Neural Information Processing Systems

Advances in Neural Information Processing Systems 25

26th Annual Conference on Neural Information Processing Systems 2012

December 3-6, 2012
Lake Tahoe, Nevada, USA

Volume 1 of 4

Printed from e-media with permission by:

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

ISBN: 978-1-62748-003-1

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

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© (2012) by Neural Information Processing Systems
All rights reserved.

Printed by Curran Associates, Inc. (2013)

For permission requests, please contact Neural Information Processing Systems
at the address below.

Neural Information Processing Systems
10010 North Torrey Pines Road
La Jolla, CA 92037

Phone: (858) 453-4100
Fax: (858) 453-8534

info@nips.cc

Additional copies of this publication are available from:

Curran Associates, Inc.
57 Morehouse Lane
Red Hook, NY 12571 USA
Phone: 845-758-0400
Fax: 845-758-2634
Email: curran@proceedings.com
Web: www.proceedings.com

Contents

Contents iii

Preface xxvii

Donors xxxi

NIPS foundation xxxii

Committees xxxiii

Reviewers xxxvi

Locally Uniform Comparison Image Descriptor, Andrew Ziegler, UCSD, Eric Christiansen, David Kriegman, University of California, San Diego, and Serge Belongie, UC San Diego 1

Learning from Distributions via Support Measure Machines, Krikamol Muandet, MPI for Intelligent Systems, Kenji Fukumizu, The Institute of Statistical Mathematics, Francesco Dinuzzo, Max Planck Institute for Intelligent Systems, and Bernhard Schölkopf, Max Planck Institute for Intelligent Systems 10

Finding Exemplars from Pairwise Dissimilarities via Simultaneous Sparse Recovery, Ehsan Elhamifar, Johns Hopkins University, Guillermo Sapiro, Minnesota, and Rene Vidal 19

Feature Clustering for Accelerating Parallel Coordinate Descent, Chad Scherrer, Ambuj Tewari, University of Michigan, Mahantesh Halappanavar, Pacific Northwest National Laboratory, and David Haglin, Pacific Northwest National Laboratory 28

Multi-scale Hyper-time Hardware Emulation of Human Motor Nervous System Based on Spiking Neurons using FPGA, C. Niu, USC, Sirish Nandyala, Won Sohn, and Terence Sanger, USC 37

Active Learning of Model Evidence Using Bayesian Quadrature , Michael Osborne, University of Oxford, David Duvenaud, University of Cambridge, Roman Garnett, Carnegie Mellon University, Carl Rasmussen, University of Cambridge, Stephen Roberts, University of Oxford, and Zoubin Ghahramani, University of Cambridge 46

Coupling Nonparametric Mixtures via Latent Dirichlet Processes, Dahua Lin, MIT, and John Fisher, MIT 55

Nonparametric Max-Margin Matrix Factorization for Collaborative Prediction, Minjie Xu, Tsinghua University, Jun Zhu, and Bo Zhang 64

Bayesian Hierarchical Reinforcement Learning, Feng Cao, Case Western Reserve Universit, and Soumya Ray, Case Western Reserve University .. 73

Dynamic Pruning of Factor Graphs for Maximum Marginal Prediction, Christoph Lampert, IST Austria 82

Local Supervised Learning through Space Partitioning, Joseph Wang, Boston University, and Venkatesh Saligrama, Boston University 91

A Generative Model for Parts-based Object Segmentation, S. Eslami, University of Edinburgh, and Chris Williams, University of Edinburgh 100

Super-Bit Locality-Sensitive Hashing, Jianqiu Ji, Tsinghua University, Jianmin Li, Shuicheng Yan, National University of Singapore, Bo Zhang, and Qi Tian, University of Texas at San Antonio 108

The Bethe Partition Function of Log-supermodular Graphical Models, Nicholas Ruozzi, EPFL 117

Random Utility Theory for Social Choice, Hossein Azari, Harvard University, David Parks, and Lirong Xia, Harvard University 126

Putting Bayes to sleep, Wouter Koolen, Royal Holloway, Dmitry Adamskiy, Royal Holloway, University of London, and Manfred Warmuth, University of California at Santa Cruz 135

A new metric on the manifold of kernel matrices with application to matrix geometric means, Suvrit Sra, Max Planck Institute 144

Mandatory Leaf Node Prediction in Hierarchical Multilabel Classification, Wei Bi, HKUST, and James Kwok, Hong Kong University of Science and Technology 153

Smooth-projected Neighborhood Pursuit for High-dimensional Nonparanormal Graph Estimation, Tuo Zhao, Johns Hopkins University, Kathryn Roeder, and Han Liu, Princeton University 162

Semiparametric Principal Component Analysis, Fang Han, Johns Hopkins University, and Han Liu, Princeton University 171

Coding efficiency and detectability of rate fluctuations with non-Poisson neuronal firing, Shinsuke Koyama, ISM 180

The representer theorem for Hilbert spaces: a necessary and sufficient condition, Francesco Dinuzzo, MPI for Intelligent Systems, and Bernhard Schölkopf, Max Planck Institute for Intelligent Systems 189

On the (Non-)existence of Convex, Calibrated Surrogate Losses for Ranking, Clément Calauzènes, UPMC-LIP6, Nicolas Usunier, Universitá Pierre & MarieCurie, and Patrick Gallinari, UPMC-LIP6 197

Exploration in Model-based Reinforcement Learning by Empirically Estimating Learning Progress, Manuel Lopes, INRIA, Tobias Lang, Marc Toussaint, FU Berlin, and Pierre-Yves Oudeyer, INRIA 206

Supervised Learning with Similarity Functions, Purushottam Kar, IIT Kanpur, and Prateek Jain, MSR Bangalore 215

Cocktail Party Processing via Structured Prediction, Yuxuan Wang, The Ohio State University, and DeLiang Wang, Dept. of CSE at The Ohio State University 224

Robustness and risk-sensitivity in Markov decision processes, Takayuki Osogami, IBM Research - Tokyo 233

Dynamical And-Or Graph Learning for Object Shape Modeling and Detection, Xiaolong Wang, Sun Yat-Sen University, and Liang Lin, Sysu.edu.cn 242

Adaptive Stratified Sampling for Monte-Carlo integration of Differentiable functions, Alexandra Carpentier, INRIA, and Remi Munos, Inria Lille 251

Distributed Non-Stochastic Experts, Varun Kanade, Harvard University, Zhenming Liu, and Bozidar Radunovic, Microsoft Research Cambridge 260

Learning Image Descriptors with the Boosting-Trick, Tomasz Trzcinski, EPFL, Mario Christoudias, Vincent Lepetit, EPF Lausanne, and Pascal Fua, Ecole Polytechnique Federal de Lausanne 269

Fast Resampling Weighted v-Statistics, Chunxiao Zhou, NIH, jiseong Park, and Yun Fu, Northeastern University 278

Multi-task Vector Field Learning, Binbin Lin, Zhejiang University, Sen Yang, Arizona State University, Chiyuan Zhang, Jieping Ye, Arizona State U., and Xiaofei He 287

Memorability of Image Regions, Aditya Khosla, MIT, Jianxiong Xiao, CSAIL MIT, Antonio Torralba, MIT, and Aude Oliva, MIT 296

Nonparametric Bayesian Inverse Reinforcement Learning for Multiple Reward Functions, Jaedeug Choi, KAIST, and Kee-Eung Kim, KAIST 305

Automatic Feature Induction for Stagewise Collaborative Filtering, Joonseok Lee, Georgia Tech, Mingxuan Sun, Seungyeon Kim, and Guy Lebanon, Google and Georgia Tech 314

Selective Labeling via Error Bound Minimization, Quanquan Gu, CS, UIUC, Tong Zhang, Rutgers University, Chris Ding, University of Texas at Arlington, and Jiawei Han, UIUC 323

Volume Regularization for Binary Classification, Koby Crammer, Technion, and Tal Wagner 332

Image Denoising and Inpainting with Deep Neural Networks, Junyuan Xie, Univ. of Sci. & Tech. of China, Linli Xu, University of Science and Tech, and Enhong Chen, University of Science and Technology of China ... 341

Max-Margin Structured Output Regression for Spatio-Temporal Action Localization, Du Tran, Nanyang Technological Univ, and Junsong Yuan, NTU 350

Transelliptical Component Analysis, Fang Han, Johns Hopkins University, and Han Liu, Princeton University 359

Action-Model Based Multi-agent Plan Recognition, Hankz Zhuo, Sun Yat-sen University, Qiang Yang, Hong Kong University of Science and Technology, and Subbarao Kambhampati 368

Visual Recognition using Embedded Feature Selection for Curvature Self-Similarity, Angela Eigenstetter, IWR University Heidelberg, and Bjorn Ommer 377

Non-parametric Approximate Dynamic Programming via the Kernel Method, Nikhil Bhat, Columbia University, Ciamac Moallemi, and Vivek Farias, Massachusetts Institute of Technology 386

Optimal Regularized Dual Averaging Methods for Stochastic Optimization, Xi Chen, Carnegie Mellon University, Qihang Lin, and Javier Pena, Carnegie Mellon University 395

The variational hierarchical EM algorithm for clustering hidden Markov models., Emanuele Coviello, UCSD, Antoni Chan, City University of Hong Kong, and Gert Lanckriet, UCSD 404

Truncation-free Online Variational Inference for Bayesian Nonparametric Models, Chong Wang, Princeton University, and David Blei, Princeton 413

3D Social Saliency from Head-mounted Cameras, Hyun Park, CMU, eakta Jain, and Yaser Sheikh 422

Context-Sensitive Decision Forests for Object Detection, Peter Kontschieder, Graz University of Technology, Samuel Bulò, Università Ca' Foscari Venezia, Italy, Antonio Criminisi, Microsoft Research, Pushmeet Kohli, Microsoft Research Cambridge, Marcello Pelillo, Universitá Ca' Foscari di Venezia, and Horst Bischof, TU Graz 431

Learning Invariant Representations of Molecules for Atomization Energy Prediction, Grégoire Montavon, TU Berlin, Katja Hansen, MPG Fritz Haber Institute, Siamac Fazli, TU Berlin, Matthias Rupp, ETH Zurich, Franziska Biegler, TU Berlin, Andreas Ziehe, Technische Universität Berlin, Alexandre Tkatchenko, MPG Fritz Haber Institute, Anatole Lilienfeld, Argonne National Laboratory, and Klaus-Robert Mueller, TU Berlin 440

Bandit Algorithms boost Brain Computer Interfaces for motor-task selection of a brain-controlled button, Joan Fruitet, INRIA, Alexandra Carpentier, Remi Munos, Inria Lille, and Maureen Clerc, INRIA 449

Multiplicative Forests for Continuous-Time Processes, Jeremy Weiss, University of Wisconsin, Sriraam Natarajan, Wake Forest University Baptist Medical Center, and David Page, University of Wisconsin 458

Patient Risk Stratification for Hospital-Associated C. diff as a Time-Series Classification Task, Jenna Wiens, MIT, John Guttag, and Eric Horvitz 467

Nyström Method vs Random Fourier Features: A Theoretical and Empirical Comparison, Tianbao Yang, Michigan State University, Yu-Feng Li, Nanjing University, Mehrdad Mahdavi, Michigan State University, Rong Jin, Michigan State, and Zhi-Hua Zhou, Nanjing University 476

Multiclass Learning Approaches: A Theoretical Comparison with Implications, Amit Daniely, Hebrew university, Sivan Sabato, and Shai Shwartz, The Hebrew University 485

Stochastic Gradient Descent with Only One Projection, Mehrdad Mahdavi, Michigan State University, Tianbao Yang, Rong Jin, Michigan State, Shenghuo Zhu, NEC Laboratories America, and Jinfeng Yi, Michigan State University 494

Neuronal spike generation mechanism as an oversampling, noise-shaping A-to-D converter, Dmitri Chklovskii, Janelia Farm, and Daniel Soudry 503

Deep Spatio-Temporal Architectures and Learning for Protein Structure Prediction, Pietro Lena, UCI, Pierre Baldi, University of California @ Irvine, and Ken Nagata, UCI 512

Assessing Blinding in Clinical Trials, Ognjen Arandjelovic, Swansea University 521

Scalable nonconvex inexact proximal splitting, Suvrit Sra, Max Planck Institute 530

Learning to Discover Social Circles in Ego Networks, Julian McAuley, Stanford University, and Jure Leskovec, Stanford 539

A Conditional Multinomial Mixture Model for Superset Label Learning, Liping Liu, Oregon State University, and Thomas Dietterich, Oregon State University 548

Majorization for CRFs and Latent Likelihoods, Tony Jebara, Columbia, and Anna Choromanska, Columbia University 557

Ensemble weighted kernel estimators for multivariate entropy estimation, Kumar Sricharan, NASA Ames Reasearch Center, and Alfred III, University of Michigan, Ann Arbor 566

Efficient high dimensional maximum entropy modeling via symmetric partition functions, Paul Vernaza, Carnegie Mellon University, and Drew Bagnell 575

Discriminatively Trained Sparse Code Gradients for Contour Detection , Ren Xiaofeng, Intel, and Liefeng Bo, Intel Labs 584

Analyzing 3D Objects in Cluttered Images, Mohsen Hejrati, University of california irvin, and Deva Ramanan 593

Nonconvex Penalization Using Laplace Exponents and Concave Conjugates, Zhihua Zhang, and Bojun Tu, Zhejiang University 602

3D Object Detection and Viewpoint Estimation with a Deformable 3D Cuboid Model, Sanja Fidler, TTI Chicago, Sven Dickinson, University of Toronto, and Raquel Urtasun, Toyota Technological Institute Chicago 611

Structured Learning of Gaussian Graphical Models, Karthik Mohan, Mike Chung, Seungyeop Han, University of Washington, Daniela Witten, Su-In Lee, University of Washington, and Maryam Fazel 620

A Polylog Pivot Steps Simplex Algorithm for Classification, Elad Hazan, Technion, and Zohar Karnin, Yahoo! 629

Shifting Weights: Adapting Object Detectors from Image to Video, Kevin Tang, Stanford University, Vignesh Ramanathan, Li Fei-Fei, and Daphne Koller, Stanford 638

A Scalable CUR Matrix Decomposition Algorithm: Lower Time Complexity and Tighter Bound, Shusen Wang, Zhejiang University, and Zhihua Zhang 647

Convolutional-Recursive Deep Learning for 3D Object Classification, Richard Socher, Stanford University, Brody Huval, Bharath Bath, Christopher Manning, and Andrew Ng, Stanford Univ. 656

Semi-Supervised Domain Adaptation with Non-Parametric Copulas, David Lopez-Paz, Max Planck Institute for IS, Jose Hernandez-Lobato, Cambridge University, and Bernhard Schölkopf, Max Planck Institute for Intelligent Systems 665

Identification of Recurrent Patterns in the Activation of Brain Networks, Firdaus Janoos, Exxon Mobil Corporate Research, weichang Li, ExxonMobil Research, Niranjan Subrahmanya, ExxonMobil Corporate Research, Istvan Morocz, Harvard Medical School, and William Wells, Harvard Medical School 674

Density-Difference Estimation, Masashi Sugiyama, Tokyo Institute of Technology, Takafumi Kanamori, Nagoya University, Taiji Suzuki, The University of Tokyo, Marthinus Plessis, Tokyo Institute of Technology, Song Liu, and Ichiro Takeuchi, Nagoya Institute of Technology 683

Variational Inference for Crowdsourcing, Qiang Liu, UCI, Jian Peng, TTI Chicago, and Alex Ihler, UC Irvine 692

MCMC for continuous-time discrete-state systems, Vinayak Rao, Gatsby Unit UCL, and Yee Teh, Gatsby 701

A P300 BCI for the Masses: Prior Information Enables Instant Unsupervised Spelling, Pieter-Jan Kindermans, Ghent University, Hannes Verschore, David Verstraeten, and Benjamin Schrauwen, Ghent University 710

Learning about Canonical Views from Internet Image Collections, Elad Mezuman, The Hebrew University of Jerus, and Yair Weiss, Hebrew University 719

Learning High-Density Regions for a Generalized Kolmogorov-Smirnov Test in High-Dimensional Data, Assaf Glazer, Technion, Michael Lindenbaoum, and Shaul Markovitch, Technion 728

Multiresolution Gaussian Processes, Emily Fox, University of Washington, and David Dunson, Duke Statistics 737

Localizing 3D cuboids in single-view images, Jianxiong Xiao, CSAIL MIT, Bryan Russell, U Washington, and Antonio Torralba, MIT 746

Newton-Like Methods for Sparse Inverse Covariance Estimation, Peder Olsen, IBM, Figen Oztoprak, Northwestern University, Jorge Nocedal, and Steven Rennie, IBM research 755

Learning to Align from Scratch, Gary Huang, UMass Amherst, Marwan Mattar, Honglak Lee, University of Michigan, and Erik Learned-Miller, UMass Amherst 764

Homeostatic plasticity in Bayesian spiking networks as Expectation Maximization with posterior constraints, Stefan Habenschuss, Graz University of Technology, Johannes Bill, and Bernhard Nessler 773

Clustering Aggregation as Maximum-Weight Independent Set, Nan Li, Temple University, and Longin Latecki, Temple University 782

Topology Constraints in Graphical Models, Marcelo Fiori, Facultad de Ingeniería, UdelaR, Pablo Musé, and Guillermo Sapiro, Minnesota 791

Transelliptical Graphical Models, Han Liu, Princeton University, Fang Han, Johns Hopkins University, and Cun-Hui Zhang, Rutgers 800

Kernel Latent SVM for Visual Recognition, Weilong Yang, Simon Fraser University, Yang Wang, UIUC, Arash Vahdat, Simon Fraser University, and Greg Mori, Simon Fraser University 809

Learning Partially Observable Models Using Temporally Abstract Decision Trees, Erik Talvitie, Franklin & Marshall College 818

Proximal Newton-type methods for convex optimization, Jason Lee, Stanford University, Yuekai Sun, and Michael Saunders, Stanford University 827

Regularized Off-Policy TD-Learning, Bo Liu, University of Massachusetts, Sridhar Mahadevan, University of Massachusetts Amherst, and Ji Liu, University Wisconsin-Madison 836

Multi-criteria Anomaly Detection using Pareto Depth Analysis, Ko-Jen Hsiao, University of Michigan, Kevin Xu, Jeff Calder, and Alfred Hero, Michigan 845

Synchronization can Control Regularization in Neural Systems via Correlated Noise Processes, Jake Bouvrie, Massachusetts Institute of Technology, and Jean-Jacques Slotine, Massachusetts Institute of Technology 854

Calibrated Elastic Regularization in Matrix Completion, Tingni Sun, Rutgers University, and Cun-Hui Zhang, Rutgers University 863

Predicting Action Content On-Line and in Real Time before Action Onset – an Intracranial Human Study, Uri Maoz, Caltech, Shengxuan Ye, University of Virginia, Ian Ross, Huntington Memorial Hospital, Adam Mamelak, Cedars Sinai Medical Center, and Christof Koch, Caltech, Allen Inst. 872

Searching for objects driven by context, Bogdan Alexe, ETH ZURICH, Nicolas Heess, Gatsby Unit, Yee Teh, and Vittorio Ferrari . 881

Timely Object Recognition, Sergey Karayev, UC Berkeley, Tobias Baumgartner, Mario Fritz, Max Planck Institute for Informatics, and Trevor Darrell, UC Berkeley EECS and ICSI 890

Nonparanormal Belief Propagation (NPNBP), Gal Elidan, Hebrew University, and Cobi Cario, Hebrew University 899

Deep Representations and Codes for Image Auto-Annotation, Ryan Kiros, University of Alberta, and Csaba Szepesvari, U. Alberta 908

A Spectral Algorithm for Latent Dirichlet Allocation, Anima Anandkumar, UC Irvine, Dean Foster, University of Pennsylvania, Daniel Hsu, Microsoft Research, Sham Kakade, and Yi-Kai Liu, NIST 917

Learning Halfspaces with the Zero-One Loss: Time-Accuracy Tradeoffs, Aharon Birnbaum, Hebrew university, and Shai Shwartz, The Hebrew University 926

Matrix reconstruction with the local max norm, Rina Foygel, University of Chicago, Nathan Srebro, Toyota Technical Intitute, and Ruslan Salakhutdinov, University of Toronto 935

Analog readout for optical reservoir computers, Anteo Smerieri, Université libre de Bruxelles, François Duport, Yvon Paquot, Benjamin Schrauwen, Univeristy of Gent, Marc Haelterman, Université libre de Bruxelles, and Serge Massar, Université libre de Bruxelles 944

Accuracy at the Top, Stephen Boyd, Stanford/Google, Corinna Cortes, Google, Mehryar Mohri, NYU, and Ana Radovanovic, Google Research 953

Minimizing Sparse High-Order Energies by Submodular Vertex-Cover, Andrew Delong, University of Western Ontario, Olga Veksler, Anton Osokin, Moscow State University, and Yuri Boykov, University of Western Ontario 962

Perfect Dimensionality Recovery by Variational Bayesian PCA, Shinichi Nakajima, Nikon Corporation, Ryota Tomioka, The University of Tokyo, Masashi Sugiyama, Tokyo Institute of Technology, and S. Babacan, Illinois University 971

Mirror Descent Meets Fixed Share (and feels no regret), Nicolò Cesa-Bianchi, Univ. degli Studi di Milano, Pierre Gaillard, Ecole normale supérieure, Gabor Lugosi, Pompeu Fabra, and Gilles Stoltz, Ecole normale superieure, Paris 980

Near-optimal Differentially Private Principal Components, Kamalika Chaudhuri, UC San Diego, Anand Sarwate, TTI Chicago, and Kaushik Sinha, UCSD 989

Random function priors for exchangeable arrays with applications to graphs and relational data, James Lloyd, University of Cambridge, Peter Orbanz, Zoubin Ghahramani, and Daniel Roy, University of Cambridge 998

Inverse Reinforcement Learning through Structured Classification, Edouard Klein, Supélec, Matthieu Geist, SUPELEC, BILAL PIOT, and Olivier Pietquin, SUPELEC - UMI 2958 1007

Augmented-SVM: Automatic space partitioning for combining multiple non-linear dynamics, Ashwini Shukla, Ecole Polytechnique Fédérale de Lausanne, and Aude Billard, EPFL 1016

Efficient Bayes-Adaptive Reinforcement Learning using Sample-Based Search, Arthur Guez, Gatsby Unit, David Silver, UCL, and Peter Dayan, UCL 1025

Dimensionality Dependent PAC-Bayes Margin Bound, Chi Jin, Peking University, and Liwei Wang, Peking University 1034

Latent Graphical Model Selection: Efficient Methods for Locally Tree-like Graphs, Anima Anandkumar, UC Irvine, and Ragupathyraj Valluvan, UCI 1043

Learning Mixtures of Tree Graphical Models, Anima Anandkumar, UC Irvine, Daniel Hsu, Microsoft Research, Furong Huang, and Sham Kakade, Microsoft Research 1052

Hamming Distance Metric Learning, Mohammad Norouzi, University of Toronto, David Fleet, and Ruslan Salakhutdinov, University of Toronto 1061

Spiking and saturating dendrites differentially expand single neuron computation capacity., Romain Cazé, Departement Sciences Cognitives ENS, Mark Humphries, INSERM U960, and Boris Gutkin, INSERM U960 1070

Clustering by Nonnegative Matrix Factorization Using Graph Random Walk, Zhirong Yang, Aalto University, Tele Hao, Onur Dikmen, Aalto University, Xi Chen, and Erkki Oja, Aalto University 1079

Delay Compensation with Dynamical Synapses, Chi Fung, HKUST, K. Wong, Department of Physics, Hong Kong University of Science and Technology, and Si Wu, State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal Univ 1088

ImageNet Classification with Deep Convolutional Neural Networks, Alex Krizhevsky, Ilya Sutskever, University of Toronto, and Geoff Hinton, University of Toronto 1097

Recognizing Activities by Attribute Dynamics , Weixin LI, University of Calif, San Diego, and Nuno Vasconcelos, UC San Diego 1106

Compressive Sensing MRI with Wavelet Tree Sparsity, Chen Chen, University of Texas at Arlington, and Junzhou Huang, University of Texas at Arlington 1115

Training sparse natural image models with a fast Gibbs sampler of an extended state space, Lucas Theis, CIN, Jascha Sohl-Dickstein, UC Berkeley, and Matthias Bethge, Max Planck Institute Tuebingen 1124

A Bayesian Approach for Policy Learning from Trajectory Preference Queries, Aaron Wilson, Oregon State University, Alan Fern, and Prasad Tadepalli, Oregon State 1133

GenDeR: A Generic Diversified Ranking Algorithm, Jingrui He, IBM Research, Hanghang Tong, Qiaozhu Mei, University of Michigan, and Boleslaw Szymanski, RPI 1142

On Multilabel Classification and Ranking with Partial Feedback, Claudio Gentile, University of Subria, and Francesco Orabona, Toyota Technological Institute 1151

The Lovasz θ function, SVMs and finding large dense subgraphs, Vinay Jethava, Chalmers University, Anders Martinsson, Chiranjib Bhattacharyya, Indian Institute of Science, and Devdatt Dubhashi, Chalmers University 1160

Multi-Task Averaging, Sergey Feldman, Maya Gupta, University of Washington, and Bela Frigyik 1169

Unsupervised Structure Discovery for Semantic Analysis of Audio, Sourish Chaudhuri, Carnegie Mellon University, and Bhiksha Raj, Carnegie Mellon University 1178

A Marginalized Particle Gaussian Process Regression, Yali Wang, Laval University, and Brahim Chaib-draa 1187

Angular Quantization-based Binary Codes for Fast Similarity Search, Yunchao Gong, UNC, Sanjiv Kumar, Google, Vishal Verma, and Svetlana Lazebnik, UIUC 1196

Optimal kernel choice for large-scale two-sample tests, Arthur Gretton, Gatsby Unit, CSML, UCL, Bharath Sriperumbudur, Gatsby, Dino Sejdinovic, Gatsby Unit, CSML, UCL, heiko Strathmann, CSML, UCL, Sivaraman Balakrishnan, CMU, Massimiliano Pontil, University College London, and Kenji Fukumizu, The Institute of Statistical Mathematics 1205

Factoring nonnegative matrices with linear programs, Ben Recht, University of Wisconsin, Christopher Re, Joel Tropp, California Institute of Technology, and Victor Bittorf, University of Wisconsin 1214

Large Scale Distributed Deep Networks, Jeffrey Dean, Greg Corrado, Google, Rajat Monga, Kai Chen, Matthieu Devin, Quoc Le, Mark Mao, Marc'Aurelio Ranzato, Google, Andrew Senior, Paul Tucker, Ke Yang, and Andrew Ng, Stanford Univ. 1223

Statistical Consistency of Ranking Methods in A Rank-Differentiable Probability Space, Yanyan Lan, ICT, Jiafeng Guo, Xueqi Cheng, ICT, and Tie-Yan Liu, Microsoft 1232

Wavelet based multi-scale shape features on arbitrary surfaces for cortical thickness discrimination , Won Kim, University of Wisconsin, Deepti Pachauri, Charles Hatt, Moo. Chung, Sterling Johnson, and Vikas Singh, University of Wisconsin 1241

A Convex Formulation for Learning Scale-Free Networks via Submodular Relaxation, Aaron Defazio, ANU, and Tiberio Caetano, NICTA Canberra 1250

Fused sparsity and robust estimation for linear models with unknown variance, Arnak Dalalyan, ENPC, and Yin Chen 1259

How Prior Probability Influences Decision Making: A Unifying Probabilistic Model, Yanping Huang, University of Washington, Abram Friesen, Timothy Hanks, Princeton University, Mike Shadlen, University of Washington, and Rajesh Rao, University of Washington 1268

High-Order Multi-Task Feature Learning to Identify Longitudinal Phenotypic Markers for Alzheimer's Disease Progression Prediction, Hua Wang, Univ. of Texas at Arlington, Feiping Nie, University of Texas at Arlington, Heng Huang, University of Texas Arlington, Jingwen Yan, IUPUI, Sungeun Kim, Shannon Risacher, Andrew Saykin, and Li Shen, IUPUI 1277

Symmetric Correspondence Topic Models for Multilingual Text Analysis, Kosuke Fukumasu, Kobe University, Koji Eguchi, and Eric Xing, CMU 1286

Effective Split-Merge Monte Carlo Methods for Nonparametric Models of Sequential Data, Michael Hughes, Brown University, Emily Fox, University of Washington, and Erik Sudderth, Brown University 1295

Efficient coding provides a direct link between prior and likelihood in perceptual Bayesian inference, Xue-Xin Wei, University of Pennsylvania, and Alan Stocker, U Penn 1304

Efficient Sampling for Bipartite Matching Problems, Maksims Volkovs, University of Toronto, and Rich Zemel, University of Toronto 1313

Learning visual motion in recurrent neural networks, Marius Pachitariu, UCL, and Maneesh Sahani, UCL 1322

Learned Prioritization for Trading Off Accuracy and Speed, Jiarong Jiang, University of Maryland, Adam Teichert, Johns Hopkins University, Hal III, University of Maryland, and Jason Eisner 1331

Value Pursuit Iteration, Amir Farahmand, McGill University, and Doina Precup, McGill Montreal 1340

Compressive neural representation of sparse, high-dimensional probabilities, Xaq Pitkow, University of Rochester 1349

Graphical Models via Generalized Linear Models, Eunho Yang, University of Texas, Austin, Pradeep Ravikumar, University of Texas-Austin, Genevera Allen, Rice University, and Zhandong Liu, Baylor College of Medicine 1358

CPRL – An Extension of Compressive Sensing to the Phase Retrieval Problem, Henrik Ohlsson, Linköping university, Allen Yang, Department of EECS, UC Berkeley, Roy Dong, and Shankar Sastry, Department of EECS, UC Berkeley 1367

Co-Regularized Hashing for Multimodal Data, Yi Zhen, HKUST, and Dit-Yan Yeung, Hong Kong University of Science and Technology 1376

Convergence and Energy Landscape for Cheeger Cut Clustering, Xavier Bresson, Thomas Laurent, UC Riverside, David Uminsky, and James Brecht 1385

Symbolic Dynamic Programming for Continuous State and Observation POMDPs, Zahra Zamani, ANU and NICTA, Scott Sanner, NICTA and The ANU, Pascal Poupart, University of Waterloo, and Kristian Kersting, Fraunhofer IAIS 1394

Bayesian Probabilistic Co-Subspace Addition, Lei Shi, Baidu.com, Inc 1403

Scaled Gradients on Grassmann Manifolds for Matrix Completion, Thanh Ngo, University of Minnesota, and Yousef Saad, University of Minnesota 1412

Q-MKL: Matrix-induced Regularization in Multi-Kernel Learning with Applications to Neuroimaging , Chris Hinrichs, University of Wisconsin, Vikas Singh, Jiming Peng, University of Illinois, and Sterling Johnson, University of Wisconsin 1421

Privacy Aware Learning, John Duchi, Berkeley, Michael Jordan, Berkley University, and Martin Wainwright, UC Berkeley 1430

Finite Sample Convergence Rates of Zero-Order Stochastic Optimization Methods, John Duchi, Berkeley, Michael Jordan, Berkley University, Martin Wainwright, UC Berkeley, and Andre Wibisono, University of California, Berkeley 1439

Hierarchical Optimistic Region Selection driven by Curiosity, Odalric-Ambrym Maillard, Leoben 1448

Sparse Prediction with the k-Support Norm, Andreas Argyriou, TTI Chicago, Rina Foygel, University of Chicago, and Nathan Srebro, Toyota Technical Intitute 1457

Active Learning of Multi-Index Function Models, Tyagi Hemant, and Volkan Cevher, EPFL 1466

Learning Multiple Tasks using Shared Hypotheses, Koby Crammer, Technion, and Yishay Mansour, Tel-Aviv University 1475

On-line Reinforcement Learning Using Incremental Kernel-Based Stochastic Factorization, Andre Barreto, McGill University, Doina Precup, McGill Montreal, and Joelle Pineau, McGill University 1484

Forward-Backward Activation Algorithm for Hierarchical Hidden Markov Models, Kei Wakabayashi, University of Tsukuba, and Takao Miura, Hosei University 1493

Communication-Efficient Algorithms for Statistical Optimization, Yuchen Zhang, UC Berkeley, John Duchi, Berkeley, and Martin Wainwright, UC Berkeley 1502

Identifiability and Unmixing of Latent Parse Trees, Daniel Hsu, Microsoft Research, Sham Kakade, and Percy Liang, Google 1511

Bayesian nonparametric models for ranked data, Francois Caron, INRIA Bordeaux, and Yee Teh, Gatsby 1520

Feature-aware Label Space Dimension Reduction for Multi-label Classification, Yao-Nan Chen, National Taiwan University, and Hsuan-Tien Lin, National Taiwan University 1529

Stochastic optimization and sparse statistical recovery: Optimal algorithms for high dimensions, Alekh Agarwal, UC Berkeley, Sahand Negahban, MIT, and Martin Wainwright, U C Berkeley 1538

Graphical Gaussian Vector for Image Categorization, Tatsuya Harada, The University of Tokyo, and Yasuo Kuniyoshi, The Univ. of Tokyo 1547

Joint Modeling of a Matrix with Associated Text via Latent Binary Features, XianXing Zhang, Duke University, and Lawrence Carin, Duke 1556

Proper losses for learning from partial labels, Jesús Cid-Sueiro, Univ. Carlos III de Madrid 1565

Iterative Thresholding Algorithm for Sparse Inverse Covariance Estimation, Benjamin Rolfs, Stanford University, Bala Rajaratnam, Dominique Guillot, Ian Wong, and Arian Maleki, Stanford University 1574

Selecting Diverse Features via Spectral Regularization, Abhimanyu Das, Yahoo!, Anirban Dasgupta, and Ravi Kumar, Yahoo! 1583

Monte Carlo Methods for Maximum Margin Supervised Topic Models, Qixia Jiang, Tsinghua University, Jun Zhu, Maosong Sun, Tsinghua University, and Eric Xing, CMU 1592

Parametric Local Metric Learning for Nearest Neighbor Classification, Jun Wang, University of Geneva, Alexandros Kalousis, and Adam Woznica, University of Geneva 1601

A Linear Time Active Learning Algorithm for Link Classification, Nicolò Cesa-Bianchi, Univ. degli Studi di Milano, Claudio Gentile, University of Subria, Fabio Vitale, University of Milan, and Giovanni Zappella, University of Milan 1610

Bayesian Warped Gaussian Processes, Miguel Lázaro-Gredilla, Univ. Carlos III de Madrid 1619

Nonparametric Reduced Rank Regression, Rina Foygel, University of Chicago, Michael Horrell, Mathias Drton, and John Lafferty, University of Chicago 1628

Multiresolution analysis on the symmetric group, Risi Kondor, University of Chicago, and Walter Dempsey, University of Chicago, Department of Statistics 1637

Isotropic Hashing, Weihao Kong, Shanghai Jiao Tong University, and Wu-Jun Li, Shanghai Jiao Tong University 1646

On Lifting the Gibbs Sampling Algorithm, Deepak Venugopal, The University of Texas at Dallas, and Vibhav Gogate, UT Dallas 1655

On the connections between saliency and tracking, Vijay Mahadevan, Yahoo! Labs, and Nuno Vasconcelos, UC San Diego 1664

Convex Multi-view Subspace Learning, Martha White, Yaoliang Yu, University of Alberta, Xinhua Zhang, and Dale Schuurmans, University of Alberta 1673

Spectral learning of linear dynamics from generalised-linear observations with application to neural population data, Lars Buesing, Gatsy Unit, UCL, Jakob Macke, MPI Biological Cybernetics, and Maneesh Sahani, UCL 1682

Mixability in Statistical Learning, Tim Erven, Université Paris-Sud, Peter Grünwald, CWI, Mark Reid, and Robert Williamson, NICTA 1691

Waveform Driven Plasticity in BiFeO3 Memristive Devices: Model and Implementation, Christian Mayr, TU Dresden, Paul Stärke, Johannes Partzsch, Love Cederstroem, ZMDI AG, Rene Schüffny, TU Dresden, Yao Shuai, Helmholtz-Zentrum Dresden-Rossendorf e.V., Nan Du, TU Chemnitz, and Heidemarie Schmidt, TU Chemnitz 1700

A lattice filter model of the visual pathway, Karol Gregor, Janelia, and Dmitri Chklovskii, Janelia Farm 1709

Semantic Kernel Forests from Multiple Taxonomies, Sung Hwang, University of Texas, Austin, Kristen Grauman, University of Texas at Austin, and Fei Sha, U. Southern California 1718

Causal discovery with scale-mixture model for spatiotemporal variance dependencies, Zhitang CHEN, Chinese Univ. of Hong Kong, Kun ZHANG, Max Planck Institute for Biological Cybernetics, and Laiwan CHAN, Department of Computer Science and Engineering, Chinese University of Hong Kong 1727

Natural Images, Gaussian Mixtures and Dead Leaves, Daniel Zoran, Hebrew Universi, and Yair Weiss, Hebrew University 1736

Dual-Space Analysis of the Sparse Linear Model, David Wipf, Microsoft Research Asia, and Yi Wu, Tsinghua University 1745

Active Comparison of Prediction Models, Christoph Sawade, University of Potsdam, Niels Landwehr, and Tobias Scheffer, Universität Potsdam 1754

Online Regret Bounds for Undiscounted Continuous Reinforcement Learning, Ronald Ortner, Leoben, and Daniil Ryabko, INRIA 1763

Semi-Crowdsourced Clustering: Generalizing Crowd Labeling by Robust Distance Metric Learning, Jinfeng Yi, Michigan State University, Rong Jin, Michigan State, Anil Jain, Michigan State University, Shaili Jain, Yale University, and Tianbao Yang, Michigan State University 1772

Learning curves for multi-task Gaussian process regression, Peter Sollich, Kings College London, and Simon Ashton, King's College London 1781

Kernel Hyperalignment, Alexander Lorbert, Princeton University, and Peter Ramadge, Princeton 1790

Multiple Choice Learning: Learning to Produce Multiple Structured Outputs, Abner Guzmán-Rivera, UIUC, Dhruv Batra, Virginia Tech, and Pushmeet Kohli, Microsoft Research Cambridge 1799

Mixing Properties of Conditional Markov Chains with Unbounded Feature Functions, Mathieu Sinn, IBM Research, and Bei Chen, McMaster University 1808

Persistent Homology for Learning Densities with Bounded Support, Florian Pokorny, KTH, Carl Ek, Hedvig Kjellström, and Danica Kragic, KTH 1817

On the Use of Non-Stationary Policies for Stationary Infinite-Horizon Markov Decision Processes, Bruno Scherrer, INRIA Loria, and Boris Lesner, INRIA 1826

Efficient Spike-Coding with Multiplicative Adaptation in a Spike Response Model, Sander Bohte, CWI 1835

MAP Inference in Chains using Column Generation, David Belanger, UMass Amherst, Alexandre Passos, Sebastian Riedel, and Andrew McCallum, UMass 1844

Bayesian Nonparametric Modeling of Suicide Attempts, Francisco Ruiz, University Carlos III, Isabel Valera, University Carlos III in Madrid, Carlos Blanco, Columbia University College of Physicians and Surgeons, and Fernando Perez-Cruz 1853

Fiedler Random Fields: A Large-Scale Spectral Approach to Statistical Network Modeling, Antonino Freno, INRIA, Mikaela Keller, and Marc Tommasi, INRIA 1862

Neurally Plausible Reinforcement Learning of Working Memory Tasks, Jaldert Rombouts, Centrum Wiskunde Informatica, Sander Bohte, and Pieter Roelfsema, Netherlands Institute for Neuroscience 1871

Efficient Monte Carlo Counterfactual Regret Minimization in Games with Many Player Actions, Richard Gibson, University of Alberta, Neil Burch, Marc Lanctot, and Duane Szafron, University of Alberta 1880

Repulsive Mixtures, FRANCESCA PETRALIA, DUKE UNIVERSITY, Vinayak Rao, Gatsby Unit UCL, and David Dunson, Duke Statistics 1889

Fully Bayesian inference for neural models with negative-binomial spiking, Jonathan Pillow, UT Austin, and James Scott 1898

Slice Normalized Dynamic Markov Logic Networks, Tivadar Papai, University of Rochester, Henry Kautz, and Daniel Stefankovic, University of Rochester 1907

Meta-Gaussian Information Bottleneck, Melanie Rey, University of Basel, and Volker Roth, University of Basel 1916

Diffusion Decision Making for Adaptive k-Nearest Neighbor Classification, Yung-Kyun Noh, Seoul National University, Frank Park, and Daniel Lee, University of Pennsylvania 1925

The Perturbed Variation, Maayan Harel, Technion, and Shie Mannor, Technion 1934

Communication/Computation Tradeoffs in Consensus-Based Distributed Optimization, Konstantinos Tsianos, McGIll University, Sean Lawlor, and Michael Rabbat 1943

The Coloured Noise Expansion and Parameter Estimation of Diffusion Processes, Simon Lyons, University of Edinburgh, Amos Storkey, and Simo Sarkka, Aalto University 1952

Online allocation and homogeneous partitioning for piecewise constant mean-approximation, Alexandra Carpentier, INRIA, and Odalric-Ambrym Maillard, Leoben 1961

Learning as MAP Inference in Discrete Graphical Models, Xianghang Liu, National ICT Australia, James Petterson, NICTA, and Tiberio Caetano, NICTA Canberra 1970

A mechanistic model of early sensory processing based on subtracting sparse representations, Shaul Druckmann, Janelia Farm Research Campus, Tao Hu, JFRC, HHMI, and Dmitri Chklovskii, JFRC, HHMI 1979

Multi-Stage Multi-Task Feature Learning, Pinghua Gong, Tsinghua University, Jieping Ye, Arizona State U., and Changshui Zhang, Tsinghua University 1988

From Deformations to Parts: Motion-based Segmentation of 3D Objects, Soumya Ghosh, Brown University, Erik Sudderth, Matthew Loper, Max Planck Institute for Intelligent Systems, and Michael Black, Max Planck Institute 1997

Phoneme Classification using Constrained Variational Gaussian Process Dynamical System, Hyunsin Park, KAIST, Sungrack Yun, Sanghyuk Park, Jongmin Kim, and Chang Yoo, KAIST 2006

Bayesian estimation of discrete entropy with mixtures of stick-breaking priors, Evan Archer, University of Texas at Austin, Il Park, and Jonathan Pillow, UT Austin 2015

A Geometric take on Metric Learning, Søren Hauberg, Max Planck Institute, Oren Freifeld, Brown University, and Michael Black, Max Planck Institute 2024

Learning the Architecture of Sum-Product Networks Using Clustering on Variables, Aaron Dennis, Brigham Young University, and Dan Ventura, Brigham Young University 2033

Pointwise Tracking the Optimal Regression Function, Yair Wiener, Technion, and Ran Yaniv, Technion 2042

Bayesian nonparametric models for bipartite graphs, Francois Caron, INRIA Bordeaux 2051

Reducing statistical time-series problems to binary classification, Daniil Ryabko, INRIA, and Jeremie MARY, Univ. Lille / INRIA 2060

Tractable Objectives for Robust Policy Optimization, Katherine Chen, University of Alberta, and Michael Bowling, Alberta 2069

Classification Calibration Dimension for General Multiclass Losses, Harish Guruprasad, Indian Institute of Science, and Shivani Agarwal, IISc Bangalore 2078

Structure estimation for discrete graphical models: Generalized covariance matrices and their inverses, Po-Ling Loh, UC Berkeley, and Martin Wainwright, UC Berkeley 2087

Collaborative Gaussian Processes for Preference Learning, Neil Houlsby, University of Cambridge, Jose Hernandez-Lobato, Cambridge University, Ferenc Huszar, University of Cambridge, and Zoubin Ghahramani, University of Cambridge 2096

Approximating Concavely Parameterized Optimization Problems, Joachim Giesen, Universitaet Jena, Jens Mueller, Soeren Laue, and Sascha Swiercy, Universitaet Jena 2105

Gradient-based kernel method for feature extraction and variable selection, Kenji Fukumizu, The Institute of Statistical Mathematics, and Chenlei Leng, National University of Singapore 2114

Strategic Impatience in Go/NoGo versus Forced-Choice Decision-Making, Pradeep Shenoy, UCSD, and Angela Yu, UCSD .. 2123

On Triangular versus Edge Representations — Towards Scalable Modeling of Networks, Qirong Ho, Carnegie Mellon University, Junming Yin, and Eric Xing, CMU 2132

Relax and Randomize : From Value to Algorithms, Sasha Rakhlin, University of Pennsylvania, Ohad Shamir, Microsoft Research, and Karthik Sridharan, University of Pennsylvania 2141

Minimax Multi-Task Learning and a Generalized Loss-Compositional Paradigm for MTL, Nishant Mehta, Georgia Institute of Technology, Dongryeol Lee, GE Global Research, and Alexander Gray, Georgia Institute of Technology 2150

Spectral Learning of General Weighted Automata via Constrained Matrix Completion, Borja Balle, UPC, and Mehryar Mohri, NYU . 2159

Optimal Neural Tuning Curves for Arbitrary Stimulus Distributions: Discrimax, Infomax and Minimum L_p Loss, Zhuo Wang, University of Pennsylvania, Alan Stocker, U Penn, and Daniel Lee, University of Pennsylvania 2168

Algorithms for Learning Markov Field Policies, Abdeslam Boularias, MPI Tuebingen, Oliver Kroemer, TU Darmstadt, and Jan Peters, Technische Universität Darmstadt 2177

Affine Independent Variational Inference, Edward Challis, University College London, and David Barber, University College London 2186

Learning from the Wisdom of Crowds by Minimax Entropy, Dengyong Zhou, Microsoft Research, John Platt, Sumit Basu, and Yi Mao, Microsoft 2195

Clustering Sparse Graphs, Yudong Chen, The University of Texas at Aus, Sujay Sanghavi, Univ. of Austin, and Huan Xu, NUS 2204

Sketch-Based Linear Value Function Approximation, Marc Bellemare, University of Alberta, Joel Veness, and Michael Bowling, Alberta 2213

Multimodal Learning with Deep Boltzmann Machines, Nitish Srivastava, University of Toronto, and Ruslan Salakhutdinov, University of Toronto 2222

Learning with Target Prior, Zuoguan Wang, Rensselaer Polytechnic Institu, Siwei Lyu, SUNY Albany, Gerwin Schalk, Wadsworth Center, and Qiang Ji 2231

Slice sampling normalized kernel-weighted completely random measure mixture models, Nicholas Foti, Dartmouth College, and Sinead Williamson, CMU 2240

Scalable Inference of Overlapping Communities, Prem Gopalan, Princeton University, David Mimno, Sean Gerrish, Michael Freedman, and David Blei, Princeton 2249

Online L1-Dictionary Learning with Application to Novel Document Detection, Shiva Kasiviswanathan, IBM Research, Huahua Wang, Uni. of Minnesota, twin cities, Arindam Banerjee, Univ. of Minnesota, and Prem Melville, IBM Watson Research 2258

A systematic approach to extracting semantic information from functional MRI data, Francisco Pereira, Siemens Corporate Research, and Matthew Botvinick, Princeton University 2267

Why MCA? Nonlinear sparse coding with spike-and-slab prior for neurally plausible image encoding , Jacquelyn Shelton, FIAS, Philip Sterne, Joerg Bornschein, Abdul-Saboor Sheikh, and Joerg Luecke, FIAS 2276

Learning optimal spike-based representations, ralph Bourdoukan, École Normale Supérieure, David Barrett, Christian Machens, Champalimaud Institute, and Sophie Deneve, ENS Paris 2285

Collaborative Ranking With 17 Parameters, Maksims Volkovs, University of Toronto, and Rich Zemel, University of Toronto 2294

Rational inference of relative preferences, Nisheeth Srivastava, University of Minnesota, and Paul Schrater, University of Minnesota 2303

The topographic unsupervised learning of natural sounds in the auditory cortex, Hiroki Terashima, The University of Tokyo / JSPS, and Masato Okada, The University of Tokyo / RIKEN BSI 2312

Approximating Equilibria in Sequential Auctions with Incomplete Information and Multi-Unit Demand, Amy Greenwald, Brown University, Jiacui Li, and Eric Sodomka, Brown University 2321

A Divide-and-Conquer Method for Sparse Inverse Covariance Estimation, Cho-Jui Hsieh, University of Texas at Austin, Inderjit Dhillon, University of Texas, Pradeep Ravikumar, University of Texas-Austin, and Arindam Banerjee, Univ. of Minnesota 2330

A Simple and Practical Algorithm for Differentially Private Data Release, Moritz Hardt, IBM Almaden Research, Katrina Ligett, Caltech, and Frank McSherry, Microsoft Research 2339

Bayesian active learning with localized priors for fast receptive field characterization, Mijung Park, UT Austin, and Jonathan Pillow, UT Austin 2348

Weighted Likelihood Policy Search with Model Selection, Tsuyoshi Ueno, Minato Discrete Structure Mani, Kohei Hayashi, The University of Tokyo, Takashi Washio, and Yoshinobu Kawahara, Osaka University 2357

Learning the Dependency Structure of Latent Factors, Yunlong He, Georgia Tech, Yanjun Qi, NEC Labs America, Koray Kavukcuoglu, NEC Laboratories America, and Haesun Park, Georgia Tech 2366

Provable ICA with Unknown Gaussian Noise, with Implications for Gaussian Mixtures and Autoencoders, Sanjeev Arora, Rong Ge, Princeton University, Ankur Moitra, IAS, and Sushant Sachdeva .. 2375

Globally Convergent Dual MAP LP Relaxation Solvers using Fenchel-Young Margins, Alex Schwing, ETH Zurich, Tamir Hazan, Toyota Technological Institute Chicago, Marc Pollefeys, ETH Zurich, and Raquel Urtasun, Toyota Technological Institute Chicago 2384

Dip-means: an incremental clustering method for estimating the number of clusters, Argyris Kalogeratos, University of Ioannina, and Aristidis Likas, University of Ioannina 2393

No-Regret Algorithms for Unconstrained Online Convex Optimization, Matt Streeter, Duolingo, and Brendan McMahan, Google 2402

Bayesian models for Large-scale Hierarchical Classification , Siddharth Gopal, Carnegie Mellon University, Yiming Yang, cmu, Bing Bai, NEC Laboratories America, and Alexandru Niculescu-Mizil, NEC Laboratories America 2411

Recovery of Sparse Probability Measures via Convex Programming, Mert Pilanci, UC Berkeley, Laurent Ghaoui, and Venkat Chandrasekaran 2420

Multiple Operator-valued Kernel Learning, Hachem Kadri, Aix-Marseille University/LIF-QARMA, Alain Rakotomamonjy, University of Rouen, Francis Bach, INRIA, and Philippe Preux, INRIA Lille 2429

Approximate Message Passing with Consistent Parameter Estimation and Applications to Sparse Learning, Ulugbek Kamilov, EPFL, Sundeep Rangan, NYU-Poly, Alyson Fletcher, UC Berkeley, and MIchael Unser, EPFL 2438

A Better Way to Pretrain Deep Boltzmann Machines, Ruslan Salakhutdinov, University of Toronto, and Geoff Hinton, University of Toronto 2447

Towards a learning-theoretic analysis of spike-timing dependent plasticity, David Balduzzi, Max Planck Institute for Intelligent Systems, and Michel Besserve, MPI for Intelligent Systems 2456

Learning Manifolds with K-Means and K-Flats, Guillermo Canas, IIT-MIT, Tomaso Poggio, MIT, and Lorenzo Rosasco, MIT ... 2465

Iterative ranking from pair-wise comparisons , Sahand Negahban, MIT, Sewoong Oh, and Devavrat Shah, MIT 2474

A Polynomial-time Form of Robust Regression, Yaoliang Yu, University of Alberta, Ozlem Aslan, and Dale Schuurmans, University of Alberta 2483

Learning Probability Measures with respect to Optimal Transport Metrics, Guillermo Canas, IIT-MIT, and Lorenzo Rosasco, MIT 2492

Label Ranking with Partial Abstention based on Thresholded Probabilistic Models, Weiwei Cheng, University of Marburg, Eyke Huellermeier, Universitaet Marburg, Willem Waegeman, Ghent University, and Volkmar Welker, University of Marburg 2501

Adaptive Learning of Smoothing Functions: Application to Electricity Load Forecasting, Amadou Ba, IBM Research, Mathieu Sinn, Yannig Goude, EDF R&D, and Pascal Pompey, IBM Research 2510

Tensor Decomposition for Fast Parsing with Latent-Variable PCFGs, Shay Cohen, Columbia University, and Michael Collins, Columbia Univ. 2519

Semi-supervised Eigenvectors for Locally-biased Learning, Toke Hansen, DTU Informatics, and Michael Mahoney, Stanford University 2528

Exponential Concentration for Mutual Information Estimation with Application to Forests, Han Liu, Princeton University, John Lafferty, University of Chicago, and Larry Wasserman 2537

Augment-and-Conquer Negative Binomial Processes, Mingyuan Zhou, Duke University, and Lawrence Carin, Duke 2546

Transferring Expectations in Model-based Reinforcement Learning, Trung Nguyen, NUS, Tomi Silander, and Tze Leong, NUS 2555

Minimization of Continuous Bethe Approximations: A Positive Variation, Jason Pacheco, Brown University, and Erik Sudderth, Brown University 2564

Non-linear Metric Learning, Dor Kedem, Washington University in St. Louis, Stephen Tyree, Washington U in St Louis, Kilian Weinberger, Washington University in St. Louis, Fei Sha, U. Southern California, and Gert Lanckriet, UCSD 2573

Factorial LDA: Sparse Multi-Dimensional Text Models, Michael Paul, Johns Hopkins University, and Mark Dredze, Johns Hopkins 2582

Ancestor Sampling for Particle Gibbs, Fredrik Lindsten, Linköping University, Michael Jordan, Berkley University, and Thomas Schön, Linköping University 2591

Modelling Reciprocating Relationships with Hawkes Processes, Charles Blundell, Gatsby Unit UCL, Katherine Heller, MIT, and Jeff Beck 2600

Expectation Propagation in Gaussian Process Dynamical Systems, Marc Deisenroth, TU Darmstadt, and Shakir Mohamed, UBC 2609

A quasi-Newton proximal splitting method, Stephen Becker, Paris 6 University, and Jalal Fadili, CNRS-ENSICAEN-Univ. Caen 2618

Exact and Stable Recovery of Sequences of Signals with Sparse Increments via Differential ℓ_1-Minimization, Demba Ba, MIT/Harvard, Behtash Babadi, Patrick Purdon, and Emery Brown, MIT/Harvard 2627

Efficient Reinforcement Learning for High Dimensional Linear Quadratic Systems, Morteza Ibrahimi, Stanford University, Adel Javanmard, and Benjamin Roy, Stanford University 2636

Multilabel Classification using Bayesian Compressed Sensing, Ashish Kapoor, MSR Redmond, Raajay Viswanathan, Microsoft Research Lab India, and Prateek Jain, MSR Bangalore 2645

Scaling MPE Inference for Constrained Continuous Markov Random Fields with Consensus Optimization, Stephen Bach, University of Maryland, Matthias Broecheler, Lise Getoor, University of Maryland College Park, and Dianne O'Leary, University of Maryland 2654

A Stochastic Gradient Method with an Exponential Convergence

Rate for Finite Training Sets, Nicolas Roux, INRIA, Mark Schmidt, and Francis Bach, INRIA 2663

Query Complexity of Derivative-Free Optimization, Kevin Jamieson, University of Wisconsin, Rob Nowak, Wisconsin, and Ben Recht, University of Wisconsin 2672

Emergence of Object-Selective Features in Unsupervised Feature Learning, Adam Coates, Stanford University, Andrej Karpathy, and Andrew Ng, Stanford Univ. 2681

Burn-in, bias, and the rationality of anchoring, Falk Lieder, ETH Zurich, Thomas Griffiths, UC Berkeley, and Noah Goodman, Stanford .. 2690

Truly Nonparametric Online Variational Inference for Hierarchical Dirichlet Processes, Michael Bryant, Brown University, and Erik Sudderth, Brown University 2699

A Neural Autoregressive Topic Model, Hugo Larochelle, Sherbrooke University, and Stanislas Lauly, Université de Sherbrooke 2708

A Unifying Perspective of Parametric Policy Search Methods for Markov Decision Processes, Thomas Furmston, University College London, and David Barber, University College London 2717

Entangled Monte Carlo, Seong-Hwan Jun, UBC Department of Statistics, Liangliang Wang, and Alexandre Bouchard-Côté, U. British Columbia 2726

Near-Optimal MAP Inference for Determinantal Point Processes, Jennifer Gillenwater, University of Pennsylvania, Alex Kulesza, and Ben Taskar, University of Pennsylvania 2735

Probabilistic Low-Rank Subspace Clustering, S. BABACAN, UNIVERSITY OF ILLINOIS, Shinichi Nakajima, Nikon Corporation, and Minh Do, University of Illinois 2744

How They Vote: Issue-Adjusted Models of Legislative Behavior, Sean Gerrish, Princeton University, and David Blei, Princeton 2753

Density Propagation and Improved Bounds on the Partition Function, Stefano Ermon, Cornell University, Carla Gomes, Ashish Sabharwal, IBM Watson Research Center, and Bart Selman, Cornell University 2762

Perceptron Learning of SAT, Alex Flint, University of Oxford, and Matthew Blaschko, Ecole Centrale Paris 2771

Learning Networks of Heterogeneous Influence, NAN DU, Georgia Tech, Le Song, Alex Smola, Yahoo! Research, and Ming Yuan 2780

Multiclass Learning with Simplex Coding, Youssef Mroueh, MIT, Tomaso Poggio, Lorenzo Rosasco, and Jean-Jacques Slotine, Massachusetts Institute of Technology 2789

FastEx: Hash Clustering with Exponential Families, Amr Ahmed, Sujith Ravi, Yahoo! Research, Shravan Narayanamurthy, Yahoo, and Alex Smola, Yahoo! Research 2798

Topic-Partitioned Multinetwork Embeddings, Peter Krafft, Massachusetts Institute of Technology, Juston Moore, University of Massachusetts Amherst, Bruce Desmarais, and Hanna Wallach, University of Massachusetts Amherst 2807

Learning Label Trees for Probabilistic Modelling of Implicit Feedback, Andriy Mnih, Gatsby Computational Neuroscience Unit, and Yee Teh, Gatsby 2816

Learning with Recursive Perceptual Representations, Oriol Vinyals, UC Berkeley, Yangqing Jia, Li Deng, Microsoft Research, and Trevor Darrell 2825

Link Prediction in Graphs with Autoregressive Features, Emile Richard, Stephane Gaiffas, Universite Paris 6, and Nicolas Vayatis, CMLA ENS Cachan 2834

Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy Images, Dan Ciresan, IDSIA, Alessandro Giusti, luca Gambardella, and Juergen Schmidhuber, IDSIA 2843

Scalable imputation of genetic data with a discrete fragmentation-coagulation process, Lloyd Elliott, Gatsby Unit UCL, and Yee Teh, Gatsby 2852

Gradient Weights help Nonparametric Regressors, Samory Kpotufe, Tuebingen, and Abdeslam Boularias, MPI Tuebingen 2861

Online Sum-Product Computation Over Trees, Mark Herbster, University College London, Stephen Pasteris, UCL, and Fabio Vitale, University of Milan 2870

Sparse Approximate Manifolds for Differential Geometric MCMC, Ben Calderhead, University College London, and Matyas Sustik, University of Texas at Austin 2879

Fast Variational Inference in the Conjugate Exponential Family , James Hensman, The University of Sheffield, Magnus Rattray, The University of Manchester, and Neil Lawrence, U. Sheffield 2888

Bayesian Pedigree Analysis using Measure Factorization, Bonnie Kirkpatrick, UBC, and Alexandre Bouchard-Côté, U. British Columbia 2897

Accelerated Training for Matrix-norm Regularization: A Boosting Approach, Xinhua Zhang, University of Alberta, Yaoliang Yu, and Dale Schuurmans, University of Alberta 2906

Controlled Recognition Bounds for Visual Learning and Exploration, Vasiliy Karasev, UCLA, Alessandro Chiuso, University of Padova, and Stefano Soatto, UCLA 2915

Distributed Probabilistic Learning for Camera Networks with Missing Data, Sejong Yoon, Rutgers University, and Vladimir Pavlovic, Rutgers University 2924

Submodular-Bregman and the Lovász-Bregman Divergences with Applications, Rishabh Iyer, University of Washington, and Jeff Bilmes, Univ. Washington 2933

Minimizing Uncertainty in Pipelines, Nilesh Dalvi, Yahoo! Research, Aditya Parameswaran, Stanford University, and Vibhor Rastogi, Yahoo! Research 2942

Practical Bayesian Optimization of Machine Learning Algorithms , Jasper Snoek, University of Toronto, Hugo Larochelle, Sherbrooke University, and Ryan Adams, Harvard University 2951

Forging The Graphs: A Low Rank and Positive Semidefinite Graph Learning Approach, Dijun Luo, UTA, Chris Ding, University of Texas at Arlington, Heng Huang, University of Texas Arlington, and Feiping Nie, University of Texas at Arlington 2960

The Time-Marginalized Coalescent Prior for Hierarchical Clustering, Levi Boyles, University of California Irvin, and Max Welling, UCI 2969

Fusion with Diffusion for Robust Visual Tracking, Yu Zhou, Huazhong Univ.of Sci. & Tech. , Xiang Bai, Huazhong Univ. of Sci. & Tech., Wenyu Liu, and Longin Latecki, Temple University 2978

A nonparametric variable clustering model, David Knowles, University of Cambridge, Konstantina Palla, and Zoubin Ghahramani, University of Cambridge 2987

Priors for Diversity in Generative Latent Variable Models, James Zou, Harvard University, and Ryan Adams, Harvard University 2996

A Nonparametric Conjugate Prior Distribution for the Maximizing Argument of a Noisy Function, Pedro Ortega, MPI for Intelligent Systems, Jordi Grau-Moya, Tim Genewein, David Balduzzi, Max Planck Institute for Intelligent Systems, and Daniel Braun . 3005

Convergence Rate Analysis of MAP Coordinate Minimization Algorithms, Ofer Meshi, Hebrew University, Tommi Jaakkola, MIT, and Amir Globerson, Hebrew University 3014

Projection Retrieval for Classification, Madalina Fiterau, Carnegie Mellon University, and Artur Dubrawski, Carnegie Mellon University 3023

Hierarchical spike coding of sound, Yan Karklin, New York University, Chaitanya Ekanadham, Courant Institute, NYU, and Eero Simoncelli, NYU 3032

Human memory search as a random walk in a semantic network, Joshua Abbott, UC Berkeley, Joseph Austerweil, and Thomas Griffiths, UC Berkeley 3041

Probabilistic n-Choose-k Models for Classification and Ranking, Kevin Swersky, Danny Tarlow, University of Toronto, Ryan Adams, Harvard University, Rich Zemel, University of Toronto, and Brendan Frey, University of Toronto 3050

Complex Inference in Neural Circuits with Probabilistic Population Codes and Topic Models, Jeff Beck, Katherine Heller, MIT, and Alexandre Pouget 3059

Cost-Sensitive Exploration in Bayesian Reinforcement Learning, Dongho Kim, University of Cambridge, Kee-Eung Kim, KAIST, and Pascal Poupart, University of Waterloo 3068

Learning with Partially Absorbing Random Walks, Xiao-Ming Wu, Columbia University, Zhenguo Li, Anthony So, John Wright, and Shih-Fu Chang, Columbia 3077

Locating Changes in Highly Dependent Data with Unknown Number of Change Points, Azadeh Khaleghi, INRIA Lille - Nord Europe, and Daniil Ryabko, INRIA 3086

Probabilistic Event Cascades for Alzheimer's disease, Jonathan Huang, Stanford University, and Daniel Alexander, University College London 3095

Efficient and direct estimation of a neural subunit model for sensory coding, Brett Vintch, New York University, Andrew Zaharia, J Movshon, and Eero Simoncelli, New York University, HHMI 3104

One Permutation Hashing, Ping Li, Cornell, Art Owen, Stanford University, and Cun-Hui Zhang, Rutgers 3113

Unsupervised Template Learning for Fine-Grained Object Recognition, Shulin Yang, University of Washington, Liefeng Bo, Jue Wang, and Linda Shapiro 3122

Risk Aversion in Markov Decision Processes via Near Optimal Chernoff Bounds , Teodor Moldovan, UC Berkeley, and Pieter Abbeel, Berkrley 3131

Fast Bayesian Inference for Non-Conjugate Gaussian Process Regression, Emtiyaz Khan, Shakir Mohamed, UBC, and Kevin Murphy, UBC 3140

Imitation Learning by Coaching, He He, Univ. of Maryland,College Park, Hal III, University of Maryland, and Jason Eisner 3149

Small-Variance Asymptotics for Exponential Family Dirichlet Process Mixture Models, Ke Jiang, The Ohio State University, Brian Kulis, Ohio State U., and Michael Jordan, Berkley University 3158

A latent factor model for highly multi-relational data, Rodolphe Jenatton, INRIA, Nicolas Roux, Antoine Bordes, Universitá de Technologie de Compiágne, and Guillaume Obozinski, ENS 3167

Entropy Estimations Using Correlated Symmetric Stable Random Projections, Ping Li, Cornell, and Cun-Hui Zhang, Rutgers 3176

Simultaneously Leveraging Output and Task Structures for Multiple-Output Regression, Piyush Rai, Utah, Abhishek Kumar, University of Maryland, College Park, and Hal III, University of Maryland 3185

Continuous Relaxations for Discrete Hamiltonian Monte Carlo, Yichuan Zhang, University of Edinburgh, Charles Sutton, Amos Storkey, and Zoubin Ghahramani, University of Cambridge 3194

Deep Learning of Invariant Features via Simulated Fixations in Video, Will Zou, Stanford University, Andrew Ng, Stanford Univ., Shenghuo Zhu, NEC Laboratories America, and Kai Yu, NEC Labs Cupertino 3203

Best Arm Identification: A Unified Approach to Fixed Budget and Fixed Confidence, Victor Gabillon, INRIA Lille, Team SequeL, Mohammad Ghavamzadeh, INRIA Lille, and Alessandro Lazaric, INRIA Lille 3212

On the Sample Complexity of Robust PCA, Matthew Coudron, MIT, and Gilad Lerman, University of Minnesota 3221

Latent Coincidence Analysis: A Hidden Variable Model for Distance Metric Learning, Matthew Der, UC San Diego, and Lawrence Saul, University of California San Diego 3230

Discriminative Learning of Sum-Product Networks, Robert Gens, University of Washington, and Pedro Domingos, University of Washington .. 3239

Trajectory-Based Short-Sighted Probabilistic Planning, Felipe Trevizan, Carnegie Mellon University, and Manuela Veloso, Carnegie Mellon University 3248

Tight Bounds on Profile Redundancy and Distinguishability, Jayadev Acharya, UCSD, Hirakendu Das, University of California, San Diego, and Alon Orlitsky 3257

Interpreting prediction markets: a stochastic approach, Nicolas Penna, ANU / NICTA, Mark Reid, and Rafael Frongillo, UC Berkley . 3266

Risk-Aversion in Multi-armed Bandits, Amir Sani, INRIA Lille - Nord Europe, Alessandro Lazaric, INRIA Lille, and Remi Munos, Inria Lille 3275

Confusion-Based Online Learning and a Passive-Aggressive Scheme, Liva Ralaivola, Aix-Marseille University 3284

Cardinality Restricted Boltzmann Machines , Kevin Swersky, Danny Tarlow, University of Toronto, Ilya Sutskever, Ruslan Salakhutdinov, Rich Zemel, and Ryan Adams, Harvard University 3293

An Integer Optimization Approach to Associative Classification, Dimitris Bertsimas, Allison Chang, Cynthia Rudin.....3302

A dynamic excitatory-inhibitory network in a VLSI chip for spiking information reregistration, Juan Huo.....3311

Generalization Bounds for Domain Adaptation, Chao Zhang, Lei Zhang, Jieping Ye.....3320

Author Index