

Neural Information Processing Systems

# Advances in Neural Information Processing Systems 20

21st Annual Conference on Neural Information  
Processing Systems 2007

December 3-6, 2007  
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](http://www.proceedings.com)

ISBN: 978-1-60560-352-0

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

Copyright (2009) by Neural Information Processing Systems.

All rights reserved.

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](mailto:info@nips.cc)

# TABLE OF CONTENTS

## VOLUME 1

|  |     |
|--|-----|
| <b>Classification Via Minimum Incremental Coding Length (MICL)</b> .....   | 1   |
| <i>John Wright, Yangyu Tao, Zhouchen Lin, Yi Ma, Heung-Yeung Shum</i>  |     |
| <b>Optimal Models of Sound Localization by Barn Owls</b> .....   | 9   |
| <i>Brian Fischer</i>   |     |
| <b>A Unified Near-Optimal Estimator for Dimension Reduction in <math>l_\alpha</math> (<math>0 &lt; \alpha \leq 2</math>) Using Stable Random Projections</b> ..... | 17  |
| <i>Ping Li, Trevor Hastie</i>  |     |
| <b>Support Vector Machine Classification with Indefinite Kernels</b> .....   | 25  |
| <i>Ronny Luss, Alexandre DAspremont</i>  |     |
| <b>Nearest-Neighbor-Based Active Learning for Rare Category Detection</b> .....  | 33  |
| <i>Jingrui He, Jaime Carbonell</i>   |     |
| <b>Learning and Using Relational Theories</b> .....  | 41  |
| <i>Charles Kemp, Noah Goodman, Joshua Tenenbaum</i>  |     |
| <b>Topmoumoute Online Natural Gradient Algorithm</b> .....   | 49  |
| <i>Nicolas Le Roux, Pierre-Antoine Manzagol, Yoshua Bengio</i>   |     |
| <b>FilterBoost: Regression and Classification on Large Datasets</b> .....  | 57  |
| <i>Joseph K Bradley, Robert Schapire</i>   |     |
| <b>Simulated Annealing: Rigorous Finite-time Guarantees for Optimization on Continuous Domains</b> .....   | 65  |
| <i>Andrea Lecchini-Visintini, John Lygeros, Jan Maciejowski</i>  |     |
| <b>Variational Inference for Markov Jump Processes</b> .....   | 73  |
| <i>Manfred Opper, Guido Sanguinetti</i>  |     |
| <b>Cooled and Relaxed Survey Propagation for MRFs</b> .....  | 81  |
| <i>Hai Leong Chieu, Wee Sun Lee, Yee Whye Teh</i>  |     |
| <b>Cluster Stability for Finite Samples</b> .....  | 89  |
| <i>Ohad Shamir, Naftali Tishby</i>   |     |
| <b>Infinite State Bayesian Networks</b> .....  | 97  |
| <i>Max Welling, Ian Porteous, Evgeniy Bart</i>   |     |
| <b>Spatial Latent Dirichlet Allocation</b> .....   | 105 |
| <i>Xiaogang Wang, Eric Grimson</i>   |     |
| <b>EEG-Based Brain-Computer Interaction: Improved Accuracy by Automatic Single-Trial Error Detection</b> .....   | 113 |
| <i>Pierre Ferrez, José del R. Millán</i>   |     |
| <b>People Tracking with the Laplacian Eigenmaps Latent Variable Model</b> .....  | 121 |
| <i>Zhengdong Lu, Miguel Carreira-Perpinan, Cristian Sminchisescu</i>   |     |
| <b>Kernels on Attributed Pointsets with Applications</b> .....   | 129 |
| <i>Mehul Parsana, Sourangshu Bhattacharya, Chiru Bhattacharya, K. R. Ramakrishnan</i>  |     |
| <b>Learning the Structure of Manifolds Using Random Projections</b> .....  | 137 |
| <i>Yoav Freund, Sanjoy Dasgupta, Mayank Kabra, Nakul Verma</i>   |     |

|  |     |
|--|-----|
| <b>Ensemble Clustering Using Semidefinite Programming</b> .....  | 145 |
| <i>Vikas Singh, Lopamudra Mukherjee, Jiming Peng, Jinhui Xu</i>  |     |
| <b>Bayesian Binning Beats Approximate Alternatives: Estimating Peri-Stimulus Time Histograms</b> .....           | 153 |
| <i>Dominik Endres, Mike Oram, Johannes Schindelin, Peter Foldiak</i>   |     |
| <b>Regularized Boost for Semi-Supervised Learning</b> .....  | 161 |
| <i>Ke Chen, Shihai Wang</i>  |     |
| <b>Inferring Elapsed Time from Stochastic Neural Processes</b> .....   | 169 |
| <i>Misha Ahrens, Maneesh Sahani</i>  |     |
| <b>An Online Hebbian Learning Rule That Performs Independent Component Analysis</b> .....                        | 177 |
| <i>Claudia Clopath, André Longtin, Wulfram Gerstner</i>  |     |
| <b>A General Agnostic Active Learning Algorithm</b> .....  | 185 |
| <i>Sanjoy Dasgupta, Daniel Hsu, Claire Monteleoni</i>  |     |
| <b>A Spectral Regularization Framework for Multi-Task Structure Learning</b> .....                               | 193 |
| <i>Andreas Argyriou, Charles A. Micchelli, Massimiliano Pontil, Yiming Ying</i>                                  |     |
| <b>HM-BiTAM: Bilingual Topic Exploration, Word Alignment, and Translation</b> .....                              | 201 |
| <i>Bing Zhao, Eric P. Xing</i>   |     |
| <b>Compressed Regression</b> .....   | 209 |
| <i>Shuheng Zhou, John Lafferty, Larry Wasserman</i>  |     |
| <b>Stability Bounds for Non-i.i.d. Processes</b> .....   | 217 |
| <i>Mehryar Mohri, Afshin Rostamizadeh</i>  |     |
| <b>Selecting Observations Against Adversarial Objectives</b> .....   | 225 |
| <i>Andreas Krause, Brendan McMahan, Carlos Guestrin, Anupam Gupta</i>  |     |
| <b>Learning with Transformation Invariant Kernels</b> .....  | 233 |
| <i>Christian Walder, Olivier Chapelle</i>  |     |
| <b>Anytime Induction of Cost-sensitive Trees</b> .....   | 241 |
| <i>Saher Esmeir, Shaul Markovitch</i>  |     |
| <b>Direct Importance Estimation with Model Selection and Its Application to Covariate Shift Adaptation</b> ..... | 249 |
| <i>Masashi Sugiyama, Shinichi Nakajima, Hisashi Kashima, Paul Von Buenau, Motoaki Kawanabe</i>                   |     |
| <b>Receptive Fields Without Spike-Triggering</b> .....   | 257 |
| <i>Jakob Macke, Guenther Zeck, Matthias Bethge</i>   |     |
| <b>On Higher-order Perceptron Algorithms</b> .....   | 265 |
| <i>Claudio Gentile, Fabio Vitale, Cristian Broto</i>   |     |
| <b>Scene Segmentation with CRFs Learned from Partially Labeled Images</b> .....                                  | 273 |
| <i>Jakob Verbeek, Bill Triggs</i>  |     |
| <b>What Makes Some POMDP Problems Easy to Approximate?</b> .....   | 281 |
| <i>David Hsu, Wee Sun Lee, Nan Rong</i>  |     |
| <b>Contraction Properties of VLSI Cooperative Competitive Neural Networks of Spiking Neurons</b> .....           | 289 |
| <i>Emre Neftci, Elisabetta Chicca, Giacomo Indiveri, Jean-Jacques Slotine, Rodney Douglas</i>                    |     |
| <b>An Analysis of Convex Relaxations for MAP Estimation</b> .....  | 297 |
| <i>Pawan Mudigonda, Vladimir Kolmogorov, Philip Torr</i>   |     |

|  |     |
|--|-----|
| <b>Predicting Brain States from fMRI Data: Incremental Functional Principal Component Regression</b> ..... | 305 |
| <i>Sennay Ghebreab, Arnold Smeulders, Pieter Adriaans</i>  |     |
| <b>Bayesian Inference for Spiking Neuron Models with a Sparsity Prior</b> .....                            | 313 |
| <i>Sebastian Gerwinn, Jakob Macke, Matthias Seeger, Matthias Bethge</i>                                    |     |
| <b>Markov Chain Monte Carlo with People</b> .....  | 321 |
| <i>Adam Sanborn, Thomas Griffiths</i>  |     |
| <b>Statistical Analysis of Semi-Supervised Regression</b> .....  | 329 |
| <i>John Lafferty, Larry Wasserman</i>  |     |
| <b>Incremental Natural Actor-Critic Algorithms</b> .....   | 337 |
| <i>Shalabh Bhatnagar, Richard Sutton, Mohammad Ghavamzadeh, Mark Lee</i>                                   |     |
| <b>Near-Maximum Entropy Models for Binary Neural Representations of Natural Images</b> .....               | 345 |
| <i>Matthias Bethge, Philipp Berens</i>   |     |
| <b>Convex Clustering with Exemplar-Based Models</b> .....  | 353 |
| <i>Danial Lashkari, Polina Golland</i>   |     |
| <b>Transfer Learning Using Kolmogorov Complexity: Basic Theory and Empirical Evaluations</b> .....         | 361 |
| <i>M. M. Mahmud, Sylvian Ray</i>   |     |
| <b>Random Sampling of States in Dynamic Programming</b> .....  | 369 |
| <i>Chris Atkeson, Benjamin Stephens</i>  |     |
| <b>The Infinite Gamma-Poisson Feature Model</b> .....  | 377 |
| <i>Michalis Titsias</i>  |     |
| <b>Comparison of Objective Functions for Estimating Linear-Nonlinear Models</b> .....                      | 385 |
| <i>Tatyana Sharpee</i>   |     |
| <b>Retrieved Context and the Discovery of Semantic Structure</b> .....                                     | 393 |
| <i>Vinayak Rao, Marc Howard</i>  |     |
| <b>A Constraint Generation Approach to Learning Stable Linear Dynamical Systems</b> .....                  | 401 |
| <i>Sajid Siddiqi, Byron Boots, Geoffrey Gordon</i>   |     |
| <b>Boosting the Area Under the ROC Curve</b> .....   | 409 |
| <i>Phil Long, Rocco Servedio</i>   |     |
| <b>Scan Strategies for Adaptive Meteorological Radars</b> .....  | 417 |
| <i>Victoria Manfredi, Jim Kurose</i>   |     |
| <b>Experience-Guided Search: A Theory of Attentional Control</b> .....                                     | 425 |
| <i>Michael Mozer, David Baldwin</i>  |     |
| <b>Theoretical Analysis of Heuristic Search Methods for Online POMDPs</b> .....                            | 433 |
| <i>Stephane Ross, Joelle Pineau, Brahim Chaib-draa</i>   |     |
| <b>Inferring Neural Firing Rates from Spike Trains Using Gaussian Processes</b> .....                      | 441 |
| <i>John Cunningham, Byron Yu, Krishna Shenoy, Maneesh Sahani</i>   |     |
| <b>Comparing Bayesian Models for Multisensory Cue Combination Without Mandatory Integration</b> .....      | 449 |
| <i>Ulrik Beierholm, Konrad Kording, Ladan Shams, Wei Ji Ma</i>   |     |
| <b>Learning Horizontal Connections in a Sparse Coding Model of Natural Images</b> .....                    | 457 |
| <i>Pierre Garrigues, Bruno Olshausen</i>   |     |

|  |     |
|--|-----|
| <b>Rapid Inference on a Novel AND/OR Graph for Object Detection, Segmentation and Parsing.....</b>                         | 465 |
| <i>Yuanhao Chen, Long Zhu, Chenxi Lin, Alan Yuille, Hongjiang Zhang</i>  |     |
| <b>SpAM: Sparse Additive Models.....</b>   | 473 |
| <i>Pradeep Ravikumar, Han Liu, John Lafferty, Larry Wasserman</i>  |     |
| <b>Simplified Rules and Theoretical Analysis for Information Bottleneck Optimization and PCA with Spiking Neurons.....</b> | 481 |
| <i>Lars Buesing, Wolfgang Maass</i>  |     |
| <b>Consistent Minimization of Clustering Objective Functions.....</b>  | 489 |
| <i>Ulrike Von Luxburg, Sebastien Bubeck, Stefanie Jegelka, Michael Kaufmann</i>  |     |
| <b>Optimal ROC Curve for a Combination of Classifiers.....</b>   | 497 |
| <i>Marco Barreno, Alvaro Cardenas, J. D. Tygar</i>   |     |
| <b>Multi-task Gaussian Process Prediction.....</b>   | 505 |
| <i>Edwin Bonilla, Kian Ming Chai, Chris Williams</i>   |     |
| <b>PSVM: Parallelizing Support Vector Machines on Distributed Computers .....</b>  | 513 |
| <i>Edward Chang, Kaihua Zhu, Hao Wang, Hongjie Bai, Jian Li, Zhihuan Qiu, Hang Cui</i>                                     |     |
| <b>Heterogeneous Component Analysis .....</b>  | 521 |
| <i>Shigeyuki Oba, Motoaki Kawanabe, Klaus-Robert Müller, Shin Ishii</i>  |     |
| <b>A Probabilistic Model for Generating Realistic Lip Movements from Speech.....</b>                                       | 529 |
| <i>Gwenn Englebienne, Tim Cootes, Magnus Rattray</i>   |     |
| <b>Multi-Task Learning Via Conic Programming.....</b>  | 537 |
| <i>Tsuyoshi Kato, Hisashi Kashima, Masashi Sugiyama, Kiyoshi Asai</i>  |     |
| <b>A Neural Network Implementing Optimal State Estimation Based on Dynamic Spike Train Decoding.....</b>                   | 545 |
| <i>Omer Bobrowski , Ron Meir, Shy Shoham, Yonina Eldar</i>   |     |
| <b>Bundle Methods for Machine Learning .....</b>   | 553 |
| <i>Alex Smola, S V N Vishwanathan, Quoc Le</i>   |     |
| <b>GRIFT: a Graphical Model for Inferring Visual Classification Features from Human Data .....</b>                         | 561 |
| <i>Michael Ross, Andrew Cohen</i>  |     |
| <b>Progressive Mixture Rules Are Deviation Suboptimal .....</b>  | 569 |
| <i>Jean-Yves Audibert</i>  |     |
| <b>Colored Maximum Variance Unfolding.....</b>   | 577 |
| <i>Le Song, Alex Smola, Karsten Borgwardt, Arthur Gretton</i>  |     |

## VOLUME 2

|  |     |
|--|-----|
| <b>How SVMs Can Estimate Quantiles and the Median .....</b>                        | 585 |
| <i>Andreas CHRISTMANN, Ingo Steinwart</i>  |     |
| <b>Bayesian Co-Training .....</b>  | 593 |
| <i>Shipeng Yu, Balaji Krishnapuram, Romer Rosales, Harald Steck, R. Bharat Rao</i> |     |
| <b>Temporal Difference Updating Without a Learning Rate.....</b>                   | 601 |
| <i>Marcus Hutter, Shane Legg</i>   |     |
| <b>A Randomized Algorithm for Large Scale Support Vector Learning.....</b>         | 609 |
| <i>Krishnan Kumar, Chiru Bhattacharya, Ramesh Hariharan</i>                        |     |
| <b>Collective Inference on Markov Models for Modeling Bird Migration .....</b>     | 617 |
| <i>Daniel Sheldon, M.A. Saleh Elmohamed, Dexter Kozen</i>                          |     |

|  |            |
|--|------------|
| <b>Efficient Convex Relaxation for Transductive Support Vector Machine .....</b>                                 | <b>625</b> |
| <i>Zenglin Xu, Rong Jin, Jianke Zhu, Irwin King, Michael Lyu</i>   |            |
| <b>On Ranking in Survival Analysis: Bounds on the Concordance Index .....</b>                                    | <b>633</b> |
| <i>Vikas Raykar, Harald Steck, Balaji Krishnapuram, Cary Dehing-Oberije, Philippe Lambin</i>                     |            |
| <b>Adaptive Embedded Subgraph Algorithms Using Walk-Sum Analysis.....</b>  | <b>641</b> |
| <i>Venkat Chandrasekaran, Jason Johnson, Alan Willsky</i>  |            |
| <b>Learning Visual Attributes.....</b>   | <b>649</b> |
| <i>Vittorio Ferrari, Andrew Zisserman</i>  |            |
| <b>The Generalized FITC Approximation.....</b>   | <b>657</b> |
| <i>Andrew Naish-Guzman, Sean Holden</i>  |            |
| <b>Continuous Time Particle Filtering for fMRI.....</b>  | <b>665</b> |
| <i>Lawrence Murray, Amos Storkey</i>   |            |
| <b>Kernel Measures of Conditional Dependence.....</b>  | <b>673</b> |
| <i>Kenji Fukumizu, Arthur Gretton, Xiaohai Sun, Bernhard Schölkopf</i>   |            |
| <b>One-Pass Boosting .....</b>   | <b>681</b> |
| <i>Zafer Barutcuoglu, Phil Long, Rocco Servedio</i>  |            |
| <b>Measuring Neural Synchrony by Message Passing .....</b>   | <b>689</b> |
| <i>Justin Dauwels, François Vialatte, Tomasz Rutkowski, Andrzej Cichocki</i>                                     |            |
| <b>Multiple-Instance Pruning for Learning Efficient Cascade Detectors .....</b>                                  | <b>697</b> |
| <i>Cha Zhang, Paul Viola</i>   |            |
| <b>Convex Relaxations of Latent Variable Training.....</b>   | <b>705</b> |
| <i>Yuhong Guo, Dale Schuurmans</i>   |            |
| <b>Stable Dual Dynamic Programming.....</b>  | <b>713</b> |
| <i>Tao Wang, Daniel Lizotte, Michael Bowling, Dale Schuurmans</i>  |            |
| <b>A Risk Minimization Principle for a Class of Parzen Estimators.....</b>                                       | <b>721</b> |
| <i>Kristiaan Pelckmans, Johan Suykens, Bart De Moor</i>  |            |
| <b>Exponential Family Predictive Representations of State.....</b>   | <b>729</b> |
| <i>David Wingate, Satinder Singh Baveja</i>  |            |
| <b>Modeling Homophily and Stochastic Equivalence in Symmetric Relational Data.....</b>                           | <b>737</b> |
| <i>Peter Hoff</i>  |            |
| <b>Agreement-Based Learning.....</b>   | <b>745</b> |
| <i>Percy Liang, Dan Klein, Michael Jordan</i>  |            |
| <b>Competition Adds Complexity .....</b>   | <b>753</b> |
| <i>Judy Goldsmith, Martin Mundhenk</i>   |            |
| <b>Learning to Classify Complex Patterns Using a VLSI Network of Spiking Neurons .....</b>                       | <b>761</b> |
| <i>Srinjoy Mitra, Giacomo Indiveri, Stefano Fusi</i>   |            |
| <b>Extending Position/Phase-Shift Tuning to Motion Energy Neurons Improves Velocity<br/>Discrimination .....</b> | <b>769</b> |
| <i>Yiu Man Lam, Bertram Shi</i>  |            |
| <b>COFI RANK - Maximum Margin Matrix Factorization for Collaborative Ranking.....</b>                            | <b>777</b> |
| <i>Markus Weimer, Alexandros Karatzoglou, Quoc Le, Alex Smola</i>  |            |
| <b>Hierarchical Penalization .....</b>   | <b>785</b> |
| <i>Marie Szafranski, Yves Grandvalet, Pierre Morizet-Mahoudeaux</i>  |            |
| <b>Efficient Multiple Hyperparameter Learning for Log-Linear Models .....</b>                                    | <b>793</b> |
| <i>Chuong Do, Chuan-Sheng Foo, Andrew Ng</i>   |            |

|  |     |
|--|-----|
| <b>Learning Bounds for Domain Adaptation</b> .....   | 801 |
| <i>John Blitzer, Koby Crammer, Alex Kulesza, Fernando Pereira, Jennifer Wortman</i>                              |     |
| <b>Discriminative Log-Linear Grammars with Latent Variables</b> .....  | 809 |
| <i>Slav Petrov, Dan Klein</i>  |     |
| <b>Online Linear Regression and Its Application to Model-Based Reinforcement Learning</b> .....                  | 817 |
| <i>Alexander Strehl, Michael Littman</i>   |     |
| <b>Mining Internet-Scale Software Repositories</b> .....   | 825 |
| <i>Erik Linstead, Paul Rigor, Sushil Bajracharya, Cristina Lopes, Pierre Baldi</i>                               |     |
| <b>Theoretical Analysis of Learning with Reward-Modulated Spike-Timing-Dependent Plasticity</b> .....            | 833 |
| <i>Robert Legenstein, Dejan Pecevski, Wolfgang Maass</i>   |     |
| <b>Configuration Estimates Improve Pedestrian Finding</b> .....  | 841 |
| <i>Duan Tran, David Forsyth</i>  |     |
| <b>Adaptive Bayesian Inference</b> .....   | 849 |
| <i>Ozgur Sumer, Umut Acar, Alexander T. Ihler, Ramgopal R. Mettu</i>   |     |
| <b>The Noisy-Logical Distribution and Its Application to Causal Inference</b> .....                              | 857 |
| <i>Alan Yuille, HongJing Lu</i>  |     |
| <b>Modelling Motion Primitives and Their Timing in Biologically Executed Movements</b> .....                     | 865 |
| <i>Ben Williams, Marc Toussaint, Amos Storkey</i>  |     |
| <b>Distributed Inference for Latent Dirichlet Allocation</b> .....   | 873 |
| <i>David Newman, Arthur Asuncion, Padhraic Smyth, Max Welling</i>  |     |
| <b>Optimistic Linear Programming Gives Logarithmic Regret for Irreducible Mdps</b> .....                         | 881 |
| <i>Ambuj Tewari, Peter Bartlett</i>  |     |
| <b>A Learning Framework for Nearest Neighbor Search</b> .....  | 889 |
| <i>Lawrence Cayton, Sanjoy Dasgupta</i>  |     |
| <b>Modeling Natural Sounds with Modulation Cascade Processes</b> .....   | 897 |
| <i>Richard Turner, Maneesh Sahani</i>  |     |
| <b>Regret Minimization in Games with Incomplete Information</b> .....  | 905 |
| <i>Martin Zinkevich, Michael Johanson, Michael Bowling, Carmelo Piccione</i>                                     |     |
| <b>Iterative Non-linear Dimensionality Reduction with Manifold Sculpting</b> .....                               | 913 |
| <i>Michael Gashler, Dan Ventura, Tony Martinez</i>   |     |
| <b>Computational Equivalence of Fixed Points and No Regret Algorithms, and Convergence to Equilibria</b> .....   | 921 |
| <i>Elad Hazan, Satyen Kale</i>   |     |
| <b>Adaptive Online Gradient Descent</b> .....  | 929 |
| <i>Peter Bartlett, Elad Hazan, Alexander Rakhlin</i>   |     |
| <b>Catching Change-points with Lasso</b> .....   | 937 |
| <i>Zaid Harchaoui, Céline LEVY-LEDUC</i>   |     |
| <b>Discriminative Keyword Selection Using Support Vector Machines</b> .....                                      | 945 |
| <i>William Campbell, Fred Richardson</i>   |     |
| <b>Sequential Hypothesis Testing Under Stochastic Deadlines</b> .....  | 953 |
| <i>Peter Frazier, Angela Yu</i>  |     |
| <b>A Configurable Analog VLSI Neural Network with Spiking Neurons and Self-Regulating Plastic Synapses</b> ..... | 961 |
| <i>Massimiliano Giulioni, Mario Pannunzi, Davide Badoni, Vittorio Dante, Paolo Del giudice</i>                   |     |

|   |      |
|---|------|
| <b>The Distribution Family of Similarity Distances</b> .....  | 968  |
| <i>Gertjan Burghouts, Arnold Smeulders, Jan-Mark Geusebroek</i>   |      |
| <b>The Tradeoffs of Large Scale Learning</b> .....  | 976  |
| <i>Leon Bottou, Olivier Bousquet</i>  |      |
| <b>A Kernel Statistical Test of Independence</b> .....  | 984  |
| <i>Arthur Gretton, Kenji Fukumizu, Choon Hui Teo, Le Song, Bernhard Schölkopf, Alex Smola</i>                 |      |
| <b>Discriminative K-means for Clustering</b> .....  | 992  |
| <i>Jieping Ye, Zheng Zhao, Mingrui Wu</i>   |      |
| <b>A General Boosting Method and Its Application to Learning Ranking Functions for Web Search</b> .....       | 1000 |
| <i>Zhaohui Zheng, Hongyuan Zha, Tong Zhang, Olivier Chapelle, Keke Chen, Gordon Sun</i>                       |      |
| <b>Hidden Common Cause Relations in Relational Learning</b> .....   | 1008 |
| <i>Ricardo Silva, Wei Chu, Zoubin Ghahramani</i>  |      |
| <b>Evaluating Search Engines by Modeling the Relationship Between Relevance and Clicks</b> .....              | 1016 |
| <i>Ben Carterette, Rosie Jones</i>  |      |
| <b>Catching Up Faster in Bayesian Model Selection and Model Averaging</b> .....                               | 1024 |
| <i>Tim Van Erven, Peter Grunwald, Steven De Rooij</i>   |      |
| <b>The Price of Bandit Information for Online Optimization</b> .....  | 1032 |
| <i>Varsha Dani, Thomas Hayes, Sham Kakade</i>   |      |
| <b>Bayesian Agglomerative Clustering with Coalescents</b> .....   | 1040 |
| <i>Yee Whye Teh, Hal Daume III, Daniel Roy</i>  |      |
| <b>Collapsed Variational Inference for HDP</b> .....  | 1048 |
| <i>Yee Whye Teh, Kenichi Kurihara, Max Welling</i>  |      |
| <b>Second Order Bilinear Discriminant Analysis for Single Trial EEG Analysis</b> .....                        | 1056 |
| <i>Christoforos Christoforou, Paul Sajda, Lucas C. Parra</i>  |      |
| <b>Object Recognition by Scene Alignment</b> .....  | 1064 |
| <i>Bryan Russell, Antonio Torralba, Ce Liu, Rob Fergus, William Freeman</i>                                   |      |
| <b>An Analysis of Inference with the Universum</b> .....  | 1072 |
| <i>Fabian Sinz, Olivier Chapelle, Alekh Agarwal, Bernhard Schölkopf</i>                                       |      |
| <b>Local Algorithms for Approximate Inference in Minor-Excluded Graphs</b> .....                              | 1080 |
| <i>Kyomin Jung, Devavrat Shah</i>   |      |
| <b>Estimating Divergence Functionals and the Likelihood Ratio by Penalized Convex Risk Minimization</b> ..... | 1088 |
| <i>XuanLong Nguyen, Martin Wainwright, Michael Jordan</i>   |      |
| <b>The Epoch-Greedy Algorithm for Multi-armed Bandits with Side Information</b> .....                         | 1096 |
| <i>John Langford, Tong Zhang</i>  |      |
| <b>Blind Channel Identification for Speech Dereverberation Using L1-Norm Sparse Learning</b> .....            | 1104 |
| <i>Yuanqing Lin, Jingdong Chen, Youngmoo Kim, Daniel Lee</i>  |      |
| <b>Unsupervised Feature Selection for Accurate Recommendation of High-Dimensional Image Data</b> .....        | 1112 |
| <i>Sabri Boutemedjet, Djemel Ziou, Nizar Bouguila</i>   |      |
| <b>Subspace-Based Face Recognition in Analog VLSI</b> .....   | 1120 |
| <i>Gonzalo Carvajal, Waldo Valenzuela, Miguel Figueroa</i>  |      |
| <b>Computing Robust Counter-Strategies</b> .....  | 1128 |
| <i>Michael Johanson, Martin Zinkevich, Michael Bowling</i>  |      |

|  |      |
|--|------|
| <b>Structured Learning with Approximate Inference</b> .....  | 1136 |
| <i>Alex Kulesza, Fernando Pereira</i>  |      |
| <b>Augmented Functional Time Series Representation and Forecasting with Gaussian Processes</b> ..... | 1144 |
| <i>Nicolas Chapados, Yoshua Bengio</i>   |      |
| <b>New Outer Bounds on the Marginal Polytope</b> .....   | 1152 |
| <i>David Sontag, Tommi Jaakkola</i>  |      |
| <b>Random Features for Large-Scale Kernel Machines</b> .....   | 1160 |
| <i>Ali Rahimi, Benjamin Recht</i>  |      |

### VOLUME 3

|  |      |
|--|------|
| <b>The Infinite Markov Model</b> .....   | 1168 |
| <i>Daichi Mochihashi, Eiichiro Sumita</i>  |      |
| <b>McRank: Learning to Rank Using Multiple Classification and Gradient Boosting</b> .....                | 1176 |
| <i>Ping Li, Christopher Burges, Qiang Wu</i>   |      |
| <b>Non-parametric Modeling of Partially Ranked Data</b> .....  | 1184 |
| <i>Guy Lebanon, Yi Mao</i>   |      |
| <b>Fast Variational Inference for Large-scale Internet Diagnosis</b> .....                               | 1192 |
| <i>John Platt, Emre Kiciman, David Maltz</i>   |      |
| <b>Efficient Inference for Distributions on Permutations</b> .....                                       | 1200 |
| <i>Jonathan Huang, Carlos Guestrin, Leonidas Guibas</i>  |      |
| <b>Fast and Scalable Training of Semi-Supervised CRFs with Application to Activity Recognition</b> ..... | 1208 |
| <i>Maryam Mahdavian, Tanzeem Choudhury</i>   |      |
| <b>Bayes-Adaptive POMDPs</b> .....   | 1216 |
| <i>Stephane Ross, Brahim Chaib-draa, Joelle Pineau</i>   |      |
| <b>DIFFRAC: A Discriminative and Flexible Framework for Clustering</b> .....                             | 1224 |
| <i>Francis Bach, Zaid Harchaoui</i>  |      |
| <b>The Discriminant Center-Surround Hypothesis for Bottom-Up Saliency</b> .....                          | 1232 |
| <i>Dashan Gao, Vijay Mahadevan, Nuno Vasconcelos</i>   |      |
| <b>A Bayesian Framework for Cross-Situational Word-Learning</b> .....                                    | 1240 |
| <i>Michael Frank, Noah Goodman, Joshua Tenenbaum</i>   |      |
| <b>Privacy-Preserving Belief Propagation and Sampling</b> .....  | 1248 |
| <i>Michael Kearns, Jinsong Tan, Jennifer Wortman</i>   |      |
| <b>A Probabilistic Approach to Language Change</b> .....   | 1256 |
| <i>Alexandre Bouchard-Côté, Percy Liang, Thomas Griffiths, Dan Klein</i>                                 |      |
| <b>Boosting Algorithms for Maximizing the Soft Margin</b> .....  | 1264 |
| <i>Manfred Warmuth, Karen Glocer, Gunnar Rätsch</i>  |      |
| <b>Automatic Generation of Social Tags for Music Recommendation</b> .....                                | 1272 |
| <i>Douglas Eck, Paul Lamere, Thierry Bertin-Mahieux, Stephen Green</i>                                   |      |
| <b>Supervised Topic Models</b> .....   | 1280 |
| <i>David Blei, Jon McAuliffe</i>   |      |
| <b>Unconstrained On-line Handwriting Recognition with Recurrent Neural Networks</b> .....                | 1288 |
| <i>Alex Graves, Santiago Fernandez, Marcus Liwicki, Horst Bunke, Juergen Schmidhuber</i>                 |      |

|   |      |
|---|------|
| <b>Predictive Matrix-Variate t Models</b> .....   | 1296 |
| <i>Shenghuo Zhu, Kai Yu, Yihong Gong</i>  |      |
| <b>Active Preference Learning with Discrete Choice Data</b> .....   | 1304 |
| <i>Brochu Eric, Nando De Freitas, Abhijeet Ghosh</i>  |      |
| <b>Linear Programming Analysis of Loopy Belief Propagation for Weighted Matching</b> .....                      | 1312 |
| <i>Sujay Sanghavi, Dmitry Malioutov, Alan Willsky</i>   |      |
| <b>Managing Power Consumption and Performance of Computing Systems Using Reinforcement Learning</b> .....       | 1320 |
| <i>Gerald Tesauro, Rajarshi Das, Hoi Chan, Jeffrey Kephart, David Levine, Freeman Rawson, Charles Lefurgy</i>   |      |
| <b>Congruence Between Model and Human Attention Reveals Unique Signatures of Critical Visual Events</b> .....   | 1328 |
| <i>Robert Peters, Laurent Itti</i>  |      |
| <b>Locality and Low-Dimensions in the Prediction of Natural Experience from fMRI</b> .....                      | 1336 |
| <i>Francois Meyer, Greg Stephens</i>  |      |
| <b>Fitted Q-iteration in Continuous Action-Space MDPs</b> .....   | 1344 |
| <i>András Antos, Remi Munos, Csaba Szepesvari</i>   |      |
| <b>Expectation Maximization and Posterior Constraints</b> .....   | 1352 |
| <i>Joao Graca, Kuzman Ganchev, Ben Taskar</i>   |      |
| <b>Regulator Discovery from Gene Expression Time Series of Malaria Parasites: a Hierarchical Approach</b> ..... | 1360 |
| <i>José Miguel Hernández-Lobato, Tjeerd Dijkstra, Tom Heskes</i>  |      |
| <b>Discriminative Batch Mode Active Learning</b> .....  | 1368 |
| <i>Yuhong Guo, Dale Schuurmans</i>  |      |
| <b>Message Passing for Max-weight Independent Set</b> .....   | 1376 |
| <i>Sujay Sanghavi, Devavrat Shah, Alan Willsky</i>  |      |
| <b>Learning the 2-D Topology of Images</b> .....  | 1384 |
| <i>Nicolas Le Roux, Yoshua Bengio, Pascal Lamblin, Marc Joliveau, Balázs Kégl</i>                               |      |
| <b>Hippocampal Contributions to Control: The Third Way</b> .....  | 1392 |
| <i>Máté Lengyel, Peter Dayan</i>  |      |
| <b>Gaussian Process Models for Link Analysis and Transfer Learning</b> .....                                    | 1400 |
| <i>Kai Yu, Wei Chu</i>  |      |
| <b>TrueSkill Through Time: Revisiting the History of Chess</b> .....  | 1408 |
| <i>Pierre Dangauthier, Ralf Herbrich, Tom Minka, Thore Graepel</i>  |      |
| <b>Sparse Deep Belief Net Model for Visual Area V2</b> .....  | 1416 |
| <i>Honglak Lee, Chaitanya Ekanadham, Andrew Ng</i>  |      |
| <b>Fixing Max-Product: Convergent Message Passing Algorithms for MAP LP-Relaxations</b> .....                   | 1424 |
| <i>Amir Globerson, Tommi Jaakkola</i>   |      |
| <b>Using Deep Belief Nets to Learn Covariance Kernels for Gaussian Processes</b> .....                          | 1432 |
| <i>Ruslan Salakhutdinov, Geoffrey Hinton</i>  |      |
| <b>A Game-Theoretic Approach to Apprenticeship Learning</b> .....   | 1440 |
| <i>Umar Syed, Robert Schapire</i>   |      |
| <b>Modeling Image Patches with a Directed Hierarchy of Markov Random Fields</b> .....                           | 1448 |
| <i>Simon Osindero, Geoffrey Hinton</i>  |      |

|  |      |
|--|------|
| <b>Reinforcement Learning in Continuous Action Spaces Through Sequential Monte Carlo Methods</b> .....               | 1456 |
| <i>Alessandro Lazaric, Marcello Restelli, Andrea Bonarini</i>  |      |
| <b>A Bayesian LDA-Based Model for Semi-Supervised Part-of-Speech Tagging</b> .....                                   | 1464 |
| <i>Kristina Toutanova, Mark Johnson</i>  |      |
| <b>A New View of Automatic Relevance Determination</b> .....   | 1472 |
| <i>David Wipf, Srikantan Nagarajan</i>   |      |
| <b>Robust Regression with Twinned Gaussian Processes</b> .....   | 1480 |
| <i>Andrew Naish-Guzman, Sean Holden</i>  |      |
| <b>Invariant Common Spatial Patterns: Alleviating Nonstationarities in Brain-Computer Interfacing</b> .....          | 1488 |
| <i>Benjamin Blankertz, Motoaki Kawanabe, Ryota Tomioka, Friederike Hohlefeld, Vadim Nikulin, Klaus-Robert Müller</i> |      |
| <b>Hierarchical Apprenticeship Learning with Application to Quadruped Locomotion</b> .....                           | 1496 |
| <i>J. Zico Kolter, Pieter Abbeel, Andrew Ng</i>  |      |
| <b>CPR for CSPs: A Probabilistic Relaxation of Constraint Propagation</b> .....                                      | 1504 |
| <i>Luis E. Ortiz</i>   |      |
| <b>Combined Discriminative and Generative Articulated Pose and Non-Rigid Shape Estimation</b> .....                  | 1512 |
| <i>Leonid Sigal, Alexandru Balan, Michael Black</i>  |      |
| <b>Neural Characterization in Partially Observed Populations of Spiking Neurons</b> .....                            | 1520 |
| <i>Jonathan Pillow, Peter Latham</i>   |      |
| <b>The Pigeon As Particle Filter</b> .....   | 1528 |
| <i>Nathaniel Daw, Aaron Courville</i>  |      |
| <b>The Value of Labeled and Unlabeled Examples When the Model is Imperfect</b> .....                                 | 1536 |
| <i>Kaushik Sinha, Mikhail Belkin</i>   |      |
| <b>Probabilistic Matrix Factorization</b> .....  | 1544 |
| <i>Ruslan Salakhutdinov, Andriy Mnih</i>   |      |
| <b>A Bayesian Model of Conditioned Perception</b> .....  | 1552 |
| <i>Alan Stocker, Eero Simoncelli</i>   |      |
| <b>Efficient Principled Learning of Thin Junction Trees</b> .....  | 1560 |
| <i>Anton Chechetka, Carlos Guestrin</i>  |      |
| <b>Semi-Supervised Multitask Learning</b> .....  | 1568 |
| <i>Qihua Liu, Xuejun Liao, Lawrence Carin</i>  |      |
| <b>Sparse Overcomplete Latent Variable Decomposition of Counts Data</b> .....  | 1576 |
| <i>Madhusudana Shashanka, Bhiksha Raj, Paris Smaragdís</i>   |      |
| <b>Learning Monotonic Transformations for Classification</b> .....   | 1584 |
| <i>Andrew Howard, Tony Jebara</i>  |      |
| <b>Convex Learning with Invariances</b> .....  | 1592 |
| <i>Choon Hui Teo, Amir Globerson, Sam Roweis, Alex Smola</i>   |      |
| <b>On Sparsity and Overcompleteness in Image Models</b> .....  | 1600 |
| <i>Pietro Berkes, Richard Turner, Maneesh Sahani</i>   |      |
| <b>Feature Selection Methods for Improving Protein Structure Prediction with Rosetta</b> .....                       | 1608 |
| <i>Ben Blum, Michael Jordan, David Kim, Rhiju Das, Philip Bradley, David Baker</i>                                   |      |

|  |      |
|--|------|
| <b>Testing for Homogeneity with Kernel Fisher Discriminant Analysis</b> .....                | 1616 |
| <i>Zaid Harchaoui, Francis Bach, Moulines Eric</i>   |      |
| <b>Density Estimation Under Independent Similarly Distributed Sampling Assumptions</b> ..... | 1624 |
| <i>Tony Jebara, Yingbo Song, Kapil Thadani</i>   |      |
| <b>Ultrafast Monte Carlo for Statistical Summations</b> .....                                | 1632 |
| <i>Michael Holmes, Alexander Gray, Charles Isbell</i>  |      |
| <b>Multiple-Instance Active Learning</b> .....   | 1640 |
| <i>Burr Settles, Mark Craven, Soumya Ray</i>   |      |
| <b>Predicting Human Gaze Using Low-Level Saliency Combined with Face Detection</b> .....     | 1648 |
| <i>Moran Cerf, Jonathan Harel, Wolfgang Einhaeuser, Christof Koch</i>                        |      |
| <b>Estimating Disparity with Confidence from Energy Neurons</b> .....                        | 1656 |
| <i>Eric Kong-Chau Tsang, Bertram Shi</i>   |      |
| <b>Loop Series and Bethe Variational Bounds in Attractive Graphical Models</b> .....         | 1664 |
| <i>Erik Sudderth, Martin Wainwright, Alan Willsky</i>  |      |
| <b>Learning with Tree-Averaged Densities and Distributions</b> .....                         | 1672 |
| <i>Sergey Kirshner</i>   |      |
| <b>Variational Inference for Diffusion Processes</b> .....                                   | 1680 |
| <i>Cédric Archambeau, Manfred Opper, Yuan Shen, Dan Cornford, John Shawe-Taylor</i>          |      |
| <b>Random Projections for Manifold Learning</b> .....  | 1688 |
| <i>Chinmay Hegde, Michael Wakin, Richard Baraniuk</i>  |      |
| <b>Receding Horizon Differential Dynamic Programming</b> .....                               | 1696 |
| <i>Yuval Tassa, Tom Erez, William Smart</i>  |      |
| <b>An In-silico Neural Model of Dynamic Routing Through Neuronal Coherence</b> .....         | 1704 |
| <i>Devarajan Sridharan, Brian Percival, John Arthur, Kwabena Boahen</i>                      |      |
| <b>Sparse Feature Learning for Deep Belief Networks</b> .....                                | 1712 |
| <i>MarcAurelio Ranzato, Y-Lan Boureau, Yann LeCun</i>  |      |
| <b>Discovering Weakly-Interacting Factors in a Complex Stochastic Process</b> .....          | 1720 |
| <i>Charlie Frogner, Avi Pfeffer</i>  |      |
| <b>Trans-Dimensional MCMC for Bayesian Policy Learning</b> .....                             | 1728 |
| <i>Matthew Hoffman, Arnaud Doucet, Nando De Freitas, Ajay Jasra</i>                          |      |

**Author Index**