FUNCTION POLYTOPE IN REINFORCEMENT LEARNING | 2675 |
| <i>Robert Dadashi, Adrien Ali Taiga, Nicolas Le Roux, Dale Schuurmans, Marc G. Bellemare</i> | |
| BAYESIAN OPTIMIZATION MEETS BAYESIAN OPTIMAL STOPPING | 2701 |
| <i>Zhongxiang Dai, Haibin Yu, Bryan Kian Hsiang Low, Patrick Jaillet</i> | |
| POLICY CERTIFICATES: TOWARDS ACCOUNTABLE REINFORCEMENT LEARNING | 2726 |
| <i>Christoph Dann, Lihong Li, Wei Wei, Emma Brunskill</i> | |
| LEARNING FAST ALGORITHMS FOR LINEAR TRANSFORMS USING BUTTERFLY FACTORIZATIONS | 2736 |
| <i>Tri Dao, Albert Gu, Matthew Eichhorn, Atri Rudra, Christopher Re</i> | |
| A KERNEL THEORY OF MODERN DATA AUGMENTATION | 2755 |
| <i>Tri Dao, Albert Gu, Alexander Ratner, Virginia Smith, Chris De Sa, Christopher Re</i> | |
| TARMAC: TARGETED MULTI-AGENT COMMUNICATION | 2776 |
| <i>Abhishek Das, Théophile Gervet, Joshua Romoff, Dhruv Batra, Devi Parikh, Mike Rabbat, Joelle Pineau</i> | |
| TEACHING A BLACK-BOX LEARNER | 2785 |
| <i>Sanjoy Dasgupta, Daniel Hsu, Stefanos Poulis, Xiaojin Zhu</i> | |
| STOCHASTIC DEEP NETWORKS | 2804 |
| <i>Gwendoline De Bie, Gabriel Peyré, Marco Cuturi</i> | |
| LEARNING-TO-LEARN STOCHASTIC GRADIENT DESCENT WITH BIASED REGULARIZATION | 2814 |
| <i>Giulia Denevi, Carlo Ciliberto, Riccardo Grazi, Massimiliano Pontil</i> | |
| A MULTITASK MULTIPLE KERNEL LEARNING ALGORITHM FOR SURVIVAL ANALYSIS WITH APPLICATION TO CANCER BIOLOGY | 2844 |
| <i>Onur Dereli, Ceyda Oğuz, Mehmet Gönen</i> | |

PART 5

| | |
|---|------|
| LEARNING TO CONVOLVE: A GENERALIZED WEIGHT-TYING APPROACH | 2854 |
| <i>Nichita Diaconu, Daniel Worrall</i> | |
| SEVER: A ROBUST META-ALGORITHM FOR STOCHASTIC OPTIMIZATION | 2867 |
| <i>Ilias Diakonikolas, Gautam Kamath, Daniel Kane, Jerry Li, Jacob Steinhardt, Alistair Stewart</i> | |
| APPROXIMATED ORACLE FILTER PRUNING FOR DESTRUCTIVE CNN WIDTH OPTIMIZATION | 2899 |
| <i>Xiaohan Ding, Guiguang Ding, Yuchen Guo, Jungong Han, Chenggang Yan</i> | |
| NOISY DUAL PRINCIPAL COMPONENT PURSUIT | 2910 |
| <i>Tianyu Ding, Zhihui Zhu, Tianjiao Ding, Yunchen Yang, Daniel Robinson, Manolis Tsakiris, Rene Vidal</i> | |
| FINITE-TIME ANALYSIS OF DISTRIBUTED TD(0) WITH LINEAR FUNCTION APPROXIMATION ON MULTI-AGENT REINFORCEMENT LEARNING | 2919 |
| <i>Thinh Doan, Siva Maguluri, Justin Romberg</i> | |
| TRAJECTORY-BASED OFF-POLICY DEEP REINFORCEMENT LEARNING | 2932 |
| <i>Andreas Doerr, Michael Volpp, Marc Toussaint, Trimpe Sebastian, Christian Daniel</i> | |
| GENERALIZED NO FREE LUNCH THEOREM FOR ADVERSARIAL ROBUSTNESS | 2944 |
| <i>Elvis Dohmatob</i> | |
| WIDTH PROVABLY MATTERS IN OPTIMIZATION FOR DEEP LINEAR NEURAL NETWORKS | 2956 |
| <i>Simon Du, Wei Hu</i> | |
| PROVABLY EFFICIENT RL WITH RICH OBSERVATIONS VIA LATENT STATE DECODING | 2971 |
| <i>Simon Du, Akshay Krishnamurthy, Nan Jiang, Alekh Agarwal, Miroslav Dudik, John Langford</i> | |
| GRADIENT DESCENT FINDS GLOBAL MINIMA OF DEEP NEURAL NETWORKS | 3003 |
| <i>Simon Du, Jason Lee, Haochuan Li, Liwei Wang, Xiyu Zhai</i> | |
| INCORPORATING GROUPING INFORMATION INTO BAYESIAN DECISION TREE ENSEMBLES | 3049 |
| <i>Junliang Du, Antonio Linero</i> | |
| TASK-AGNOSTIC DYNAMICS PRIORS FOR DEEP REINFORCEMENT LEARNING | 3063 |
| <i>Yilun Du, Karthic Narasimhan</i> | |
| OPTIMAL AUCTIONS THROUGH DEEP LEARNING | 3079 |
| <i>Paul Duetting, Zhe Feng, Harikrishna Narasimhan, David Parkes, Sai Srivatsa Ravindranath</i> | |
| WASSERSTEIN OF WASSERSTEIN LOSS FOR LEARNING GENERATIVE MODELS | 3099 |
| <i>Yonatan Dukler, Wuchen Li, Alex Lin, Guido Montufar</i> | |
| LEARNING INTERPRETABLE CONTINUOUS-TIME MODELS OF LATENT STOCHASTIC DYNAMICAL SYSTEMS | 3114 |
| <i>Lea Düncker, Gergo Bohner, Julien Boussard, Maneesh Sahani</i> | |
| AUTOREGRESSIVE ENERGY MACHINES | 3127 |
| <i>Conor Durkan, Charlie Nash</i> | |
| BAND-LIMITED TRAINING AND INFERENCE FOR CONVOLUTIONAL NEURAL NETWORKS | 3139 |
| <i>Adam Dziedzic, John Paparrizos, Sanjay Krishnan, Aaron Elmore, Michael Franklin</i> | |

| | |
|--|------|
| IMITATING LATENT POLICIES FROM OBSERVATION | 3156 |
| <i>Ashley Edwards, Himanshu Sahni, Yannick Schroecker, Charles Isbell</i> | |
| SEMI-CYCLIC STOCHASTIC GRADIENT DESCENT | 3165 |
| <i>Hubert Eichner, Tomer Koren, Brendan McMahan, Nathan Srebro, Kunal Talwar</i> | |
| GDPP: LEARNING DIVERSE GENERATIONS USING DETERMINANTAL POINT PROCESSES | 3178 |
| <i>Mohamed Elfeki, Camille Couprie, Morgane Riviere, Mohamed Elhoseiny</i> | |
| SEQUENTIAL FACILITY LOCATION: APPROXIMATE SUBMODULARITY AND GREEDY ALGORITHM | 3194 |
| <i>Ehsan Elhamifar</i> | |
| IMPROVED CONVERGENCE FOR L_1 AND L_∞ REGRESSION VIA ITERATIVELY REWEIGHTED LEAST SQUARES | 3204 |
| <i>Alina Ene, Adrian Vladu</i> | |
| EXPLORING THE LANDSCAPE OF SPATIAL ROBUSTNESS | 3218 |
| <i>Logan Engstrom, Brandon Tran, Dimitris Tsipras, Ludwig Schmidt, Aleksander Madry</i> | |
| CROSS-DOMAIN 3D EQUIVARIANT IMAGE EMBEDDINGS | 3239 |
| <i>Carlos Esteves, Avneesh Sud, Zhengyi Luo, Kostas Daniilidis, Ameesh Makadia</i> | |
| ON THE CONNECTION BETWEEN ADVERSARIAL ROBUSTNESS AND SALIENCY MAP INTERPRETABILITY | 3255 |
| <i>Christian Etmann, Sebastian Lunz, Peter Maass, Carola Schoenlieb</i> | |
| NON-MONOTONE SUBMODULAR MAXIMIZATION WITH NEARLY OPTIMAL ADAPTIVITY AND QUERY COMPLEXITY | 3267 |
| <i>Matthew Fahrback, Vahab Mirrokni, Morteza Zadimoghaddam</i> | |
| MULTI-FREQUENCY VECTOR DIFFUSION MAPS | 3279 |
| <i>Yijeng Fan, Zhizhen Zhao</i> | |
| STABLE-PREDICTIVE OPTIMISTIC COUNTERFACTUAL REGRET MINIMIZATION | 3289 |
| <i>Gabriele Farina, Christian Kroer, Noam Brown, Tuomas Sandholm</i> | |
| REGRET CIRCUITS: COMPOSABILITY OF REGRET MINIMIZERS | 3304 |
| <i>Gabriele Farina, Christian Kroer, Tuomas Sandholm</i> | |
| DEAD-ENDS AND SECURE EXPLORATION IN REINFORCEMENT LEARNING | 3315 |
| <i>Mehdi Fatemi, Shikhar Sharma, Harm Van Seijen, Samira Ebrahimi Kahou</i> | |
| INVARIANT-EQUIVARIANT REPRESENTATION LEARNING FOR MULTI-CLASS DATA | 3329 |
| <i>Ilya Feige</i> | |
| THE ADVANTAGES OF MULTIPLE CLASSES FOR REDUCING OVERFITTING FROM TEST SET REUSE | 3341 |
| <i>Vitaly Feldman, Roy Frostig, Moritz Hardt</i> | |
| DECENTRALIZED EXPLORATION IN MULTI-ARMED BANDITS | 3368 |
| <i>Raphael Feraud, Reda Alami, Romain Laroche</i> | |
| ALMOST SURELY CONSTRAINED CONVEX OPTIMIZATION | 3380 |
| <i>Olivier Fercoq, Ahmet Alacaoglu, Ion Necoara, Volkan Cevher</i> | |
| ONLINE META-LEARNING | 3398 |
| <i>Chelsea Finn, Aravind Rajeswaran, Sham Kakade, Sergey Levine</i> | |
| DL2: TRAINING AND QUERYING NEURAL NETWORKS WITH LOGIC | 3411 |
| <i>Marc Fischer, Mislav Balunovic, Dana Drachler-Cohen, Timon Gehr, Ce Zhang, Martin Vechev</i> | |
| BAYESIAN ACTION DECODER FOR DEEP MULTI-AGENT REINFORCEMENT LEARNING | 3428 |
| <i>Jakob Foerster, Francis Song, Edward Hughes, Neil Burch, Iain Dunning, Shimon Whiteson, Matthew Botvinick, Michael Bowling</i> | |
| SCALABLE NONPARAMETRIC SAMPLING FROM MULTIMODAL POSTERiors WITH THE POSTERIOR BOOTSTRAP | 3443 |
| <i>Edwin Fong, Simon Lyddon, Chris Holmes</i> | |
| ON DISCRIMINATIVE LEARNING OF PREDICTION UNCERTAINTY | 3465 |
| <i>Vojtech Franc, Daniel Prusa</i> | |
| LEARNING DISCRETE STRUCTURES FOR GRAPH NEURAL NETWORKS | 3481 |
| <i>Luca Franceschi, Mathias Niepert, Massimiliano Pontil, Xiao He</i> | |
| DISTRIBUTIONAL MULTIVARIATE POLICY EVALUATION AND EXPLORATION WITH THE BELLMAN GAN | 3494 |
| <i>Dror Freirich, Tzahi Shimkin, Ron Meir, Aviv Tamar</i> | |
| APPROXIMATING ORTHOGONAL MATRICES WITH EFFECTIVE GIVENS FACTORIZATION | 3509 |
| <i>Thomas Frerix, Joan Bruna</i> | |
| FAST AND FLEXIBLE INFERENCE OF JOINT DISTRIBUTIONS FROM THEIR MARGINALS | 3518 |
| <i>Charlie Frogner, Tomaso Poggio</i> | |
| ANALYZING AND IMPROVING REPRESENTATIONS WITH THE SOFT NEAREST NEIGHBOR LOSS | 3531 |
| <i>Nicholas Frosst, Nicolas Papernot, Geoffrey Hinton</i> | |

PART 6

| | |
|---|------|
| DIAGNOSING BOTTLENECKS IN DEEP Q-LEARNING ALGORITHMS | 3548 |
| <i>Justin Fu, Aviral Kumar, Matthew Soh, Sergey Levine</i> | |
| METRICGAN: GENERATIVE ADVERSARIAL NETWORKS BASED BLACK-BOX METRIC SCORES OPTIMIZATION FOR SPEECH ENHANCEMENT | 3566 |
| <i>Szu-Wei Fu, Chien-Feng Liao, Yu Tsao, Shou-De Lin</i> | |

| | |
|---|------|
| BEYOND ADAPTIVE SUBMODULARITY: APPROXIMATION GUARANTEES OF GREEDY POLICY WITH ADAPTIVE SUBMODULARITY RATIO | 3577 |
| <i>Kaito Fujii, Shinsaku Sakaue</i> | |
| OFF-POLICY DEEP REINFORCEMENT LEARNING WITHOUT EXPLORATION | 3599 |
| <i>Scott Fujimoto, David Meger, Doina Precup</i> | |
| TRANSFER LEARNING FOR RELATED REINFORCEMENT LEARNING TASKS VIA IMAGE-TO-IMAGE TRANSLATION | 3623 |
| <i>Shani Gamrian, Yoav Goldberg</i> | |
| BREAKING THE SOFTMAX BOTTLENECK VIA LEARNABLE MONOTONIC POINTWISE NON-LINEARITIES | 3635 |
| <i>Octavian Ganea, Sylvain Gelly, Gary Becigneul, Aliaksei Severyn</i> | |
| GRAPH U-NETS | 3651 |
| <i>Hongyang Gao, Shuiwang Ji</i> | |
| DEEP GENERATIVE LEARNING VIA VARIATIONAL GRADIENT FLOW | 3661 |
| <i>Yuan Gao, Yuling Jiao, Yang Wang, Yao Wang, Can Yang, Shunkang Zhang</i> | |
| RATE DISTORTION FOR MODEL COMPRESSION: FROM THEORY TO PRACTICE | 3674 |
| <i>Weihao Gao, Yu-Han Liu, Chong Wang, Sewoong Oh</i> | |
| DEMYSTIFYING DROPOUT | 3692 |
| <i>Hongchang Gao, Jian Pei, Heng Huang</i> | |
| GEOMETRIC SCATTERING FOR GRAPH DATA ANALYSIS | 3702 |
| <i>Feng Gao, Guy Wolf, Matthew Hirn</i> | |
| MULTI-FREQUENCY PHASE SYNCHRONIZATION | 3718 |
| <i>Tingran Gao, Zhizhen Zhao</i> | |
| OPTIMAL MINI-BATCH AND STEP SIZES FOR SAGA | 3734 |
| <i>Nidham Gazagnadou, Robert Gower, Joseph Salmon</i> | |
| SELECTIVENET: A DEEP NEURAL NETWORK WITH AN INTEGRATED REJECT OPTION | 3768 |
| <i>Yonatan Geifman, Ran El-Yaniv</i> | |
| A THEORY OF REGULARIZED MARKOV DECISION PROCESSES | 3777 |
| <i>Mathieu Geist, Bruno Scherrer, Olivier Pietquin</i> | |
| DEEPMDP: LEARNING CONTINUOUS LATENT SPACE MODELS FOR REPRESENTATION LEARNING | 3802 |
| <i>Carles Gelada, Saurabh Kumar, Jacob Buckman, Ofir Nachum, Marc G. Bellemare</i> | |
| PARTIALLY LINEAR ADDITIVE GAUSSIAN GRAPHICAL MODELS | 3827 |
| <i>Sinong Geng, Minhao Yan, Mladen Kolar, Sami Koyejo</i> | |
| LEARNING AND DATA SELECTION IN BIG DATASETS | 3848 |
| <i>Hossein Shokri Ghadikolaei, Hadi Ghauch, Carlo Fischione, Mikael Skoglund</i> | |
| IMPROVED PARALLEL ALGORITHMS FOR DENSITY-BASED NETWORK CLUSTERING | 3861 |
| <i>Mohsen Ghaffari, Silvio Lattanzi, Slobodan Mitrovic</i> | |
| RECURSIVE SKETCHES FOR MODULAR DEEP LEARNING | 3884 |
| <i>Badih Ghazi, Rina Panigrahy, Joshua Wang</i> | |
| AN INSTABILITY IN VARIATIONAL INFERENCE FOR TOPIC MODELS | 3936 |
| <i>Behrooz Ghorbani, Hamid Javadi, Andrea Montanari</i> | |
| AN INVESTIGATION INTO NEURAL NET OPTIMIZATION VIA HESSIAN EIGENVALUE DENSITY | 4039 |
| <i>Behrooz Ghorbani, Shankar Krishnan, Ying Xiao</i> | |
| DATA SHAPLEY: EQUITABLE VALUATION OF DATA FOR MACHINE LEARNING | 4053 |
| <i>Amirata Ghorbani, James Zou</i> | |
| EFFICIENT DICTIONARY LEARNING WITH GRADIENT DESCENT | 4066 |
| <i>Dar Gilboa, Sam Buchanan, John Wright</i> | |
| A TREE-BASED METHOD FOR FAST REPEATED SAMPLING OF DETERMINANTAL POINT PROCESSES | 4092 |
| <i>Jennifer Gillenwater, Alex Kulesza, Zelda Mariet, Sergei Vassilvtiskii</i> | |
| LEARNING TO GROOVE WITH INVERSE SEQUENCE TRANSFORMATIONS | 4104 |
| <i>Jon Gillick, Adam Roberts, Jesse Engel, Douglas Eck, David Bamman</i> | |
| ADVERSARIAL EXAMPLES ARE A NATURAL CONSEQUENCE OF TEST ERROR IN NOISE | 4115 |
| <i>Justin Gilmer, Nicolas Ford, Nicholas Carlini, Ekin Cubuk</i> | |
| DISCOVERING CONDITIONALLY SALIENT FEATURES WITH STATISTICAL GUARANTEES | 4140 |
| <i>Jaime Roquero Gimenez, James Zou</i> | |
| ESTIMATING INFORMATION FLOW IN DEEP NEURAL NETWORKS | 4153 |
| <i>Ziv Goldfeld, Ewout Van Den Berg, Kristjan Greenewald, Igor Melnyk, Nam Nguyen, Brian Kingsbury, Yury Polyanskiy</i> | |
| AMORTIZED MONTE CARLO INTEGRATION | 4163 |
| <i>Adam Golinski, Frank Wood, Tom Rainforth</i> | |
| ONLINE ALGORITHMS FOR RENT-OR-BUY WITH EXPERT ADVICE | 4181 |
| <i>Sreenivas Gollapudi, Debmalaya Panigrahi</i> | |
| THE INFORMATION-THEORETIC VALUE OF UNLABELED DATA IN SEMI-SUPERVISED LEARNING | 4190 |
| <i>Alexander Golovnev, David Pal, Balazs Szorenyi</i> | |
| EFFICIENT TRAINING OF BERT BY PROGRESSIVELY STACKING | 4202 |
| <i>Linyuan Gong, Di He, Zhuohan Li, Tao Qin, Liwei Wang, Tiejian Liu</i> | |
| QUANTILE STEIN VARIATIONAL GRADIENT DESCENT FOR BATCH BAYESIAN OPTIMIZATION | 4212 |
| <i>Chengyue Gong, Jian Peng, Qiang Liu</i> | |
| OBTAINING FAIRNESS USING OPTIMAL TRANSPORT THEORY | 4222 |
| <i>Paula Gordaliza, Eustasio Del Barrio, Gamboa Fabrice, Jean-Michel Loubes</i> | |

| | |
|--|------|
| COMBINING PARAMETRIC AND NONPARAMETRIC MODELS FOR OFF-POLICY EVALUATION | 4236 |
| <i>Omer Gottesman, Yao Liu, Scott Sussex, Emma Brunskill, Finale Doshi-Velez</i> | |

PART 7

| | |
|---|------|
| COUNTERFACTUAL VISUAL EXPLANATIONS | 4254 |
| <i>Yash Goyal, Ziyang Wu, Jan Ernst, Dhruv Batra, Devi Parikh, Stefan Lee</i> | |
| ADAPTIVE SENSOR PLACEMENT FOR CONTINUOUS SPACES | 4264 |
| <i>James Grant, Alexis Boukouvalas, Ryan-Rhys Griffiths, David Leslie, Sattar Vakili, Enrique Munoz De Cote</i> | |
| A STATISTICAL INVESTIGATION OF LONG MEMORY IN LANGUAGE AND MUSIC | 4278 |
| <i>Alexander Greaves-Tunnell, Zaid Harchaoui</i> | |
| AUTOMATIC POSTERIOR TRANSFORMATION FOR LIKELIHOOD-FREE INFERENCE | 4288 |
| <i>David Greenberg, Marcel Nonnenmacher, Jakob Macke</i> | |
| LEARNING TO OPTIMIZE MULTIGRID PDE SOLVERS | 4305 |
| <i>Daniel Greenfeld, Meirav Galun, Ronen Basri, Irad Yavneh, Ron Kimmel</i> | |
| MULTI-OBJECT REPRESENTATION LEARNING WITH ITERATIVE VARIATIONAL INFERENCE | 4317 |
| <i>Klaus Greff, Raphaël Lopez Kaufman, Rishabh Kabra, Nick Watters, Christopher Burgess, Daniel Zoran, Loic Matthey, Matthew Botvinick, Alexander Lerchner</i> | |
| GRAPHITE: ITERATIVE GENERATIVE MODELING OF GRAPHS | 4344 |
| <i>Aditya Grover, Aaron Zweig, Stefano Ermon</i> | |
| FAST ALGORITHM FOR GENERALIZED MULTINOMIAL MODELS WITH RANKING DATA | 4357 |
| <i>Jiaqi Gu, Guosheng Yin</i> | |
| TOWARDS A DEEP AND UNIFIED UNDERSTANDING OF DEEP NEURAL MODELS IN NLP | 4366 |
| <i>Chaoyu Guan, Xiting Wang, Quanshi Zhang, Runjin Chen, Di He, Xing Xie</i> | |
| AN INVESTIGATION OF MODEL-FREE PLANNING | 4376 |
| <i>Arthur Guez, Mehdi Mirza, Karol Gregor, Rishabh Kabra, Sebastien Racaniere, Theophane Weber, David Raposo, Adam Santoro, Laurent Orseau, Tom Eccles, Greg Wayne, David Silver, Timothy Lillicrap</i> | |
| HUMOR IN WORD EMBEDDINGS: COCKAMAMIE GOBBLEDEGOOK FOR NINCOMPOOPS | 4397 |
| <i>Limor Gultchin, Genevieve Patterson, Nancy Baym, Nathaniel Swinger, Adam Kalai</i> | |
| SIMPLE BLACK-BOX ADVERSARIAL ATTACKS | 4410 |
| <i>Chuan Guo, Jacob Gardner, Yurong You, Andrew Gordon Wilson, Kilian Weinberger</i> | |
| EXPLORING INTERPRETABLE LSTM NEURAL NETWORKS OVER MULTI-VARIABLE DATA | 4424 |
| <i>Tian Guo, Tao Lin, Nino Antulov-Fantulin</i> | |
| LEARNING TO EXPLOIT LONG-TERM RELATIONAL DEPENDENCIES IN KNOWLEDGE GRAPHS | 4441 |
| <i>Lingbing Guo, Zequn Sun, Wei Hu</i> | |
| MEMORY-OPTIMAL DIRECT CONVOLUTIONS FOR MAXIMIZING CLASSIFICATION ACCURACY IN EMBEDDED APPLICATIONS | 4454 |
| <i>Albert Gural, Boris Murmann</i> | |
| IMEXNET A FORWARD STABLE DEEP NEURAL NETWORK | 4473 |
| <i>Eldad Haber, Keegan Lensink, Eran Treister, Lars Ruthotto</i> | |
| ON THE POWER OF CURRICULUM LEARNING IN TRAINING DEEP NETWORKS | 4483 |
| <i>Guy Hacohen, Daphna Weinshall</i> | |
| TRADING REDUNDANCY FOR COMMUNICATION: SPEEDING UP DISTRIBUTED SGD FOR NON-CONVEX OPTIMIZATION | 4497 |
| <i>Farzin Haddadpour, Mohammad Mahdi Kamani, Mehrdad Mahdavi, Viveck Cadambe</i> | |
| LEARNING LATENT DYNAMICS FOR PLANNING FROM PIXELS | 4528 |
| <i>Danijar Hafner, Timothy Lillicrap, Ian Fischer, Ruben Villegas, David Ha, Honglak Lee, James Davidson</i> | |
| NEURAL SEPARATION OF OBSERVED AND UNOBSERVED DISTRIBUTIONS | 4548 |
| <i>Tavi Halperin, Ariel Ephrat, Yedid Hoshen</i> | |
| GRID-WISE CONTROL FOR MULTI-AGENT REINFORCEMENT LEARNING IN VIDEO GAME AI | 4558 |
| <i>Lei Han, Peng Sun, Yali Du, Jiechao Xiong, Qing Wang, Xinghai Sun, Han Liu, Tong Zhang</i> | |
| DIMENSION-WISE IMPORTANCE SAMPLING WEIGHT CLIPPING FOR SAMPLE-EFFICIENT REINFORCEMENT LEARNING | 4572 |
| <i>Seungyul Han, Youngchul Sung</i> | |
| COMPLEXITY OF LINEAR REGIONS IN DEEP NETWORKS | 4585 |
| <i>Boris Hanin, David Rolnick</i> | |
| IMPORTANCE SAMPLING POLICY EVALUATION WITH AN ESTIMATED BEHAVIOR POLICY | 4601 |
| <i>Josiah Hanna, Scott Niekum, Peter Stone</i> | |
| DOUBLY-COMPETITIVE DISTRIBUTION ESTIMATION | 4621 |
| <i>Yi Hao, Alon Orlitsky</i> | |
| RANDOM SHUFFLING BEATS SGD AFTER FINITE EPOCHS | 4651 |
| <i>Jeff Haochen, Suvrit Sra</i> | |
| SUBMODULAR MAXIMIZATION BEYOND NON-NEGATIVITY: GUARANTEES, FAST ALGORITHMS, AND APPLICATIONS | 4684 |
| <i>Chris Harshaw, Moran Feldman, Justin Ward, Amin Karbasi</i> | |
| PER-DECISION OPTION DISCOUNTING | 4706 |
| <i>Anna Harutyunyan, Peter Vrancx, Philippe Hamel, Ann Nowe, Doina Precup</i> | |
| SUBMODULAR OBSERVATION SELECTION AND INFORMATION GATHERING FOR QUADRATIC MODELS | 4719 |
| <i>Abolfazl Hashemi, Mahsa Ghasemi, Haris Vikalo, Ufuk Topcu</i> | |

| | |
|---|------|
| UNDERSTANDING AND CONTROLLING MEMORY IN RECURRENT NEURAL NETWORKS | 4733 |
| <i>Doron Haviv, Alexander Rivkind, Omri Barak</i> | |
| ON THE IMPACT OF THE ACTIVATION FUNCTION ON DEEP NEURAL NETWORKS TRAINING | 4746 |
| <i>Soufiane Hayou, Arnaud Doucet, Judith Rousseau</i> | |
| PROVABLY EFFICIENT MAXIMUM ENTROPY EXPLORATION | 4774 |
| <i>Elad Hazan, Sham Kakade, Karan Singh, Abby Van Soest</i> | |
| ON THE LONG-TERM IMPACT OF ALGORITHMIC DECISION POLICIES: EFFORT UNFAIRNESS AND FEATURE SEGREGATION THROUGH SOCIAL LEARNING | 4787 |
| <i>Hoda Heidari, Vedant Nanda, Krishna Gummadi</i> | |
| GRAPH RESISTANCE AND LEARNING FROM PAIRWISE COMPARISONS | 4800 |
| <i>Julien Hendrickx, Alexander Olshevsky, Venkatesh Saligrama</i> | |
| USING PRE-TRAINING CAN IMPROVE MODEL ROBUSTNESS AND UNCERTAINTY | 4815 |
| <i>Dan Hendrycks, Kimin Lee, Mantas Mazeika</i> | |
| FLOW++: IMPROVING FLOW-BASED GENERATIVE MODELS WITH VARIATIONAL DEQUANTIZATION AND ARCHITECTURE DESIGN | 4827 |
| <i>Jonathan Ho, Xi Chen, Aravind Srinivas, Yan Duan, Pieter Abbeel</i> | |
| POPULATION BASED AUGMENTATION: EFFICIENT LEARNING OF AUGMENTATION POLICY SCHEDULES | 4843 |
| <i>Daniel Ho, Eric Liang, Xi Chen, Ion Stoica, Pieter Abbeel</i> | |
| COLLECTIVE MODEL FUSION FOR MULTIPLE BLACK-BOX EXPERTS | 4857 |
| <i>Minh Hoang, Nghia Hoang, Bryan Kian Hsiang Low, Carleton Kingsford</i> | |
| CONNECTIVITY-OPTIMIZED REPRESENTATION LEARNING VIA PERSISTENT HOMOLOGY | 4868 |
| <i>Christoph Hofer, Roland Kwitt, Marc Niethammer, Mandar Dixit</i> | |
| BETTER GENERALIZATION WITH LESS DATA USING ROBUST GRADIENT DESCENT | 4884 |
| <i>Matthew Holland, Kazushi Ikeda</i> | |
| EMERGING CONVOLUTIONS FOR GENERATIVE NORMALIZING FLOWS | 4903 |
| <i>Emiel Hoogeboom, Rianne Van Den Berg, Max Welling</i> | |
| NONCONVEX VARIANCE REDUCED OPTIMIZATION WITH ARBITRARY SAMPLING | 4913 |
| <i>Samuel Horváth, Peter Richtarik</i> | |
| PARAMETER-EFFICIENT TRANSFER LEARNING FOR NLP | 4944 |
| <i>Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin De Laroussilhe, Andrea Gesmundo, Mona Attariyan, Sylvain Gelly</i> | |

PART 8

| | |
|--|------|
| STAY WITH ME: LIFETIME MAXIMIZATION THROUGH HETEROSCEDASTIC LINEAR BANDITS WITH RENEGING | 4957 |
| <i>Ping-Chun Hsieh, Xi Liu, Anirban Bhattacharya, P R Kumar</i> | |
| FINDING MIXED NASH EQUILIBRIA OF GENERATIVE ADVERSARIAL NETWORKS | 4972 |
| <i>Ya-Ping Hsieh, Chen Liu, Volkan Cevher</i> | |
| CLASSIFICATION FROM POSITIVE, UNLABELED AND BIASED NEGATIVE DATA | 5001 |
| <i>Yu-Guan Hsieh, Gang Niu, Masashi Sugiyama</i> | |
| BAYESIAN DECONDITIONAL KERNEL MEAN EMBEDDINGS | 5021 |
| <i>Kelvin Hsu, Fabio Ramos</i> | |
| FASTER STOCHASTIC ALTERNATING DIRECTION METHOD OF MULTIPLIERS FOR NONCONVEX OPTIMIZATION | 5043 |
| <i>Feihu Huang, Songcan Chen, Heng Huang</i> | |
| UNSUPERVISED DEEP LEARNING BY NEIGHBOURHOOD DISCOVERY | 5090 |
| <i>Jiabo Huang, Qi Dong, Shaogang Gong, Xiatian Zhu</i> | |
| DETECTING OVERLAPPING AND CORRELATED COMMUNITIES WITHOUT PURE NODES: IDENTIFIABILITY AND ALGORITHM | 5100 |
| <i>Kejun Huang, Xiao Fu</i> | |
| HIERARCHICAL IMPORTANCE WEIGHTED AUTOENCODERS | 5113 |
| <i>Chin-Wei Huang, Kris Sankaran, Eeshan Dhekane, Alexandre Lacoste, Aaron Courville</i> | |
| STABLE AND FAIR CLASSIFICATION | 5130 |
| <i>Lingxiao Huang, Nisheeth Vishnoi</i> | |
| ADDRESSING THE LOSS-METRIC MISMATCH WITH ADAPTIVE LOSS ALIGNMENT | 5145 |
| <i>Chen Huang, Shuangfei Zhai, Walter Talbott, Miguel Bautista Martin, Shih-Yu Sun, Carlos Guestrin, Josh Susskind</i> | |
| CAUSAL DISCOVERY AND FORECASTING IN NONSTATIONARY ENVIRONMENTS WITH STATE-SPACE MODELS | 5159 |
| <i>Biwei Huang, Kun Zhang, Mingming Gong, Clark Glymour</i> | |
| COMPOSING ENTROPIC POLICIES USING DIVERGENCE CORRECTION | 5183 |
| <i>Jonathan Hunt, Andre Barreto, Timothy Lillicrap, Nicolas Heess</i> | |
| HEXAGON: GENERATIVE ADVERSARIAL NETS FOR REAL WORLD CLASSIFICATION | 5206 |
| <i>Uiwon Hwang, Dahuin Jung, Sungroh Yoon</i> | |
| OVERCOMING MEAN-FIELD APPROXIMATIONS IN RECURRENT GAUSSIAN PROCESS MODELS | 5223 |
| <i>Alessandro Davide Ialongo, Mark Van Der Wilk, James Hensman, Carl Edward Rasmussen</i> | |
| LEARNING STRUCTURED DECISION PROBLEMS WITH UNAWARENESS | 5233 |
| <i>Craig Innes, Alex Lascarides</i> | |

| | |
|--|------|
| PHASE TRANSITION IN PCA WITH MISSING DATA: REDUCED SIGNAL-TO-NOISE RATIO, NOT SAMPLE SIZE! | 5248 |
| <i>Niels Ipsen, Lars Kai Hansen</i> | |
| ACTOR-ATTENTION-CRITIC FOR MULTI-AGENT REINFORCEMENT LEARNING | 5261 |
| <i>Shariq Iqbal, Fei Sha</i> | |
| COMPLEMENTARY-LABEL LEARNING FOR ARBITRARY LOSSES AND MODELS | 5275 |
| <i>Takashi Ishida, Gang Niu, Aditya Menon, Masashi Sugiyama</i> | |
| CAUSAL IDENTIFICATION UNDER MARKOV EQUIVALENCE: COMPLETENESS RESULTS | 5286 |
| <i>Amin Jaber, Jiji Zhang, Elias Bareinboim</i> | |
| LEARNING FROM A LEARNER | 5295 |
| <i>Alexis Jacq, Matthieu Geist, Ana Paiva, Olivier Pietquin</i> | |
| DIFFERENTIALLY PRIVATE FAIR LEARNING | 5306 |
| <i>Matthew Jagielski, Michael Kearns, Jieming Mao, Alina Oprea, Aaron Roth, Saeed Sharifi -Malvajerdi, Jonathan Ullman</i> | |
| SUM-OF-SQUARES POLYNOMIAL FLOW | 5335 |
| <i>Priyank Jaini, Kira A. Selby, Yaoliang Yu</i> | |
| DBSCAN++: TOWARDS FAST AND SCALABLE DENSITY CLUSTERING | 5348 |
| <i>Jennifer Jang, Heinrich Jiang</i> | |
| LEARNING WHAT AND WHERE TO TRANSFER | 5360 |
| <i>Yunhun Jang, Hankook Lee, Sung Ju Hwang, Jinwoo Shin</i> | |
| SOCIAL INFLUENCE AS INTRINSIC MOTIVATION FOR MULTI-AGENT DEEP REINFORCEMENT LEARNING | 5372 |
| <i>Natasha Jaques, Angeliki Lazaridou, Edward Hughes, Caglar Gulcehre, Pedro Ortega, Dj Strouse, Joel Z. Leibo, Nando De Freitas</i> | |
| A DEEP REINFORCEMENT LEARNING PERSPECTIVE ON INTERNET CONGESTION CONTROL | 5390 |
| <i>Nathan Jay, Noga Rotman, Brighton Godfrey, Michael Schapira, Aviv Tamar</i> | |
| GRAPH NEURAL NETWORK FOR MUSIC SCORE DATA AND MODELING EXPRESSIVE PIANO PERFORMANCE | 5400 |
| <i>Dasaem Jeong, Taegyun Kwon, Yoojin Kim, Juhan Nam</i> | |
| LADDER CAPSULE NETWORK | 5417 |
| <i>Taewon Jeong, Youngmin Lee, Heeyoung Kim</i> | |
| TRAINING CNNs WITH SELECTIVE ALLOCATION OF CHANNELS | 5426 |
| <i>Jongheon Jeong, Jinwoo Shin</i> | |
| LEARNING DISCRETE AND CONTINUOUS FACTORS OF DATA VIA ALTERNATING DISENTANGLEMENT | 5441 |
| <i>Yeonwoo Jeong, Hyun Oh Song</i> | |
| IMPROVED ZERO-TH-ORDER VARIANCE REDUCED ALGORITHMS AND ANALYSIS FOR NONCONVEX OPTIMIZATION | 5453 |
| <i>Kaiyi Ji, Zhe Wang, Yi Zhou, Yingbin Liang</i> | |
| NEURAL LOGIC REINFORCEMENT LEARNING | 5497 |
| <i>Zhengyao Jiang, Shan Luo</i> | |
| FINDING OPTIONS THAT MINIMIZE PLANNING TIME | 5509 |
| <i>Yuu Jimnai, David Abel, David Hershkowitz, Michael Littman, George Konidaris</i> | |
| DISCOVERING OPTIONS FOR EXPLORATION BY MINIMIZING COVER TIME | 5528 |
| <i>Yuu Jimnai, Jee Won Park, David Abel, George Konidaris</i> | |
| KERNEL MEAN MATCHING FOR CONTENT ADDRESSABILITY OF GANS | 5538 |
| <i>Wittawat Jitkrittum, Patsorn Sangkloy, Muhammad Waleed Gondal, Amit Raj, James Hays, Bernhard Schölkopf</i> | |
| GOODE: A GAUSSIAN OFF-THE-SHELF ORDINARY DIFFERENTIAL EQUATION SOLVER | 5551 |
| <i>David John, Vincent Heuveline, Michael Schober</i> | |
| BILINEAR BANDITS WITH LOW-RANK STRUCTURE | 5566 |
| <i>Kwang-Sung Jun, Rebecca Willett, Stephen Wright, Robert Nowak</i> | |
| STATISTICAL FOUNDATIONS OF VIRTUAL DEMOCRACY | 5580 |
| <i>Anson Kahng, Min Kyung Lee, Ritesh Noothigattu, Ariel Procaccia, Christos-Alexandros Psomas</i> | |
| MOLECULAR HYPERGRAPH GRAMMAR WITH ITS APPLICATION TO MOLECULAR OPTIMIZATION | 5590 |
| <i>Hiroshi Kajino</i> | |
| ROBUST INFLUENCE MAXIMIZATION FOR HYPERPARAMETRIC MODELS | 5606 |
| <i>Dimitris Kalimeris, Gal Kaplun, Yaron Singer</i> | |
| CLASSIFYING TREATMENT RESPONDERS UNDER CAUSAL EFFECT MONOTONICITY | 5618 |
| <i>Nathan Kallus</i> | |
| TRAINABLE DECODING OF SETS OF SEQUENCES FOR NEURAL SEQUENCE MODELS | 5628 |
| <i>Ashwin Kalyan, Peter Anderson, Stefan Lee, Dhruv Batra</i> | |
| MYOPIC POSTERIOR SAMPLING FOR ADAPTIVE GOAL ORIENTED DESIGN OF EXPERIMENTS | 5640 |
| <i>Kirthevasan Kandasamy, Willie Neiswanger, Reed Zhang, Akshay Krishnamurthy, Jeff Schneider, Barnabas Poczos</i> | |
| DIFFERENTIALLY PRIVATE LEARNING OF GEOMETRIC CONCEPTS | 5659 |
| <i>Haim Kaplan, Yishay Mansour, Yossi Matias, Uri Stemmer</i> | |
| POLICY CONSOLIDATION FOR CONTINUAL REINFORCEMENT LEARNING | 5668 |
| <i>Christos Kaplanis, Murray Shanahan, Claudia Clopath</i> | |

PART 9

| | |
|--|------|
| ERROR FEEDBACK FIXES SIGNSGD AND OTHER GRADIENT COMPRESSION SCHEMES | 5681 |
| <i>Sai Praneeth Karimireddy, Quentin Rebjock, Sebastian Stich, Martin Jaggi</i> | |
| RIEMANNIAN ADAPTIVE STOCHASTIC GRADIENT ALGORITHMS ON MATRIX MANIFOLDS | 5699 |
| <i>Hiroyuki Kasai, Pratik Jawanpuria, Bamdev Mishra</i> | |
| NEURAL INVERSE KNITTING: FROM IMAGES TO MANUFACTURING INSTRUCTIONS | 5709 |
| <i>Alexandre Kaspar, Tae-Hyun Oh, Liane Makatura, Petr Kellnhofer, Wojciech Matusik</i> | |
| PROCESSING MEGAPIXEL IMAGES WITH DEEP ATTENTION-SAMPLING MODELS | 5724 |
| <i>Angelos Katharopoulos, Francois Fleuret</i> | |
| ROBUST ESTIMATION OF TREE STRUCTURED GAUSSIAN GRAPHICAL MODELS | 5742 |
| <i>Ashish Katiyar, Jessica Hoffmann, Constantine Caramanis</i> | |
| SHALLOW-DEEP NETWORKS: UNDERSTANDING AND MITIGATING NETWORK OVERTHINKING | N/A |
| <i>Yigitcan Kaya, Sanghyun Hong, Tudor Dumitras</i> | |
| SUBMODULAR STREAMING IN ALL ITS GLORY: TIGHT APPROXIMATION, MINIMUM MEMORY AND LOW ADAPTIVE COMPLEXITY | 5767 |
| <i>Ehsan Kazemi, Marko Mitrovic, Morteza Zadimoghaddam, Silvio Lattanzi, Amin Karbasi</i> | |
| ADAPTIVE SCALE-INVARIANT ONLINE ALGORITHMS FOR LEARNING LINEAR MODELS | 5785 |
| <i>Michal Kempka, Wojciech Kotlowski, Manfred K. Warmuth</i> | |
| CHIVE: VARYING PROSODY IN SPEECH SYNTHESIS WITH A LINGUISTICALLY DRIVEN DYNAMIC HIERARCHICAL CONDITIONAL VARIATIONAL NETWORK | 5806 |
| <i>Tom Kenter, Vincent Wan, Chun-An Chan, Rob Clark, Jakub Vit</i> | |
| COLLABORATIVE EVOLUTIONARY REINFORCEMENT LEARNING | 5816 |
| <i>Shauharda Khadka, Somdeb Majumdar, Tarek Nassar, Zach Dwiell, Evren Tumer, Santiago Miret, Yinyin Liu, Kagan Tumer</i> | |
| GEOMETRY AWARE CONVOLUTIONAL FILTERS FOR OMNIDIRECTIONAL IMAGES REPRESENTATION | 5828 |
| <i>Renata Khasanova, Pascal Frossard</i> | |
| EMI: EXPLORATION WITH MUTUAL INFORMATION | 5837 |
| <i>Hyungseok Kim, Jaekyeom Kim, Yeonwoo Jeong, Sergey Levine, Hyun Oh Song</i> | |
| FLOWAVENET : A GENERATIVE FLOW FOR RAW AUDIO | 5852 |
| <i>Sungwon Kim, Sang-Gil Lee, Jongyoon Song, Jaehyeon Kim, Sungroh Yoon</i> | |
| CURIOSITY-BOTTLENECK: EXPLORATION BY DISTILLING TASK-SPECIFIC NOVELTY | 5861 |
| <i>Youngjin Kim, Wontae Nam, Hyunwoo Kim, Ji-Hoon Kim, Gunhee Kim</i> | |
| CONTEXTUAL MULTI-ARMED BANDIT ALGORITHM FOR SEMIPARAMETRIC REWARD MODEL | 5875 |
| <i>Gi-Soo Kim, Myunghee Cho Paik</i> | |
| UNIFORM CONVERGENCE RATE OF THE KERNEL DENSITY ESTIMATOR ADAPTIVE TO INTRINSIC VOLUME DIMENSION | 5890 |
| <i>Jisu Kim, Jaehyeok Shin, Alessandro Rinaldo, Larry Wasserman</i> | |
| BIT-SWAP: RECURSIVE BITS-BACK CODING FOR LOSSLESS COMPRESSION WITH HIERARCHICAL LATENT VARIABLES | 5925 |
| <i>Friso Kingma, Pieter Abbeel, Jonathan Ho</i> | |
| COMPILE: COMPOSITIONAL IMITATION LEARNING AND EXECUTION | 5941 |
| <i>Thomas Kipf, Yujia Li, Hanjun Dai, Vinicius Zambaldi, Alvaro Sanchez-Gonzalez, Edward Grefenstette, Pushmeet Kohli, Peter Battaglia</i> | |
| ADAPTIVE AND SAFE BAYESIAN OPTIMIZATION IN HIGH DIMENSIONS VIA ONE-DIMENSIONAL SUBSPACES | 5959 |
| <i>Johannes Kirschner, Mojmir Mutny, Nicole Hiller, Rasmus Ischebeck, Andreas Krause</i> | |
| AUC_μ: A PERFORMANCE METRIC FOR MULTI-CLASS MACHINE LEARNING MODELS | 5972 |
| <i>Ross Kleiman, David Page</i> | |
| FAIR K-CENTER CLUSTERING FOR DATA SUMMARIZATION | 5984 |
| <i>Mathäus Kleindessner, Pranjal Awasthi, Jamie Morgenstern</i> | |
| GUARANTEES FOR SPECTRAL CLUSTERING WITH FAIRNESS CONSTRAINTS | 6004 |
| <i>Mathäus Kleindessner, Samira Samadi, Pranjal Awasthi, Jamie Morgenstern</i> | |
| POPORN: QUANTIFYING ROBUSTNESS OF RECURRENT NEURAL NETWORKS | 6031 |
| <i>Ching-Yun Ko, Zhaoyang Lyu, Lily Weng, Luca Daniel, Ngai Wong, Dahua Lin</i> | |
| DECENTRALIZED STOCHASTIC OPTIMIZATION AND GOSSIP ALGORITHMS WITH COMPRESSED COMMUNICATION | 6088 |
| <i>Anastasia Koloskova, Sebastian Stich, Martin Jaggi</i> | |
| ROBUST LEARNING FROM UNTRUSTED SOURCES | 6112 |
| <i>Nikola Konstantinov, Christoph Lampert</i> | |
| STOCHASTIC BEAMS AND WHERE TO FIND THEM: THE GUMBEL-TOP-K TRICK FOR SAMPLING SEQUENCES WITHOUT REPLACEMENT | 6133 |
| <i>Wouter Kool, Herke Van Hoof, Max Welling</i> | |
| LIT: LEARNED INTERMEDIATE REPRESENTATION TRAINING FOR MODEL COMPRESSION | 6146 |
| <i>Animesh Koratana, Daniel Kang, Peter Bailis, Matei Zaharia</i> | |
| SIMILARITY OF NEURAL NETWORK REPRESENTATIONS REVISITED | 6156 |
| <i>Simon Kornblith, Mohammad Norouzi, Honglak Lee, Geoffrey Hinton</i> | |
| ON THE COMPLEXITY OF APPROXIMATING WASSERSTEIN BARYCENTERS | 6176 |
| <i>Alexey Kroshnin, Nazarii Tupitsa, Darina Dvinskikh, Pavel Dvurechensky, Alexander Gasnikov, Cesar Uribe</i> | |

| | |
|--|------|
| ESTIMATE SEQUENCES FOR VARIANCE-REDUCED STOCHASTIC COMPOSITE OPTIMIZATION | 6209 |
| <i>Andrei Kulunchakov, Julien Mairal</i> | |
| FASTER ALGORITHMS FOR BINARY MATRIX FACTORIZATION | 6233 |
| <i>Ravi Kumar, Rina Panigrahy, Ali Rahimi, David Woodruff</i> | |
| LOSS LANDSCAPES OF REGULARIZED LINEAR AUTOENCODERS | 6245 |
| <i>Daniel Kumin, Jonathan Bloom, Aleksandrina Goeva, Cotton Seed</i> | |
| GEOMETRY AND SYMMETRY IN SHORT-AND-SPARSE DECONVOLUTION | 6259 |
| <i>Han-Wen Kuo, Yenson Lau, Yuqian Zhang, John Wright</i> | |
| A LARGE-SCALE STUDY ON REGULARIZATION AND NORMALIZATION IN GANS | 6350 |
| <i>Karol Kurach, Mario Lucic, Xiaohua Zhai, Marcin Michalski, Sylvain Gelly</i> | |

PART 10

| | |
|--|------|
| MAKING DECISIONS THAT REDUCE DISCRIMINATORY IMPACTS | 6368 |
| <i>Matt Kusner, Chris Russell, Joshua Loftus, Ricardo Silva</i> | |
| GARBAGE IN, REWARD OUT: BOOTSTRAPPING EXPLORATION IN MULTI-ARMED BANDITS | 6380 |
| <i>Branislav Kveton, Csaba Szepesvari, Sharan Vaswani, Zheng Wen, Tor Lattimore, Mohammad Ghavamzadeh</i> | |
| CHARACTERIZING WELL-BEHAVED VS. PATHOLOGICAL DEEP NEURAL NETWORKS | 6396 |
| <i>Antoine Labatie</i> | |
| STATE-REIFICATION NETWORKS: IMPROVING GENERALIZATION BY MODELING THE DISTRIBUTION OF HIDDEN REPRESENTATIONS | 6446 |
| <i>Alex Lamb, Jonathan Binas, Anirudh Goyal, Sandeep Subramanian, Ioannis Mitliagkas, Yoshua Bengio, Michael Mozer</i> | |
| A RECURRENT NEURAL CASCADE-BASED MODEL FOR CONTINUOUS-TIME DIFFUSION | 6456 |
| <i>Sylvain Lamprier</i> | |
| PROJECTION ONTO MINKOWSKI SUMS WITH APPLICATION TO CONSTRAINED LEARNING | 6470 |
| <i>Joong-Ho Won, Jason Xu, Kenneth Lange</i> | |
| SAFE POLICY IMPROVEMENT WITH BASELINE BOOTSTRAPPING | 6487 |
| <i>Romain Laroche, Paul Trichelair, Remi Tachet Des Combes</i> | |
| A BETTER K-MEANS++ ALGORITHM VIA LOCAL SEARCH | 6521 |
| <i>Silvio Lattanzi, Christian Sohler</i> | |
| LORENTZIAN DISTANCE LEARNING FOR HYPERBOLIC REPRESENTATIONS | 6531 |
| <i>Marc Law, Renjie Liao, Jake Snell, Richard Zemel</i> | |
| DP-GP-LVM: A BAYESIAN NON-PARAMETRIC MODEL FOR LEARNING MULTIVARIATE DEPENDENCY STRUCTURES | 6550 |
| <i>Andrew Lawrence, Carl Henrik Ek, Neill Campbell</i> | |
| POLITEX: REGRET BOUNDS FOR POLICY ITERATION USING EXPERT PREDICTION | 6570 |
| <i>Yasin Abbasi-Yadkori, Peter Bartlett, Kush Bhatia, Nevena Lazic, Csaba Szepesvari, Gellert Weisz</i> | |
| BATCH POLICY LEARNING UNDER CONSTRAINTS | 6589 |
| <i>Hoang Le, Cameron Voloshin, Yisong Yue</i> | |
| TARGET-BASED TEMPORAL-DIFFERENCE LEARNING | 6619 |
| <i>Donghwan Lee, Niao He</i> | |
| FUNCTIONAL TRANSPARENCY FOR STRUCTURED DATA: A GAME-THEORETIC APPROACH | 6646 |
| <i>Guang-He Lee, Wengong Jin, David Alvarez-Melis, Tommi Jaakkola</i> | |
| SELF-ATTENTION GRAPH POOLING | 6661 |
| <i>Junhyun Lee, Inyeop Lee, Jaewoo Kang</i> | |
| SET TRANSFORMER: A FRAMEWORK FOR ATTENTION-BASED PERMUTATION-INVARIANT NEURAL NETWORKS | 6671 |
| <i>Juho Lee, Yoonho Lee, Jungtaek Kim, Adam Kosiorek, Seungjin Choi, Yee Whye Teh</i> | |
| FIRST-ORDER ALGORITHMS CONVERGE FASTER THAN $O(1/K)$ ON CONVEX PROBLEMS | 6678 |
| <i>Ching-Pei Lee, Stephen Wright</i> | |
| ROBUST INFERENCE VIA GENERATIVE CLASSIFIERS FOR HANDLING NOISY LABELS | 6688 |
| <i>Kimin Lee, Sukmin Yun, Kibok Lee, Honglak Lee, Bo Li, Jinwoo Shin</i> | |
| SUBLINEAR TIME NEAREST NEIGHBOR SEARCH OVER GENERALIZED WEIGHTED SPACE | 6705 |
| <i>Yifan Lei, Qiang Huang, Mohan Kankanhalli, Anthony Tung</i> | |
| MONK OUTLIER-ROBUST MEAN EMBEDDING ESTIMATION BY MEDIAN-OF-MEANS | 6717 |
| <i>Mathieu Lerasle, Zoltan Szabo, Timothée Mathieu, Guillaume Lecue</i> | |
| CHEAP ORTHOGONAL CONSTRAINTS IN NEURAL NETWORKS: A SIMPLE PARAMETRIZATION OF THE ORTHOGONAL AND UNITARY GROUP | 6734 |
| <i>Mario Lezcano-Casado, David Marti'nez-Rubio</i> | |
| ARE GENERATIVE CLASSIFIERS MORE ROBUST TO ADVERSARIAL ATTACKS? | 6754 |
| <i>Yingzhen Li, John Bradshaw, Yash Sharma</i> | |
| SUBLINEAR QUANTUM ALGORITHMS FOR TRAINING LINEAR AND KERNEL-BASED CLASSIFIERS | 6784 |
| <i>Tongyang Li, Shouvanik Chakrabarti, Xiaodi Wu</i> | |
| LGM-NET: LEARNING TO GENERATE MATCHING NETWORKS FOR FEW-SHOT LEARNING | 6805 |
| <i>Huaiyu Li, Weiming Dong, Xing Mei, Chongyang Ma, Feiyue Huang, Bao-Gang Hu</i> | |
| GRAPH MATCHING NETWORKS FOR LEARNING THE SIMILARITY OF GRAPH STRUCTURED OBJECTS | 6815 |
| <i>Yujia Li, Chenjie Gu, Thomas Dullien, Oriol Vinyals, Pushmeet Kohli</i> | |

| | |
|---|------|
| AREA ATTENTION | 6833 |
| <i>Yang Li, Lukasz Kaiser, Samy Bengio, Si Si</i> | |
| ONLINE LEARNING TO RANK WITH FEATURES | 6847 |
| <i>Shuai Li, Tor Lattimore, Csaba Szepesvari</i> | |
| NATTACK: LEARNING THE DISTRIBUTIONS OF ADVERSARIAL EXAMPLES FOR AN IMPROVED BLACK-BOX ATTACK ON DEEP NEURAL NETWORKS | 6860 |
| <i>Yandong Li, Lijun Li, Liqiang Wang, Tong Zhang, Boqing Gong</i> | |
| BAYESIAN JOINT SPIKE-AND-SLAB GRAPHICAL LASSO | 6874 |
| <i>Zehang Li, Tyler McCormick, Samuel Clark</i> | |
| EXPLOITING WORKER CORRELATION FOR LABEL AGGREGATION IN CROWDSOURCING | 6897 |
| <i>Yuan Li, Benjamin Rubinstein, Trevor Cohn</i> | |
| ADVERSARIAL CAMERA STICKERS: A PHYSICAL CAMERA-BASED ATTACK ON DEEP LEARNING SYSTEMS | 6907 |
| <i>Juncheng Li, Frank Schmidt, Zico Kolter</i> | |
| TOWARDS A UNIFIED ANALYSIS OF RANDOM FOURIER FEATURES | 6916 |
| <i>Zhu Li, Jean-Francois Ton, Dino Oglic, Dino Sejdinovic</i> | |
| FEATURE-CRITIC NETWORKS FOR HETEROGENEOUS DOMAIN GENERALIZATION | N/A |
| <i>Yiyi Li, Yongxin Yang, Wei Zhou, Timothy Hospedales</i> | |
| LEARN TO GROW: A CONTINUAL STRUCTURE LEARNING FRAMEWORK FOR OVERCOMING CATASTROPHIC FORGETTING | 6937 |
| <i>Xilai Li, Yingbo Zhou, Tianfu Wu, Richard Socher, Caiming Xiong</i> | |
| ALTERNATING MINIMIZATIONS CONVERGE TO SECOND-ORDER OPTIMAL SOLUTIONS | 6949 |
| <i>Qiuwei Li, Zhihui Zhu, Gongguo Tang</i> | |
| CAUTIOUS REGRET MINIMIZATION: ONLINE OPTIMIZATION WITH LONG-TERM BUDGET CONSTRAINTS | 6958 |
| <i>Nikolaos Liakopoulos, Apostolos Destounis, Georgios Paschos, Thrasyvoulos Spyropoulos, Panayotis Mertikopoulos</i> | |
| REGULARIZATION IN DIRECTABLE ENVIRONMENTS WITH APPLICATION TO TETRIS | 6974 |
| <i>Jan Malte Lichtenberg, Özgür Simsek</i> | |
| INFERENCE AND SAMPLING OF K_{33}-FREE ISING MODELS | 6996 |
| <i>Valerii Likhoshesterov, Yury Maximov, Misha Chertkov</i> | |
| KERNEL-BASED REINFORCEMENT LEARNING IN ROBUST MARKOV DECISION PROCESSES | 7015 |
| <i>Shiau Hong Lim, Arnaud Autef</i> | |
| ON EFFICIENT OPTIMAL TRANSPORT: AN ANALYSIS OF GREEDY AND ACCELERATED MIRROR DESCENT ALGORITHMS | 7024 |
| <i>Tianyi Lin, Nhat Ho, Michael Jordan</i> | |

PART 11

| | |
|--|------|
| FAST AND SIMPLE NATURAL-GRADIENT VARIATIONAL INFERENCE WITH MIXTURE OF EXPONENTIAL-FAMILY APPROXIMATIONS | 7044 |
| <i>Wu Lin, Mohammad Emtiyaz Khan, Mark Schmidt</i> | |
| ACCELERATION OF SVRG AND KATYUSHA X BY INEXACT PRECONDITIONING | 7085 |
| <i>Yanli Liu, Fei Feng, Wotao Yin</i> | |
| TRANSFERABLE ADVERSARIAL TRAINING: A GENERAL APPROACH TO ADAPTING DEEP CLASSIFIERS | 7103 |
| <i>Hong Liu, Mingsheng Long, Jianmin Wang, Michael Jordan</i> | |
| RAO-BLACKWELLIZED STOCHASTIC GRADIENTS FOR DISCRETE DISTRIBUTIONS | 7113 |
| <i>Runjing Liu, Jeffrey Regier, Nilesh Tripuraneni, Michael Jordan, Jon McAuliffe</i> | |
| SPARSE EXTREME MULTI-LABEL LEARNING WITH ORACLE PROPERTY | 7124 |
| <i>Weiwei Liu, Xiaobo Shen</i> | |
| DATA POISONING ATTACKS ON STOCHASTIC BANDITS | 7134 |
| <i>Fang Liu, Ness Shroff</i> | |
| THE IMPLICIT FAIRNESS CRITERION OF UNCONSTRAINED LEARNING | 7146 |
| <i>Lydia T. Liu, Max Simchowitz, Moritz Hardt</i> | |
| TAMING MAML: EFFICIENT UNBIASED META-REINFORCEMENT LEARNING | 7156 |
| <i>Hao Liu, Richard Socher, Caiming Xiong</i> | |
| ON CERTIFYING NON-UNIFORM BOUNDS AGAINST ADVERSARIAL ATTACKS | 7170 |
| <i>Chen Liu, Ryota Tomioka, Volkan Cevher</i> | |
| UNDERSTANDING AND ACCELERATING PARTICLE-BASED VARIATIONAL INFERENCE | 7187 |
| <i>Chang Liu, Jingwei Zhuo, Pengyu Cheng, Ruiyi Zhang, Jun Zhu</i> | |
| UNDERSTANDING MCMC DYNAMICS AS FLOWS ON THE WASSERSTEIN SPACE | 7206 |
| <i>Chang Liu, Jingwei Zhuo, Jun Zhu</i> | |
| SLICED-WASSERSTEIN FLOWS: NONPARAMETRIC GENERATIVE MODELING VIA OPTIMAL TRANSPORT AND DIFFUSIONS | 7221 |
| <i>Antoine Liutkus, Umur Simsekli, Szymon Majewski, Alain Durmus, Fabian-Robert Stöter</i> | |
| CHALLENGING COMMON ASSUMPTIONS IN THE UNSUPERVISED LEARNING OF DISENTANGLED REPRESENTATIONS | 7247 |
| <i>Francesco Locatello, Stefan Bauer, Mario Lucic, Gunnar Raetsch, Sylvain Gelly, Bernhard Schölkopf, Olivier Bachem</i> | |

| | |
|--|------|
| BAYESIAN COUNTERFACTUAL RISK MINIMIZATION | 7284 |
| <i>Ben London, Ted Sandler</i> | |
| PA-GD: ON THE CONVERGENCE OF PERTURBED ALTERNATING GRADIENT DESCENT TO SECOND-ORDER STATIONARY POINTS FOR STRUCTURED NONCONVEX OPTIMIZATION | 7301 |
| <i>Songtao Lu, Mingyi Hong, Zhengdao Wang</i> | |
| NEURALLY-GUIDED STRUCTURE INFERENCE | 7328 |
| <i>Sidi Lu, Jiayuan Mao, Joshua Tenenbaum, Jiajun Wu</i> | |
| OPTIMAL ALGORITHMS FOR LIPSCHITZ BANDITS WITH HEAVY-TAILED REWARDS | 7339 |
| <i>Shiyin Lu, Guanghui Wang, Yao Hu, Lijun Zhang</i> | |
| COT: COOPERATIVE TRAINING FOR GENERATIVE MODELING OF DISCRETE DATA | 7358 |
| <i>Sidi Lu, Lantao Yu, Siyuan Feng, Yaoming Zhu, Weinan Zhang</i> | |
| GENERALIZED APPROXIMATE SURVEY PROPAGATION FOR HIGH-DIMENSIONAL ESTIMATION | 7370 |
| <i>Carlo Lucibello, Luca Saglietti, Yue Lu</i> | |
| HIGH-FIDELITY IMAGE GENERATION WITH FEWER LABELS | 7392 |
| <i>Mario Lucic, Michael Tschamen, Marvin Ritter, Xiaohua Zhai, Olivier Bachem, Sylvain Gelly</i> | |
| LEVERAGING LOW-RANK RELATIONS BETWEEN SURROGATE TASKS IN STRUCTURED PREDICTION | 7415 |
| <i>Giulia Luise, Dimitrios Stamos, Massimiliano Pontil, Carlo Ciliberto</i> | |
| DIFFERENTIABLE DYNAMIC NORMALIZATION FOR LEARNING DEEP REPRESENTATION | 7445 |
| <i>Ping Luo, Peng Zhanglin, Shao Wenqi, Zhang Ruimao, Ren Jiamin, Wu Lingyun</i> | |
| DISENTANGLED GRAPH CONVOLUTIONAL NETWORKS | 7454 |
| <i>Jianxin Ma, Peng Cui, Kun Kuang, Xin Wang, Wenwu Zhu</i> | |
| VARIATIONAL IMPLICIT PROCESSES | 7464 |
| <i>Chao Ma, Yingzhen Li, Jose Miguel Hernandez-Lobato</i> | |
| EDDI: EFFICIENT DYNAMIC DISCOVERY OF HIGH-VALUE INFORMATION WITH PARTIAL VAE | 7483 |
| <i>Chao Ma, Sebastian Tschiatschek, Konstantina Palla, Jose Miguel Hernandez-Lobato, Sebastian Nowozin, Cheng Zhang</i> | |
| BAYESIAN LEAVE-ONE-OUT CROSS-VALIDATION FOR LARGE DATA | 7505 |
| <i>Måns Magnusson, Michael Andersen, Johan Jonasson, Aki Vehtari</i> | |
| COMPOSABLE CORE-SETS FOR DETERMINANT MAXIMIZATION: A SIMPLE NEAR-OPTIMAL ALGORITHM | 7526 |
| <i>Sepeideh Mahabadi, Piotr Indyk, Shayan Oveis Gharan, Alireza Rezaei</i> | |
| GUIDED EVOLUTIONARY STRATEGIES: AUGMENTING RANDOM SEARCH WITH SURROGATE GRADIENTS | 7538 |
| <i>Niru Maheswaranathan, Luke Metz, George Tucker, Dami Choi, Jascha Sohl-Dickstein</i> | |
| DATA POISONING ATTACKS IN MULTI-PARTY LEARNING | 7553 |
| <i>Saeed Mahloujifar, Mohammad Mahmoody, Ameer Mohammed</i> | |
| TRADITIONAL AND HEAVY TAILED SELF REGULARIZATION IN NEURAL NETWORK MODELS | 7563 |
| <i>Michael Mahoney, Charles Martin</i> | |
| CURVATURE-EXPLOITING ACCELERATION OF ELASTIC NET COMPUTATIONS | 7573 |
| <i>Vien Mai, Mikael Johansson</i> | |
| BREAKING THE GRIDLOCK IN MIXTURE-OF-EXPERTS: CONSISTENT AND EFFICIENT ALGORITHMS | 7595 |
| <i>Ashok Makkua, Pramod Viswanath, Sreeram Kannan, Sewoong Oh</i> | |
| CALIBRATED MODEL-BASED DEEP REINFORCEMENT LEARNING | 7624 |
| <i>Ali Malik, Volodymyr Kuleshov, Jiaming Song, Danny Nemer, Harlan Seymour, Stefano Ermon</i> | |
| LEARNING FROM DELAYED OUTCOMES VIA PROXIES WITH APPLICATIONS TO RECOMMENDER SYSTEMS | 7638 |
| <i>Timothy Arthur Mann, Sven Gowal, Andras Gyorgy, Huiyi Hu, Ray Jiang, Balaji Lakshminarayanan, Prav Srinivasan</i> | |
| PASSED & SPURIOUS: DESCENT ALGORITHMS AND LOCAL MINIMA IN SPIKED MATRIX-TENSOR MODELS | 7651 |
| <i>Stefano Sarao Mannelli, Florent Krzakala, Pierfrancesco Urbani, Lenka Zdeborova</i> | |
| A BASELINE FOR ANY ORDER GRADIENT ESTIMATION IN STOCHASTIC COMPUTATION GRAPHS | 7665 |
| <i>Jingkai Mao, Jakob Foerster, Tim Rocktäschel, Maruan Al-Shedivat, Gregory Farquhar, Shimon Whiteson</i> | |
| ADVERSARIAL GENERATION OF TIME-FREQUENCY FEATURES WITH APPLICATION IN AUDIO SYNTHESIS | 7680 |
| <i>Andrés Marafioti, Nathanaël Perraudin, Nicki Holighaus, Piotr Majdak</i> | |
| ON THE UNIVERSALITY OF INVARIANT NETWORKS | 7693 |
| <i>Haggai Maron, Ethan Fetaya, Nimrod Segol, Yaron Lipman</i> | |
| DECOMPOSING FEATURE-LEVEL VARIATION WITH COVARIATE GAUSSIAN PROCESS LATENT VARIABLE MODELS | 7703 |
| <i>Kaspar Märtens, Kieran Campbell, Christopher Yau</i> | |
| FAIRNESS-AWARE LEARNING FOR CONTINUOUS ATTRIBUTES AND TREATMENTS | 7719 |
| <i>Jeremie Mary, Clément Calauzènes, Noureddine El Karoui</i> | |

PART 12

| | |
|---|------|
| OPTIMAL MINIMAL MARGIN MAXIMIZATION WITH BOOSTING | 7734 |
| <i>Alexander Mathiasen, Kasper Green Larsen, Allan Grønlund</i> | |

| | |
|--|------|
| DISENTANGLING DISENTANGLEMENT IN VARIATIONAL AUTOENCODERS | 7744 |
| <i>Emile Mathieu, Tom Rainforth, N Siddharth, Yee Whye Teh</i> | |
| MIWAE: DEEP GENERATIVE MODELLING AND IMPUTATION OF INCOMPLETE DATA SETS | 7762 |
| <i>Pierre-Alexandre Mattei, Jes Frelsen</i> | |
| DISTRIBUTIONAL REINFORCEMENT LEARNING FOR EFFICIENT EXPLORATION | 7775 |
| <i>Borislav Mavrin, Hengshuai Yao, Linglong Kong, Kaiwen Wu, Yaoliang Yu</i> | |
| GRAPHICAL-MODEL BASED ESTIMATION AND INFERENCE FOR DIFFERENTIAL PRIVACY | 7786 |
| <i>Ryan McKenna, Daniel Sheldon, Jerome Miklau</i> | |
| EFFICIENT AMORTISED BAYESIAN INFERENCE FOR HIERARCHICAL AND NONLINEAR DYNAMICAL SYSTEMS | 7800 |
| <i>Ted Meeds, Geoffrey Roeder, Paul Grant, Andrew Phillips, Neil Dalchau</i> | |
| TOWARD CONTROLLING DISCRIMINATION IN ONLINE AD AUCTIONS | 7827 |
| <i>Elisa Celis, Anay Mehrotra, Nisheeth Vishnoi</i> | |
| STOCHASTIC BLOCKMODELS MEET GRAPH NEURAL NETWORKS | 7849 |
| <i>Nikhil Mehta, Lawrence Carin Duke, Piyush Rai</i> | |
| IMPUTING MISSING EVENTS IN CONTINUOUS-TIME EVENT STREAMS | 7861 |
| <i>Hongyuan Mei, Guanghui Qin, Jason Eisner</i> | |
| SAME, SAME BUT DIFFERENT: RECOVERING NEURAL NETWORK QUANTIZATION ERROR THROUGH WEIGHT FACTORIZATION | 7886 |
| <i>Eldad Meller, Alexander Finkelstein, Uri Almog, Mark Grobman</i> | |
| THE WASSERSTEIN TRANSFORM | 7897 |
| <i>Facundo Memoli, Zane Smith, Zhengchao Wan</i> | |
| ITHEMAL: ACCURATE, PORTABLE AND FAST BASIC BLOCK THROUGHPUT ESTIMATION USING DEEP NEURAL NETWORKS | 7908 |
| <i>Charith Mendis, Alex Renda, Dr. Saman Amarasinghe, Michael Carbin</i> | |
| GEOMETRIC LOSSES FOR DISTRIBUTIONAL LEARNING | 7923 |
| <i>Arthur Mensch, Mathieu Blondel, Gabriel Peyré</i> | |
| SPECTRAL CLUSTERING OF SIGNED GRAPHS VIA MATRIX POWER MEANS | 7940 |
| <i>Pedro Mercado, Francesco Tudisco, Matthias Hein</i> | |
| SIMPLE STOCHASTIC GRADIENT METHODS FOR NON-SMOOTH NON-CONVEX REGULARIZED OPTIMIZATION | 7968 |
| <i>Michael Metel, Akiko Takeda</i> | |
| REINFORCEMENT LEARNING IN CONFIGURABLE CONTINUOUS ENVIRONMENTS | 7984 |
| <i>Alberto Maria Metelli, Emanuele Ghelfi, Marcello Restelli</i> | |
| UNDERSTANDING AND CORRECTING PATHOLOGIES IN THE TRAINING OF LEARNED OPTIMIZERS | 8016 |
| <i>Luke Metz, Niru Maheswaranathan, Jeremy Nixon, Daniel Freeman, Jascha Sohl-Dickstein</i> | |
| OPTIMALITY IMPLIES KERNEL SUM CLASSIFIERS ARE STATISTICALLY EFFICIENT | 8036 |
| <i>Raphael Meyer, Jean Honorio</i> | |
| ON DROPOUT AND NUCLEAR NORM REGULARIZATION | 8052 |
| <i>Poorya Mianjy, Raman Arora</i> | |
| DISCRIMINATIVE REGULARIZATION FOR LATENT VARIABLE MODELS WITH APPLICATIONS TO ELECTROCARDIOGRAPHY | 8072 |
| <i>Andrew Miller, Ziad Obermeyer, John Cunningham, Sendhil Mullainathan</i> | |
| FORMAL PRIVACY FOR FUNCTIONAL DATA WITH GAUSSIAN PERTURBATIONS | 8082 |
| <i>Ardalan Mirshani, Matthew Reimherr, Aleksandra Slavkovic</i> | |
| CO-MANIFOLD LEARNING WITH MISSING DATA | 8097 |
| <i>Gal Mishne, Eric Chi, Ronald Coifman</i> | |
| AGNOSTIC FEDERATED LEARNING | 8114 |
| <i>Mehryar Mohri, Gary Sivek, Ananda Theertha Suresh</i> | |
| FLAT METRIC MINIMIZATION WITH APPLICATIONS IN GENERATIVE MODELING | 8137 |
| <i>Thomas Möllenhoff, Daniel Cremers</i> | |
| PARSIMONIOUS BLACK-BOX ADVERSARIAL ATTACKS VIA EFFICIENT COMBINATORIAL OPTIMIZATION | 8149 |
| <i>Seungyong Moon, Gaon An, Hyun Oh Song</i> | |
| PARAMETER EFFICIENT TRAINING OF DEEP CONVOLUTIONAL NEURAL NETWORKS BY DYNAMIC SPARSE REPARAMETERIZATION | 8163 |
| <i>Hesham Mostafa, Xin Wang</i> | |
| A DYNAMICAL SYSTEMS PERSPECTIVE ON NESTEROV ACCELERATION | 8181 |
| <i>Michael Muehlebach, Michael Jordan</i> | |
| RELATIONAL POOLING FOR GRAPH REPRESENTATIONS | 8192 |
| <i>Ryan Murphy, Balasubramaniam Srinivasan, Vinayak Rao, Bruno Ribeiro</i> | |
| LEARNING OPTIMAL FAIR POLICIES | 8211 |
| <i>Razieh Nabi, Daniel Malinsky, Ilya Shpitser</i> | |
| LEXICOGRAPHIC AND DEPTH-SENSITIVE MARGINS IN HOMOGENEOUS AND NON-HOMOGENEOUS DEEP MODELS | 8224 |
| <i>Mor Shpigel Nacson, Suriya Gunasekar, Jason Lee, Nathan Srebro, Daniel Soudry</i> | |
| A WRAPPED NORMAL DISTRIBUTION ON HYPERBOLIC SPACE FOR GRADIENT-BASED LEARNING | 8242 |
| <i>Yoshihiro Nagano, Shoichiro Yamaguchi, Yasuhiro Fujita, Masanori Koyama</i> | |
| SGD WITHOUT REPLACEMENT: SHARPER RATES FOR GENERAL SMOOTH CONVEX FUNCTIONS | 8261 |
| <i>Dheeraj Nagaraj, Prateek Jain, Praneeth Netrapalli</i> | |

| | |
|--|------|
| DROPOUT AS A STRUCTURED SHRINKAGE PRIOR | 8273 |
| <i>Eric Nalisnick, Jose Miguel Hernandez-Lobato, Padhraic Smyth</i> | |
| HYBRID MODELS WITH DEEP AND INVERTIBLE FEATURES | 8295 |
| <i>Eric Nalisnick, Akihiro Matsukawa, Yee Whye Teh, Dilan Gorur, Balaji Lakshminarayanan</i> | |
| LEARNING CONTEXT-DEPENDENT LABEL PERMUTATIONS FOR MULTI-LABEL CLASSIFICATION | 8307 |
| <i>Jinseok Nam, Young-Bum Kim, Eneldo Loza Mencia, Sunghyun Park, Ruhi Sarikaya, Johannes Fürnkranz</i> | |
| ZERO-SHOT KNOWLEDGE DISTILLATION IN DEEP NETWORKS | 8317 |
| <i>Gaurav Kumar Nayak, Konda Reddy Mopuri, Vaisakh Shaj, Venkatesh Babu Radhakrishnan, Anirban Chakraborty</i> | |
| A FRAMEWORK FOR BAYESIAN OPTIMIZATION IN EMBEDDED SUBSPACES | 8334 |
| <i>Amin Nayebi, Alexander Munteanu, Matthias Poloczek</i> | |
| PHASELESS PCA: LOW-RANK MATRIX RECOVERY FROM COLUMN-WISE PHASELESS MEASUREMENTS | 8353 |
| <i>Seyedehsara Nayer, Praneeth Narayanamurthy, Namrata Vaswani</i> | |
| SAFE GRID SEARCH WITH OPTIMAL COMPLEXITY | 8362 |
| <i>Eugene Ndiaye, Tam Le, Olivier Fercoq, Joseph Salmon, Ichiro Takeuchi</i> | |
| LEARNING TO BID IN REVENUE-MAXIMIZING AUCTIONS | 8380 |
| <i>Thomas Nedelec, Noureddine El Karoui, Vianney Perchet</i> | |
| ON CONNECTED SUBLEVEL SETS IN DEEP LEARNING | 8402 |
| <i>Quynh Nguyen</i> | |

PART 13

| | |
|--|------|
| ANOMALY DETECTION WITH MULTIPLE-HYPOTHESES PREDICTIONS | 8418 |
| <i>Duc Tam Nguyen, Zhongyu Lou, Michael Klar, Thomas Brox</i> | |
| NON-ASYMPTOTIC ANALYSIS OF FRACTIONAL LANGEVIN MONTE CARLO FOR NON-CONVEX OPTIMIZATION | 8433 |
| <i>Than Huy Nguyen, Umut Simsekli, Gael Richard</i> | |
| ROTATION INVARIANT HOUSEHOLDER PARAMETERIZATION FOR BAYESIAN PCA | 8466 |
| <i>Rajbir Nirwan, Nils Bertschinger</i> | |
| LOSSLESS OR QUANTIZED BOOSTING WITH INTEGER ARITHMETIC | 8475 |
| <i>Richard Nock, Robert Williamson</i> | |
| TRAINING NEURAL NETWORKS WITH LOCAL ERROR SIGNALS | 8529 |
| <i>Arild Nøklund, Lars Hiller Eidnes</i> | |
| REMEMBER AND FORGET FOR EXPERIENCE REPLAY | N/A |
| <i>Guido Novati, Petros Koumoutsakos</i> | |
| LEARNING TO INFER PROGRAM SKETCHES | 8543 |
| <i>Maxwell Nye, Luke Hewitt, Joshua Tenenbaum, Armando Solar-Lezama</i> | |
| TENSOR VARIABLE ELIMINATION FOR PLATED FACTOR GRAPHS | 8560 |
| <i>Fritz Obermeyer, Eli Bingham, Martin Jankowiak, Neeraj Pradhan, Justin Chiu, Alexander Rush, Noah Goodman</i> | |
| COUNTERFACTUAL OFF-POLICY EVALUATION WITH GUMBEL-MAX STRUCTURAL CAUSAL MODELS | 8577 |
| <i>Michael Oberst, David Sontag</i> | |
| MODEL FUNCTION BASED CONDITIONAL GRADIENT METHOD WITH ARMIJO-LIKE LINE SEARCH | 8591 |
| <i>Peter Ochs, Yura Malitsky</i> | |
| TENSORFUZZ: DEBUGGING NEURAL NETWORKS WITH COVERAGE-GUIDED FUZZING | 8603 |
| <i>Augustus Odena, Catherine Olsson, David Andersen, Ian Goodfellow</i> | |
| SCALABLE LEARNING IN REPRODUCING KERNEL KREIN SPACES | 8614 |
| <i>Dino Ogljic, Thomas Gärtnner</i> | |
| APPROXIMATION AND NON-PARAMETRIC ESTIMATION OF RESNET-TYPE CONVOLUTIONAL NEURAL NETWORKS | 8626 |
| <i>Kenta Oono, Taiji Suzuki</i> | |
| ORTHOGONAL RANDOM FOREST FOR CAUSAL INFERENCE | 8655 |
| <i>Miruna Oprescu, Vasilis Syrgkanis, Zhiwei Steven Wu</i> | |
| INFERRING HETEROGENEOUS CAUSAL EFFECTS IN PRESENCE OF SPATIAL CONFOUNDING | 8697 |
| <i>Muhammad Osama, Dave Zachariah, Thomas B. Schön</i> | |
| OVERPARAMETERIZED NONLINEAR LEARNING: GRADIENT DESCENT TAKES THE SHORTEST PATH? | 8707 |
| <i>Samet Oymak, Mahdi Soltanolkotabi</i> | |
| MULTIPLICATIVE WEIGHTS UPDATES AS A DISTRIBUTED CONSTRAINED OPTIMIZATION ALGORITHM: CONVERGENCE TO SECOND-ORDER STATIONARY POINTS ALMOST ALWAYS | 8748 |
| <i>Ioannis Panageas, Georgios Piliouras, Xiao Wang</i> | |
| IMPROVING ADVERSARIAL ROBUSTNESS VIA PROMOTING ENSEMBLE DIVERSITY | 8759 |
| <i>Tianyu Pang, Kun Xu, Chao Du, Ning Chen, Jun Zhu</i> | |
| NONPARAMETRIC BAYESIAN DEEP NETWORKS WITH LOCAL COMPETITION | 8772 |
| <i>Konstantinos Panousis, Sotirios Chatzis, Sergios Theodoridis</i> | |
| OPTIMISTIC POLICY OPTIMIZATION VIA MULTIPLE IMPORTANCE SAMPLING | 8784 |
| <i>Matteo Papini, Alberto Maria Metelli, Lorenzo Lupo, Marcello Restelli</i> | |
| DEEP RESIDUAL OUTPUT LAYERS FOR NEURAL LANGUAGE GENERATION | 8807 |
| <i>Nikolaos Pappas, James Henderson</i> | |

| | |
|--|------|
| MEASUREMENTS OF THREE-LEVEL HIERARCHICAL STRUCTURE IN THE OUTLIERS IN THE SPECTRUM OF DEEPNET HESSIANS | 8819 |
| <i>Vardan Papayan</i> | |
| GENERALIZED MAJORIZATION-MINIMIZATION | 8831 |
| <i>Sobhan Naderi Parizi, Kun He, Reza Aghajani, Stan Sclaroff, Pedro Felzenszwalb</i> | |
| VARIATIONAL LAPLACE AUTOENCODERS | 8844 |
| <i>Yookoon Park, Chris Kim, Gunhee Kim</i> | |
| THE EFFECT OF NETWORK WIDTH ON STOCHASTIC GRADIENT DESCENT AND GENERALIZATION: AN EMPIRICAL STUDY | 8857 |
| <i>Daniel Park, Jascha Sohl-Dickstein, Quoc Le, Samuel Smith</i> | |
| SPECTRAL APPROXIMATE INFERENCE | 8874 |
| <i>Sejun Park, Eunho Yang, Se-Young Yun, Jinwoo Shin</i> | |
| SELF-SUPERVISED EXPLORATION VIA DISAGREEMENT | 8887 |
| <i>Deepak Pathak, Dhiraj Gandhi, Abhinav Gupta</i> | |
| SUBSPACE ROBUST WASSERSTEIN DISTANCES | 8897 |
| <i>François-Pierre Paty, Marco Cuturi</i> | |
| FINGERPRINT POLICY OPTIMISATION FOR ROBUST REINFORCEMENT LEARNING | 8913 |
| <i>Supratik Paul, Michael A. Osborne, Shimon Whiteson</i> | |
| COMIC: MULTI-VIEW CLUSTERING WITHOUT PARAMETER SELECTION | 8925 |
| <i>Xi Peng, Zhenyu Huang, Jiancheng Lv, Hongyuan Zhu, Joey Tianyi Zhou</i> | |
| DOMAIN AGNOSTIC LEARNING WITH DISENTANGLED REPRESENTATIONS | 8935 |
| <i>Xingchao Peng, Zijun Huang, Ximeng Sun, Kate Saenko</i> | |
| COLLABORATIVE CHANNEL PRUNING FOR DEEP NETWORKS | 8947 |
| <i>Hanyu Peng, Jiaxiang Wu, Shifeng Chen, Junzhou Huang</i> | |
| EXPLOITING STRUCTURE OF UNCERTAINTY FOR EFFICIENT MATROID SEMI-BANDITS | 8968 |
| <i>Pierre Perrault, Vianney Perchet, Michal Valko</i> | |
| COGNITIVE MODEL PRIORS FOR PREDICTING HUMAN DECISIONS | 8984 |
| <i>David D. Bourgin, Joshua C. Peterson, Daniel Reichman, Stuart J. Russell, Thomas L. Griffiths</i> | |
| TOWARDS UNDERSTANDING KNOWLEDGE DISTILLATION | 8993 |
| <i>Mary Phuong, Christoph Lampert</i> | |
| TEMPORAL GAUSSIAN MIXTURE LAYER FOR VIDEOS | 9008 |
| <i>Aj Piergiovanni, Michael Ryoo</i> | |
| VORONOI BOUNDARY CLASSIFICATION: A HIGH-DIMENSIONAL GEOMETRIC APPROACH VIA WEIGHTED MONTE CARLO INTEGRATION | 9024 |
| <i>Vladislav Polianskii, Florian T. Pokorny</i> | |
| ON VARIATIONAL BOUNDS OF MUTUAL INFORMATION | 9036 |
| <i>Ben Poole, Sherjil Ozair, Aaron Van Den Oord, Alex Alemi, George Tucker</i> | |
| HIRING UNDER UNCERTAINTY | 9050 |
| <i>Manish Purohit, Sreenivas Gollapudi, Manish Raghavan</i> | |
| SAGA WITH ARBITRARY SAMPLING | 9061 |
| <i>Xun Qian, Zheng Qu, Peter Richtárik</i> | |
| SGD: GENERAL ANALYSIS AND IMPROVED RATES | 9090 |
| <i>Robert Mansel Gower, Nicolas Loizou, Xun Qian, Alibek Sailanbayev, Egor Shulgin, Peter Richtárik</i> | |
| AUTOVC: ZERO-SHOT VOICE STYLE TRANSFER WITH ONLY AUTOENCODER LOSS | 9113 |
| <i>Kaizhi Qian, Yang Zhang, Shiyu Chang, Xuesong Yang, Mark Hasegawa-Johnson</i> | |

PART 14

| | |
|---|------|
| FAULT TOLERANCE IN ITERATIVE-CONVERGENT MACHINE LEARNING | 9125 |
| <i>Aurick Qiao, Bryon Aragam, Bingjing Zhang, Eric Xing</i> | |
| IMPERCEPTIBLE, ROBUST, AND TARGETED ADVERSARIAL EXAMPLES FOR AUTOMATIC SPEECH RECOGNITION | 9141 |
| <i>Yao Qin, Nicholas Carlini, Garrison Cottrell, Ian Goodfellow, Colin Raffel</i> | |
| GMNN: GRAPH MARKOV NEURAL NETWORKS | 9155 |
| <i>Meng Qu, Yoshua Bengio, Jian Tang</i> | |
| NONLINEAR DISTRIBUTIONAL GRADIENT TEMPORAL-DIFFERENCE LEARNING | 9168 |
| <i>Chao Qu, Shie Mannor, Huan Xu</i> | |
| LEARNING TO COLLABORATE IN MARKOV DECISION PROCESSES | 9178 |
| <i>Goran Radanovic, Rati Devidze, David Parkes, Adish Singla</i> | |
| META-LEARNING NEURAL BLOOM FILTERS | 9188 |
| <i>Jack Rae, Sergey Bartunov, Timothy Lillicrap</i> | |
| DIRECT UNCERTAINTY PREDICTION FOR MEDICAL SECOND OPINIONS | 9202 |
| <i>Maitra Raghu, Katy Blumer, Rory Sayres, Ziad Obermeyer, Bobby Kleinberg, Sendhil Mullainathan, Jon Kleinberg</i> | |
| GAME THEORETIC OPTIMIZATION VIA GRADIENT-BASED NIKAIIDO-ISODA FUNCTION | 9217 |
| <i>Arvind Raghunathan, Anoop Cherian, Devesh Jha</i> | |
| ON THE SPECTRAL BIAS OF NEURAL NETWORKS | 9230 |
| <i>Nasim Rahaman, Aristide Baratin, Devansh Arpit, Felix Draxler, Min Lin, Fred Hamprecht, Yoshua Bengio, Aaron Courville</i> | |
| LOOK MA, NO LATENT VARIABLES: ACCURATE CUTSET NETWORKS VIA COMPILATION | 9253 |
| <i>Tahrima Rahman, Shasha Jin, Vibhav Gogate</i> | |

| | |
|---|------|
| DOES DATA AUGMENTATION LEAD TO POSITIVE MARGIN? | 9273 |
| <i>Shashank Rajput, Zhili Feng, Zachary Charles, Po-Ling Loh, Dimitris Papailopoulos</i> | |
| EFFICIENT OFF-POLICY META-REINFORCEMENT LEARNING VIA PROBABILISTIC CONTEXT VARIABLES | 9291 |
| <i>Kate Rakelly, Aurick Zhou, Chelsea Finn, Sergey Levine, Deirdre Quillen</i> | |
| SCREENING RULES FOR LASSO WITH NON-CONVEX SPARSE REGULARIZERS | 9302 |
| <i>Alain Rakotomamonjy, Gilles Gasso, Joseph Salmon</i> | |
| TOPOLOGICAL DATA ANALYSIS OF DECISION BOUNDARIES WITH APPLICATION TO MODEL SELECTION | 9316 |
| <i>Karthikeyan Natesan Ramamurthy, Kush Varshney, Krishnan Mody</i> | |
| HYPERGAN: A GENERATIVE MODEL FOR DIVERSE, PERFORMANT NEURAL NETWORKS | 9371 |
| <i>Neale Ratzlaff, Li Fuxin</i> | |
| EFFICIENT ON-DEVICE MODELS USING NEURAL PROJECTIONS | 9382 |
| <i>Sujith Ravi</i> | |
| A BLOCK COORDINATE DESCENT PROXIMAL METHOD FOR SIMULTANEOUS FILTERING AND PARAMETER ESTIMATION | 9395 |
| <i>Ramin Raziperchikolaei, Harish Bhat</i> | |
| DO IMAGENET CLASSIFIERS GENERALIZE TO IMAGENET? | 9413 |
| <i>Benjamin Recht, Rebecca Roelofs, Ludwig Schmidt, Vaishaal Shankar</i> | |
| FAST RATES FOR A KNN CLASSIFIER ROBUST TO UNKNOWN ASYMMETRIC LABEL NOISE | 9472 |
| <i>Henry Reeve, Ata Kaban</i> | |
| ALMOST UNSUPERVISED TEXT TO SPEECH AND AUTOMATIC SPEECH RECOGNITION | 9483 |
| <i>Yi Ren, Xu Tan, Tao Qin, Sheng Zhao, Zhou Zhao, Tie-Yan Liu</i> | |
| ADAPTIVE ANTITHETIC SAMPLING FOR VARIANCE REDUCTION | 9493 |
| <i>Hongyu Ren, Shengjia Zhao, Stefano Ermon</i> | |
| ADVERSARIAL ONLINE LEARNING WITH NOISE | 9506 |
| <i>Alon Resler, Yishay Mansour</i> | |
| A POLYNOMIAL TIME MCMC METHOD FOR SAMPLING FROM CONTINUOUS DETERMINANTAL POINT PROCESSES | 9532 |
| <i>Alireza Rezaei, Shayan Oveis Gharan</i> | |
| A PERSISTENT WEISFELER-LEHMAN PROCEDURE FOR GRAPH CLASSIFICATION | 9556 |
| <i>Bastian Rieck, Christian Bock, Karsten Borgwardt</i> | |
| EFFICIENT LEARNING OF SMOOTH PROBABILITY FUNCTIONS FROM BERNOULLI TESTS WITH GUARANTEES | 9569 |
| <i>Paul Rolland, Ali Kavis, Alexander Immer, Adish Singla, Volkan Cevher</i> | |
| SEPARATING VALUE FUNCTIONS ACROSS TIME-SCALES | 9589 |
| <i>Joshua Romoff, Peter Henderson, Ahmed Touati, Emma Brunskill, Joelle Pineau, Yann Ollivier</i> | |
| ONLINE CONVEX OPTIMIZATION IN ADVERSARIAL MARKOV DECISION PROCESSES | 9643 |
| <i>Aviv Rosenberg, Yishay Mansour</i> | |
| GOOD INITIALIZATIONS OF VARIATIONAL BAYES FOR DEEP MODELS | 9659 |
| <i>Simone Rossi, Pietro Michiardi, Maurizio Filippone</i> | |
| THE ODDS ARE ODD: A STATISTICAL TEST FOR DETECTING ADVERSARIAL EXAMPLES | 9677 |
| <i>Kevin Roth, Yannic Kilcher, Thomas Hofmann</i> | |
| NEURON BIRTH-DEATH DYNAMICS ACCELERATES GRADIENT DESCENT AND CONVERGES ASYMPTOTICALLY | 9689 |
| <i>Grant Rotskoff, Samy Jelassi, Joan Bruna, Eric Vanden-Eijnden</i> | |
| ITERATIVE LINEARIZED CONTROL: STABLE ALGORITHMS AND COMPLEXITY GUARANTEES | 9717 |
| <i>Vincent Roulet, Dmitry Drusvyatskiy, Siddhartha Srinivasa, Zaid Harchaoui</i> | |
| STATISTICS AND SAMPLES IN DISTRIBUTIONAL REINFORCEMENT LEARNING | 9727 |
| <i>Mark Rowland, Robert Dadashi, Saurabh Kumar, Remi Munos, Marc G. Bellemare, Will Dabney</i> | |
| A CONTRASTIVE DIVERGENCE FOR COMBINING VARIATIONAL INFERENCE AND MCMC | 9751 |
| <i>Francisco Ruiz, Michalis Titsias</i> | |
| PLUG-AND-PLAY METHODS PROVABLY CONVERGE WITH PROPERLY TRAINED DENOISERS | 9764 |
| <i>Ernest Ryu, Jialin Liu, Sicheng Wang, Xiaohan Chen, Zhangyang Wang, Wotao Yin</i> | |
| WHITE-BOX VS BLACK-BOX: BAYES OPTIMAL STRATEGIES FOR MEMBERSHIP INFERENCE | 9780 |
| <i>Alexandre Sablayrolles, Matthijs Douze, Cordelia Schmid, Yann Ollivier, Herve Jegou</i> | |
| AN OPTIMAL PRIVATE STOCHASTIC-MAB ALGORITHM BASED ON OPTIMAL PRIVATE STOPPING RULE | 9791 |
| <i>Touqir Sajed, Or Sheffet</i> | |
| DEEP GAUSSIAN PROCESSES WITH IMPORTANCE-WEIGHTED VARIATIONAL INFERENCE | 9801 |
| <i>Hugh Salimbeni, Vincent Dutordoir, James Hensman, Marc Deisenroth</i> | |
| MULTIVARIATE SUBMODULAR OPTIMIZATION | 9820 |
| <i>Richard Santiago, F. Bruce Shepherd</i> | |

PART 15

| | |
|--|------|
| NEAR OPTIMAL FINITE TIME IDENTIFICATION OF ARBITRARY LINEAR DYNAMICAL SYSTEMS | 9846 |
| <i>Tuhin Sarkar, Alexander Rakhlin</i> | |

| | |
|--|-------|
| BREAKING INTER-LAYER CO-ADAPTATION BY CLASSIFIER ANONYMIZATION | 9895 |
| <i>Ikuro Sato, Kohta Ishikawa, Guoqing Liu, Masayuki Tanaka</i> | |
| A THEORETICAL ANALYSIS OF CONTRASTIVE UNSUPERVISED REPRESENTATION LEARNING | 9904 |
| <i>Nikunj Saunshi, Orestis Plevrakis, Sanjeev Arora, Mikhail Khodak, Hrishikesh Khandeparkar</i> | |
| LOCALLY PRIVATE BAYESIAN INFERENCE FOR COUNT MODELS | 9924 |
| <i>Aaron Schein, Zhiwei Steven Wu, Alexandra Schofield, Mingyuan Zhou, Hanna Wallach</i> | |
| WEAKLY-SUPERVISED TEMPORAL LOCALIZATION VIA OCCURRENCE COUNT LEARNING | 9939 |
| <i>Julien Schroeter, Kirill Sidorov, David Marshall</i> | |
| DISCOVERING CONTEXT EFFECTS FROM RAW CHOICE DATA | 9951 |
| <i>Arjun Seshadri, Alex Peysakhovich, Johan Ugander</i> | |
| ON THE FEASIBILITY OF LEARNING, RATHER THAN ASSUMING, HUMAN BIASES FOR REWARD INFERENCE | 9974 |
| <i>Rohin Shah, Noah Gundotra, Pieter Abbeel, Anca Dragan</i> | |
| EXPLORATION CONSCIOUS REINFORCEMENT LEARNING REVISITED | 9986 |
| <i>Lior Shani, Yonathan Efroni, Shie Mannor</i> | |
| COMPRESSED FACTORIZATION: FAST AND ACCURATE LOW-RANK FACTORIZATION OF COMPRESSIVELY-SENSED DATA | 10013 |
| <i>Vatsal Sharan, Kai Sheng Tai, Peter Bailis, Gregory Valiant</i> | |
| CONDITIONAL INDEPENDENCE IN TESTING BAYESIAN NETWORKS | 10031 |
| <i>Yujia Shen, Haiying Huang, Arthur Choi, Adnan Darwiche</i> | |
| LEARNING TO CLEAR THE MARKET | 10041 |
| <i>Weiran Shen, Sebastien Lahaie, Renato Paes Leme</i> | |
| MIXTURE MODELS FOR DIVERSE MACHINE TRANSLATION: TRICKS OF THE TRADE | 10050 |
| <i>Tianxiao Shen, Myle Ott, Michael Auli, Marc Aurelio Ranzato</i> | |
| HESSIAN AIDED POLICY GRADIENT | 10062 |
| <i>Zebang Shen, Alejandro Ribeiro, Hamed Hassani, Hui Qian, Chao Mi</i> | |
| LEARNING WITH BAD TRAINING DATA VIA ITERATIVE TRIMMED LOSS MINIMIZATION | 10075 |
| <i>Yanyao Shen, Sujay Sanghavi</i> | |
| REPLICA CONDITIONAL SEQUENTIAL MONTE CARLO | 10098 |
| <i>Alex Shestopaloff, Arnaud Doucet</i> | |
| SCALABLE TRAINING OF INFERENCE NETWORKS FOR GAUSSIAN-PROCESS MODELS | 10108 |
| <i>Jiaxin Shi, Mohammad Emteyaz Khan, Jun Zhu</i> | |
| FAST DIRECT SEARCH IN AN OPTIMALLY COMPRESSED CONTINUOUS TARGET SPACE FOR EFFICIENT MULTI-LABEL ACTIVE LEARNING | 10124 |
| <i>Weishi Shi, Qi Yu</i> | |
| MODEL-BASED ACTIVE EXPLORATION | 10136 |
| <i>Pranav Shyam, Wojciech Jaskowski, Faustino Gomez</i> | |
| REHASHING KERNEL EVALUATION IN HIGH DIMENSIONS | 10153 |
| <i>Paris Siminelakis, Kexin Rong, Peter Bailis, Moses Charikar, Philip Levis</i> | |
| REVISITING PRECISION RECALL DEFINITION FOR GENERATIVE MODELING | 10174 |
| <i>Loic Simon, Ryan Webster, Julien Rabin</i> | |
| FIRST-ORDER ADVERSARIAL VULNERABILITY OF NEURAL NETWORKS AND INPUT DIMENSION | 10184 |
| <i>Carl-Johann Simon-Gabriel, Yann Ollivier, Leon Bottou, Bernhard Schölkopf, David Lopez-Paz</i> | |
| REFINED COMPLEXITY OF PCA WITH OUTLIERS | 10204 |
| <i>Kirill Simonov, Fedor Fomin, Petr Golovach, Fahad Panolan</i> | |
| A TAIL-INDEX ANALYSIS OF STOCHASTIC GRADIENT NOISE IN DEEP NEURAL NETWORKS | 10214 |
| <i>Umut Simsekli, Levent Sagun, Mert Gurbuzbalaban</i> | |
| NON-PARAMETRIC PRIORS FOR GENERATIVE ADVERSARIAL NETWORKS | 10225 |
| <i>Rajhans Singh, Pavan Turaga, Suren Jayasuriya, Ravi Garg, Martin Braun</i> | |
| UNDERSTANDING IMPACTS OF HIGH-ORDER LOSS APPROXIMATIONS AND FEATURES IN DEEP LEARNING INTERPRETATION | 10248 |
| <i>Sahil Singla, Eric Wallace, Shi Feng, Soheil Feizi</i> | |
| KERNELPSI: A POST-SELECTION INFERENCE FRAMEWORK FOR NONLINEAR VARIABLE SELECTION | 10268 |
| <i>Lofci Slim, Clément Chatelain, Chloe-Agathe Azencott, Jean-Philippe Vert</i> | |
| GEOMETRICS: EXPLOITING GEOMETRIC STRUCTURE FOR GRAPH-ENCODED OBJECTS | 10297 |
| <i>Edward Smith, Scott Fujimoto, Adriana Romero, David Meger</i> | |
| THE EVOLVED TRANSFORMER | 10315 |
| <i>David So, Quoc Le, Chen Liang</i> | |
| QTRAN: LEARNING TO FACTORIZE WITH TRANSFORMATION FOR COOPERATIVE MULTI-AGENT REINFORCEMENT LEARNING | 10329 |
| <i>Kyunghwan Son, Daewoo Kim, Wan Ju Kang, David Earl Hostallero, Yung Yi</i> | |
| DISTRIBUTION CALIBRATION FOR REGRESSION | 10347 |
| <i>Hao Song, Tom Diethe, Meelis Kull, Peter Flach</i> | |
| SELFIE: REFURBISHING UNCLEAN SAMPLES FOR ROBUST DEEP LEARNING | 10357 |
| <i>Hwanjun Song, Minseok Kim, Jae-Gil Lee</i> | |
| REVISITING THE SOFTMAX BELLMAN OPERATOR: NEW BENEFITS AND NEW PERSPECTIVE | 10368 |
| <i>Zhao Song, Ron Parr, Lawrence Carin</i> | |
| MASS: MASKED SEQUENCE TO SEQUENCE PRE-TRAINING FOR LANGUAGE GENERATION | 10384 |
| <i>Kaitao Song, Xu Tan, Tao Qin, Jianfeng Lu, Tie-Yan Liu</i> | |

| | |
|---|-------|
| DUAL ENTANGLED POLYNOMIAL CODE: THREE-DIMENSIONAL CODING FOR DISTRIBUTED MATRIX MULTIPLICATION | 10395 |
| <i>Pedro Soto, Jun Li, Xiaodi Fan</i> | |
| COMPRESSING GRADIENT OPTIMIZERS VIA COUNT-SKETCHES | 10404 |
| <i>Ryan Spring, Anastasios Kyrillidis, Vijai Mohan, Anshumali Shrivastava</i> | |
| ESCAPING SADDLE POINTS WITH ADAPTIVE GRADIENT METHODS | 10420 |
| <i>Matthew Staib, Sashank Reddi, Satyen Kale, Sanjiv Kumar, Suvrit Sra</i> | |
| FASTER ATTEND-INFER-REPEAT WITH TRACTABLE PROBABILISTIC MODELS | 10455 |
| <i>Karl Stelzner, Robert Peharz, Kristian Kersting</i> | |
| INSERTION TRANSFORMER: FLEXIBLE SEQUENCE GENERATION VIA INSERTION OPERATIONS | 10467 |
| <i>Mitchell Stern, William Chan, Jamie Kiros, Jakob Uszkoreit</i> | |
| BERT AND PALS: PROJECTED ATTENTION LAYERS FOR EFFICIENT ADAPTATION IN MULTI-TASK LEARNING | 10477 |
| <i>Asa Cooper Stickland, Iain Murray</i> | |
| LEARNING OPTIMAL LINEAR REGULARIZERS | 10489 |
| <i>Matthew Streeter</i> | |
| CAB: CONTINUOUS ADAPTIVE BLENDING FOR POLICY EVALUATION AND LEARNING | 10499 |
| <i>Yi Su, Lequn Wang, Michele Santacatterina, Thorsten Joachims</i> | |
| LEARNING DISTANCE FOR SEQUENCES BY LEARNING A GROUND METRIC | 10517 |
| <i>Bing Su, Ying Wu</i> | |
| CONTEXTUAL MEMORY TREES | 10535 |
| <i>Wen Sun, Alina Beygelzimer, Hal Daumé III, John Langford, Paul Mineiro</i> | |
| PROBABLY EFFICIENT IMITATION LEARNING FROM OBSERVATION ALONE | 10553 |
| <i>Wen Sun, Anirudh Vemula, Byron Boots, Drew Bagnell</i> | |

PART 16

| | |
|---|-------|
| ACTIVE LEARNING FOR DECISION-MAKING FROM IMBALANCED OBSERVATIONAL DATA | 10578 |
| <i>Iris Sundin, Peter Schulam, Eero Siivola, Aki Vehtari, Suchi Saria, Samuel Kaski</i> | |
| ROBUSTLY DISENTANGLED CAUSAL MECHANISMS: VALIDATING DEEP REPRESENTATIONS FOR INTERVENTIONAL ROBUSTNESS | 10593 |
| <i>Raphael Suter, Djordje Miladinovic, Bernhard Schölkopf, Stefan Bauer</i> | |
| HYPERBOLIC DISK EMBEDDINGS FOR DIRECTED ACYCLIC GRAPHS | 10617 |
| <i>Ryota Suzuki, Ryusuke Takahama, Shun Onoda</i> | |
| ACCELERATED FLOW FOR PROBABILITY DISTRIBUTIONS | 10632 |
| <i>Amirhossein Taghvaei, Prashant Mehta</i> | |
| EQUIVARIANT TRANSFORMER NETWORKS | 10649 |
| <i>Kai Sheng Tai, Peter Bailis, Gregory Valiant</i> | |
| MAKING DEEP Q-LEARNING METHODS ROBUST TO TIME DISCRETIZATION | 10664 |
| <i>Corentin Tallec, Léonard Blier, Yann Ollivier</i> | |
| EFFICIENTNET: RETHINKING MODEL SCALING FOR CONVOLUTIONAL NEURAL NETWORKS | 10691 |
| <i>Mingxing Tan, Quoc Le</i> | |
| HIERARCHICAL DECOMPOSITIONAL MIXTURES OF VARIATIONAL AUTOENCODERS | 10701 |
| <i>Ping Liang Tan, Robert Peharz</i> | |
| MALLOWS RANKING MODELS: MAXIMUM LIKELIHOOD ESTIMATE AND REGENERATION | 10712 |
| <i>Wenpin Tang</i> | |
| CORRELATED VARIATIONAL AUTO-ENCODERS | 10722 |
| <i>Da Tang, Dawen Liang, Tony Jebara, Nicholas Ruozi</i> | |
| THE VARIATIONAL PREDICTIVE NATURAL GRADIENT | 10735 |
| <i>Da Tang, Rajesh Ranganath</i> | |
| DOUBLESQUEEZE: PARALLEL STOCHASTIC GRADIENT DESCENT WITH DOUBLE-PASS ERROR-COMPENSATED COMPRESSION | 10747 |
| <i>Hanlin Tang, Chen Yu, Xiangru Lian, Tong Zhang, Ji Liu</i> | |
| ADAPTIVE NEURAL TREES | 10761 |
| <i>Ryutaro Tanno, Kai Arulkumaran, Daniel Alexander, Antonio Criminisi, Aditya Nori</i> | |
| VARIATIONAL ANNEALING OF GANS: A LANGEVIN PERSPECTIVE | 10777 |
| <i>Chenyang Tao, Shuyang Dai, Liqun Chen, Ke Bai, Junya Chen, Chang Liu, Ruiyi Zhang, Georgiy Bobashev, Lawrence Carin Duke</i> | |
| PREDICATE EXCHANGE: INFERENCE WITH DECLARATIVE KNOWLEDGE | 10792 |
| <i>Zenna Tavares, Javier Burroni, Edgar Minasyan, Armando Solar-Lezama, Rajesh Ranganath</i> | |
| THE NATURAL LANGUAGE OF ACTIONS | 10802 |
| <i>Guy Tennenholtz, Shie Mannor</i> | |
| KERNEL NORMALIZED CUT: A THEORETICAL REVISIT | 10817 |
| <i>Yoshikazu Terada, Michio Yamamoto</i> | |
| ACTION ROBUST REINFORCEMENT LEARNING AND APPLICATIONS IN CONTINUOUS CONTROL | 10846 |
| <i>Chen Tessler, Yonathan Efroni, Shie Mannor</i> | |
| CONCENTRATION INEQUALITIES FOR CONDITIONAL VALUE AT RISK | 10876 |
| <i>Philip Thomas, Erik Learned-Miller</i> | |
| COMBATING LABEL NOISE IN DEEP LEARNING USING ABSTENTION | N/A |
| <i>Sunil Thulasidasan, Tanmoy Bhattacharya, Jeff Bilmes, Gopinath Chennupati, Jamal Mohd-Yusof</i> | |

| | |
|---|-------|
| ELF OPENGO: AN ANALYSIS AND OPEN REIMPLEMENTATION OF ALPHAZERO | 10885 |
| <i>Yuandong Tian, Jerry Ma, Qucheng Gong, Shubho Sengupta, Zhuoyuan Chen, James Pinkerton, Larry Zitnick</i> | |
| RANDOM MATRIX IMPROVED COVARIANCE ESTIMATION FOR A LARGE CLASS OF METRICS | 10898 |
| <i>Malik Tiomoko, Romain Couillet, Florent Bouchard, Guillaume Ginolhac</i> | |
| TRANSFER OF SAMPLES IN POLICY SEARCH VIA MULTIPLE IMPORTANCE SAMPLING | 10912 |
| <i>Andrea Tirinzoni, Mattia Salvini, Marcello Restelli</i> | |
| OPTIMAL TRANSPORT FOR STRUCTURED DATA WITH APPLICATION ON GRAPHS | 10940 |
| <i>Vayer Titouan, Nicolas Courty, Romain Tavenard, Chapel Laetitia, Rémi Flamary</i> | |
| DISCOVERING LATENT COVARIANCE STRUCTURES FOR MULTIPLE TIME SERIES | 10956 |
| <i>Anh Tong, Jaesik Choi</i> | |
| BAYESIAN GENERATIVE ACTIVE DEEP LEARNING | 10969 |
| <i>Toan Tran, Thanh-Toan Do, Ian Reid, Gustavo Carneiro</i> | |
| DEEPROSE: USING ARTIFICIAL NEURAL NETWORKS TO REPRESENT THE SPACE OF ODORANTS | 10979 |
| <i>Ngoc Tran, Daniel Kepple, Sergey Shuvaev, Alexei Koulakov</i> | |
| LR-GLM: HIGH-DIMENSIONAL BAYESIAN INFERENCE USING LOW-RANK DATA APPROXIMATIONS | 10989 |
| <i>Brian Trippe, Jonathan Huggins, Raj Agrawal, Tamara Broderick</i> | |
| LEARNING HAWKES PROCESSES UNDER SYNCHRONIZATION NOISE | 11022 |
| <i>William Trouleau, Jalal Etesami, Matthias Grossglauser, Negar Kiyavash, Patrick Thiran</i> | |
| HOMOMORPHIC SENSING | 11034 |
| <i>Manolis Tsakiris, Liangzu Peng</i> | |
| METROPOLIS-HASTINGS GENERATIVE ADVERSARIAL NETWORKS | 11044 |
| <i>Ryan Turner, Jane Hung, Eric Frank, Yunus Saatchi, Jason Yosinski</i> | |
| DISTRIBUTED, EGOCENTRIC REPRESENTATIONS OF GRAPHS FOR DETECTING CRITICAL STRUCTURES | 11053 |
| <i>Ruo-Chun Tzeng, Shan-Hung Wu</i> | |
| SUBLINEAR SPACE PRIVATE ALGORITHMS UNDER THE SLIDING WINDOW MODEL | 11065 |
| <i>Jalaj Upadhyay</i> | |
| FAIRNESS WITHOUT HARM: DECOUPLED CLASSIFIERS WITH PREFERENCE GUARANTEES | 11080 |
| <i>Berk Ustun, Yang Liu, David Parkes</i> | |
| LARGE-SCALE SPARSE KERNEL CANONICAL CORRELATION ANALYSIS | 11094 |
| <i>Viivi Uurtio, Sahely Bhadra, Juho Rousu</i> | |
| CHARACTERIZATION OF CONVEX OBJECTIVE FUNCTIONS AND OPTIMAL EXPECTED CONVERGENCE RATES FOR SGD | 11106 |
| <i>Marten Van Dijk, Lam Nguyen, Phuong Ha Nguyen, Dzong Phan</i> | |
| COMPOSING VALUE FUNCTIONS IN REINFORCEMENT LEARNING | 11132 |
| <i>Benjamin Van Niekerk, Steven James, Adam Earle, Benjamin Rosman</i> | |
| MODEL COMPARISON FOR SEMANTIC GROUPING | 11147 |
| <i>Francisco Vargas, Kamen Brestnichki, Nils Hammerla</i> | |
| LEARNING DEPENDENCY STRUCTURES FOR WEAK SUPERVISION MODELS | 11160 |
| <i>Paroma Varma, Frederic Sala, Ann He, Alexander Ratner, Christopher Re</i> | |
| PROBABILISTIC NEURAL SYMBOLIC MODELS FOR INTERPRETABLE VISUAL QUESTION ANSWERING | 11181 |
| <i>Ramakrishna Vedantam, Karan Desai, Stefan Lee, Marcus Rohrbach, Dhruv Batra, Devi Parikh</i> | |
| MANIFOLD MIXUP: BETTER REPRESENTATIONS BY INTERPOLATING HIDDEN STATES | 11196 |
| <i>Vikas Verma, Alex Lamb, Christopher Beckham, Amir Najafi, Ioannis Mitliagkas, David Lopez-Paz, Yoshua Bengio</i> | |
| MAXIMUM LIKELIHOOD ESTIMATION FOR LEARNING POPULATIONS OF PARAMETERS | 11217 |
| <i>Ramya Korlakai Vinayak, Weihao Kong, Gregory Valiant, Sham Kakade</i> | |

PART 17

| | |
|--|-------|
| UNDERSTANDING PRIORS IN BAYESIAN NEURAL NETWORKS AT THE UNIT LEVEL | 11248 |
| <i>Mariia Vladimirova, Jakob Verbeek, Pablo Mesejo, Julian Arbel</i> | |
| ON THE DESIGN OF ESTIMATORS FOR BANDIT OFF-POLICY EVALUATION | 11261 |
| <i>Nikos Vlassis, Aurelien Bibaut, Maria Dimakopoulou, Tony Jebara</i> | |
| LEARNING TO SELECT FOR A PREDEFINED RANKING | 11270 |
| <i>Aleksei Ustimenko, Aleksandr Vorobev, Gleb Gusev, Pavel Serdyukov</i> | |
| ON THE LIMITATIONS OF REPRESENTING FUNCTIONS ON SETS | 11285 |
| <i>Edward Wagstaff, Fabian Fuchs, Martin Engelcke, Ingmar Posner, Michael A. Osborne</i> | |
| GRAPH CONVOLUTIONAL GAUSSIAN PROCESSES | 11299 |
| <i>Ian Walker, Ben Glocker</i> | |
| GAINING FREE OR LOW-COST INTERPRETABILITY WITH INTERPRETABLE PARTIAL SUBSTITUTE | 11309 |
| <i>Tong Wang</i> | |
| CONVOLUTIONAL POISSON GAMMA BELIEF NETWORK | 11321 |
| <i>Chaojie Wang, Bo Chen, Sucheng Xiao, Mingyuan Zhou</i> | |
| DIFFERENTIALLY PRIVATE EMPIRICAL RISK MINIMIZATION WITH NON-CONVEX LOSS FUNCTIONS | 11334 |
| <i>Di Wang, Changyou Chen, Jinhui Xu</i> | |

| | |
|--|-------|
| RANDOM EXPERT DISTILLATION: IMITATION LEARNING VIA EXPERT POLICY SUPPORT ESTIMATION | 11364 |
| <i>Ruohan Wang, Carlo Ciliberto, Pierluigi Vito Amadori, Yiannis Demiris</i> | |
| SATNET: BRIDGING DEEP LEARNING AND LOGICAL REASONING USING A DIFFERENTIABLE SATISFIABILITY SOLVER | 11373 |
| <i>Po-Wei Wang, Priya Donti, Bryan Wilder, Zico Kolter</i> | |
| IMPROVING NEURAL LANGUAGE MODELING VIA ADVERSARIAL TRAINING | 11387 |
| <i>Dilin Wang, Chengyue Gong, Qiang Liu</i> | |
| EIGENDAMAGE: STRUCTURED PRUNING IN THE KRONECKER-FACTORED EIGENBASIS | 11398 |
| <i>Chaoqi Wang, Roger Grosse, Sanja Fidler, Guodong Zhang</i> | |
| NONLINEAR STEIN VARIATIONAL GRADIENT DESCENT FOR LEARNING DIVERSIFIED MIXTURE MODELS | 11413 |
| <i>Dilin Wang, Qiang Liu</i> | |
| ON THE CONVERGENCE AND ROBUSTNESS OF ADVERSARIAL TRAINING | 11426 |
| <i>Yisen Wang, Xingjun Ma, James Bailey, Jinfeng Yi, Bowen Zhou, Quanquan Gu</i> | |
| STATE-REGULARIZED RECURRENT NEURAL NETWORKS | 11439 |
| <i>Cheng Wang, Mathias Niepert</i> | |
| DEEP FACTORS FOR FORECASTING | 11460 |
| <i>Yuyang Wang, Alex Smola, Danielle Maddix, Jan Gasthaus, Dean Foster, Tim Januschowski</i> | |
| REPAIRING WITHOUT RETRAINING: AVOIDING DISPARATE IMPACT WITH COUNTERFACTUAL DISTRIBUTIONS | 11476 |
| <i>Hao Wang, Berk Ustun, Flavio Calmon</i> | |
| ON SPARSE LINEAR REGRESSION IN THE LOCAL DIFFERENTIAL PRIVACY MODEL | 11497 |
| <i>Di Wang, Jinhui Xu</i> | |
| DOUBLY ROBUST JOINT LEARNING FOR RECOMMENDATION ON DATA MISSING NOT AT RANDOM | 11526 |
| <i>Xiaojie Wang, Rui Zhang, Yu Sun, Jianzhong Qi</i> | |
| ON THE GENERALIZATION GAP IN REPARAMETERIZABLE REINFORCEMENT LEARNING | 11539 |
| <i>Huan Wang, Stephan Zheng, Caiming Xiong, Richard Socher</i> | |
| BIAS ALSO MATTERS: BIAS ATTRIBUTION FOR DEEP NEURAL NETWORK EXPLANATION | 11553 |
| <i>Shengjie Wang, Tianyi Zhou, Jeff Bilmes</i> | |
| JUMPOUT : IMPROVED DROPOUT FOR DEEP NEURAL NETWORKS WITH RELUS | 11565 |
| <i>Shengjie Wang, Tianyi Zhou, Jeff Bilmes</i> | |
| ADAGRAD STEPSIZES: SHARP CONVERGENCE OVER NONCONVEX LANDSCAPES | 11574 |
| <i>Rachel Ward, Xiaoxia Wu, Leon Bottou</i> | |
| GENERALIZED LINEAR RULE MODELS | 11589 |
| <i>Dennis Wei, Sanjeeb Dash, Tian Gao, Oktay Gunluk</i> | |
| ON THE STATISTICAL RATE OF NONLINEAR RECOVERY IN GENERATIVE MODELS WITH HEAVY-TAILED DATA | 11606 |
| <i>Xiaohan Wei, Zhuoran Yang, Zhaoran Wang</i> | |
| CAPSANDRUNS: AN IMPROVED METHOD FOR APPROXIMATELY OPTIMAL ALGORITHM CONFIGURATION | 11644 |
| <i>Gellert Weisz, Andras Gyorgy, Csaba Szepesvari</i> | |
| NON-MONOTONIC SEQUENTIAL TEXT GENERATION | 11656 |
| <i>Sean Welleck, Kianté Brantley, Hal Daumé III, Kyunghyun Cho</i> | |
| PROVEN: VERIFYING ROBUSTNESS OF NEURAL NETWORKS WITH A PROBABILISTIC APPROACH | 11677 |
| <i>Lily Weng, Pin-Yu Chen, Lam Nguyen, Mark Squillante, Akhilan Boopathy, Ivan Oseledets, Luca Daniel</i> | |
| LEARNING DEEP KERNELS FOR EXPONENTIAL FAMILY DENSITIES | 11693 |
| <i>Li Wenliang, Dougal Sutherland, Heiko Strathmann, Arthur Gretton</i> | |
| IMPROVING MODEL SELECTION BY EMPLOYING THE TEST DATA | 11711 |
| <i>Max Westphal, Werner Brannath</i> | |
| AUTOMATIC CLASSIFIERS AS SCIENTIFIC INSTRUMENTS: ONE STEP FURTHER AWAY FROM GROUND-TRUTH | 11721 |
| <i>Jacob Whitehill, Anand Ramakrishnan</i> | |
| MOMENT-BASED VARIATIONAL INFERENCE FOR MARKOV JUMP PROCESSES | 11734 |
| <i>Christian Wildner, Heinz Koeppl</i> | |
| END-TO-END PROBABILISTIC INFERENCE FOR NONSTATIONARY AUDIO ANALYSIS | 11751 |
| <i>William Wilkinson, Michael Andersen, Joshua D. Reiss, Dan Stowell, Arno Solin</i> | |
| FAIRNESS RISK MEASURES | 11763 |
| <i>Robert Williamson, Aditya Menon</i> | |
| PARTIALLY EXCHANGEABLE NETWORKS AND ARCHITECTURES FOR LEARNING SUMMARY STATISTICS IN APPROXIMATE BAYESIAN COMPUTATION | 11795 |
| <i>Samuel Wqivist, Pierre-Alexandre Mattei, Umberto Picchini, Jes Frelsen</i> | |
| WASSERSTEIN ADVERSARIAL EXAMPLES VIA PROJECTED SINKHORN ITERATIONS | 11812 |
| <i>Eric Wong, Frank Schmidt, Zico Kolter</i> | |
| IMITATION LEARNING FROM IMPERFECT DEMONSTRATION | 11826 |
| <i>Yueh-Hua Wu, Nontawat Charoenphakdee, Han Bao, Voot Tangkaratt, Masashi Sugiyama</i> | |
| LEARNING A COMPRESSED SENSING MEASUREMENT MATRIX VIA GRADIENT UNROLLING | 11845 |
| <i>Shanshan Wu, Alex Dimakis, Sujay Sanghavi, Felix Yu, Daniel Holmann-Rice, Dmitry Storcheus, Afshin Rostamizadeh, Sanjiv Kumar</i> | |

| | |
|---|-------|
| HETEROGENEOUS MODEL REUSE VIA OPTIMIZING MULTIPARTY MULTICLASS MARGIN | 11862 |
| <i>Xi-Zhu Wu, Song Liu, Zhi-Hua Zhou</i> | |
| DEEP COMPRESSED SENSING | 11872 |
| <i>Yan Wu, Mihaela Rosca, Timothy Lillicrap</i> | |
| SIMPLIFYING GRAPH CONVOLUTIONAL NETWORKS | 11884 |
| <i>Felix Wu, Amauri Souza, Tianyi Zhang, Christopher Fifty, Tao Yu, Kilian Weinberger</i> | |
| DOMAIN ADAPTATION WITH ASYMMETRICALLY-RELAXED DISTRIBUTION ALIGNMENT | 11899 |
| <i>Yifan Wu, Ezra Winston, Divyansh Kaushik, Zachary Lipton</i> | |
| ON SCALABLE AND EFFICIENT COMPUTATION OF LARGE SCALE OPTIMAL TRANSPORT | 11915 |
| <i>Yujia Xie, Minshuo Chen, Haoming Jiang, Tuo Zhao, Hongyuan Zha</i> | |
| ZENO: DISTRIBUTED STOCHASTIC GRADIENT DESCENT WITH SUSPICION-BASED FAULT-TOLERANCE | 11928 |
| <i>Cong Xie, Sanmi Koyejo, Indranil Gupta</i> | |
| DIFFERENTIABLE LINEARIZED ADMM | 11945 |
| <i>Xingyu Xie, Jianlong Wu, Guangcan Liu, Zhisheng Zhong, Zhouchen Lin</i> | |
| CALIBRATED APPROXIMATE BAYESIAN INFERENCE | 11961 |
| <i>Hanwen Xing, Geoff Nicholls, Jeong Lee</i> | |

PART 18

| | |
|--|-------|
| POWER K-MEANS CLUSTERING | 11977 |
| <i>Jason Xu, Kenneth Lange</i> | |
| GROMOV-WASSERSTEIN LEARNING FOR GRAPH MATCHING AND NODE EMBEDDING | 11992 |
| <i>Hongteng Xu, Dixin Luo, Hongyuan Zha, Lawrence Carin Duke</i> | |
| STOCHASTIC OPTIMIZATION FOR DC FUNCTIONS AND NON-SMOOTH NON-CONVEX REGULARIZERS WITH NON-ASYMPTOTIC CONVERGENCE | 12008 |
| <i>Yi Xu, Qi Qi, Qihang Lin, Rong Jin, Tianbao Yang</i> | |
| LEARNING A PRIOR OVER INTENT VIA META-INVERSE REINFORCEMENT LEARNING | 12036 |
| <i>Kelvin Xu, Ellis Ratner, Anca Dragan, Sergey Levine, Chelsea Finn</i> | |
| VARIATIONAL RUSSIAN ROULETTE FOR DEEP BAYESIAN NONPARAMETRICS | 12051 |
| <i>Kai Xu, Akash Srivastava, Charles Sutton</i> | |
| SUPERVISED HIERARCHICAL CLUSTERING WITH EXPONENTIAL LINKAGE | 12067 |
| <i>Nishant Yadav, Ari Kobren, Nicholas Monath, Andrew McCallum</i> | |
| LEARNING TO PROVE THEOREMS VIA INTERACTING WITH PROOF ASSISTANTS | 12079 |
| <i>Kaiyu Yang, Jia Deng</i> | |
| SAMPLE-OPTIMAL PARAMETRIC Q-LEARNING USING LINEARLY ADDITIVE FEATURES | 12095 |
| <i>Lin Yang, Mengdi Wang</i> | |
| LEGONET: EFFICIENT CONVOLUTIONAL NEURAL NETWORKS WITH LEGO FILTERS | 12115 |
| <i>Zhaohui Yang, Yunhe Wang, Chuanjian Liu, Hanting Chen, Chunjing Xu, Boxin Shi, Chao Xu, Chang Xu</i> | |
| SWALP : STOCHASTIC WEIGHT AVERAGING IN LOW PRECISION TRAINING | 12125 |
| <i>Guandao Yang, Tianyi Zhang, Polina Kirichenko, Junwen Bai, Andrew Gordon Wilson, Chris De Sa</i> | |
| ME-NET: TOWARDS EFFECTIVE ADVERSARIAL ROBUSTNESS WITH MATRIX ESTIMATION | 12152 |
| <i>Yuzhe Yang, Guo Zhang, Dina Katabi, Zhi Xu</i> | |
| EFFICIENT NONCONVEX REGULARIZED TENSOR COMPLETION WITH STRUCTURE-AWARE PROXIMAL ITERATIONS | 12174 |
| <i>Quanming Yao, James Tin-Yau Kwok, Bo Han</i> | |
| HIERARCHICALLY STRUCTURED META-LEARNING | 12189 |
| <i>Huaxiu Yao, Ying Wei, Junzhou Huang, Zhenhui Li</i> | |
| TIGHT KERNEL QUERY COMPLEXITY OF KERNEL RIDGE REGRESSION AND KERNEL K -MEANS CLUSTERING | 12209 |
| <i>Taisuke Yasuda, David Woodruff, Manuel Fernandez</i> | |
| UNDERSTANDING GEOMETRY OF ENCODER-DECODER CNNs | 12245 |
| <i>Jong Chul Ye, Woon Kyoung Sung</i> | |
| DEFENDING AGAINST SADDLE POINT ATTACK IN BYZANTINE-ROBUST DISTRIBUTED LEARNING | 12270 |
| <i>Dong Yin, Yudong Chen, Ramchandran Kannan, Peter Bartlett</i> | |
| RADEMACHER COMPLEXITY FOR ADVERSARIAL ROBUST GENERALIZATION | 12296 |
| <i>Dong Yin, Ramchandran Kannan, Peter Bartlett</i> | |
| ARSM: AUGMENT-REINFORCE-SWAP-MERGE ESTIMATOR FOR GRADIENT BACKPROPAGATION THROUGH CATEGORICAL VARIABLES | 12314 |
| <i>Mingchang Yin, Yuguang Yue, Mingyuan Zhou</i> | |
| NAS-BENCH-101: TOWARDS REPRODUCIBLE NEURAL ARCHITECTURE SEARCH | 12334 |
| <i>Chris Ying, Aaron Klein, Eric Christiansen, Esteban Real, Kevin Murphy, Frank Hutter</i> | |
| TAPNET: NEURAL NETWORK AUGMENTED WITH TASK-ADAPTIVE PROJECTION FOR FEW-SHOT LEARNING | 12349 |
| <i>Sung Whan Yoon, Jun Seo, Jaekyun Moon</i> | |
| TOWARDS ACCURATE MODEL SELECTION IN DEEP UNSUPERVISED DOMAIN ADAPTATION | 12362 |
| <i>Kaichao You, Ximei Wang, Mingsheng Long, Michael Jordan</i> | |
| POSITION-AWARE GRAPH NEURAL NETWORKS | 12373 |
| <i>Jiaxuan You, Rex Ying, Jure Leskovec</i> | |

| | |
|--|-------|
| LEARNING NEUROSymbOLIC GENERATIVE MODELS VIA PROGRAM SYNTHESIS | 12382 |
| <i>Halley Young, Osbert Bastani, Mayur Naik</i> | |
| DAG-GNN: DAG STRUCTURE LEARNING WITH GRAPH NEURAL NETWORKS | 12395 |
| <i>Yue Yu, Jie Chen, Tian Gao, Mo Yu</i> | |
| HOW DOES DISAGREEMENT HELP GENERALIZATION AGAINST LABEL CORRUPTION? | 12407 |
| <i>Xingrui Yu, Bo Han, Jiangchao Yao, Gang Niu, Ivor Tsang, Masashi Sugiyama</i> | |
| ON THE COMPUTATION AND COMMUNICATION COMPLEXITY OF PARALLEL SGD WITH DYNAMIC BATCH SIZES FOR STOCHASTIC NON-CONVEX OPTIMIZATION | 12418 |
| <i>Hao Yu, Rong Jin</i> | |
| ON THE LINEAR SPEEDUP ANALYSIS OF COMMUNICATION EFFICIENT MOMENTUM SGD FOR DISTRIBUTED NON-CONVEX OPTIMIZATION | 12431 |
| <i>Hao Yu, Rong Jin, Sen Yang</i> | |
| MULTI-AGENT ADVERSARIAL INVERSE REINFORCEMENT LEARNING | 12468 |
| <i>Lantao Yu, Jiaming Song, Stefano Ermon</i> | |
| DISTRIBUTED LEARNING OVER UNRELIABLE NETWORKS | 12481 |
| <i>Chen Yu, Hanlin Tang, Cedric Renggli, Simon Kassing, Ankit Singla, Dan Alistarh, Ce Zhang, Ji Liu</i> | |
| ONLINE ADAPTIVE PRINCIPAL COMPONENT ANALYSIS AND ITS EXTENSIONS | 12513 |
| <i>Jianjun Yuan, Andrew Lamperski</i> | |
| GENERATIVE MODELING OF INFINITE OCCLUDED OBJECTS FOR COMPOSITIONAL SCENE REPRESENTATION | 12526 |
| <i>Jinyang Yuan, Bin Li, Xiangyang Xue</i> | |
| DIFFERENTIAL INCLUSIONS FOR MODELING NONSMOOTH ADMM VARIANTS: A CONTINUOUS LIMIT THEORY | 12546 |
| <i>Huizhuo Yuan, Yuren Zhou, Chris Junchi Li, Qingyun Sun</i> | |
| TRIMMING THE L1 REGULARIZER: STATISTICAL ANALYSIS, OPTIMIZATION, AND APPLICATIONS TO DEEP LEARNING | 12566 |
| <i>Jihun Yun, Peng Zheng, Eunho Yang, Aurelie Lozano, Aleksandr Aravkin</i> | |
| BAYESIAN NONPARAMETRIC FEDERATED LEARNING OF NEURAL NETWORKS | 12583 |
| <i>Mikhail Yurochkin, Mayank Agarwal, Soumya Ghosh, Kristjan Greenewald, Nghia Hoang, Yasaman Khazaeni</i> | |
| DIRICHLET SIMPLEX NEST AND GEOMETRIC INFERENCE | 12598 |
| <i>Mikhail Yurochkin, Aritra Guha, Yuekai Sun, Xuanlong Nguyen</i> | |
| A CONDITIONAL-GRADIENT-BASED AUGMENTED LAGRANGIAN FRAMEWORK | 12617 |
| <i>Alp Yurtsever, Olivier Fercoq, Volkan Cevher</i> | |
| CONDITIONAL GRADIENT METHODS VIA STOCHASTIC PATH-INTEGRATED DIFFERENTIAL ESTIMATOR | 12640 |
| <i>Alp Yurtsever, Suvrit Sra, Volkan Cevher</i> | |
| CONTEXT-AWARE ZERO-SHOT LEARNING FOR OBJECT RECOGNITION | 12663 |
| <i>Eloi Zablocki, Patrick Bordes, Laure Soulier, Benjamin Piwowarski, Patrick Gallinari</i> | |

PART 19

| | |
|---|-------|
| TIGHTER PROBLEM-DEPENDENT REGRET BOUNDS IN REINFORCEMENT LEARNING WITHOUT DOMAIN KNOWLEDGE USING VALUE FUNCTION BOUNDS | 12676 |
| <i>Andrea Zanette, Emma Brunskill</i> | |
| GLOBAL CONVERGENCE OF BLOCK COORDINATE DESCENT IN DEEP LEARNING | 12685 |
| <i>Jinshan Zeng, Tim Tsz-Kit Lau, Shaobo Lin, Yuan Yao</i> | |
| MAKING CONVOLUTIONAL NETWORKS SHIFT-INVARIANT AGAIN | 12712 |
| <i>Richard Zhang</i> | |
| WARM-STARTING CONTEXTUAL BANDITS: ROBUSTLY COMBINING SUPERVISED AND BANDIT FEEDBACK | 12723 |
| <i>Chicheng Zhang, Alekh Agarwal, Hal Daumé III, John Langford, Sahand Negahban</i> | |
| WHEN SAMPLES ARE STRATEGICALLY SELECTED | 12733 |
| <i>Hanrui Zhang, Yu Cheng, Vincent Conitzer</i> | |
| SELF-ATTENTION GENERATIVE ADVERSARIAL NETWORKS | 12744 |
| <i>Han Zhang, Ian Goodfellow, Dimitris Metaxas, Augustus Odena</i> | |
| CIRCUIT-GNN: GRAPH NEURAL NETWORKS FOR DISTRIBUTED CIRCUIT DESIGN | 12754 |
| <i>Guo Zhang, Hao He, Dina Katabi</i> | |
| LATENTGNN: LEARNING EFFICIENT NON-LOCAL RELATIONS FOR VISUAL RECOGNITION | 12767 |
| <i>Songyang Zhang, Xuming He, Shipeng Yan</i> | |
| NEURAL COLLABORATIVE SUBSPACE CLUSTERING | 12777 |
| <i>Tong Zhang, Pan Ji, Mehrtaash Harandi, Wenbing Huang, Hongdong Li</i> | |
| INCREMENTAL RANDOMIZED SKETCHING FOR ONLINE KERNEL LEARNING | 12787 |
| <i>Xiao Zhang, Shizhong Liao</i> | |
| BRIDGING THEORY AND ALGORITHM FOR DOMAIN ADAPTATION | 12805 |
| <i>Yuchen Zhang, Tianle Liu, Mingsheng Long, Michael Jordan</i> | |
| ADAPTIVE REGRET OF CONVEX AND SMOOTH FUNCTIONS | 12824 |
| <i>Lijun Zhang, Tie-Yan Liu, Zhi-Hua Zhou</i> | |
| RANDOM FUNCTION PRIORS FOR CORRELATION MODELING | 12834 |
| <i>Aonan Zhang, John Paisley</i> | |

| | |
|--|-------|
| CO-REPRESENTATION NETWORK FOR GENERALIZED ZERO-SHOT LEARNING | 12847 |
| <i>Fei Zhang, Guangming Shi</i> | |
| SOLAR: DEEP STRUCTURED REPRESENTATIONS FOR MODEL-BASED REINFORCEMENT LEARNING | 12861 |
| <i>Marvin Zhang, Sharad Vikram, Laura Smith, Pieter Abbeel, Matthew Johnson, Sergey Levine</i> | |
| A COMPOSITE RANDOMIZED INCREMENTAL GRADIENT METHOD | 12875 |
| <i>Junyu Zhang, Lin Xiao</i> | |
| FAST AND STABLE MAXIMUM LIKELIHOOD ESTIMATION FOR INCOMPLETE MULTINOMIAL MODELS | 12898 |
| <i>Chenyang Zhang, Guosheng Yin</i> | |
| THEORETICALLY PRINCIPLED TRADE-OFF BETWEEN ROBUSTNESS AND ACCURACY | 12907 |
| <i>Hongyang Zhang, Yaodong Yu, Jiantao Jiao, Eric Xing, Laurent El Ghaoui, Michael Jordan</i> | |
| LEARNING NOVEL POLICIES FOR TASKS | 12930 |
| <i>Yunbo Zhang, Wenhao Yu, Greg Turk</i> | |
| GREEDY ORTHOGONAL PIVOTING ALGORITHM FOR NON-NEGATIVE MATRIX FACTORIZATION | 12942 |
| <i>Kai Zhang, Sheng Zhang, Jun Liu, Jun Wang, Jie Zhang</i> | |
| INTERPRETING ADVERSARIALY TRAINED CONVOLUTIONAL NEURAL NETWORKS | 12951 |
| <i>Tianyuan Zhang, Zhanxing Zhu</i> | |
| ADAPTIVE MONTE CARLO MULTIPLE TESTING VIA MULTI-ARMED BANDITS | 12967 |
| <i>Martin Zhang, James Zou, David Tse</i> | |
| ON LEARNING INVARIANT REPRESENTATIONS FOR DOMAIN ADAPTATION | 12985 |
| <i>Han Zhao, Remi Tachet Des Combes, Kun Zhang, Geoffrey Gordon</i> | |
| METRIC-OPTIMIZED EXAMPLE WEIGHTS | 13000 |
| <i>Sen Zhao, Mahdi Milani Fard, Harikrishna Narasimhan, Maya Gupta</i> | |
| IMPROVING NEURAL NETWORK QUANTIZATION WITHOUT RETRAINING USING OUTLIER CHANNEL SPLITTING | 13012 |
| <i>Ritchie Zhao, Yuwei Hu, Jordan Dotzel, Chris De Sa, Zhiru Zhang</i> | |
| MAXIMUM ENTROPY-REGULARIZED MULTI-GOAL REINFORCEMENT LEARNING | 13022 |
| <i>Rui Zhao, Xudong Sun, Volker Tresp</i> | |
| STOCHASTIC ITERATIVE HARD THRESHOLDING FOR GRAPH-STRUCTURED SPARSITY OPTIMIZATION | 13036 |
| <i>Baojian Zhou, Feng Chen, Yiming Ying</i> | |
| LOWER BOUNDS FOR SMOOTH NONCONVEX FINITE-SUM OPTIMIZATION | 13059 |
| <i>Dongruo Zhou, Quanquan Gu</i> | |
| LIPSCHITZ GENERATIVE ADVERSARIAL NETS | 13073 |
| <i>Zhiming Zhou, Jiadong Liang, Yuxuan Song, Lantao Yu, Hongwei Wang, Weinan Zhang, Yong Yu, Zhihua Zhang</i> | |
| TOWARD UNDERSTANDING THE IMPORTANCE OF NOISE IN TRAINING NEURAL NETWORKS | 13083 |
| <i>Mo Zhou, Tianyi Liu, Yan Li, Dachao Lin, Enlu Zhou, Tuo Zhao</i> | |
| BAYESNAS: A BAYESIAN APPROACH FOR NEURAL ARCHITECTURE SEARCH | 13116 |
| <i>Hongpeng Zhou, Minghao Yang, Jun Wang, Wei Pan</i> | |
| TRANSFERABLE CLEAN-LABEL POISONING ATTACKS ON DEEP NEURAL NETS | 13141 |
| <i>Chen Zhu, W. Ronny Huang, Hengduo Li, Gavin Taylor, Christoph Studer, Tom Goldstein</i> | |
| IMPROVED DYNAMIC GRAPH LEARNING THROUGH FAULT-TOLERANT SPARSIFICATION | 13155 |
| <i>Chunjiang Zhu, Sabine Storandt, Kam-Yiu Lam, Song Han, Jinbo Bi</i> | |
| POISSON SUBSAMPLED RÉNYI DIFFERENTIAL PRIVACY | 13169 |
| <i>Yuqing Zhu, Yu-Xiang Wang</i> | |
| LEARNING CLASSIFIERS FOR TARGET DOMAIN WITH LIMITED OR NO LABELS | 13188 |
| <i>Pengkai Zhu, Hanxiao Wang, Venkatesh Saligrama</i> | |
| THE ANISOTROPIC NOISE IN STOCHASTIC GRADIENT DESCENT: ITS BEHAVIOR OF ESCAPING FROM SHARP MINIMA AND REGULARIZATION EFFECTS | 13199 |
| <i>Zhanxing Zhu, Jingfeng Wu, Bing Yu, Lei Wu, Jinwen Ma</i> | |
| SURROGATE LOSSES FOR ONLINE LEARNING OF STEPSIZES IN STOCHASTIC NON-CONVEX OPTIMIZATION | 13215 |
| <i>Zhenxun Zhuang, Ashok Cutkosky, Francesco Orabona</i> | |
| LATENT NORMALIZING FLOWS FOR DISCRETE SEQUENCES | 13227 |
| <i>Zachary Ziegler, Alexander Rush</i> | |
| BEATING STOCHASTIC AND ADVERSARIAL SEMI-BANDITS OPTIMALLY AND SIMULTANEOUSLY | 13240 |
| <i>Julian Zimmert, Haipeng Luo, Chen-Yu Wei</i> | |
| FAST CONTEXT ADAPTATION VIA META-LEARNING | 13262 |
| <i>Luisa Zintgraf, Kyriacos Siarlis, Vitaly Kurin, Katja Hofmann, Shimon Whiteson</i> | |
| NATURAL ANALYSTS IN ADAPTIVE DATA ANALYSIS | 13277 |
| <i>Tijana Zrnica, Moritz Hardt</i> | |
| Author Index | |