

31st International Conference on Machine Learning

(ICML 2014)

Journal of Machine Learning Research Volume 32

**Beijing, China
21-26 June 2014**

Volume 1 of 5

ISBN: 978-1-63439-397-3

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© (2014) by the International Machine Learning Society (IMLS)
All rights reserved.

Printed by Curran Associates, Inc. (2014)

For permission requests, please contact the 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

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: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

TABLE OF CONTENTS

Volume 1

A DISCRIMINATIVE LATENT VARIABLE MODEL FOR ONLINE CLUSTERING	1
<i>Rajhans Samdani, Kai-Wei Chang, Dan Roth</i>	
KERNEL MEAN ESTIMATION AND STEIN EFFECT	12
<i>Krikamol Muandet, Kenji Fukumizu, Bharath Sriperumbudur, Arthur Gretton, Bernhard Scholkopf</i>	
DEMYSTIFYING INFORMATION-THEORETIC CLUSTERING	37
<i>Greg Ver Steeg, Aram Galstyan, Fei Sha, Simon Dedeo</i>	
COVERING NUMBER FOR EFFICIENT HEURISTIC-BASED POMDP PLANNING	48
<i>Zongzhang Zhang, David Hsu, Wee Sun Lee</i>	
THE COHERENT LOSS FUNCTION FOR CLASSIFICATION	61
<i>Wenzhuo Yang, Melyvn Sim, Huan Xu</i>	
FAST STOCHASTIC ALTERNATING DIRECTION METHOD OF MULTIPLIERS	78
<i>Leon Wenliang Zhong, James T. Kwok</i>	
ACTIVE DETECTION VIA ADAPTIVE SUBMODULARITY	87
<i>Yuxin Chen, Hiroaki Shioi, César Antonio Fuentes Montesinos, Lian Pin Koh, Serge Wich, Andreas Krause</i>	
ACCELERATED PROXIMAL STOCHASTIC DUAL COORDINATE ASCENT FOR REGULARIZED LOSS MINIMIZATION	111
<i>Shai Shalev-Shwartz, Tong Zhang</i>	
AN ADAPTIVE ACCELERATED PROXIMAL GRADIENT METHOD AND ITS HOMOTOPY CONTINUATION FOR SPARSE OPTIMIZATION	120
<i>Qihang Lin, Lin Xiao</i>	
RECURRENT CONVOLUTIONAL NEURAL NETWORKS FOR SCENE LABELING	151
<i>Pedro O. Pinheiro, Ronan Collobert</i>	
A STATISTICAL PERSPECTIVE ON ALGORITHMIC LEVERAGING	160
<i>Ping Ma, Michael W. Mahoney, Bin Yu</i>	
THOMPSON SAMPLING FOR COMPLEX ONLINE PROBLEMS	169
<i>Aditya Gopalan, Shie Mannor, Yishay Mansour</i>	
BOOSTING MULTI-STEP AUTOREGRESSIVE FORECASTS	187
<i>Souhaib Ben Taieb, Rob J Hyndman</i>	
A STATISTICAL CONVERGENCE PERSPECTIVE OF ALGORITHMS FOR RANK AGGREGATION FROM PAIRWISE DATA	206
<i>Arun Rajkumar, Shivani Agarwal</i>	
SCALING UP APPROXIMATE VALUE ITERATION WITH OPTIONS: BETTER POLICIES WITH FEWER ITERATIONS	228
<i>Timothy A. Mann, Shie Mannor</i>	
LATENT BANDITS	251
<i>Odalric-Ambrym Maillard, Shie Mannor</i>	
FAST ALLOCATION OF GAUSSIAN PROCESS EXPERTS	260
<i>Trung V. Nguyen, Edwin V. Bonilla</i>	
VON MISES-FISHER CLUSTERING MODELS	269
<i>Siddharth Gopal, Yiming Yang</i>	
CONVERGENCE RATES FOR PERSISTENCE DIAGRAM ESTIMATION IN TOPOLOGICAL DATA ANALYSIS	303
<i>Frédéric Chazal, Marc Glisse, Catherine Labruère, Bertrand Michel</i>	
BUFFER K-D TREES: PROCESSING MASSIVE NEAREST NEIGHBOR QUERIES ON GPU	312
<i>Fabian Gieseke, Justin Heineremann, Cosmin Oancea, Christian Igel</i>	
AUSTERITY IN MCMC LAND: CUTTING THE METROPOLIS-HASTINGS BUDGET	321
<i>Anoop Korattikara, Yutian Chen, Max Welling</i>	
UNDERSTANDING THE LIMITING FACTORS OF TOPIC MODELING VIA POSTERIOR CONTRACTION ANALYSIS	337
<i>Jian Tang, Zhaoshi Meng, Xuanlong Nguyen, Qiaozhu Mei, Ming Zhang</i>	
THE INVERSE REGRESSION TOPIC MODEL	346
<i>Maxim Rabinovich, David M. Blei</i>	
A CONSISTENT HISTOGRAM ESTIMATOR FOR EXCHANGEABLE GRAPH MODELS	360
<i>Stanley H. Chan, Edoardo M. Airoldi</i>	

LATENT VARIABLE COPULA INFERENCE FOR BUNDLE PRICING FROM RETAIL TRANSACTION DATA	369
<i>Benjamin Letham, Wei Sun, Anshul Sheopuri</i>	
TOWARDS MINIMAX ONLINE LEARNING WITH UNKNOWN TIME HORIZON	378
<i>Haipeng Luo, Robert E. Schapire</i>	
FACTORIZED POINT PROCESS INTENSITIES: A SPATIAL ANALYSIS OF PROFESSIONAL BASKETBALL	398
<i>Andrew Miller, Luke Bornn, Ryan Adams, Kirk Goldsberry</i>	
MARGINS, KERNELS AND NON-LINEAR SMOOTHED PERCEPTRONS	415
<i>Aaditya Ramdas, Javier Pena</i>	
ROBUST REGBAYES: SELECTIVELY INCORPORATING FIRST-ORDER LOGIC DOMAIN KNOWLEDGE INTO BAYESIAN MODELS	425
<i>Shike Mei, Jun Zhu, Xiaojin Zhu</i>	
LEARNING THEORY AND ALGORITHMS FOR REVENUE OPTIMIZATION IN SECOND-PRICE AUCTIONS WITH RESERVE	443
<i>Mehryar Mohri, Andres Munoz Medina</i>	
LOW-DENSITY PARITY CONSTRAINTS FOR HASHING-BASED DISCRETE INTEGRATION	452
<i>Stefano Ermon, Carla P. Gomes, Ashish Sabharwal, Bart Selman</i>	
PREDICTION WITH LIMITED ADVICE AND MULTIARMED BANDITS WITH PAID OBSERVATIONS	463
<i>Yevgeny Seldin, Peter Bartlett, Koby Crammer, Yasin Abbasi-Yadkori</i>	
BAYESIAN NONPARAMETRIC MULTILEVEL CLUSTERING WITH GROUP-LEVEL CONTEXTS	481
<i>Vu Nguyen, Dinh Phung, Xuanlong Nguyen, Svetha Venkatesh, Hung Hai Bui</i>	
LARGE-MARGIN METRIC LEARNING FOR CONSTRAINED PARTITIONING PROBLEMS	502
<i>Rémi Lajugie, Sylvain Arlot, Francis Bach</i>	
WASSERSTEIN PROPAGATION FOR SEMI-SUPERVISED LEARNING	511
<i>Justin Solomon, Raif M. Rustamov, Leonidas Guibas, Adrian Butscher</i>	
MAX-MARGIN INFINITE HIDDEN MARKOV MODELS	520
<i>Aonan Zhang, Jun Zhu, Bo Zhang</i>	
EFFICIENT APPROXIMATION OF CROSS-VALIDATION FOR KERNEL METHODS USING BOULIGAND INFLUENCE FUNCTION	529
<i>Yong Liu, Shali Jiang, Shizhong Liao</i>	
GENERALIZED EXPONENTIAL CONCENTRATION INEQUALITY FOR RÉNYI DIVERGENCE ESTIMATION	538
<i>Shashank Singh, Barnabás Póczos</i>	
BOOSTING WITH ONLINE BINARY LEARNERS FOR THE MULTICLASS BANDIT PROBLEM	547
<i>Shang-Tse Chen, Hsuan-Tien Lin, Chi-Jen Lu</i>	
OPTIMAL BUDGET ALLOCATION: THEORETICAL GUARANTEE AND EFFICIENT ALGORITHM	556
<i>Tasuku Soma, Naonori Kakimura, Kazuhiro Inaba, Ken-Ichi Kawarabayashi</i>	
COMPUTING PARAMETRIC RANKING MODELS VIA RANK-BREAKING	569
<i>Hossein Azari Soufiani, David C. Parkes, Lirong Xia</i>	
TRACKING ADVERSARIAL TARGETS	581
<i>Yasin Abbasi-Yadkori, Peter Bartlett, Varun Kanade</i>	
ONLINE BAYESIAN PASSIVE-AGGRESSIVE LEARNING	595
<i>Tianlin Shi, Jun Zhu</i>	
DETERMINISTIC POLICY GRADIENT ALGORITHMS	605
<i>David Silver, Guy Lever, Nicolas Heess, Thomas Degris, Daan Wierstra, Martin Riedmiller</i>	
MODELING CORRELATED ARRIVAL EVENTS WITH LATENT SEMI-MARKOV PROCESSES	620
<i>Wenzhao Lian, Vinayak Rao, Brian Eriksson, Lawrence Carin</i>	
TOWARDS SCALING UP MARKOV CHAIN MONTE CARLO: AN ADAPTIVE SUBSAMPLING APPROACH	630
<i>Rémi Bardenet, Arnaud Doucet, Chris Holmes</i>	
DIAGNOSIS DETERMINATION: DECISION TREES OPTIMIZING SIMULTANEOUSLY WORST AND EXPECTED TESTING COST	654
<i>Ferdinando Cicalese, Eduardo Laber, Aline Medeiros Saettler</i>	
CONDENSED FILTER TREE FOR COST-SENSITIVE MULTI-LABEL CLASSIFICATION	663
<i>Chun-Liang Li, Hsuan-Tien Lin</i>	

ON MEASURE CONCENTRATION OF RANDOM MAXIMUM A-POSTERIORI PERTURBATIONS.....	674
<i>Francesco Orabona, Tamir Hazan, Anand D. Sarwate, Tommi S. Jaakkola</i>	
BIAS IN NATURAL ACTOR-CRITIC ALGORITHMS	693
<i>Philip S. Thomas</i>	
DIMENSION-FREE CONCENTRATION BOUNDS ON HANKEL MATRICES FOR SPECTRAL LEARNING.....	701
<i>Francois Denis, Mattias Gybels, Amaury Habrard</i>	
ON MODELLING NON-LINEAR TOPICAL DEPENDENCIES.....	710
<i>Zhixing Li, Siqiang Wen, Juanzi Li, Peng Zhang, Jie Tang</i>	
A DEEP AND TRACTABLE DENSITY ESTIMATOR.....	719
<i>Benigno Uria, Iain Murray, Hugo Larochelle</i>	
(NEAR) DIMENSION INDEPENDENT RISK BOUNDS FOR DIFFERENTIALLY PRIVATE LEARNING.....	728
<i>Prateek Jain, Abhradeep Thakurta</i>	
QUASI-MONTE CARLO FEATURE MAPS FOR SHIFT-INVARIANT KERNELS.....	740
<i>Jiyan Yang, Vikas Sindhwani, Haim Avron, Michaelw. Mahoney</i>	
DISCRIMINATIVE FEATURES VIA GENERALIZED EIGENVECTORS	756
<i>Nikos Karampatziakis, Paul Mineiro</i>	
FORWARD-BACKWARD GREEDY ALGORITHMS FOR GENERAL CONVEX SMOOTH FUNCTIONS OVER A CARDINALITY CONSTRAINT.....	765
<i>Ji Liu, Ryohei Fujimaki, Jieping Ye</i>	
ONLINE LEARNING IN MARKOV DECISION PROCESSES WITH CHANGING COST SEQUENCES.....	774
<i>Travis Dick, Andras Gyorgy, Csaba Szepesvári</i>	
UNIMODAL BANDITS: REGRET LOWER BOUNDS AND OPTIMAL ALGORITHMS	799
<i>Richard Combes, Alexandre Proutiere</i>	
MAXIMUM MEAN DISCREPANCY FOR CLASS RATIO ESTIMATION: CONVERGENCE BOUNDS AND KERNEL SELECTION	808
<i>Arun Iyer, Saketh Nath, Sunita Sarawagi</i>	
ASYMPTOTICALLY CONSISTENT ESTIMATION OF THE NUMBER OF CHANGE POINTS IN HIGHLY DEPENDENT TIME SERIES	824
<i>Azadeh Khaleghi, Daniil Ryabko</i>	
COORDINATE-DESCENT FOR LEARNING ORTHOGONAL MATRICES THROUGH GIVENS ROTATIONS.....	833
<i>Uri Shalit, Gal Chechik</i>	
DENSIFYING ONE PERMUTATION HASHING VIA ROTATION FOR FAST NEAR NEIGHBOR SEARCH	846
<i>Anshumali Shrivastava, Ping Li</i>	
A DIVIDE-AND-CONQUER SOLVER FOR KERNEL SUPPORT VECTOR MACHINES	855
<i>Cho-Jui Hsieh, Inderjit S. Dhillon</i>	
NUCLEAR NORM MINIMIZATION VIA ACTIVE SUBSPACE SELECTION	871
<i>Cho-Jui Hsieh, Peder A. Olsen</i>	
PROVABLE BOUNDS FOR LEARNING SOME DEEP REPRESENTATIONS.....	883
<i>Sanjeev Arora, Aditya Bhaskara, Rong Ge, Tengyu Ma</i>	
LARGE-SCALE MULTI-LABEL LEARNING WITH MISSING LABELS	892
<i>Hsiang-Fu Yu, Prateek Jain, Purushottam Kar, Inderjit S. Dhillon</i>	
LEARNING GRAPHS WITH A FEW HUBS	920
<i>Rashish Tandon, Pradeep Ravikumar</i>	

Volume 2

AGNOSTIC BAYESIAN LEARNING OF ENSEMBLES.....	934
<i>Alexandre Lacoste, Hugo Larochelle, Mario Marchand, François Laviolette</i>	
TOWARDS AN OPTIMAL STOCHASTIC ALTERNATING DIRECTION METHOD OF MULTIPLIERS	944
<i>Samaneh Azadi, Suvrit Sra</i>	
SPHERICAL HAMILTONIAN MONTE CARLO FOR CONSTRAINED TARGET DISTRIBUTIONS	960
<i>Shiwei Lan, Bo Zhou, Babak Shahbaba</i>	

EFFICIENT CONTINUOUS-TIME MARKOV CHAIN ESTIMATION	969
<i>Monir Hajiaghayi, Bonnie Kirkpatrick, Liangliang Wang, Alexandre Bouchard-Cote</i>	
DECAF: A DEEP CONVOLUTIONAL ACTIVATION FEATURE FOR GENERIC VISUAL RECOGNITION	988
<i>Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, Trevor Darrell</i>	
MAKING THE MOST OF BAG OF WORDS: SENTENCE REGULARIZATION WITH ALTERNATING DIRECTION METHOD OF MULTIPLIERS	997
<i>Dani Yogatama, Noah A. Smith</i>	
NARROWING THE GAP: RANDOM FORESTS IN THEORY AND IN PRACTICE	1006
<i>Misha Denil, David Matheson, Nando De Freitas</i>	
COHERENT MATRIX COMPLETION	1017
<i>Yudong Chen, Srinadh Bhojanapalli, Sujay Sanghavi, Rachel Ward</i>	
ADMIXTURE OF POISSON MRFs: A TOPIC MODEL WITH WORD DEPENDENCIES	1037
<i>David I. Inouye, Pradeep Ravikumar, Inderjit S. Dhillon</i>	
TRUE ONLINE TD(λ)	1048
<i>Harm Van Seijen, Richard S. Sutton</i>	
MEMORY EFFICIENT KERNEL APPROXIMATION	1057
<i>Cho-Jui Hsieh, Inderjit S. Dhillon</i>	
LEARNING SUM-PRODUCT NETWORKS WITH DIRECT AND INDIRECT VARIABLE INTERACTIONS	1070
<i>Amirmohammad Rooshenas, Daniel Lowd</i>	
HAMILTONIAN MONTE CARLO WITHOUT DETAILED BALANCE	1081
<i>Jascha Sohl-Dickstein, Mayur Mudigonda, Michael R. Dewese</i>	
FILTERING WITH ABSTRACT PARTICLES	1089
<i>Jacob Steinhardt, Percy Liang</i>	
STOCHASTIC DUAL COORDINATE ASCENT WITH ALTERNATING DIRECTION METHOD OF MULTIPLIERS	1100
<i>Taiji Suzuki</i>	
DEEP SUPERVISED AND CONVOLUTIONAL GENERATIVE STOCHASTIC NETWORK FOR PROTEIN SECONDARY STRUCTURE PREDICTION	1121
<i>Jian Zhou, Olga G. Troyanskaya</i>	
AN EFFICIENT APPROACH FOR ASSESSING HYPERPARAMETER IMPORTANCE	1130
<i>Frank Hutter, Holger Hoos, Kevin Leyton-Brown</i>	
AN INFORMATION GEOMETRY OF STATISTICAL MANIFOLD LEARNING	1145
<i>Ke Sun, Stéphane Marchand-Maillet</i>	
RELATIVE UPPER CONFIDENCE BOUND FOR THE K-ARMED DUELING BANDIT PROBLEM	1162
<i>Masrour Zoghi, Shimon Whiteson, Remi Munos, Maarten De Rijke</i>	
COMPACT RANDOM FEATURE MAPS	1176
<i>Raffay Hamid, Ying Xiao, Alex Gittens, Dennis Decoste</i>	
CONCENTRATION IN UNBOUNDED METRIC SPACES AND ALGORITHMIC STABILITY	1185
<i>Aryeh Kontorovich</i>	
HEAVY-TAILED REGRESSION WITH A GENERALIZED MEDIAN-OF-MEANS	1196
<i>Daniel Hsu, Sivan Sabato</i>	
SPECTRAL BANDITS FOR SMOOTH GRAPH FUNCTIONS	1205
<i>Michal Valko, Rémi Munos, Branislav Kveton, Tomas Kocak</i>	
ROBUST PRINCIPAL COMPONENT ANALYSIS WITH COMPLEX NOISE	1216
<i>Qian Zhao, Deyu Meng, Zongben Xu, Wangmeng Zuo, Lei Zhang</i>	
SCALABLE SEMIDEFINITE RELAXATION FOR MAXIMUM A POSTERIOR ESTIMATION	1227
<i>Qixing Huang, Yuxin Chen, Leonidas Guibas</i>	
SQUARE DEAL: LOWER BOUNDS AND IMPROVED RELAXATIONS FOR TENSOR RECOVERY	1242
<i>Cun Mu, Bo Huang, John Wright, Donald Goldfarb</i>	
AUTOMATED INFERENCE OF POINT OF VIEW FROM USER INTERACTIONS IN COLLECTIVE INTELLIGENCE VENUES	1251
<i>Sanmay Das, Allen Lavoie</i>	
RANK-ONE MATRIX PURSUIT FOR MATRIX COMPLETION	1260
<i>Zheng Wang, Ming-Jun Lai, Zhaosong Lu, Wei Fan, Hasan Davulcu, Jieping Ye</i>	
NEAR-OPTIMAL JOINT OBJECT MATCHING VIA CONVEX RELAXATION	1269
<i>Yuxin Chen, Leonidas Guibas, Qixing Huang</i>	
CONVEX TOTAL LEAST SQUARES	1301
<i>Dmitry Malioutov, Nikolai Slavov</i>	

ON P-NORM PATH FOLLOWING IN MULTIPLE KERNEL LEARNING FOR NON-LINEAR FEATURE	1310
<i>Pratik Jawanpuria, Manik Varma, J. Saketha Nath</i>	
GRADIENT HARD THRESHOLDING PURSUIT FOR SPARSITY-CONSTRAINED OPTIMIZATION	1322
<i>Xiao-Tong Yuan, Ping Li, Tong Zhang</i>	
A UNIFIED FRAMEWORK FOR CONSISTENCY OF REGULARIZED LOSS MINIMIZERS	1331
<i>Jean Honorio, Tommi Jaakkola</i>	
GEODESIC DISTANCE FUNCTION LEARNING VIA HEAT FLOW ON VECTOR FIELDS	1346
<i>Binbin Lin, Ji Yang, Xiaofei He, Jieping Ye</i>	
NEAR-OPTIMALLY TEACHING THE CROWD TO CLASSIFY	1355
<i>Adish Singla, Ilija Bogunovic, Gábor Bartók, Amin Karbasi, Andreas Krause</i>	
ON THE CONVERGENCE OF NO-REGRET LEARNING IN SELFISH ROUTING	1379
<i>Walid Krichene, Benjamin Drighès, Alexandre Bayen</i>	
IMPROVING OFFLINE EVALUATION OF CONTEXTUAL BANDIT ALGORITHMS VIA BOOTSTRAPPING TECHNIQUES	1390
<i>Olivier Nicol, Jérémie Mary, Philippe Preux</i>	
SCALING UP ROBUST MDPS USING FUNCTION APPROXIMATION	1401
<i>Aviv Tamar, Shie Mannor, Huan Xu</i>	
MARGINAL STRUCTURED SVM WITH HIDDEN VARIABLES	1416
<i>Wei Ping, Qiang Liu, Alexander Ihler</i>	
LINEAR AND PARALLEL LEARNING OF MARKOV RANDOM FIELDS	1428
<i>Yariv Dror Mizrahi, Misha Denil, Nando De Freitas</i>	
PITFALLS IN THE USE OF PARALLEL INFERENCE FOR THE DIRICHLET PROCESS	1437
<i>Yarin Gal, Zoubin Ghahramani</i>	
OPTIMAL PAC MULTIPLE ARM IDENTIFICATION WITH APPLICATIONS TO CROWDSOURCING	1446
<i>Yuan Zhou, Carnegie Mellon U, Xi Chen, Uc Berkeley, Jian Li, U. Tsinghua</i>	
DEEP GENERATIVE STOCHASTIC NETWORKS TRAINABLE BY BACKPROP	1470
<i>Yoshua Bengio, Éric Thibodeau-Laufer, Guillaume Alain, Jason Yosinski</i>	
A HIGHLY SCALABLE PARALLEL ALGORITHM FOR ISOTROPIC TOTAL VARIATION MODELS	1486
<i>Jie Wang, Qingyang Li, Sen Yang, Wei Fan, Peter Wonka, Jieping Ye</i>	
SUPPLEMENTARY MATERIALS FOR “STATISTICAL-COMPUTATIONAL PHASE TRANSITIONS IN PLANTED MODELS: THE HIGH-DIMENSIONAL SETTING”	1502
<i>Yudong Chen, Jiaming Xu, Jie Wang, Qingyang Li, Sen Yang, Wei Fan, Peter Wonka, Jieping Ye</i>	
GAUSSIAN PROCESS OPTIMIZATION WITH MUTUAL INFORMATION	1515
<i>Emile Contal, Vianney Perchet, Nicolas Vayatis</i>	
AGGREGATING ORDINAL LABELS FROM CROWDS BY MINIMAX CONDITIONAL ENTROPY	1524
<i>Dengyong Zhou, Qiang Liu, John C. Platt, Christopher Meek</i>	
EXCHANGEABLE VARIABLE MODELS	1533
<i>Mathias Niepert, Pedro Domingos</i>	
CLUSTERING IN THE PRESENCE OF BACKGROUND NOISE	1545
<i>Shai Ben-David, Nika Haghtalab</i>	
SAFE SCREENING WITH VARIATIONAL INEQUALITIES AND ITS APPLICATION TO LASSO	1556
<i>Jun Liu, Zheng Zhao, Jie Wang, Jieping Ye</i>	
LEARNING THE CONSISTENT BEHAVIOR OF COMMON USERS FOR TARGET NODE PREDICTION ACROSS SOCIAL NETWORKS	1572
<i>Shan-Hung Wu, Hao-Heng Chien, Kuan-Hua Lin, Philip S. Yu</i>	
SIGNAL RECOVERY FROM POOLING REPRESENTATIONS	1585
<i>Joan Bruna, Arthur Szlam, Yann Lecun</i>	
PAC-INSPIRED OPTION DISCOVERY IN LIFELONG REINFORCEMENT LEARNING	1599
<i>Emma Brunskill, Lihong Li</i>	
MULTI-LABEL CLASSIFICATION VIA FEATURE-AWARE IMPLICIT LABEL SPACE ENCODING	1611
<i>Zijia Lin, Guiguang Ding, Mingqing Hu, Jianmin Wang</i>	
SCALABLE GAUSSIAN PROCESS STRUCTURED PREDICTION FOR GRID FACTOR GRAPH APPLICATIONS	1625
<i>Sébastien Bratières, Novi Quadrianto, Sebastian Nowozin, Zoubin Ghahramani</i>	

ANOMALY RANKING AS SUPERVISED BIPARTITE RANKING	1637
<i>Stephan Clemençon, Sylvain Robbiano</i>	
HIERARCHICAL QUASI-CLUSTERING METHODS FOR ASYMMETRIC NETWORKS	1646
<i>Gunnar Carlsson, Facundo Mémoli, Alejandro Ribeiro, Santiago Segarra</i>	
RECTANGULAR TILING PROCESS	1665
<i>Masahiro Nakano, Katsuhiko Ishiguro, Akisato Kimura, Takeshi Yamada, Naonori Ueda</i>	
TWO-STAGE METRIC LEARNING	1683
<i>Jun Wang, Ke Sun, Fei Sha, Stéphane Marchand-Maillet, Alexandros Kalousis</i>	
STOCHASTIC INFERENCE FOR SCALABLE PROBABILISTIC MODELING OF BINARY MATRICES	1693
<i>José Miguel Hernández-Lobato, Neil Houlsby, Zoubin Ghahramani</i>	
ELEMENTARY ESTIMATORS FOR HIGH-DIMENSIONAL LINEAR REGRESSION	1711
<i>Eunho Yang, Aurélie C. Lozano, Pradeep Ravikumar</i>	
ELEMENTARY ESTIMATORS FOR SPARSE COVARIANCE MATRICES AND OTHER STRUCTURED MOMENTS	1723
<i>Eunho Yang, Aurélie C. Lozano, Pradeep Ravikumar</i>	
GRAPH-BASED SEMI-SUPERVISED LEARNING: REALIZING POINTWISE SMOOTHNESS PROBABILISTICALLY	1736
<i>Yuan Fang, Kevin Chen-Chuan Chang, Hady W. Lauw</i>	
BAYESIAN MAX-MARGIN MULTI-TASK LEARNING WITH DATA AUGMENTATION	1755
<i>Chengtao Li, Jun Zhu, Jianfei Chen</i>	
SPARSE REINFORCEMENT LEARNING VIA CONVEX OPTIMIZATION	1764
<i>Zhiwei Qin, Weichang Li, Firdaus Janoos</i>	
GAUSSIAN PROCESS CLASSIFICATION AND ACTIVE LEARNING WITH MULTIPLE ANNOTATORS	1774
<i>Filipe Rodrigues, Francisco C. Pereira, Bernardete Ribeiro</i>	
STRUCTURED PREDICTION OF NETWORK RESPONSE	1787
<i>Hongyu Su, Aristides Gionis, Juho Rousu</i>	
AN ANALYSIS OF STATE-RELEVANCE WEIGHTS AND SAMPLING DISTRIBUTIONS ON L1-REGULARIZED APPROXIMATE LINEAR PROGRAMMING APPROXIMATION ACCURACY	1799
<i>Gavin Taylor, Connor Geer, David Piekut</i>	
OPTIMIZATION EQUIVALENCE OF DIVERGENCES IMPROVES NEIGHBOR EMBEDDING	1808
<i>Zhirong Yang, Jaakko Peltonen, Samuel Kaski</i>	
AN ASYNCHRONOUS PARALLEL STOCHASTIC COORDINATE DESCENT ALGORITHM	1840
<i>Ji Liu, Stephen J. Wright, Christopher Ré, Victor Bittorf, Srikrishna Sridhar</i>	
CONSISTENCY OF CAUSAL INFERENCE UNDER THE ADDITIVE NOISE MODEL	1849
<i>Samory Kpotufe, Eleni Sgouritsa, Dominik Janzing, Bernhard Scholkopf</i>	

Volume 3

 Globally Convergent Parallel Map LP Relaxation Solver Using the Frank-Wolfe Algorithm	1858
<i>Alexander G. Schwing, Tamir Hazan, Marc Pollefeys, Raquel Urtasun</i>	
Linear Programming for Large-Scale Markov Decision Problems	1867
<i>Yasin Abbasi-Yadkori, Peter L. Bartlett, Alan Malek</i>	
New Primal SVM Solver with Linear Computational Cost for Big Data Classifications	1883
<i>Feiping Nie, Yizhen Huang, Xiaoqian Wang, Heng Huang</i>	
Memory (and Time) Efficient Sequential Monte Carlo	1892
<i>Seong-Hwan Jun, Alexandre Bouchard-Cote</i>	
Scaling SVM and Least Absolute Deviations via Exact Data Reduction	1912
<i>Jie Wang, Peter Wonka, Jieping Ye</i>	
Latent Semantic Representation Learning for Scene Classification	1928
<i>Xin Li, Yuhong Guo</i>	
Least Squares Revisited: Scalable Approaches for Multi-Class Prediction	1937
<i>Alekh Agarwal, Sham M. Kakade, Nikos Karampatziakis, Le Song, Gregory Valiant</i>	
Local Algorithms for Interactive Clustering	1946
<i>Pranjal Awasthi, Maria-Florina Balcan, Konstantin Voevodski</i>	
Model-Based Relational RL when Object Existence is Partially Observable	1955
<i>Ngo Anh Vien, Marc Toussaint</i>	

COLD-START ACTIVE LEARNING WITH ROBUST ORDINAL MATRIX FACTORIZATION	1964
<i>Neil Houlsby, José Miguel Hernández-Lobato, Zoubin Ghahramani</i>	
A NEW $Q(\lambda)$ WITH INTERIM FORWARD VIEW AND MONTE CARLO EQUIVALENCE	1973
<i>Richard S. Sutton, A. Rupam Mahmood, Doina Precup, Hado Van Hasselt</i>	
ON ROBUSTNESS AND REGULARIZATION OF STRUCTURAL SUPPORT VECTOR MACHINES	1989
<i>Mohamadali Torkamani, Daniel Lowd</i>	
GUESS-AVERSE LOSS FUNCTIONS FOR COST-SENSITIVE MULTICLASS BOOSTING.....	2000
<i>Oscar Beijbom, Mohammad Saberian, David Kriegman, Nuno Vasconcelos</i>	
MULTIMODAL NEURAL LANGUAGE MODELS.....	2012
<i>Ryan Kiros, Ruslan Salakhutdinov, Richard Zemel</i>	
FAST LARGE-SCALE OPTIMIZATION BY UNIFYING STOCHASTIC GRADIENT AND QUASI-NEWTON METHODS.....	2026
<i>Jascha Sohl-Dickstein, Ben Poole, Surya Ganguli</i>	
ALTERNATING MINIMIZATION FOR MIXED LINEAR REGRESSION	2035
<i>Xinyang Yi, Constantine Caramanis, Sujay Sanghavi</i>	
STOCHASTIC NEIGHBOR COMPRESSION.....	2051
<i>Matt J. Kusner, Stephen Tyree, Kilian Weinberger, Kunal Agrawal</i>	
ROBUST LEARNING UNDER UNCERTAIN TEST DISTRIBUTIONS: RELATING COVARIATE SHIFT TO MODEL MISSPECIFICATION.....	2060
<i>Junfeng Wen, Chun-Nam Yu, Russell Greiner</i>	
NONPARAMETRIC ESTIMATION OF MULTI-VIEW LATENT VARIABLE MODELS	2082
<i>Le Song, Animashree Anandkumar, Bo Dai, Bo Xie</i>	
STRUCTURED GENERATIVE MODELS OF NATURAL SOURCE CODE	2098
<i>Chris J. Maddison, Daniel Tarlow</i>	
A SINGLE-PASS ALGORITHM FOR EFFICIENTLY RECOVERING SPARSE CLUSTER CENTERS OF HIGH-DIMENSIONAL DATA.....	2112
<i>Jinfeng Yi, Lijun Zhang, Jun Wang, Rong Jin, Anil K. Jain</i>	
STATISTICAL ANALYSIS OF STOCHASTIC GRADIENT METHODS FOR GENERALIZED LINEAR MODELS	2128
<i>Panos Toulis, Jason Rennie, Edoardo M. Airoldi</i>	
CODING FOR RANDOM PROJECTIONS.....	2137
<i>Ping Li, Michael Mitzenmacher, Anshumali Shrivastava</i>	
FAST COMPUTATION OF WASSERSTEIN BARYCENTERS	2146
<i>Marco Cuturi, Arnaud Doucet</i>	
GLOBAL GRAPH KERNELS USING GEOMETRIC EMBEDDINGS	2155
<i>Fredrik D. Johansson, Vinay Jethava, Devdatt Dubhashi, Chiranjib Bhattacharyya</i>	
TOPIC MODELING USING TOPICS FROM MANY DOMAINS, LIFELONG LEARNING AND BIG DATA	2168
<i>Zhiyuan Chen, Bing Liu</i>	
K-MEANS RECOVERS ICA FILTERS WHEN INDEPENDENT COMPONENTS ARE SPARSE	2178
<i>Alon Vinnikov, Shai Shalev-Shwartz</i>	
LEARNING MIXTURES OF LINEAR CLASSIFIERS.....	2199
<i>Yuekai Sun, Stratis Ioannidis, Andrea Montanari</i>	
THE FALLING FACTORIAL BASIS AND ITS STATISTICAL APPLICATIONS	2234
<i>Yu-Xiang Wang, Alex Smola, Ryan J. Tibshirani</i>	
NONMYOPIC -BAYES-OPTIMAL ACTIVE LEARNING OF GAUSSIAN PROCESSES	2254
<i>Trong Nghia Hoang, Kian Hsiang Low, Patrick Jaillet, Mohan Kankanhalli</i>	
A UNIFYING VIEW OF REPRESENTER THEOREMS.....	2284
<i>Andreas Argyriou, Francesco Dinuzzo</i>	
ONLINE CLUSTERING OF BANDITS	2296
<i>Claudio Gentile, Shuai Li, Giovanni Zappella</i>	
COLD-START ACTIVE LEARNING WITH ROBUST ORDINAL MATRIX FACTORIZATION	2316
<i>Neil Houlsby, José Miguel Hernández-Lobato, Zoubin Ghahramani</i>	
MULTIVARIATE MAXIMAL CORRELATION ANALYSIS	2346
<i>Hoang Vu Nguyen, Emmanuel Muller, Jilles Vreeken, Pavel Efros, Klemens Bohm</i>	
EFFICIENT LABEL PROPAGATION	2355
<i>Yasuhiro Fujiwara, Go Irie</i>	
ESTIMATING DIFFUSION NETWORK STRUCTURES: RECOVERY CONDITIONS, SAMPLE COMPLEXITY & SOFT-THRESHOLDING ALGORITHM.....	2364
<i>Hadi Daneshmand, Manuel Gomez-Rodriguez, Le Song, Bernhard Scholkopf</i>	

COUPLED GROUP LASSO FOR WEB-SCALE CTR PREDICTION IN DISPLAY ADVERTISING	2379
<i>Ling Yan, Wu-Jun Li, Gui-Rong Xue, Dingyi Han</i>	
PUTTING MRFS ON A TENSOR TRAIN	2388
<i>Alexander Novikov, Anton Rodomanov, Anton Osokin, Dmitry Vetrov</i>	
EFFICIENT ALGORITHMS FOR ROBUST ONE-BIT COMPRESSIVE SENSING	2400
<i>Lijun Zhang, Jinfeng Yi, Rong Jin</i>	
LEARNING COMPLEX NEURAL NETWORK POLICIES WITH TRAJECTORY OPTIMIZATION	2411
<i>Sergey Levine, Vladlen Koltun</i>	
COMPOSITE QUANTIZATION FOR APPROXIMATE NEAREST NEIGHBOR SEARCH	2421
<i>Ting Zhang, Chao Du, Jingdong Wang</i>	
LOCAL ORDINAL EMBEDDING	2440
<i>Yoshikazu Terada, Ulrike Von Luxburg</i>	
REDUCING DUELING BANDITS TO CARDINAL BANDITS	2459
<i>Nir Ailon, Zohar Karnin, Thorsten Joachims</i>	
LARGE-MARGIN WEAKLY SUPERVISED DIMENSIONALITY REDUCTION	2472
<i>Chang Xu, Dacheng Tao, Chao Xu, Yong Rui</i>	
JOINT INFERENCE OF MULTIPLE LABEL TYPES IN LARGE NETWORKS	2483
<i>Deepayan Chakrabarti, Stanislaw Funiak, Jonathan Chang, Sofus A. Macskassy</i>	
HARD-MARGIN ACTIVE LINEAR REGRESSION	2492
<i>Elad Hazan, Zohar Karnin</i>	
MAXIMUM MARGIN MULTICLASS NEAREST NEIGHBORS	2501
<i>Aryeh Kontorovich, Roi Weiss</i>	
COMBINATORIAL PARTIAL MONITORING GAME WITH LINEAR FEEDBACK AND ITS APPLICATIONS	2512
<i>Tian Lin, Bruno Abrahao, Robert Kleinberg, John C. S. Lui, Wei Chen</i>	
SPARSE META-GAUSSIAN INFORMATION BOTTLENECK	2538
<i>Mélanie Rey, Thomas J. Fuchs, Volker Roth</i>	
NONPARAMETRIC ESTIMATION OF R'ENYI DIVERGENCE AND FRIENDS	2550
<i>Akshay Krishnamurthy, Kirthevasan Kandasamy, Barnabas Poczos, Larry Wasserman</i>	
ROBUST INVERSE COVARIANCE ESTIMATION UNDER NOISY MEASUREMENTS	2572
<i>Jun-Kun Wang, Shou-De Lin</i>	
BAYESIAN OPTIMIZATION WITH INEQUALITY CONSTRAINTS	2581
<i>Jacob R. Gardner, Matt J. Kusner, Zhixiang Xu, Kilian Q. Weinberger, John P. Cunningham</i>	
CIRCULANT BINARY EMBEDDING	2592
<i>Felix X. Yu, Sanjiv Kumar, Yunchao Gong, Shih-Fu Chang</i>	
MULTIPLE TESTING UNDER DEPENDENCE VIA SEMIPARAMETRIC GRAPHICAL MODELS	2601
<i>Jie Liu, Chunming Zhang, Elizabeth Burnside, David Page</i>	
MAKING FISHER DISCRIMINANT ANALYSIS SCALABLE	2614
<i>Bojun Tu, Zhihua Zhang, Shusen Wang, Hui Qian</i>	
HIERARCHICAL DIRICHLET SCALING PROCESS	2633
<i>Dongwoo Kim, Alice Oh</i>	
APPROXIMATION ANALYSIS OF STOCHASTIC GRADIENT LANGEVIN DYNAMICS BY USING FOKKER-PLANCK EQUATION AND ITO PROCESS	2647
<i>Issei Sato, Hiroshi Nakagawa</i>	
A PAC-BAYESIAN BOUND FOR LIFELONG LEARNING	2656
<i>Anastasia Pentina, Christoph H. Lampert</i>	
COMMUNICATION-EFFICIENT DISTRIBUTED OPTIMIZATION USING AN APPROXIMATE NEWTON-TYPE METHOD	2665
<i>Ohad Shamir, Nathan Srebro, Tong Zhang</i>	
CONCEPT DRIFT DETECTION THROUGH RESAMPLING	2682
<i>Maayan Harel, Koby Crammer, Ran El-Yaniv, Shie Mannor</i>	
ANTI-DIFFERENTIATING APPROXIMATION ALGORITHMS: A CASE STUDY WITH MIN-CUTS, SPECTRAL, AND FLOW	2695
<i>David F. Gleich, Michael W. Mahoney</i>	
A BAYESIAN WILCOXON SIGNED-RANK TEST BASED ON THE DIRICHLET PROCESS	2703
<i>A. Benavoli, F. Mangili, G. Corani, M. Zaffalon, F. Ruggeri</i>	
MIN-MAX PROBLEMS ON FACTOR-GRAPHS	2714
<i>Siamak Ravanbakhsh, Christopher Srinivasa, Brendan Frey, Russell Greiner</i>	
DISTRIBUTED STOCHASTIC GRADIENT MCMC	2735
<i>Sungjin Ahn, Babak Shahbaba, Max Welling</i>	

NEAREST NEIGHBORS USING COMPACT SPARSE CODES	2746
<i>Anoop Cherian</i>	

Volume 4

OPTIMAL MEAN ROBUST PRINCIPAL COMPONENT ANALYSIS	2755
<i>Feiping Nie, Jianjun Yuan, Heng Huang</i>	
PREFERENCE-BASED RANK ELICITATION USING STATISTICAL MODELS: THE CASE OF MALLOWS	2764
<i>Róbert Busa-Fekete, Eyke Hullermeier, Balázs Szorényi</i>	
HIERARCHICAL CONDITIONAL RANDOM FIELDS FOR OUTLIER DETECTION: AN APPLICATION TO DETECTING EPILEPTOGENIC CORTICAL MALFORMATIONS	2779
<i>Bilal Ahmed, Thomas Thesen, Karen E. Blackmon, Yijun Zhao, Orrin Devinsky, Ruben Kuzniecky, Carla E. Brodley</i>	
A PHYSICS-BASED MODEL PRIOR FOR OBJECT-ORIENTED MDPS	2788
<i>Jonathan Scholz, Martin Leivhn, Charles L. Isbell, David Wingate</i>	
OUTLIER PATH: A HOMOTOPY ALGORITHM FOR ROBUST SVM	2797
<i>Shinya Suzumura, Kohei Ogawa, Masashi Sugiyama, Ichiro Takeuchi</i>	
ENSEMBLE-BASED TRACKING: AGGREGATING CROWDSOURCED STRUCTURED TIME SERIES DATA	2807
<i>Naiyan Wang, Dit-Yan Yeung</i>	
LATENT CONFUSION ANALYSIS BY NORMALIZED GAMMA CONSTRUCTION	2818
<i>Issei Sato, Hisashi Kashima, Hiroshi Nakagawa</i>	
FINITO: A FASTER, PERMUTABLE INCREMENTAL GRADIENT METHOD FOR BIG DATA PROBLEMS	2839
<i>Aaron J. Defazio, Tibério S. Caetano, Justin Domke</i>	
ENSEMBLE METHODS FOR STRUCTURED PREDICTION	2856
<i>Corinna Cortes, Vitaly Kuznetsov, Mehryar Mohri</i>	
STANDARDIZED MUTUAL INFORMATION FOR CLUSTERING COMPARISONS: ONE STEP FURTHER IN ADJUSTMENT FOR CHANCE	2873
<i>Simone Romano, James Bailey, Nguyen Xuan Vinh, Karin Verspoor</i>	
PRESERVING MODES AND MESSAGES VIA DIVERSE PARTICLE SELECTION	2883
<i>Jason Pacheco, Silvia Zuffi, Michael J. Black, Erik B. Sudderth</i>	
NONLINEAR INFORMATION-THEORETIC COMPRESSIVE MEASUREMENT DESIGN	2896
<i>Liming Wang, Abolfazl Razi, Miguel Dias Rodrigues, Robert Calderbank, Lawrence Carin</i>	
DUAL QUERY: PRACTICAL PRIVATE QUERY RELEASE FOR HIGH DIMENSIONAL DATA	2908
<i>Marco Gaboardi, Emilio Jesús Gallego Arias, Justin Hsu, Aaron Roth, Zhiwei Steven Wu</i>	
DEEP BOOSTING	2917
<i>Corinna Cortes, Mehryar Mohri, Umar Syed</i>	
DISTRIBUTED REPRESENTATIONS OF SENTENCES AND DOCUMENTS	2931
<i>Quoc Le, Tomas Mikolov</i>	
UNDERSTANDING PROTEIN DYNAMICS WITH L1-REGULARIZED REVERSIBLE HIDDEN MARKOV MODELS	2940
<i>Robert T. McGibbon, Bharath Ramsundar, Mohammad M. Sultan, Gert Kiss, Vijay S. Pande</i>	
ONLINE MULTI-TASK LEARNING FOR POLICY GRADIENT METHODS	2949
<i>Haitham Bou Ammar, Eric Eaton, Paul Ruvolo, Matthew E. Taylor</i>	
AFFINITY WEIGHTED EMBEDDING	2958
<i>Jason Weston, Ron Weiss, Hector Yee</i>	
LEARNING THE PARAMETERS OF DETERMINANTAL POINT PROCESS KERNELS	2967
<i>Raja Hafiz Affandi, Emily B. Fox, Ryan P. Adams, Ben Taskar</i>	
DISCRETE CHEBYSHEV CLASSIFIERS	2982
<i>Elad Eban, Elad Mezuman, Amir Globerson</i>	
DEEP AUTOREGRESSIVE NETWORKS	2991
<i>Karol Gregor, Ivo Danihelka, Andriy Mnih, Charles Blundell, Daan Wierstra</i>	
A CONVERGENCE RATE ANALYSIS FOR LOGITBOOST, MART AND THEIR VARIANT	3001
<i>Peng Sun, Tong Zhang, Jie Zhou</i>	
INFERNING WITH HIGH GIRTH GRAPHICAL MODELS	3022
<i>Uri Heinemann, Amir Globerson</i>	
LEARNING LATENT VARIABLE GAUSSIAN GRAPHICAL MODELS	3042
<i>Zhaoshi Meng, Brian Eriksson, Alfred O. Hero</i>	

STOCHASTIC BACKPROPAGATION AND APPROXIMATE INFERENCE IN DEEP GENERATIVE MODELS	3057
<i>Danilo Jimenez Rezende, Shakir Mohamed, Daan Wierstra</i>	
ONE PRACTICAL ALGORITHM FOR BOTH STOCHASTIC AND ADVERSARIAL BANDITS	3071
<i>Yevgeny Seldin, Aleksandrs Slivkins</i>	
ROBUST AND EFFICIENT KERNEL HYPERPARAMETER PATHS WITH GUARANTEES	3093
<i>Joachim Giesen, Sören Laue, Patrick Wieschollek</i>	
ACTIVE TRANSFER LEARNING UNDER MODEL SHIFT	3102
<i>Xuezhi Wang, Tzu-Kuo Huang, Jeff Schneider</i>	
APPROXIMATE POLICY ITERATION SCHEMES: A COMPARISON	3112
<i>Bruno Scherrer</i>	
STABLE AND EFFICIENT REPRESENTATION LEARNING WITH NONNEGATIVITY CONSTRAINTS	3127
<i>Tsung-Han Lin, H. T. Kung</i>	
SAMPLE EFFICIENT REINFORCEMENT LEARNING WITH GAUSSIAN PROCESSES	3136
<i>Robert C. Grande, Thomas J. Walsh, Jonathan P. How</i>	
MEMORY AND COMPUTATION EFFICIENT PCA VIA VERY SPARSE RANDOM PROJECTIONS	3151
<i>Farhad Pourkamali-Anaraki, Shannon M. Hughes</i>	
TIME-REGULARIZED INTERRUPTING OPTIONS	3175
<i>Daniel J. Mankowitz, Timothy A. Mann, Shie Mannor</i>	
RANDOMIZED NONLINEAR COMPONENT ANALYSIS	3196
<i>David Lopez-Paz, Suvrit Sra, Alexander J. Smola, Zoubin Ghahramani, Bernhard Scholkopf</i>	
HIGH ORDER REGULARIZATION FOR SEMI-SUPERVISED LEARNING OF STRUCTURED OUTPUT PROBLEMS	3205
<i>Yujia Li, Richard Zemel</i>	
TRANSDUCTIVE LEARNING WITH MULTI-CLASS VOLUME APPROXIMATION	3218
<i>Gang Niu, Bo Dai, Marthinus Christoffel Du Plessis, Masashi Sugiyama</i>	
METHODS OF MOMENTS FOR LEARNING STOCHASTIC LANGUAGES: UNIFIED PRESENTATION AND EMPIRICAL COMPARISON	3231
<i>Borja Balle, William L Hamilton, Joelle Pineau</i>	
EFFECTIVE BAYESIAN MODELING OF GROUPS OF RELATED COUNT TIME SERIES	3248
<i>Nicolas Chapados</i>	
VARIATIONAL INFERENCE FOR SEQUENTIAL DISTANCE DEPENDENT CHINESE RESTAURANT PROCESS	3259
<i>Sergey Bartunov, Dmitry P. Vetrov</i>	
DISCOVERING LATENT NETWORK STRUCTURE IN POINT PROCESS DATA	3268
<i>Scott W. Linderman, Ryan P. Adams</i>	
A KERNEL INDEPENDENCE TEST FOR RANDOM PROCESSES	3282
<i>Kacper Chwialkowski, Arthur Gretton</i>	
LEARNING TO DISENTANGLE FACTORS OF VARIATION WITH MANIFOLD INTERACTION	3291
<i>Scott Reed, Kihyuk Sohn, Yuting Zhang, Honglak Lee</i>	
LEARNING MODULAR STRUCTURES FROM NETWORK DATA AND NODE VARIABLES	3300
<i>Elham Azizi, Edoardo M. Airolidi, James E. Galagan</i>	
PROBABILISTIC PARTIAL CANONICAL CORRELATION ANALYSIS	3309
<i>Yusuke Mukuta, Tatsuya Harada</i>	
SKIP CONTEXT TREE SWITCHING	3322
<i>Marc G. Bellemare, Joel Veness, Erik Talvitie</i>	
LOWER BOUNDS FOR THE GIBBS SAMPLER OVER MIXTURES OF GAUSSIANS	3333
<i>Christopher Tosh, Sanjoy Dasgupta</i>	
MARGINALIZED DENOISING AUTO-ENCODERS FOR NONLINEAR REPRESENTATIONS	3342
<i>Minmin Chen, Kilian Weinberger, Fei Sha, Yoshua Bengio</i>	
GAUSSIAN PROCESSES FOR BAYESIAN ESTIMATION IN ORDINARY DIFFERENTIAL EQUATIONS	3351
<i>Yali Wang, David Barber</i>	
FAST MULTI-STAGE SUBMODULAR MAXIMIZATION	3360
<i>Kai Wei, Rishabh Iyer, Jeff Bilmes</i>	
PROGRAMMING BY FEEDBACK	3385
<i>Riad Akrou, Marc Schoenauer, Jean-Christophe Souplet, Michele Sebag</i>	
PROBABILISTIC MATRIX FACTORIZATION WITH NON-RANDOM MISSING DATA	3394
<i>José Miguel Hernández-Lobato, Neil Houlsby, Zoubin Ghahramani</i>	

PURSUIT-EVASION WITHOUT REGRET, WITH AN APPLICATION TO TRADING	3437
<i>Lili Dworkin, Michael Kearns, Yuriy Nevmyyaka</i>	
THE F-ADJUSTED GRAPH LAPLACIAN: A DIAGONAL MODIFICATION WITH A GEOMETRIC INTERPRETATION	3446
<i>Sven Kurras, Ulrike Von Luxburg, Gilles Blanchard</i>	
RIEMANNIAN PURSUIT FOR BIG MATRIX RECOVERY	3467
<i>Mingkui Tan, Ivor W. Tsang, Li Wang, Bart Vandereycken, Sinno Jialin Pan</i>	
DYNAMIC PROGRAMMING BOOSTING FOR DISCRIMINATIVE MACRO-ACTION DISCOVERY	3484
<i>Leonidas Lefakis, François Fleuret</i>	
ONLINE STOCHASTIC OPTIMIZATION UNDER CORRELATED BANDIT FEEDBACK	3493
<i>Mohammad Gheshlaghi Azar, Alessandro Lazaric, Emma Brunskill</i>	

Volume 5

WEIGHTED GRAPH CLUSTERING WITH NON-UNIFORM UNCERTAINTIES	3520
<i>Yudong Chen, Shiao Hong Lim, Huan Xu</i>	
GENGA: A GENERALIZATION OF NATURAL GRADIENT ASCENT WITH POSITIVE AND NEGATIVE CONVERGENCE RESULTS	3533
<i>Philip S. Thomas</i>	
A BAYESIAN FRAMEWORK FOR ONLINE CLASSIFIER ENSEMBLE	3542
<i>Qinxun Bai, Henry Lam, Stan Sclaroff</i>	
ADAPTIVITY AND OPTIMISM: AN IMPROVED EXPONENTIATED GRADIENT ALGORITHM	3555
<i>Jacob Steinhardt, Percy Liang</i>	
GAUSSIAN APPROXIMATION OF COLLECTIVE GRAPHICAL MODELS	3572
<i>Li-Ping Liu, Daniel Sheldon, Thomas G. Dietterich</i>	
ON LEARNING TO LOCALIZE OBJECTS WITH MINIMAL SUPERVISION	3582
<i>Hyun Oh Song, Ross Girshick, Stefanie Jegelka, Julien Mairal, Zaid Harchaoui, Trevor Darrell</i>	
MULTIRESOLUTION MATRIX FACTORIZATION	3591
<i>Risi Kondor, Nedelina Teneva, Vikas K. Garg</i>	
LEARNABILITY OF THE SUPERSET LABEL LEARNING PROBLEM	3602
<i>Li-Ping Liu, Thomas G. Dietterich</i>	
TAMING THE MONSTER: A FAST AND SIMPLE ALGORITHM FOR CONTEXTUAL BANDITS	3611
<i>Alekh Agarwal, Daniel Hsu, Satyen Kale, John Langford, Lihong Li, Robert E. Schapire</i>	
STRUCTURED RECURRENT TEMPORAL RESTRICTED BOLTZMANN MACHINES	3620
<i>Roni Mittelman, Benjamin Kuipers, Silvio Savarese, Honglak Lee</i>	
SCALABLE AND ROBUST BAYESIAN INFERENCE VIA THE MEDIAN POSTERIOR	3629
<i>Stanislav Minsker, Sanvesh Srivastava, Lizhen Lin, David B. Dunson</i>	
KERNEL ADAPTIVE METROPOLIS-HASTINGS	3640
<i>Dino Sejdinovic, Heiko Strathmann, Maria Lomeli Garcia, Christophe Andrieu, Arthur Gretton</i>	
INPUTWARPING FOR BAYESIAN OPTIMIZATION OF NON-STATIONARY FUNCTIONS	3654
<i>Jasper Snoek, Kevin Swersky, Richard Zemel, Ryan P. Adams</i>	
STOCHASTIC GRADIENT HAMILTONIAN MONTE CARLO	3663
<i>Tianqi Chen, Emily B. Fox, Carlos Guestrin</i>	
A DEEP SEMI-NMF MODEL FOR LEARNING HIDDEN REPRESENTATIONS	3677
<i>George Trigeorgis, Konstantinos Bousmalis, Stefanos Zafeiriou, Björn W. Schuller</i>	
ASYNCHRONOUS DISTRIBUTED ADMM FOR CONSENSUS OPTIMIZATION	3689
<i>Ruiliang Zhang, James T. Kwok</i>	
SPECTRAL REGULARIZATION FOR MAX-MARGIN SEQUENCE TAGGING	3698
<i>Ariadna Quattoni, Borja Balle, Xavier Carreras, Amir Globerson</i>	
LEARNING BY STRETCHING DEEP NETWORKS	3707
<i>Gaurav Pandey, Ambedkar Dukkipati</i>	
NONNEGATIVE SPARSE PCA WITH PROVABLE GUARANTEES	3717
<i>Megasthenis Asteris, Dimitris S. Papailiopoulos, Alexandros G. Dimakis</i>	
ACTIVE LEARNING OF PARAMETERIZED SKILLS	3736
<i>Bruno Castro Da Silva, George Konidaris, Andrew Barto</i>	
LEARNING ORDERED REPRESENTATIONS WITH NESTED DROPOUT	3746
<i>Oren Rippel, Michael A. Gelbart, Ryan P. Adams</i>	
LEARNING THE IRREDUCIBLE REPRESENTATIONS OF COMMUTATIVE LIE GROUPS	3757
<i>Taco Cohen, Max Welling</i>	

TOWARDS END-TO-END SPEECH RECOGNITION WITH RECURRENT NEURAL NETWORKS	3771
<i>Alex Graves, Navdeep Jaitly</i>	
MULTI-PERIOD TRADING PREDICTION MARKETS WITH CONNECTIONS TO MACHINE LEARNING	3780
<i>Jinli Hu, Amos Storkey</i>	
EFFICIENT GRADIENT-BASED INFERENCE THROUGH TRANSFORMATIONS BETWEEN BAYES NETS AND NEURAL NETS	3791
<i>Diederik P. Kingma, Max Welling</i>	
NEURAL VARIATIONAL INFERENCE AND LEARNING IN BELIEF NETWORKS	3800
<i>Andriy Mnih, Karol Gregor</i>	
SCALABLE BAYESIAN LOW-RANK DECOMPOSITION OF INCOMPLETE MULTIWAY TENSORS	3810
<i>Piyush Rai, Yingjian Wang, Shengbo Guo, Gary Chen, David Dunson, Lawrence Carin</i>	
BETA DIFFUSION TREES	3821
<i>Creighton Heaukulani, David A. Knowles, Zoubin Ghahramani</i>	
LEARNING CHARACTER-LEVEL REPRESENTATIONS FOR PART-OF-SPEECH TAGGING	3830
<i>Cicero Nogueira Dos Santos, Bianca Zadrozny</i>	
SADDLE POINTS AND ACCELERATED PERCEPTRON ALGORITHMS	3839
<i>Adams Wei Yu, Fatma K I nc-Karzan, Jaime G. Carbonell</i>	
ROBUST DISTANCE METRIC LEARNING VIA SIMULTANEOUS 1-NORM MINIMIZATION AND MAXIMIZATION	3853
<i>Hua Wang, Feiping Nie, Heng Huang</i>	
LEARNING FROM CONTAGION (WITHOUT TIMESTAMPS)	3862
<i>Kareem Amin, Hoda Heidari, Michael Kearns</i>	
STOCHASTIC VARIATIONAL INFERENCE FOR BAYESIAN TIME SERIES MODELS	3872
<i>Matthew James Johnson, Alan S. Willsky</i>	
A CLOCKWORK RNN	3881
<i>Jan Koutnik, Klaus Greff, Faustino Gomez, Jurgen Schmidhuber</i>	
ESTIMATING LATENT-VARIABLE GRAPHICAL MODELS USING MOMENTS AND LIKELIHOODS	3890
<i>Arun Tejasvi Chaganty, Percy Liang</i>	
UNIVERSAL MATRIX COMPLETION	3902
<i>Srinadh Bhojanapalli, Prateek Jain</i>	
FINDING DENSE SUBGRAPHS VIA LOW-RANK BILINEAR OPTIMIZATION	3918
<i>Dimitris S. Papailiopoulos, Ioannis Miliagkas, Alexandros G. Dimakis, Constantine Caramanis</i>	
COMPOSITIONAL MORPHOLOGY FOR WORD REPRESENTATIONS AND LANGUAGE MODELLING	3946
<i>Jan A. Botha, Phil Blunsom</i>	
LEARNING POLYNOMIALS WITH NEURAL NETWORKS	3955
<i>Alexandr Andoni, Rina Panigrahy, Gregory Valiant, Li Zhang</i>	
EXPONENTIAL FAMILY MATRIX COMPLETION UNDER STRUCTURAL CONSTRAINTS	3964
<i>Suriya Gunasekar, Pradeep Ravikumar, Joydeep Ghosh</i>	
SAMPLE-BASED APPROXIMATE REGULARIZATION	3977
<i>Philip Bachman, Amir-Massoud Farahmand, Doina Precup</i>	
A COMPILATION TARGET FOR PROBABILISTIC PROGRAMMING LANGUAGES	4001
<i>Brooks Paige, Frank Wood</i>	
ADAPTIVE MONTE CARLO VIA BANDIT ALLOCATION	4010
<i>James Neufeld, Andras Gyorgy, Dale Schuurmans, Csaba Szepesvári</i>	
EFFICIENT DIMENSIONALITY REDUCTION FOR HIGH-DIMENSIONAL NETWORK ESTIMATION	4038
<i>Safiye Celik, Benjamin A. Logsdon, Su-In Lee</i>	
DETERMINISTIC ANYTIME INFERENCE FOR STOCHASTIC CONTINUOUS-TIME MARKOV PROCESSES	4047
<i>E. Busra Celikkaya, Christian R. Shelton</i>	
DOUBLY STOCHASTIC VARIATIONAL BAYES FOR NON-CONJUGATE INFERENCE	4056
<i>Michalis K. Titsias, Miguel Lázaro-Gredilla</i>	
EFFICIENT LEARNING OF MAHALANOBIS METRICS FOR RANKING	4070
<i>Daryl K. H. Lim, Gert Lanckriet</i>	
GEV-CANONICAL REGRESSION FOR ACCURATE BINARY CLASS PROBABILITY ESTIMATION WHEN ONE CLASS IS RARE	4080
<i>Arpit Agarwal, Harikrishna Narasimhan, Shivaram Kalyanakrishnan, Shivani Agarwal</i>	

A REVERSIBLE INFINITE HMM USING NORMALISED RANDOM MEASURES	4090
<i>Konstantina Palla, David A. Knowles, Zoubin Ghahramani</i>	
STRUCTURED LOW-RANK MATRIX FACTORIZATION: OPTIMALITY, ALGORITHM, AND APPLICATIONS TO IMAGE PROCESSING	4108
<i>Benjamin D. Haefele, Eric D. Young, René Vidal</i>	
INFLUENCE FUNCTION LEARNING IN INFORMATION DIFFUSION NETWORKS.....	4118
<i>Nan Du, Yingyu Liang, Maria-Florina Balcan, Le Song</i>	
Author Index	