

Table of Contents

1. Archipelago: Nonparametric Bayesian Semi-Supervised Learning	1
Ryan Adams and Zoubin Ghahramani	
2. Tractable Nonparametric Bayesian Inference in Poisson Processes with Gaussian Process Intensities	9
Ryan Adams, Iain Murray, and David MacKay	
3. Route Kernels for Trees	17
Fabio Aioli, Giovanni Da San Martino, and Alessandro Sperduti	
4. Incorporating Domain Knowledge into Topic Modeling via Dirichlet Forest Priors	25
David Andrzejewski, Xiaojin Zhu, and Mark Craven	
5. Grammatical Inference as a Principal Component Analysis Problem	33
Raphaël Bailly, François Denis, and Liva Ralaivola	
6. Curriculum Learning	41
Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston	
7. Importance Weighted Active Learning	49
Alina Beygelzimer, Sanjoy Dasgupta, and John Langford	
8. Split Variational Inference	57
Guillaume Bouchard and Onno Zoeter	
9. Predictive Representations for Policy Gradient in POMDPs	65
Abdeslam Boularias and Brahim Chaib-draa	
10. Online Feature Elicitation in Interactive Optimization	73
Craig Boutilier, Kevin Regan, and Paolo Viappiani	
11. Spectral Clustering Based on the Graph p-Laplacian	81
Thomas Bühler and Matthias Hein	
12. Active Learning for Directed Exploration of Complex Systems	89
Michael Burl and Esther Wang	
13. Optimized Expected Information Gain for Nonlinear Dynamical Systems	97
Alberto Giovanni Busetto, Cheng Soon Ong, and Joachim M. Buhmann	
14. Probabilistic Dyadic Data Analysis with Local and Global Consistency	105
Deng Cai, Xuanhui Wang, and Xiaofei He	
15. Structure Learning of Bayesian Networks using Constraints	113
Cassio de Campos, Zhi Zeng, and Qiang Ji	
16. Robust Bounds for Classification via Selective Sampling	121
Nicolò Cesa-Bianchi, Claudio Gentile, and Francesco Orabona	
17. Multi-View Clustering via Canonical Correlation Analysis	129
Kamalika Chaudhuri, Sham Kakade, Karen Livescu, and Karthik Sridharan	
18. A Convex Formulation for Learning Shared Structures from Multiple Tasks	137
Jianhui Chen, Lei Tang, Jun Liu, and Jieping Ye	

19. Learning Kernels from Indefinite Similarities	145
Yihua Chen, Maya Gupta, and Benjamin Recht	
20. Matrix Updates for Perceptron Training of Continuous Density Hidden Markov Models	153
Chih-Chieh Cheng, Fei Sha, and Lawrence Saul	
21. Decision Tree and Instance-Based Learning for Label Ranking	161
Weiwei Cheng, Jens Hühn, and Eyke Hüllermeier	
22. Learning Dictionaries of Stable Autoregressive Models for Audio Scene Analysis	169
Youngmin Cho and Lawrence Saul	
23. Exploiting Sparse Markov and Covariance Structure in Multiresolution Models.....	177
Myung Jin Choi, Venkat Chandrasekaran, and Alan Willsky	
24. Nonparametric Estimation of the Precision-Recall Curve	185
Stéphane Cléménçon and Nicolas Vayatis	
25. EigenTransfer: A Unified Framework for Transfer Learning	193
Wenyuan Dai, Ou Jin, Gui-Rong Xue, Qiang Yang, and Yong Yu	
26. Fitting a Graph to Vector Data	201
Samuel Daitch, Jonathan Kelner, and Daniel Spielman	
27. Unsupervised Search-Based Structured Prediction	209
Hal Daumé III	
28. Deep Transfer via Second-Order Markov Logic.....	217
Jesse Davis and Pedro Domingos	
29. Analytic Moment-Based Gaussian Process Filtering	225
Marc Peter Deisenroth, Marco F. Huber, and Uwe D. Hanebeck	
30. Good Learners for Evil Teachers	233
Ofar Dekel and Ohad Shamir	
31. A Scalable Framework for Discovering Coherent Co-clusters in Noisy Data.....	241
Meghana Deodhar, Joydeep Ghosh, Gunjan Gupta, Hyuk Cho, and Inderjit Dhillon	
32. The Adaptive k-Meteorologists Problem and Its Application to Structure Learning and Feature Selection in Reinforcement Learning	249
Carlos Diuk, Lihong Li, and Bethany Leffler	
33. Proximal Regularization for Online and Batch Learning.....	257
Chuong Do, Quoc Le, and Chuan-Sheng Foo	
34. Large Margin Training for Hidden Markov Models with Partially Observed States	265
Trinh-Minh-Tri Do and Thierry Artières	
35. Accelerated Gibbs Sampling for the Indian Buffet Process	273
Finale Doshi-Velez and Zoubin Ghahramani	
36. Accounting for Burstiness in Topic Models	281
Gabriel Doyle and Charles Elkan	
37. Domain Adaptation from Multiple Sources via Auxiliary Classifiers.....	289
Lixin Duan, Ivor W. Tsang, Dong Xu, and Tat-Seng Chua	

38. Boosting with Structural Sparsity	297
John Duchi and Yoram Singer	
39. Learning to Segment from a Few Well-Selected Training Images	305
Alireza Farhangfar, Russ Greiner, and Csaba Szepesvári	
40. GAODE and HAODE: Two Proposals Based on AODE to Deal with Continuous Variables	313
M. Julia Flores, José A. Gámez, Ana M. Martínez, and José M. Puerta	
41. A Majorization-Minimization Algorithm for (Multiple) Hyperparameter Learning	321
Chuan-Sheng Foo, Chuong Do, and Andrew Ng	
42. Dynamic Mixed Membership Blockmodel for Evolving Networks	329
Wenjie Fu, Le Song, and Eric Xing	
43. Gradient Descent with Sparsification: An Iterative Algorithm for Sparse Recovery with Restricted Isometry Property	337
Rahul Garg and Rohit Khandekar	
44. Sequential Bayesian Prediction in the Presence of Changepoints	345
Roman Garnett, Michael Osborne, and Stephen Roberts	
45. PAC-Bayesian Learning of Linear Classifiers	353
Pascal Germain, Alexandre Lacasse, François Laviolette, and Mario Marchand	
46. Fast Evolutionary Maximum Margin Clustering	361
Fabian Gieseke, Tapio Pahikkala, and Oliver Kramer	
47. Dynamic Analysis of Multiagent Q-learning with ϵ-greedy Exploration	369
Eduardo Rodrigues Gomes and Ryszard Kowalczyk	
48. Bayesian Inference for Plackett-Luce Ranking Models	377
John Guiver and Edward Snelson	
49. Bayesian Clustering for Email Campaign Detection	385
Peter Haider and Tobias Scheffer	
50. Efficient Learning Algorithms for Changing Environments	393
Elad Hazan and C. Seshadhri	
51. Hoeffding and Bernstein Races for Selecting Policies in Evolutionary Direct Policy Search	401
Verena Heidrich-Meisner and Christian Igel	
52. Partially Supervised Feature Selection with Regularized Linear Models	409
Thibault Helleputte and Pierre Dupont	
53. Learning with Structured Sparsity	417
Junzhou Huang, Tong Zhang, and Dimitris Metaxas	
54. Learning Linear Dynamical Systems without Sequence Information	425
Tzu-Kuo Huang and Jeff Schneider	
55. Group Lasso with Overlaps and Graph Lasso	433
Laurent Jacob, Guillaume Obozinski, and Jean-Philippe Vert	
56. Graph Construction and b-Matching for Semi-Supervised Learning	441
Tony Jebara, Jun Wang, and Shih-Fu Chang	

57. Trajectory Prediction: Learning to Map Situations to Robot Trajectories	449
Nikolay Jetchev and Marc Toussaint	
58. An Accelerated Gradient Method for Trace Norm Minimization.....	457
Shuiwang Ji and Jieping Ye	
59. A Novel Lexicalized HMM-Based Learning Framework for Web Opinion Mining	465
Wei Jin and Hung Hay Ho	
60. Orbit-Product Representation and Correction of Gaussian Belief Propagation	473
Jason Johnson, Vladimir Chernyak, and Michael Chertkov	
61. A Bayesian Approach to Protein Model Quality Assessment.....	481
Hetunandan Kamisetty and Christopher James Langmead	
62. Learning Prediction Suffix Trees with Winnow	489
Nikos Karampatziakis and Dexter Kozen	
63. Boosting Products of Base Classifiers	497
Balázs Kégl and Róbert Busa-Fekete	
64. Learning Markov Logic Network Structure via Hypergraph Lifting	505
Stanley Kok and Pedro Domingos	
65. Near-Bayesian Exploration in Polynomial Time.....	513
J. Zico Kolter and Andrew Ng	
66. Regularization and Feature Selection in Least-Squares Temporal Difference Learning.....	521
J. Zico Kolter and Andrew Ng	
67. The Graphlet Spectrum	529
Risi Kondor, Nino Shervashidze, and Karsten Borgwardt	
68. Rule Learning with Monotonicity Constraints	537
Wojciech Kotlowski and Roman Slowinski	
69. Multiple Indefinite Kernel Learning with Mixed Norm Regularization.....	545
Matthieu Kowalski, Marie Szafranski, and Liva Ralaivola	
70. On Sampling-Based Approximate Spectral Decomposition	553
Sanjiv Kumar, Mehryar Mohri, and Ameet Talwalkar	
71. Learning Spectral Graph Transformations for Link Prediction	561
Jérôme Kunegis and Andreas Lommatzsch	
72. Block-Wise Construction of Acyclic Relational Features with Monotone Irreducibility and Relevancy Properties.....	569
Ondřej Kuželka and Filip Železný	
73. Generalization Analysis of Listwise Learning-to-Rank Algorithms	577
Yanyan Lan, Tie-Yan Liu, Zhiming Ma, and Hang Li	
74. Approximate Inference for Planning in Stochastic Relational Worlds	585
Tobias Lang and Marc Toussaint	
75. Learning Nonlinear Dynamic Models	593
John Langford, Ruslan Salakhutdinov, and Tong Zhang	

76. Non-Linear Matrix Factorization with Gaussian Processes	601
Neil D. Lawrence and Raquel Urtasun	
77. Convolutional Deep Belief Networks for Scalable Unsupervised Learning of Hierarchical Representations.....	609
Honglak Lee, Roger Grosse, Rajesh Ranganath, and Andrew Ng	
78. Transfer Learning for Collaborative Filtering via a Rating-Matrix Generative Model	617
Bin Li, Qiang Yang, and Xiangyang Xue	
79. ABC-Boost: Adaptive Base Class Boost for Multi-Class Classification	625
Ping Li	
80. Semi-Supervised Learning Using Label Mean.....	633
Yu-Feng Li, James Kwok, and Zhi-Hua Zhou	
81. Learning from Measurements in Exponential Families	641
Percy Liang, Dan Klein, and Michael Jordan	
82. Blockwise Coordinate Descent Procedures for the Multi-Task Lasso, with Applications to Neural Semantic Basis Discovery	649
Han Liu, Mark Palatucci, and Jian Zhang	
83. Efficient Euclidean Projections in Linear Time	657
Jun Liu and Jieping Ye	
84. Topic-Link LDA: Joint Models of Topic and Author Community.....	665
Yan Liu, Alexandru Niculescu-mizil, and Wojciech Gryc	
85. Geometry-Aware Metric Learning.....	673
Zhengdong Lu, Prateek Jain, and Inderjit Dhillon	
86. Identifying Suspicious URLs: An Application of Large-Scale Online Learning	681
Justin Ma, Lawrence K. Saul, Stefan Savage, and Geoffrey M. Voelker	
87. Online Dictionary Learning for Sparse Coding	689
Julien Mairal, Francis Bach, Jean Ponce, and Guillermo Sapiro	
88. Proto-Predictive Representation of States with Simple Recurrent Temporal-Difference Networks	697
Takaki Makino	
89. Sparse Gaussian Graphical Models with Unknown Block Structure	705
Benjamin M. Marlin and Kevin P. Murphy	
90. Polyhedral Outer Approximations with Application to Natural Language Parsing	713
André Martins, Noah Smith, and Eric Xing	
91. Partial Order Embedding with Multiple Kernels.....	721
Brian McFee and Gert Lanckriet	
92. Bandit-Based Optimization on Graphs with Application to Library Performance Tuning.....	729
Frédéric de Mesmay, Arpad Rimmel, Yevgen Voronenko, and Markus Püschel	
93. Deep Learning from Temporal Coherence in Video	737
Hossein Mobahi, Ronan Collobert, and Jason Weston	

94. Regression by Dependence Minimization and its Application to Causal Inference	745
Joris Mooij, Dominik Janzing, Jonas Peters, and Bernhard Schölkopf	
95. Learning Complex Motions by Sequencing Simpler Motion Templates	753
Gerhard Neumann, Wolfgang Maass, and Jan Peters	
96. Convex Variational Bayesian Inference for Large Scale Generalized Linear Models	761
Hannes Nickisch and Matthias Seeger	
97. Solution Stability in Linear Programming Relaxations: Graph Partitioning and Unsupervised Learning.....	769
Sebastian Nowozin and Stefanie Jegelka	
98. Nonparametric Factor Analysis with Beta Process Priors.....	777
John Paisley and Lawrence Carin	
99. Unsupervised Hierarchical Modeling of Locomotion Styles	785
Wei Pan and Lorenzo Torresani	
100. Binary Action Search for Learning Continuous-Action Control Policies	793
Jason Pavis and Michail Lagoudakis	
101. Detecting the Direction of Causal Time Series	801
Jonas Peters, Dominik Janzing, Arthur Gretton, and Bernhard Schölkopf	
102. Constraint Relaxation in Approximate Linear Programs	809
Marek Petrik and Shlomo Zilberstein	
103. Multi-Class Image Segmentation using Conditional Random Fields and Global Classification	817
Nils Plath, Marc Toussaint, and Shinichi Nakajima	
104. Learning When to Stop Thinking and Do Something!.....	825
Barnabás Póczos, Yasin Abbasi-Yadkori, Csaba Szepesvári, Russell Greiner, and Nathan Sturtevant	
105. Independent Factor Topic Models.....	833
Duangmanee Putthividhya, Hagai Attias, and Srikantan Nagarajan	
106. An Efficient Sparse Metric Learning in High-Dimensional Space via ℓ_1-Penalized Log-Determinant Regularization.....	841
Guo-Jun Qi, Jinhui Tang, Zheng-Jun Zha, Tat-Seng Chua, and Hong-Jiang Zhang	
107. Sparse Higher Order Conditional Random Fields for Improved Sequence Labeling.....	849
Xian Qian, Xiaoqian Jiang, Zhang Qi, Xuanjing Huang, and Lide Wu	
108. An Efficient Projection for $l_{1,\infty}$ Regularization	857
Ariadna Quattoni, Xavier Carreras, Michael Collins, and Trevor Darrell	
109. Nearest Neighbors in High-Dimensional Data: The Emergence and Influence of Hubs	865
Milos Radovanovic, Alexandros Nanopoulos, and Mirjana Ivanovic	
110. Large-Scale Deep Unsupervised Learning using Graphics Processors	873
Rajat Raina, Anand Madhavan, and Andrew Ng	
111. The Bayesian Group-Lasso for Analyzing Contingency Tables.....	881
Sudhir Raman, Thomas Fuchs, Peter Wild, Edgar Dahl, and Volker Roth	

112. Supervised Learning from Multiple Experts: Whom to Trust when Everyone Lies a Bit	889
Vikas Raykar, Shipeng Yu, Linda Zhao, Anna Jerebko, Charles Florin, Gerardo Valadez, Luca Bogoni, and Linda Moy	
113. Surrogate Regret Bounds for Proper Losses	897
Mark Reid and Robert Williamson	
114. Learning Structurally Consistent Undirected Probabilistic Graphical Models	905
Sushmita Roy, Terran Lane, and Margaret Werner-Washburne	
115. Ranking Interesting Subgroups	913
Stefan Rueding	
116. Function Factorization using Warped Gaussian Processes	921
Mikkel N. Schmidt	
117. Stochastic Methods for ℓ_1 Regularized Loss Minimization	929
Shai Shalev-Shwartz and Ambuj Tewari	
118. Structure Preserving Embedding	937
Blake Shaw and Tony Jebara	
119. Monte-Carlo Simulation Balancing	945
David Silver and Gerald Tesauro	
120. Uncertainty Sampling and Transductive Experimental Design for Active Dual Supervision	953
Vikas Sindhvani, Prem Melville, and Richard Lawrence	
121. Hilbert Space Embeddings of Conditional Distributions with Applications to Dynamical Systems	961
Le Song, Jonathan Huang, Alex Smola, and Kenji Fukumizu	
122. Multi-Assignment Clustering for Boolean Data	969
Andreas P. Streich, Mario Frank, David Basin, and Joachim M. Buhmann	
123. A Least Squares Formulation for a Class of Generalized Eigenvalue Problems in Machine Learning	977
Liang Sun, Shuiwang Ji, and Jieping Ye	
124. A Simpler Unified Analysis of Budget Perceptrons	985
Ilya Sutskever	
125. Fast Gradient-Descent Methods for Temporal-Difference Learning with Linear Function Approximation	993
Richard Sutton, Hamid Maei, Doina Precup, Shalabh Bhatnagar, David Silver, Csaba Szepesvari and Eric Wiewiora	
126. Optimistic Initialization and Greediness Lead to Polynomial Time Learning in Factored MDPs	1001
István Szita and András Lörincz	
127. Discriminative k Metrics	1009
Arthur Szlam and Guillermo Sapiro	
128. Kernelized Value Function Approximation for Reinforcement Learning	1017
Gavin Taylor and Ronald Parr	

129. Factored Conditional Restricted Boltzmann Machines for Modeling Motion Style	1025
Graham Taylor and Geoffrey Hinton	
130. Using Fast Weights to Improve Persistent Contrastive Divergence	1033
Tijmen Tieleman and Geoffrey Hinton	
131. Structure Learning with Independent Non-Identically Distributed Data.....	1041
Robert Tillman	
132. Robot Trajectory Optimization using Approximate Inference	1049
Marc Toussaint	
133. Ranking with Ordered Weighted Pairwise Classification	1057
Nicolas Usunier, David Buffoni, and Patrick Gallinari	
134. More Generality in Efficient Multiple Kernel Learning.....	1065
Manik Varma and Rakesh Babu	
135. Information Theoretic Measures for Clusterings Comparison: Is a Correction for Chance Necessary?	1073
Nguyen Xuan Vinh, Julien Epps, and James Bailey	
136. Model-Free Reinforcement Learning as Mixture Learning.....	1081
Nikos Vlassis and Marc Toussaint	
137. BoltzRank: Learning to Maximize Expected Ranking Gain	1089
Maksims Volkovs and Richard Zemel	
138. K-means in Space: A Radiation Sensitivity Evaluation.....	1097
Kiri Wagstaff and Benjamin Bornstein	
139. Evaluation Methods for Topic Models	1105
Hanna Wallach, Iain Murray, Ruslan Salakhutdinov, and David Mimno	
140. Feature Hashing for Large Scale Multitask Learning.....	1113
Kilian Weinberger, Anirban Dasgupta, John Langford, Alex Smola, and Josh Attenberg	
141. Herding Dynamic Weights to Learn	1121
Max Welling	
142. A Stochastic Memoizer for Sequence Data	1129
Frank Wood, Cédric Archambeau, Jan Gasthaus, Lancelot James, and Yee Whye Teh	
143. Optimal Reverse Prediction: A Unified Perspective on Supervised, Unsupervised and Semi-supervised Learning	1137
Linli Xu, Martha White, and Dale Schuurmans	
144. Non-Monotonic Feature Selection	1145
Zenglin Xu, Rong Jin, Jieping Ye, Michael R. Lyu, and Irwin King	
145. Online Learning by Ellipsoid Method	1153
Liu Yang, Rong Jin, and Jieping Ye	
146. Stochastic Search using the Natural Gradient.....	1161
Sun Yi, Daan Wierstra, Tom Schaul, and Juergen Schmidhuber	
147. Learning Structural SVMs with Latent Variables.....	1169
Chun-Nam Yu and Thorsten Joachims	

148. Piecewise-Stationary Bandit Problems with Side Observations	1177
Jia Yuan Yu and Shie Mannor	
149. Large-scale Collaborative Prediction Using a Nonparametric Random Effects Model.....	1185
Kai Yu, John Lafferty, Shenghuo Zhu, and Yihong Gong	
150. Robust Feature Extraction via Information Theoretic Learning	1193
Xiao-Tong Yuan and Bao-Gang Hu	
151. Interactively Optimizing Information Retrieval Systems as a Dueling Bandits Problem.....	1201
Yisong Yue and Thorsten Joachims	
152. Compositional Noisy-Logical Learning.....	1209
Alan Yuille and Songfeng Zheng	
153. Discovering Options from Example Trajectories.....	1217
Peng Zang, Peng Zhou, David Minnen, and Charles Isbell	
154. Learning Instance Specific Distances Using Metric Propagation.....	1225
De-Chuan Zhan, Ming Li, Yu-Feng Li, and Zhi-Hua Zhou	
155. Prototype Vector Machine for Large Scale Semi-supervised Learning	1233
Kai Zhang, James T. Kwok, and Bahram Parvin	
156. Learning Non-Redundant Codebooks for Classifying Complex Objects.....	1241
Wei Zhang, Akshat Surve, Xiaoli Fern, and Thomas Dietterich	
157. Multi-Instance Learning by Treating Instances As Non-I.I.D. Samples.....	1249
Zhi-Hua Zhou, Yu-Yin Sun, and Yu-Feng Li	
158. MedLDA: Maximum Margin Supervised Topic Models for Regression and Classification	1257
Jun Zhu, Amr Ahmed, and Eric Xing	
159. On Primal and Dual Sparsity of Markov Networks	1265
Jun Zhu and Eric Xing	
160. SimpleNPKL: Simple Non-Parametric Kernel Learning	1273
Jinfeng Zhuang, Ivor W. Tsang, and Steven C.H. Hoi	