Neural Information Processing Systems

Advances in Neural Information Processing Systems 23

24th Annual Conference on Neural Information Processing Systems 2010

December 6-9, 2010
Vancouver, B.C., Canada

Volume 1 of 3

Printed from e-media with permission by:

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

ISBN: 978-1-61782-380-0

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© (2010) by Neural Information Processing Systems
All rights reserved.

Printed by Curran Associates, Inc. (2011)

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

ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 23

Proceedings of the 2010 Conference

http://books.nips.cc/nips23.html

edited by

John Lafferty, Chris Williams, John Shawe-Taylor,
Richard Zemel and Aron Culotta

©2010 NIPS Foundation — All Rights Reserved.

Contents

Contents iii

Preface xxiii

Donors xxvi

NIPS foundation xxvii

Committees xxviii

Reviewers xxxi

Repeated Games against Budgeted Adversaries, Jacob Abernethy, UC Berkeley, and Manfred Warmuth, UC santa Cruz 1

Towards Property-Based Classification of Clustering Paradigms, Margareta Ackerman, Shai Ben-David, and David R Loker, University of Waterloo 10

Tree-Structured Stick Breaking for Hierarchical Data, Ryan Adams, University of Toronto, Zoubin Ghahramani, Cambridge, and Michael I Jordan, University of California Berkeley ..... N/A

Sparse Instrumental Variables (SPIV) for Genome-Wide Studies, Felix V Agakov, University of Edinburgh, Paul McKeigue, , Jon Krohn, WTCHG Oxford, and Amos J Storkey, University of Edinburgh 28

Fast global convergence rates of gradient methods for high-dimensional statistical recovery, Alekh Agarwal, UC Berkeley, Sahand Negahban, University of California at Berkeley, and Martin Wainwright, UC Berkeley 37

Learning Multiple Tasks using Manifold Regularization, Arvind Agarwal, Hal Daume III, and Samuel Gerber, University of Utah 46

Switched Latent Force Models for Movement Segmentation, Mauricio Alvarez, University of Manchester, Jan Peters, MPI for biological cybernetics, Bernhard Schoelkopf, MPI for Biological Cybernetics, and Neil D Lawrence, University of Manchester 55

A POMDP Extension with Belief-dependent Rewards, Mauricio A Araya, Nancy University/INRIA/CNRS, Olivier Buffet, INRIA / Nancy University / CNRS, Vincent Thomas, Nancy University / INRIA / CNRS, and François Charpillet, INRIA 64

Global seismic monitoring as probabilistic inference, Nimar S Arora, Stuart Russell, UC Berkeley, Paul Kidwell, Lawrence Livermore National Lab, and Erik Sudderth, Brown University 73

Learning invariant features using the Transformed Indian Buffet Process, Joseph L Austerweil, UC Berkeley, and Tom Griffiths, University of California-Berkeley 82

Supervised Clustering, Pranjal Awasthi, CMU, and Reza Bosagh Zadeh, Stanford University 91

Occlusion Detection and Motion Estimation with Convex Optimization, Alper Ayvaci, Michalis Raptis, and Stefano Soatto, UCLA 100

Batch Bayesian Optimization via Simulation Matching, Javad Azimi, Alan Fern, and Xiaoli Fern, Oregon State University 109

Structured sparsity-inducing norms through submodular functions, Francis Bach, Ecole Normale Superieure 118

A Bayesian Approach to Concept Drift, Stephen H Bach, and Mark Maloof, Georgetown University 127

Auto-Regressive HMM Inference with Incomplete Data for Short-Horizon Wind Forecasting, Chris Barber, Joseph Bockhorst, and Paul Roebber, UW Milwaukee 136

The LASSO risk: asymptotic results and real world examples, Mohsen Bayati, José Bento Ayres Pereira, and Andrea Montanari, Stanford University 145

Extensions of Generalized Binary Search to Group Identification and Exponential Costs, Gowtham Bellala, University of Michigan Ann Arbor, Suresh Bhavnani, University of Texas, and Clayton Scott, University of Michigan 154

Label Embedding Trees for Large Multi-Class Tasks, Samy Bengio, Jason Weston, Google, and David Grangier, NEC Labs America 163

Learning Networks of Stochastic Differential Equations, José Bento Ayres Pereira, Morteza Ibrahimi, and Andrea Montanari, Stanford University 172

Exploiting weakly-labeled Web images to improve object classification: a domain adaptation approach, Alessandro Bergamo, and Lorenzo Torresani, Dartmouth College 181

Online Classification with Specificity Constraints, Andrey Bernstein, Shie Mannor, and Nahum Shimkin, Technion 190

Agnostic Active Learning Without Constraints, Alina Beygelzimer, IBM Research, Daniel Hsu, Rutgers University and University of Pennsylvania, John Langford, Yahoo Research, and Zhang Tong, Rutgers University 199

Inference with Multivariate Heavy-Tails in Linear Models, Danny Bickson, CMU, and Carlos Guestrin, Carnegie Mellon University 208

CUR from a Sparse Optimization Viewpoint, Jacob Bien, Ya Xu, Stanford University, and Michael Mahoney, Stanford 217

Optimal learning rates for Kernel Conjugate Gradient regression, Gilles Blanchard, Weierstrass Institute Berlin / University of Potsdam, and Nicole Krämer, Weierstrass Institute Berlin 226

Simultaneous Object Detection and Ranking with Weak Supervision, Matthew B Blaschko, University of Oxford, Andrea Vedaldi, Oxford University, and Andrew Zisserman, University of Oxford .. 235

Kernel Descriptors for Visual Recognition, Liefeng Bo, University of Washington, Xiaofeng Ren, Intel, and Dieter Fox, HASH(0xb20114) ... 244

Fractionally Predictive Spiking Neurons, Sander Bohte, and Jaldert Rombouts, CWI ... 253

Gaussian Process Preference Elicitation, Edwin V Bonilla, Shengbo Guo, NICTA, and Scott Sanner, Nicta ... 262

Predictive State Temporal Difference Learning, Byron Boots, and Geoff Gordon, Carnegie Mellon University ... 271

Variational Inference over Combinatorial Spaces, Alexandre Bouchard-Cote, Berkeley, and Michael I Jordan, University of California Berkeley ... 280

Bootstrapping Apprenticeship Learning, Abdeslam Boularias, MPI for Biological Cybernetics, and Brahim Chaib-draa, ... 289

Random Projections for k-means Clustering, Christos Boutsidis, RPI, Anastasios Zouzias, University of Toronto, and Petros Drineas, Rensselaer Polytechnic Institute ... 298

Segmentation as Maximum-Weight Independent Set, William Brendel, OSU, and Sinisa Todorovic, Oregon State University ... 307

Computing Marginal Distributions over Continuous Markov Networks for Statistical Relational Learning, Matthias Broecheler, University of Maryland CP, and Lise Getoor, HASH(0x98d734) ... 316

Multi-label Multiple Kernel Learning by Stochastic Approximation: Application to Visual Object Recognition, Serhat S Bucak, Rong Jin, Michigan State University, and Anil K Jain, ... 325

Learning concept graphs from text with stick-breaking priors, America Chambers, University of California Irvi, Padhraic Smyth, UCI, and Mark Steyvers, University of California-Irvine ... 334

Rates of convergence for the cluster tree, Kamalika Chaudhuri, and Sanjoy Dasgupta, UC San Diego ... 343

Evidence-Specific Structures for Rich Tractable CRFs, Anton Chechetka, and Carlos Guestrin, Carnegie Mellon University ... 352

Predictive Subspace Learning for Multi-view Data: a Large Margin Approach, Ning Chen, Tsinghua University, Jun Zhu, and Eric Xing, Carnegie Mellon University ... 361

Two-Layer Generalization Analysis for Ranking Using Rademacher Average, Wei Chen, Chinese Academy of Sciences, Tie-Yan Liu, Microsoft Research, and Zhi-Ming Ma, Chinese Academy of Sciences ... 370

SpikeAnts, a spiking neuron network modelling the emergence of organization in a complex system, Sylvain Chevallier, INRIA-Saclay, Hélène Paugam-Moisy, Univ. Lyon 2, and MicheLe Sebag, Laboratoire de Recherche en Informatique CNRS ... 379

Movement extraction by detecting dynamics switches and repetitions, Silvia Chiappa, Statistical Laboratory Cambridge University UK, and Jan Peters, MPI for biological cybernetics 388

Learning sparse dynamic linear systems using stable spline kernels and exponential hyperpriors, Alessandro Chiuso, and Gianluigi Pillonetto, University of Padova 397

Universal Kernels on Non-Standard Input Spaces, Andreas Christmann, University of Bayreuth, and Ingo Steinwart, University of Stuttgart 406

Causal discovery in multiple models from different experiments, Tom Claassen, and Tom Heskes, Radboud University Nijmegen 415

Empirical Risk Minimization with Approximations of Probabilistic Grammars, Shay Cohen, and Noah Smith, Carnegie Mellon University .. 424

Mixture of time-warped trajectory models for movement decoding, Elaine A Corbett, Eric J Perreault, Northwestern University, and Konrad Koerding, Northwestern 433

Learning Bounds for Importance Weighting, Corinna Cortes, Google Inc, Yishay Mansour, Tel-Aviv, and Mehryar Mohri, ... 442

Learning via Gaussian Herding, Koby Crammer, Technion, and Daniel Lee, University of Pennsylvania 451

Spatial and anatomical regularization of SVM for brain image analysis, Remi Cuingnet, Marie Chupin, CRICM, Habib Benali, INSERM LIF, and Olivier Colliot, CRICM / CNRS / Universite Pierre et Marie Curie / Paris France 460

Phone Recognition with the Mean-Covariance Restricted Boltzmann Machine, George E Dahl, Marc'Aurelio Ranzato, Abdel-rahman Mohamed, and Geoffrey Hinton, University of Toronto 469

Co-regularization Based Semi-supervised Domain Adaptation, Hal Daume III, Abhishek Kumar, University of Utah, and Avishek Saha, ... 478

Spectral Regularization for Support Estimation, Ernesto De Vito, DIMA, Lorenzo Rosasco, MIT and IIT, and alessandro toigo, Unimi 487

Random Projection Trees Revisited, aman Dhesi, Princeton University, and Purushottam Kar, Indian Institute of Technology 496

Throttling Poisson Processes, Uwe Dick, University of Potsdam, Peter Haider, , Thomas Vanck, , Michael Brückner, University of Potsdam, and Tobias Scheffer, Uni Potsdam 505

t-logistic regression, Nan Ding, and S.V.N. Vishwanathan, Purdue University 514

Implicit Differentiation by Perturbation, Justin Domke, Rochester Institute of Technology 523

Nonparametric Bayesian Policy Priors for Reinforcement Learning, Finale Doshi-Velez, MIT, David Wingate, Massachusetts Institute of Technology, Nicholas Roy, MIT, and Joshua Tenenbaum, Massachusetts Institute of Technology 532

Over-complete representations on recurrent neural networks can support persistent percepts, Shaul Druckmann, and Dmitri B Chklovskii, JFRC 541

Distributed Dual Averaging In Networks, John Duchi, Alekh Agarwal, and Martin Wainwright, UC Berkeley 550

Copula Bayesian Networks, Gal Elidan, Hebrew University 559

Error Propagation for Approximate Policy and Value Iteration, Amir-massoud Farahmand, University of Alberta, Remi Munos, INRIA Lille - Nord Europe, and Csaba Szepesvari, University of Alberta 568

A Computational Decision Theory for Interactive Assistants, Alan Fern, and Prasad Tadepalli, Oregon State University 577

Parametric Bandits: The Generalized Linear Case, Sarah Filippi, LTCI-CNRS Télécom ParisTech, Olivier Cappe, Telecom ParisTech stochastic optimization, Aurélien Garivier, LTCI -CNRS Télécom ParisTech, and Csaba Szepesvari, University of Alberta 586

A Novel Kernel for Learning a Neuron Model from Spike Train Data, Nicholas K Fisher, and Arunava Banerjee, University of Florida 595

Extended Bayesian Information Criteria for Gaussian Graphical Models, Foygel Rina, and Mathias Drton, University of Chicago 604

Shadow Dirichlet for Restricted Probability Modeling, Bela A Frigyik, , Maya Gupta, University of Washington, and Yihua Chen, ... 613

Size Matters: Metric Visual Search Constraints from Monocular Metadata, Mario J Fritz, UC Berkeley & ICSI, Kate Saenko, MIT, and Trevor Darrell, UC Berkeley 622

A Bayesian Framework for Figure-Ground Interpretation, Vicky Froyen, Rutgers, Jacob Feldman, Rutgers University, and Manish Singh, ... 631

Attractor Dynamics with Synaptic Depression, C. C. Alan Fung, K. Y. Michael Wong, The Hong Kong University of Science and Technology, He Wang, Tsinghua University, and Si Wu, Institute of Neuroscience Chinese Academy of Sciences 640

Learning Kernels with Radiuses of Minimum Enclosing Balls, Kun Gai, Guangyun Chen, and Changshui Zhang, Tsinghua University 649

Implicit encoding of prior probabilities in optimal neural populations, Deep Ganguli, and Eero P Simoncelli, New York University 658

Short-term memory in neuronal networks through dynamical compressed sensing, Surya Ganguli, UCSF, and Haim Sompolinsky, Hebrew University and Harvard University 667

Group Sparse Coding with a Laplacian Scale Mixture Prior, Pierre J Garrigues, IQ Engines Inc., and Bruno A Olshausen, UC Berkeley 676

Improvements to the Sequence Memoizer, Jan Gasthaus, and Yee Whye Teh, Gatsby Unit UCL 685

On Herding and the Perceptron Cycling Theorem, Andrew E Gelfand, Univ. of California Irvine, Yutian Chen, , Laurens van der Maaten, University of California San Diego, and Max Welling, UC Irvine ... 694

Rescaling, thinning or complementing? On goodness-of-fit procedures for point process models and Generalized Linear Models, Felipe Gerhard, and Wulfram Gerstner, Brain Mind Institute EPFL ... 703

The Neural Costs of Optimal Control, Samuel Gershman, and Robert Wilson, Princeton University ... 712

LSTD with Random Projections, Mohammad Ghavamzadeh, Alessandro Lazaric, INRIA Lille, Odalric Maillard, INRIA Lille Team SequeL, and Remi Munos, INRIA Lille - Nord Europe ... 721

Humans Learn Using Manifolds, Reluctantly, Bryan R Gibson, Univ. of Wisonsin-Madison, Xiaojin (Jerry) Zhu, U. Wisconsin-Madison, Tim Rogers, UW-Madison, Chuck Kalish, , and Joseph Harrison, ... 730

Universal Consistency of Multi-Class Support Vector Classification, Tobias Glasmachers, IDSIA ... 739

Learning Efficient Markov Networks, Vibhav G Gogate, University of Washington Seat, William Webb, University of Washington Seattle, and Pedro Domingos, University of Washington ... 748

Transduction with Matrix Completion: Three Birds with One Stone, Andrew B Goldberg, UW-Madison, Xiaojin (Jerry) Zhu, U. Wisconsin-Madison, Ben Recht, UW-Madison, Junming Xu, and Rob Nowak, University of Wisconsin-Madison ... 757

Near-Optimal Bayesian Active Learning with Noisy Observations, Daniel Golovin, Andreas Krause, and Debajyoti Ray, Caltech ... 766

Discriminative Clustering by Regularized Information Maximization, Ryan G Gomes, Andreas Krause, and Pietro Perona, Caltech ... 775

Learning to localise sounds with spiking neural networks, Dan F Goodman, Ecole Normale Superieure, and Romain Brette, ENS ... 784

Feature Set Embedding for Incomplete Data, David Grangier, NEC Labs America, and Iain Melvin, NEC Laboratories America ... 793

Active Instance Sampling via Matrix Partition, Yuhong Guo, Temple University ... 802

Avoiding False Positive in Multi-Instance Learning, Yanjun Han, Institute of Automation CAS, Qing Tao, , and Jue Wang, ... 811

Nonparametric Density Estimation for Stochastic Optimization with an Observable State Variable, Lauren Hannah, Warren B Powell, and David Blei, Princeton University ... 820

Space-Variant Single-Image Blind Deconvolution for Removing Camera Shake, Stefan Harmeling, MPI for Biological Cybernetics, Hirsch Michael, Max Planck Institute, and Bernhard Schoelkopf, MPI for Biological Cybernetics ... 829

A Primal-Dual Message-Passing Algorithm for Approximated Large Scale Structured Prediction, Tamir Hazan, and Raquel Urtasun, TTI Chicago ... 838

An Inverse Power Method for Nonlinear Eigenproblems with Applications in 1-Spectral Clustering and Sparse PCA, Matthias Hein, and Thomas Bühler, Saarland University 847

Online Learning for Latent Dirichlet Allocation, Matthew Hoffman, David Blei, Princeton University, and Francis Bach, Ecole Normale Superieure 856

Latent Variable Models for Predicting File Dependencies in Large-Scale Software Development, Diane Hu, UC San Diego, Laurens van der Maaten, University of California San Diego, Youngmin Cho, UC San Diego, Lawrence Saul, University of California San Diego, and Sorin Lerner, UC San Diego 865

Exact inference and learning for cumulative distribution functions on loopy graphs, Jim C Huang, Nebojsa Jojic, and Chris Meek, Microsoft Research 874

Predicting Execution Time of Computer Programs Using Sparse Polynomial Regression, Ling Huang, Intel, Jinzhu Jia, Bin Yu, UC Berkeley, Byung-Gon Chun, Petros Maniatis, and Mayur Naik, Intel 883

Active Learning by Querying Informative and Representative Examples, Sheng-Jun Huang, Nanjing University, Rong Jin, Michigan State University, and Zhi-Hua Zhou, Nanjing University 892

Inter-time segment information sharing for non-homogeneous dynamic Bayesian networks, Dirk Husmeier, BioSS, Frank Dondelinger, Biomathematics & Statistics Scotland (Bioss), and Sophie Lebre, University of Strasbourg 901

Deciphering subsampled data: adaptive compressive sampling as a principle of brain communication, Guy Isely, Christopher J Hillar, UC Berkeley, and Fritz Sommer, Berkeley 910

Dynamic Infinite Relational Model for Time-varying Relational Data Analysis, Katsuhiko Ishiguro, Tomoharu Iwata, NTT, Naonori Ueda, NTT Communication Science Laboratories, and Joshua Tenenbaum, Massachusetts Institute of Technology 919

Hashing Hyperplane Queries to Near Points with Applications to Large-Scale Active Learning, Prateek Jain, Microsoft Research India Lab, Sudheendra Vijayanarasimhan, , and Kristen Grauman, University of Texas at Austin 928

Guaranteed Rank Minimization via Singular Value Projection, Prateek Jain, Microsoft Research India Lab, Raghu Meka, University of Texas Austin, and Inderjit Dhillon, University of Texas 937

Inductive Regularized Learning of Kernel Functions, Prateek Jain, Microsoft Research India Lab, Brian Kulis, University of California Berkeley, and Inderjit Dhillon, University of Texas 946

MAP estimation in Binary MRFs via Bipartite Multi-cuts, Sashank Jakkam Reddi, Sunita Sarawagi, and Sundar Vishwanathan, IIT Bombay 955

A Dirty Model for Multi-task Learning, Ali Jalali, Pradeep Ravikumar, Sujay Sanghavi, and Chao Ruan, University of Texas at Austin 964

Lifted Inference Seen from the Other Side : The Tractable Features, Abhay Jha, University of Washington, Vibhav G Gogate, University of Washington Seat, Alexandra Meliou, , and Dan Suciu, ... 973

Factorized Latent Spaces with Structured Sparsity, Yangqing Jia, UC Berkeley, Mathieu Salzmann, TTI Chicago, and Trevor Darrell, UC Berkeley 982

Bayesian Action-Graph Games, Albert Xin Jiang, and Kevin Leyton-Brown, University of British Columbia 991

On a Connection between Importance Sampling and the Likelihood Ratio Policy Gradient, Tang Jie, UC Berkeley, and Pieter Abbeel, University of California 1000

Linear Complementarity for Regularized Policy Evaluation and Improvement, Jeffrey T Johns, Christopher Painter-Wakefield, and Ronald Parr, Duke University 1009

Synergies in learning words and their referents, Mark Johnson, Katherine Demuth, Macquarie University, Michael Frank, MIT, and Bevan K Jones, University of Edinburgh 1018

Structural epitome: a way to summarize ones visual experience, Nebojsa Jojic, Microsoft Research, Alessandro Perina, Universtiy of Verona, and Vittorio Murino, University of Verona / Italian Institute of Technology 1027

Probabilistic Belief Revision with Structural Constraints, Peter Jones, MIT, Venkatesh Saligrama, Boston University, and Sanjoy K Mitter, MIT 1036

Efficient Optimization for Discriminative Latent Class Models, Armand Joulin, INRIA / ENS, Francis Bach, Ecole Normale Superieure, and Jean Ponce, ... 1045

Non-Stochastic Bandit Slate Problems, Satyen Kale, Yahoo!, Lev Reyzin, Georgia Institute of Technology, and Robert E Schapire, Princeton University 1054

Static Analysis of Binary Executables Using Structural SVMs, Nikos Karampatziakis, Cornell University 1063

Using body-anchored priors for identifying actions in single images, Leonid Karlinsky, Michael Dinerstein, and Shimon Ullman, Weizmann Institute of Science 1072

Effects of Synaptic Weight Diffusion on Learning in Decision Making Networks, Kentaro Katahira, JST ERATO, Kazuo Okanoya, JST ERATO Okanoya Emotional Information Project, and Masato Okada, The University of TOkyo / RIKEN 1081

Learning Convolutional Feature Hierarchies for Visual Recognition, Koray Kavukcuoglu, Pierre Sermanet, New York University, Y-Lan Boureau, INRIA/New York University, Karol Gregor, , Michael Mathieu, , and Yann Le Cun, New York U 1090

Accounting for network effects in neuronal responses using L1 regularized point process models, Ryan C Kelly, Carnegie Mellon University, Matthew A Smith, University of Pittsburgh, Robert E Kass, CMU, and Tai Sing Lee, ... 1099

Variational bounds for mixed-data factor analysis, Mohammad Emtiyaz Khan, University of British Columbia, Benjamin Marlin, UBC, Guillaume Bouchard, XRCE, and Kevin Murphy, UBC 1108

Sparse Coding for Learning Interpretable Spatio-Temporal Primitives, Taehwan Kim, TTI Chicago, Gregory Shakhnarovich, TTIC, and Raquel Urtasun, TTI Chicago 1117

Regularized estimation of image statistics by Score Matching, Diederik P Kingma, Universiteit Utrecht, and Yann Le Cun, New York U 1126

Random Conic Pursuit for Semidefinite Programming, Ariel Kleiner, UC Berkeley, Ali Rahimi, Intel, and Michael I Jordan, University of California Berkeley 1135

Generalized roof duality and bisubmodular functions, Vladimir Kolmogorov, UC London 1144

Energy Disaggregation via Discriminative Sparse Coding, J. Zico Kolter, MIT, Siddharth Batra, Stanford University, and Andrew Ng, Stanford 1153

Constructing Skill Trees for Reinforcement Learning Agents from Demonstration Trajectories, George D Konidaris, Scott R Kuindersma, UMass Amherst, Andrew Barto, University of Massachusetts Amherst, and Roderic A Grupen, UMass Amherst 1162

Structured Determinantal Point Processes, Alex Kulesza, University of Pennsylvania, and Ben Taskar, University of Pennsylvania 1171

MAP Estimation for Graphical Models by Likelihood Maximization, Akshat Kumar, UMass Amherst, and Shlomo Zilberstein, UMass 1180

Self-Paced Learning for Latent Variable Models, M. Pawan Kumar, Stanford University, Benjamin Packer, and Daphne Koller, Stanford 1189

Efficient algorithms for learning kernels from multiple similarity matrices with general convex loss functions, Achintya Kundu, Indian Institute of Science, vikram M Tankasali, Indian Institute of Science Bangalore, Chiranjib Bhattacharyya, I.I.Sc. Bangalore, and Aharon Ben-Tal, Technion- Israel Institute of Technology 1198

Evaluation of Rarity of Fingerprints in Forensics, Chang Su LEE, University at Buffalo, and Sargur N Srihari, University at Buffalo The State University of New York 1207

Beyond Actions: Discriminative Models for Contextual Group Activities, Tian Lan, Yang Wang, Weilong Yang, Simon Fraser University, and Greg Mori, Simon Fraser U 1216

Functional Geometry Alignment and Localization of Brain Areas, Georg Langs, MIT, Yanmei Tie, Laura Rigolo, Alexandra Golby, Harvard Medical School, and Polina Golland, MIT 1225

Efficient Relational Learning with Hidden Variable Detection, Ni Lao, Jun Zhu, Liu Xinwang, Yandong Liu, and William W Cohen, Carnegie Mellon University 1234

Learning to combine foveal glimpses with a third-order Boltzmann machine, Hugo Larochelle, and Geoffrey Hinton, University of Toronto 1243

Categories and Functional Units: An Infinite Hierarchical Model for Brain Activations, Danial Lashkari, Ramesh Sridharan, and Polina Golland, MIT 1252

Identifying Dendritic Processing, Aurel A. Lazar, and Yevgeniy B Slutskiy, Columbia University 1261

Cross Species Expression Analysis using a Dirichlet Process Mixture Model with Latent Matchings, Hai-Son P Le, and Ziv Bar-Joseph, Carnegie Mellon University 1270

Tiled convolutional neural networks, Quoc V Le, Jiquan Ngiam, Zhenghao Chen, Daniel Jin hao Chia, Pang Wei Koh, Stanford University, and Andrew Ng, Stanford 1279

Estimating Spatial Layout of Rooms using Volumetric Reasoning about Objects and Surfaces, David C Lee, Carnegie Mellon University, Abhinav Gupta, Martial Hebert, CMU, and Takeo Kanade, Carnegie Mellon University 1288

Practical Large-Scale Optimization for Max-norm Regularization, Jason Lee, Toyota Technological Institute at Chicago, Ben Recht, UW-Madison, Ruslan Salakhutdinov, MIT, Nathan Srebro, TTI, and Joel Tropp, CalTech 1297

Adaptive Multi-Task Lasso: with Application to eQTL Detection, Seunghak Lee, Jun Zhu, and Eric Xing, Carnegie Mellon University ... 1306

Joint Cascade Optimization Using A Product Of Boosted Classifiers, Leonidas Lefakis, IDIAP, and Francois Fleuret, Idiap Research Institute 1315

Learning To Count Objects in Images, Victor Lempitsky, U. Oxford, and Andrew Zisserman, University of Oxford 1324

Optimal Web-Scale Tiering as a Flow Problem, Gilbert Leung, eBay Inc., Novi Quadrianto, SML-NICTA and RSISE-ANU, Alexander J Smola, Yahoo! Research, and Kostas Tsioutsiouliklis, Yahoo! Labs ... 1333

Feature Construction for Inverse Reinforcement Learning, Sergey Levine, Stanford University, Zoran Popovic, University of Washington, and Vladlen Koltun, Stanford University 1342

Towards Holistic Scene Understanding: Feedback Enabled Cascaded Classification Models, Congcong Li, Adarsh P Kowdle, Ashutosh Saxena, and Tsuhan Chen, Cornell University 1351

Convex Multiple-Instance Learning by Estimating Likelihood Ratio, Fuxin Li, and Cristian Sminchisescu, University of Bonn 1360

Individualized ROI Optimization via Maximization of Group-wise Consistency of Structural and Functional Profiles, Kaiming Li, nwpu, Lei Guo, , Carlos Faraco, , Dajiang Zhu, , Fan Deng, University of Georgia, Tuo Zhang, , Xi Jiang, , Degang Zhang, , Hanbo Chen, Northwestern Polytechnical Unv, Xintao Hu, , Steve Miller, , and tianming Liu, ... 1369

Object Bank: A High-Level Image Representation for Scene Classification & Semantic Feature Sparsification, Li-Jia Li, Stanford University, Hao Su, Stanford, Eric Xing, Carnegie Mellon University, and Li Fei-Fei, Stanford University 1378

b-Bit Minwise Hashing for Estimating Three-Way Similarities, Ping J Li, Cornell, Arnd C Konig, Microsoft Research, and Wenhao Gui, 1387

Construction of Dependent Dirichlet Processes based on Poisson Processes, Dahua Lin, Eric Grimson, and John Fisher, MIT 1396

Deep Coding Network, Yuanqing Lin, NEC Labs, Zhang Tong, Rutgers University, Shenghuo Zhu, , and Kai Yu, NEC Laboratories America 1405

Robust Clustering as Ensembles of Affinity Relations, Hairong Liu, National University of Singapo, Longin Jan Latecki, Temple University, and Shuicheng Yan, National University of Singapore 1414

Graph-Valued Regression, Han Liu, Xi Chen, John Lafferty, and Larry Wasserman, Carnegie Mellon University 1423

Stability Approach to Regularization Selection (StARS) for High Dimensional Graphical Models, Han Liu, Kathryn Roeder, and Larry Wasserman, Carnegie Mellon University 1432

Multivariate Dyadic Regression Trees for Sparse Learning Problems, Han Liu, and Xi Chen, Carnegie Mellon University 1441

Multi-Stage Dantzig Selector, Ji Liu, Peter Wonka, and Jieping Ye, Arizona State University 1450

Moreau-Yosida Regularization for Grouped Tree Structure Learning, Jun Liu, and Jieping Ye, Arizona State University 1459

Decoding Ipsilateral Finger Movements from ECoG Signals in Humans, Yuzong Liu, Washington University, Mohit Sharma, Washington University St. Louis, Charles M Gaona, Washington University in St. Louis, Jonathan D Breshears, jarod Roland, Washington University School of Medicine, Zachary V Freudenburg, Kilian Q Weinberger, Washington University in St. Louis, and Eric C Leuthardt, Washington University School of Medicine 1468

Approximate Inference by Compilation to Arithmetic Circuits, Daniel Lowd, University of Oregon, and Pedro Domingos, University of Washington 1477

Block Variable Selection in Multivariate Regression and High-dimensional Causal Inference, Aurelie C Lozano, and Vikas Sindhwani, IBM Research 1486

Functional form of motion priors in human motion perception, Hongjing Lu, UCLA, Tungyou Lin, , Alan L Lee, , Luminita Vese, , and Alan L Yuille, UCLA 1495

Learning from Candidate Labeling Sets, Jie Luo, Idiap/EPF Lausanne, and Francesco Orabona, University of Milano 1504

Decomposing Isotonic Regression for Efficiently Solving Large Problems, Ronny Luss, Saharon Rosset, and Moni Shahar, Tel Aviv University 1513

Divisive Normalization: Justification and Effectiveness as Efficient Coding Transform, Siwei Lyu, University at Albany SUNY 1522

Permutation Complexity Bound on Out-Sample Error, Malik Magdon-Ismail, RPI 1531

Basis Construction from Power Series Expansions of Value Functions, Sridhar Mahadevan, and Bo Liu, University of Massachusetts Amherst 1540

Scrambled Objects for Least-Squares Regression, Odalric Maillard, INRIA Lille Team SequeL, and Remi Munos, INRIA Lille - Nord Europe 1549

Network Flow Algorithms for Structured Sparsity, Julien Mairal, Rodolphe Jenatton, Guillaume R Obozinski, INRIA, and Francis Bach, Ecole Normale Superieure 1558

Sphere Embedding: An Application to Part-of-Speech Induction, Yariv Maron, Bar Ilan University, Michael Lamar, Saint Louis University, and Elie Bienenstock, Brown University 1567

Variable margin losses for classifier design, Hamed Masnadi-Shirazi, UCSD, and Nuno Vasconcelos, UC San Diego 1576

Why are some word orders more common than others? A uniform information density account, Luke Maurits, Dan Navarro, and Amy Perfors, University of Adelaide 1585

Direct Loss Minimization for Structured Prediction, David A McAllester, TTI-Chicago, Tamir Hazan, and Joseph Keshet, TTI Chicago 1594

Gated Softmax Classification, Roland Memisevic, Christopher Zach, ETH Zurich, Geoffrey Hinton, University of Toronto, and Marc Pollefeys, ETH Zurich 1603

A Family of Penalty Functions for Structured Sparsity, Charles A Micchelli, City Univ. of Hong Kong, Jean Morales, UNiv. College London, and Massi Pontil, University College London 1612

PAC-Bayesian Model Selection for Reinforcement Learning, Mahdi Milani Fard, and Joelle Pineau, McGill University 1624

Subgraph Detection Using Eigenvector L1 Norms, Benjamin A Miller, Nadya T Bliss, MIT Lincoln Laboratory, and Patrick Wolfe, Harvard University 1633

A VLSI Implementation of the Adaptive Exponential Integrate-and-Fire Neuron Model, Sebastian C Millner, Kirchhoff-Institut for Physics, Andreas Grübl, , Karlheinz Meier, Kirchhoff-Institut für Physik, Johannes Schemmel, , and Marc-Olivier Schwartz, Kirchhoff-Institut für Physik 1642

Large-Scale Matrix Factorization with Missing Data under Additional Constraints, Kaushik Mitra, University of Maryland College Park, Sameer Sheorey, TTIC, and Rama Chellappa, University of Maryland College Park 1651

Natural Policy Gradient Methods with Parameter-based Exploration for Control Tasks, Atsushi Miyamae, Tokyo Institute

of Technology/JSPS Research Fellow, YUICHI NAGATA, ISAO ONO, and SHIGENOBU KOBAYASHI, Tokyo Institute of Technology 1660

An analysis on negative curvature induced by singularity in multi-layer neural-network learning, EIJI MIZUTANI, Taiwan Univ. Science and Tech., and STUART DREYFUS, UC Berkeley 1669

Layer-wise analysis of deep networks with Gaussian kernels, GRÉGOIRE MONTAVON, MIKIO L BRAUN, and KLAUS-ROBERT MÜLLER, TU Berlin 1678

Probabilistic latent variable models for distinguishing between cause and effect, JORIS MOOIJ, OLIVER STEGLE, Max Planck Institute for Biological Cybernetics, DOMINIK JANZING, MPI for Biological Cybernetics, KUN ZHANG, Max Planck Institute for Biological Cybernetics, and BERNHARD SCHOELKOPF, MPI for Biological Cybernetics 1687

Epitome driven 3-D Diffusion Tensor image segmentation: on extracting specific structures, KAMIYA MOTWANI, Department of Computer Sciences University of Wisconsin Madison, NAGESH ADLURU, CHRIS HINRICHS, Univ. of Wisconsin Madison, ANDREW L ALEXANDER, Waisman Laboratory for Brain Imaging and Behavior Univ. of Wisconsin-Madison, and VIKAS SINGH, Univ. of Wisconsin Madison 1696

Improving Human Judgments by Decontaminating Sequential Dependencies, MICHAEL MOZER, University of Colorado at Boulder, HAROLD PASHLER, , MATTHEW WILDER, ROBERT LINDSEY, University of Colorado, MATT JONES, University of Colorado at Boulder, and MICHAEL JONES, Indiana University 1705

A Theory of Multiclass Boosting, INDRANEEL MUKHERJEE, and ROBERT E SCHAPIRE, Princeton University 1714

A biologically plausible network for the computation of orientation dominance, KRITIKA MURALIDHARAN, SVCL, and NUNO VASCONCELOS, UC San Diego 1723

Slice sampling covariance hyperparameters of latent Gaussian models, IAIN MURRAY, University of Edinburgh, and RYAN ADAMS, University of Toronto 1732

On the Convexity of Latent Social Network Inference, SETH MYERS, and JURE LESKOVEC, Stanford University 1741

Infinite Relational Modeling of Functional Connectivity in Resting State fMRI, MORTEN MØRUP, DTU Informatics, KRISTOFFER H MADSEN, ANNE-MARIE DOGONOWSKI, HARTWIG R SIEBNER, Danish Research Centre for Magnetic Resonance, and LARS K HANSEN, DTU Informatics 1750

Minimum Average Cost Clustering, KIYOHITO NAGANO, University of Tokyo, YOSHINOBU KAWAHARA, Osaka University, and SATORU IWATA, Kyoto University 1759

Global Analytic Solution for Variational Bayesian Matrix Factorization, SHINICHI NAKAJIMA, Nikon Corporation, MASASHI SUGIYAMA, Tokyo Institute of Technology, and RYOTA TOMIOKA, University of Tokyo 1768

Random Walk Approach to Regret Minimization, HARIHARAN NARAYANAN, MIT, and ALEXANDER RAKHLIN, University of Pennsylvania ... 1777

Sample Complexity of Testing the Manifold Hypothesis, Hariharan Narayanan, and Sanjoy K Mitter, MIT 1786

Learning the context of a category, Dan Navarro, University of Adelaide 1795

Online Markov Decision Processes under Bandit Feedback, Gergely Neu, Budapest U. of Tech. and Econ., Andras Gyorgy, MTA SZTAKI Institute for Computer Science and Control, Csaba Szepesvari, University of Alberta, and Andras Antos, MTA SZTAKI Institute for Computer Science and Control 1804

Efficient and Robust Feature Selection via Joint $l_{2,1}$-Norms Minimization, Feiping Nie, University of Texas Arlington, Heng Huang, Computer Science and Engineering University of Texas Arlington, Xiao Cai, UTA, and Chris Ding, University of Texas at Arlington 1813

Generative Local Metric Learning for Nearest Neighbor Classification, Yung-Kyun Noh, University of Pennsylvania, Byoung-Tak Zhang, Seoul National University, and Daniel Lee, University of Pennsylvania 1822

Approximate inference in continuous time Gaussian-Jump processes, Manfred Opper, Technische Universitaet Berlin, Andreas Ruttor, TU Berlin, and Guido Sanguinetti, University of Edinburgh 1831

New Adaptive Algorithms for Online Classification, Francesco Orabona, University of Milano, and Koby Crammer, Technion 1840

Estimation of Renyi Entropy and Mutual Information Based on Generalized Nearest-Neighbor Graphs, David Pal, University of Alberta, Barnabas Poczos, Carnegie Mellon University, and Csaba Szepesvari, University of Alberta 1849

Gaussian sampling by local perturbations, George Papandreou, and Alan L Yuille, UCLA 1858

Large Margin Multi-Task Metric Learning, Shibin Parameswaran, UCSD, and Kilian Q Weinberger, Washington University in St. Louis 1867

Multiparty Differential Privacy via Aggregation of Locally Trained Classifiers, Manas A Pathak, CMU, Shantanu Rane, Mitsubishi Electric Research Labs, and Bhiksha Raj, CMU 1876

$(RF)^2$ – Random Forest Random Field, Nadia C Payet, and Sinisa Todorovic, Oregon State University 1885

On the Theory of Learnining with Privileged Information, Dmitry Pechyony, NEC Labs, and Vladimir Vapnik, ... 1894

Empirical Bernstein Inequalities for U-Statistics, Thomas Peel, LIF Aix-Marseille Université, Sandrine Anthoine, Université Aix-Marseille I / CNRS, and Liva Ralaivola, LIF Aix-Marseille Université 1903

Reverse Multi-Label Learning, James Petterson, NICTA, and Tiberio Caetano, NICTA and ANU 1912

Word Features for Latent Dirichlet Allocation, James Petterson, NICTA, Alexander J Smola, Yahoo! Research, Tiberio Caetano, NICTA and ANU, Wray Buntine, NICTA, and Shravan M Narayanamurthy, Yahoo! Labs Bangalore 1921

Probabilistic Deterministic Infinite Automata, David Pfau, Nicholas Bartlett, and Frank Wood, Columbia University 1930

The Maximal Causes of Natural Scenes are Edge Filters, Jose G Puertas, FIAS, Jorg Bornschein, FIAS Goethe-University Frankurt, and Jorg Lucke, University of Frankfurt 1939

A New Probabilistic Model for Rank Aggregation, Tao Qin, , Xiubo Geng, Chinese Academy of Sciences, and Tie-Yan Liu, Microsoft Research 1948

Multitask Learning without Label Correspondences, Novi Quadrianto, SML-NICTA and RSISE-ANU, Alexander J Smola, Yahoo! Research, Tiberio Caetano, NICTA and ANU, S.V.N. Vishwanathan, Purdue University, and James Petterson, NICTA 1957

Link Discovery using Graph Feature Tracking, Emile RICHARD, ENS Cachan, Nicolas Baskiotis, CMLA/ENS Cachan, Theodoros Evgeniou, INSEAD, and Nicolas Vayatis, CMLA/ENS Cachan 1966

Inferring Stimulus Selectivity from the Spatial Structure of Neural Network Dynamics, Kanaka Rajan, Princeton University, L F Abbott, Columbia University, and Haim Sompolinsky, Hebrew University and Harvard University 1975

Online Learning: Random Averages, Combinatorial Parameters, and Learnability, Alexander Rakhlin, University of Pennsylvania, Karthik Sridharan, Toyota Technological Institute at Chicago, and Ambuj Tewari, UT Austin 1984

Evaluating neuronal codes for inference using Fisher information, Haefner M Ralf, Max Planck Institute for Biological Cybernetics, and Matthias Bethge, ... 1993

Generating more realistic images using gated MRF's, Marc'Aurelio Ranzato, Volodymyr Mnih, and Geoffrey Hinton, University of Toronto 2002

An Approximate Inference Approach to Temporal Optimization in Optimal Control, Konrad C Rawlik, University of Edinburgh, Marc Toussaint, TU Berlin, and Sethu Vijayakumar, University of Edinburgh . 2011

Hallucinations in Charles Bonnet Syndrome Induced by Homeostasis: a Deep Boltzmann Machine Model, David P Reichert, Peggy Series, and Amos J Storkey, University of Edinburgh 2020

An Alternative to Low-level-Sychrony-Based Methods for Speech Detection, Paul L Ruvolo, UC San Diego, and Javier R Movellan, Machine Perception Laboratory 2029

Tight Sample Complexity of Large-Margin Learning, Sivan Sabato, The Hebrew University, Nathan Srebro, TTI, and Naftali Tishby, The Hebrew University Jerusalem 2038

Boosting Classifier Cascades, Mohammad Saberian, and Nuno Vasconcelos, UC San Diego 2047

Collaborative Filtering in a Non-Uniform World: Learning with the Weighted Trace Norm, Ruslan Salakhutdinov, MIT, and Nathan Srebro, TTI 2056

Implicitly Constrained Gaussian Process Regression for Monocular Non-Rigid Pose Estimation, Mathieu Salzmann, and Raquel Urtasun, TTI Chicago 2065

Deterministic Single-Pass Algorithm for LDA, Issei Sato, Tokyo University, Kenichi Kurihara, Google, and Hiroshi Nakagawa, ... 2074

Active Estimation of F-Measures, Christoph Sawade, Niels Landwehr, University of Potsdam, and Tobias Scheffer, Uni Potsdam 2083

Trading off Mistakes and Don't-Know Predictions, Amin Sayedi, Carnegie Mellon University, Morteza Zadimoghaddam, MIT, and Avrim Blum, Carnegie Mellon University 2092

Sparse Inverse Covariance Selection via Alternating Linearization Methods, Katya Scheinberg, Shiqian Ma, Columbia University, and Donald Goldfarb, ... 2101

Spike timing-dependent plasticity as dynamic filter, Joscha T Schmiedt, University of Bremen, Christian Albers, Institute for Theoretical Physiks University of Bremen, and Klaus Pawelzik, ... 2110

A novel family of non-parametric cumulative based divergences for point processes, Sohan Seth, Park Il, Austin Brockmeier, University of Florida, Mulugeta Semework, SUNY Downstate Medical Center, John Choi, Joseph T Francis, SUNY Downstate and NYU-Poly, and Jose Principe, University of Florida at Gainesville 2119

Online Learning in The Manifold of Low-Rank Matrices, Uri Shalit, Hebrew University of Jerusalem, Daphna Weinshall, , and Gal Chechik, Google 2128

Identifying graph-structured activation patterns in networks, James L Sharpnack, CMU Machine Learning, and Aarti Singh, CMU 2137

A rational decision making framework for inhibitory control, Pradeep Shenoy, UCSD, Rajesh Rao, U. Washington, and Angela J Yu, UCSD 2146

Penalized Principal Component Regression on Graphs for Analysis of Subnetworks, Ali Shojaie, and George Michailidis, University of Michigan 2155

Monte-Carlo Planning in Large POMDPs, David Silver, MIT, and Joel Veness, University of New South Wales 2164

Sodium entry efficiency during action potentials: A novel single-parameter family of Hodgkin-Huxley models, Anand Singh, Renaud Jolivet, University of Zurich, Pierre J Magistretti, EPFL, and Bruno Weber, University of Zurich 2173

More data means less inference: A pseudo-max approach to structured learning, David Sontag, MIT, Ofer Meshi, Hebrew University, Tommi Jaakkola, MIT, and Amir Globerson, Hebrew University 2181

Reward Design via Online Gradient Ascent, Jonathan D Sorg, Satinder Singh, and Richard L Lewis, University of Michigan 2190

Smoothness, Low Noise and Fast Rates, Nathan Srebro, TTI, Karthik Sridharan, Toyota Technological Institute at Chicago, and Ambuj Tewari, UT Austin 2199

Efficient Minimization of Decomposable Submodular Functions, Peter G Stobbe, and Andreas Krause, Caltech 2208

Learning from Logged Implicit Exploration Data, Alex Strehl, Facebook, John Langford, Yahoo Research, Lihong Li, Yahoo! Research, and Sham Kakade, University of Pennsylvania 2217

Layered image motion with explicit occlusions, temporal consistency, and depth ordering, Deqing Sun, Erik Sudderth, and Michael Black, Brown University 2226

Improving the Asymptotic Performance of Markov Chain Monte-Carlo by Inserting Vortices, Yi Sun, Faustino Gomez, and Juergen Schmidhuber, IDSIA 2235

Semi-Supervised Learning with Adversarially Missing Label Information, Umar Syed, and Ben Taskar, University of Pennsylvania . 2244

A Reduction from Apprenticeship Learning to Classification, Umar Syed, University of Pennsylvania, and Robert E Schapire, Princeton University 2253

Identifying Patients at Risk of Major Adverse Cardiovascular Events Using Symbolic Mismatch, Zeeshan Syed, University of Michigan, and John Guttag, MIT 2262

Switching state space model for simultaneously estimating state transitions and nonstationary firing rates, Ken Takiyama, The University of Tokyo, and Masato Okada, The University of TOkyo / RIKEN 2271

Pose-Sensitive Embedding by Nonlinear NCA Regression, Graham Taylor, NYU, Rob Fergus, George Williams, Ian Spiro, and Christoph Bregler, New York University 2280

Fast Large-scale Mixture Modeling with Component-specific Data Partitions, Bo Thiesson, Microsoft Research, and Chong Wang, Princeton University 2289

Policy gradients in linearly-solvable MDPs, Emanuel Todorov, University of Washington 2298

Phoneme Recognition with Large Hierarchical Reservoirs, Fabian Triefenbach, Ghent University ELIS, Azarakhsh Jalalvand, Benjamin Schrauwen, and Jean-Pierre Martens, UGent 2307

Exact learning curves for Gaussian process regression on large random graphs, Matthew J Urry, King's College London, and Peter Sollich, Kings College London 2316

Worst-case bounds on the quality of max-product fixed-points, MERITXELL VINYALS, Jesús Cerquides, IIIA-CSIC, Alessandro Farinelli, University of Verona, and Juan Antonio Rodríguez-Aguilar, IIIA-CSIC 2325

Brain covariance selection: better individual functional connectivity models using population prior, Gael Varoquaux, Alexandre Gramfort, INRIA, Jean-Baptiste Poline, CEA, and Bertrand Thirion, INRIA 2334

Fast detection of multiple change-points shared by many signals using group LARS, Jean-Philippe Vert, Mines ParisTech, and Kevin Bleakley, Institut Curie 2343

Optimal Bayesian Recommendation Sets and Myopically Optimal Choice Query Sets, Paolo Viappiani, and Craig Boutilier, University of Toronto 2352

Multiple Kernel Learning and the SMO Algorithm, S.V.N. Vishwanathan, Zhaonan sun, Nawanol T Ampornpunt, Purdue University, and Manik Varma, Microsoft Research 2361

Joint Analysis of Time-Evolving Binary Matrices and Associated Documents, Eric X Wang, Dehong Liu, Jorge G Silva, Duke University, David Dunson, Duke University, and Lawrence Carin, Duke ECE 2370

Unsupervised Kernel Dimension Reduction, Meihong Wang, University of Southern California, Fei Sha, University of Southern Califor, and Michael I Jordan, University of California Berkeley 2379

Multi-View Active Learning in the Non-Realizable Case, Wei Wang, and Zhi-Hua Zhou, Nanjing University 2388

A Discriminative Latent Model of Image Region and Object Tag Correspondence, Yang Wang, Simon Fraser University, and Greg Mori, Simon Fraser U 2397

Heavy-Tailed Process Priors for Selective Shrinkage, Fabian L Wauthier, and Michael I Jordan, University of California Berkeley 2406

Sidestepping Intractable Inference with Structured Ensemble Cascades, David Weiss, University of Pennsylvania, Benjamin J Sapp, and Ben Taskar, University of Pennsylvania 2415

The Multidimensional Wisdom of Crowds, Peter Welinder, Caltech, Steve Branson, University of California San Diego, Serge Belongie, UC San Diego, and Pietro Perona, Caltech 2424

Interval Estimation for Reinforcement-Learning Algorithms in Continuous-State Domains, Martha White, and Adam M White, University of Alberta 2433

Active Learning Applied to Patient-Adaptive Heartbeat Classification, Jenna Wiens, and John Guttag, MIT 2442

Probabilistic Inference and Differential Privacy, Oliver Williams, and Frank McSherry, Microsoft Research 2451

Copula Processes, Andrew G Wilson, University of Cambridge, and Zoubin Ghahramani, Cambridge 2460

Linear readout from a neural population with partial correlation data, Adrien Wohrer, Ecole Normale Superieure, Ranulfo Romo, Universidad Autonoma de Mexico, and Christian K Machens, Ecole Normale Superieure 2469

A unified model of short-range and long-range motion perception, Shuang Wu, Xuming He, Hongjing Lu, and Alan L Yuille, UCLA . 2478

A Log-Domain Implementation of the Diffusion Network in Very Large Scale Integration, Yi-Da Wu, Shi-Jie Lin, and Hsin Chen, National Tsing Hua University 2487

Robust PCA via Outlier Pursuit, Huan Xu, Constantine Caramanis, The University of Texas, and Sujay Sanghavi, University of Texas at Austin 2496

Distributionally Robust Markov Decision Processes, Huan Xu, The University of Texas, and Shie Mannor, Technion 2505

Inference and communication in the game of Password, Yang Xu, CMU, and Charles Kemp, Carnegie Mellon University 2514

Feature Transitions with Saccadic Search: Size, Color, and Orientation Are Not Alike, Stella X Yu, Boston College 2523

Relaxed Clipping: A Global Training Method for Robust Regression and Classification, Yaoliang Yu, Min Yang, University of Alberta, Linli Xu, University of Science and Technology of China, Martha White, and Dale Schuurmans, University of Alberta 2532

Lower Bounds on Rate of Convergence of Cutting Plane Methods, Xinhua Zhang, AICML, Ankan Saha, University of Chicago, and S.V.N. Vishwanathan, Purdue University 2541

Learning Multiple Tasks with a Sparse Matrix-Normal Penalty, Yi Zhang, and Jeff Schneider, Carnegie Mellon University 2550

Probabilistic Multi-Task Feature Selection, Yu Zhang, Dit-Yan Yeung, HKUST, and Qian Xu, Hong Kong University of Science and Technology 2559

Worst-Case Linear Discriminant Analysis, Yu Zhang, and Dit-Yan Yeung, HKUST 2568

Sufficient Conditions for Generating Group Level Sparsity in a Robust Minimax Framework, Hongbo Zhou, Southern Illinois University, and Qiang Cheng, Southern Illinois Univeristy 2577

Large Margin Learning of Upstream Scene Understanding Models, Jun Zhu, Carnegie Mellon University, Li-Jia Li, Li Fei-Fei, Stanford University, and Eric Xing, Carnegie Mellon University 2586

Parallelized Stochastic Gradient Descent, Martin Zinkevich, Yahoo!, Markus Weimer, , Alex Smola, , and Lihong Li, Yahoo! Research 2595

A Primal-Dual Algorithm for Group Sparse Regularization with Overlapping Groups, sofia mosci, disi, Silvia Villa, DIMA Universita' di Genova, Alessandro Verri, DISI Universita' di Genova, and Lorenzo Rosasco, MIT and IIT 2604

Double Q-learning, Hado P van Hasselt, Centrum Wiskunde & Informatica (Center for Mathematics and Computer Science) 2613

Getting lost in space: Large sample analysis of the resistance distance, Ulrike von Luxburg, Max Planck for Biological Cybernetics, Agnes Radl, Max Planck Institute for Biological Cybernetics, and Matthias Hein, Saarland University 2622

Subject Index 2631

Author Index