Preface

Advisory Committee, Area Chairs, Program Committee, Auxiliary Reviewers, Sponsors, ICML-2001 Tutorials, ICML-2001 Workshops

CONTRIBUTED PAPERS

Multiple-Instance Learning of Real-Valued Data
Robert A. Amar, Daniel R. Dooly, Sally A. Goldman, and Qi Zhang

Efficient Algorithms for Decision Tree Cross-Validation
Hendrik Blockeel and Jan Struyf

Learning from Labeled and Unlabeled Data Using Graph Mincuts
Avrim Blum and Shuchi Chawla

Convergence of Gradient Dynamics with a Variable Learning Rate
Michael Bowling and Manuela Veloso

Learning an Agent's Utility Function by Observing Behavior
Urszula Chajewska, Daphne Koller, and Dirk Ormoneit

A Generalized Kalman Filter for Fixed Point Approximation and Efficient Temporal-Difference Learning
David Choi and Benjamin Van Roy

A Unified Loss Function in Bayesian Framework for Support Vector Regression
Wei Chu, S. Sathiya Keerthi, and Chong Jin Ong

Boosting with Confidence Information
Craig W. Codrington

Latent Semantic Kernels
Nello Cristianini, John Shawe-Taylor, and Huma Lodhi

Filters, Wrappers and a Boosting-Based Hybrid for Feature Selection
Sanmay Das

Structured Prioritized Sweeping
Richard Dearden

Bias Correction in Classification Tree Construction
Alin Dobra and Johannes Gehrke

An Efficient Approach for Approximating Multi-Dimensional Range Queries and Nearest Neighbor Classification in Large Datasets
Carlotta Domeniconi and Dimitrios Gunopoulos

A General Method for Scaling Up Machine Learning Algorithms and its Application to Clustering
Pedro Domingos and Geoff Hulten

Visual Development and the Acquisition of Binocular Disparity Sensitivities
Melissa Dominguez and Robert A. Jacobs

Relevance Feedback using Support Vector Machines
Harris Drucker, Behzad Shahrary, and David C. Gibbon

A Theory-Refinement Approach to Information Extraction
Tina Eliassi-Rad and Jude Shavlik

Learning Embedded Maps of Markov Processes
Yaakov Engel and Shie Mannor
Round Robin Rule Learning .................................................. 146
Johannes Fürnkranz

WBCSVM: Weighted Bayesian Classification Based on Support Vector Machines ........................................... 154
Thomas Gärtner and Peter A. Flach

Reinforcement Learning with Bounded Risk .......................................................... 162
Peter Geibel

Learning Probabilistic Models of Relational Structure ........................................... 170
Lise Getoor, Nir Friedman, Daphne Koller, and Benjamin Taskar

Hypertext Categorization Using Hyperlink Patterns and Meta Data ....................... 178
Rayid Ghani, Seán Slattery, and Yiming Yang

Continuous-Time Hierarchical Reinforcement Learning ........................................ 186
Mohammad Ghavamzadeh and Sridhar Mahadevan

Evolutionary Search, Stochastic Policies with Memory, and Reinforcement Learning with Hidden State ........... 194
Matthew R. Glickman and Katia Sycara

Bayesian Approaches to Failure Prediction for Disk Drives .................................... 202
Greg Hamerly and Charles Elkan

General Loss Bounds for Universal Sequence Prediction ........................................ 210
Marcus Hutter

Expectation Maximization for Weakly Labeled Data ........................................... 218
Yuri Ivanov, Bruce Blumberg, and Alex Pentland

On No-Regret Learning, Fictitious Play, and Nash Equilibrium ....................... 226
Amir Jafari, Amy Greenwald, David Gondek, and Gunes Ercal

Some Theoretical Aspects of Boosting in the Presence of Noisy Data .................. 234
Wenxin Jiang

Learning to Select Good Title Words: A New Approach Based on Reverse Information Retrieval ................... 242
Rong Jin and Alexander G. Hauptmann

Composite Kernels for Hypertext Categorisation ........................................... 250
Thorsten Joachims, Nello Cristianini, and John Shawe-Taylor

Feature Construction with Version Spaces for Biochemical Applications .................. 258
Stefan Kramer and Luc De Raedt

Pairwise Comparison of Hypotheses in Evolutionary Learning ............................. 266
Krzysztof Krawiec

Boosting Noisy Data ........................................................................... 274
Abba Krieger, Chuan Long, and Abraham Wyner

John Lafferty, Andrew McCallum, and Fernando Pereira

An Improved Predictive Accuracy Bound for Averaging Classifiers .......................... 290
John Langford, Matthias Seeger, and Nimrod Megiddo

Adjusting the Outputs of a Classifier to New a Priori Probabilities May Significantly Improve Classification Accuracy: Evidence from a Multi-Class Problem in Remote Sensing .......................... 298
Patrice Latinne, Marco Saerens, and Christine Decaestecker

Estimating a Kernel Fisher Discriminant in the Presence of Label Noise .................. 306
Neil D. Lawrence and Bernhard Schölkopf

Collaborative Learning for Recommender Systems ........................................... 314
Wee Sun Lee
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Friend-or-Foe Q-learning in General-Sum Games</td>
<td>322</td>
</tr>
<tr>
<td>Michael L. Littman</td>
<td></td>
</tr>
<tr>
<td>Using EM to Learn 3D Models of Indoor Environments with Mobile Robots</td>
<td>329</td>
</tr>
<tr>
<td>Yufeng Liu, Rosemary Emery, Deepayan Chakrabarti, Wolfram Burgard, and Sebastian Thrun</td>
<td></td>
</tr>
<tr>
<td>Inducing Partially-Defined Instances with Evolutionary Algorithms</td>
<td>337</td>
</tr>
<tr>
<td>Xavier Llorà and Josep M. Garrell</td>
<td></td>
</tr>
<tr>
<td>Learning with the Set Covering Machine</td>
<td>345</td>
</tr>
<tr>
<td>Mario Marchand and John Shawe-Taylor</td>
<td></td>
</tr>
<tr>
<td>Coupled Clustering: A Method for Detecting Structural Correspondence</td>
<td>353</td>
</tr>
<tr>
<td>Zvika Marx, Ido Dagan, and Joachim Buhmann</td>
<td></td>
</tr>
<tr>
<td>Automatic Discovery of Subgoals in Reinforcement Learning Using Diverse Density</td>
<td>361</td>
</tr>
<tr>
<td>Amy McGovern and Andrew G. Barto</td>
<td></td>
</tr>
<tr>
<td>Some Greedy Algorithms for Sparse Nonlinear Regression</td>
<td>369</td>
</tr>
<tr>
<td>Prasanth B. Nair, Arindam Choudhury, and Andy J. Keane</td>
<td></td>
</tr>
<tr>
<td>Convergence Rates of the Voting Gibbs Classifier, with Application to Bayesian Feature Selection</td>
<td>377</td>
</tr>
<tr>
<td>Andrew Y. Ng and Michael I. Jordan</td>
<td></td>
</tr>
<tr>
<td>Ridge Regression Confidence Machine</td>
<td>385</td>
</tr>
<tr>
<td>Iliia Nouretdinov, Tom Melluish, and Volodya Vovk</td>
<td></td>
</tr>
<tr>
<td>Breeding Decision Trees Using Evolutionary Techniques</td>
<td>393</td>
</tr>
<tr>
<td>Athanasios Papagelis and Dimitris Kalles</td>
<td></td>
</tr>
<tr>
<td>Mixtures of Rectangles: Interpretable Soft Clustering</td>
<td>401</td>
</tr>
<tr>
<td>Dan Pelleg and Andrew Moore</td>
<td></td>
</tr>
<tr>
<td>Lyapunov-Constrained Action Sets for Reinforcement Learning</td>
<td>409</td>
</tr>
<tr>
<td>Theodore J. Perkins and Andrew G. Barto</td>
<td></td>
</tr>
<tr>
<td>Off-Policy Temporal-Difference Learning with Function Approximation</td>
<td>417</td>
</tr>
<tr>
<td>Doina Precup, Richard S. Sutton, and Sanjoy Dasgupta</td>
<td></td>
</tr>
<tr>
<td>Multiple Instance Regression</td>
<td>425</td>
</tr>
<tr>
<td>Soumya Ray and David Page</td>
<td></td>
</tr>
<tr>
<td>Comprehensible Interpretation of Relief’s Estimates</td>
<td>433</td>
</tr>
<tr>
<td>Marko Robnik-Šikonja and Igor Kononenko</td>
<td></td>
</tr>
<tr>
<td>Toward Optimal Active Learning through Sampling Estimation of Error Reduction</td>
<td>441</td>
</tr>
<tr>
<td>Nicholas Roy and Andrew McCallum</td>
<td></td>
</tr>
<tr>
<td>Using the Genetic Algorithm to Reduce the Size of a Nearest-Neighbor Classifier and to Select Relevant Attributes</td>
<td>449</td>
</tr>
<tr>
<td>Antonin Rozsypal and Miroslav Kubat</td>
<td></td>
</tr>
<tr>
<td>Repairing Faulty Mixture Models Using Density Estimation</td>
<td>457</td>
</tr>
<tr>
<td>Peter Sand and Andrew W. Moore</td>
<td></td>
</tr>
<tr>
<td>Application of Fuzzy Similarity-Based Fractal Dimensions to Characterize Medical Time Series</td>
<td>465</td>
</tr>
<tr>
<td>Manish Sarkar and Tze-Yun Leong</td>
<td></td>
</tr>
<tr>
<td>Average-Reward Reinforcement Learning for Variance Penalized Markov Decision Problems</td>
<td>473</td>
</tr>
<tr>
<td>Makoto Sato and Shigenobu Kobayashi</td>
<td></td>
</tr>
<tr>
<td>Incremental Maximization of Non-Instance-Averaging Utility Functions with Applications to Knowledge Discovery Problems</td>
<td>481</td>
</tr>
<tr>
<td>Tobias Scheffer and Stefan Wrobel</td>
<td></td>
</tr>
<tr>
<td>Discovering Communicable Scientific Knowledge from Spatio-Temporal Data</td>
<td>489</td>
</tr>
<tr>
<td>Mark Schwabacher and Pat Langley</td>
<td></td>
</tr>
</tbody>
</table>
Contents

Clustering Continuous Time Series .................................................. 497
Paola Sebastiani and Marco Ramoni

Boosting Neighborhood-Based Classifiers ........................................ 505
Marc Sebban, Richard Nock, and Stéphane Lallich

Unsupervised Sequence Segmentation by a Mixture of Switching Variable Memory Markov Sources ............... 513
Yevgeny Seldin, Gill Bejerano, and Naftali Tishby

Smoothed Bootstrap and Statistical Data Cloning for Classifier Evaluation ........................................... 521
Gregory Shakhnarovich, Ran El-Yaniv, and Yoram Baram

Learning to Generate Fast Signal Processing Implementations ................................................................. 529
Bryan Singer and Manuela Veloso

Scaling Reinforcement Learning toward RoboCup Soccer ........................................................................... 537
Peter Stone and Richard S. Sutton

Direct Policy Search Using Paired Statistical Tests .......................................................................................... 545
Malcolm J. A. Strens and Andrew W. Moore

A Multi-Agent, Policy-Gradient Approach to Network Routing ................................................................. 553
Nigel Tao, Jonathan Baxter, and Lex Weaver

Improving Probabilistic Grammatical Inference Core Algorithms with Post-Processing Techniques ............. 561
Franck Thollard

A Procedure for Unsupervised Lexicon Learning .......................................................................................... 569
Anand Venkataraman

Constrained K-means Clustering with Background Knowledge ......................................................................... 577
Kiri Wagstaff, Claire Cardie, Seth Rogers, and Stefan Schroedl

Reinforcement Learning in Dynamic Environments Using Instantiated Information ..................................... 585
Marco A. Wiering

Exploration Control in Reinforcement Learning Using Optimistic Model Selection ..................................... 593
Jeremy L. Wyatt

Feature Selection for High-Dimensional Genomic Microarray Data .............................................................. 601
Eric P. Xing, Michael I. Jordan, and Richard M. Karp

Obtaining Calibrated Probability Estimates from Decision Trees and Naive Bayesian Classifiers ...................... 609
Bianca Zadrozny and Charles Elkan

Learnability of Augmented Naive Bayes in Nominal Domains ....................................................................... 617
Huajie Zhang and Charles X. Ling

Some Sparse Approximation Bounds for Regression Problems ...................................................................... 624
Tong Zhang

Symmetry in Markov Decision Processes and its Implications for Single Agent and Multiagent Learning ............ 632
Martin Zinkevich and Tucker Balch

INVITED TALK ABSTRACTS

Successes, Failures, and New Directions in Natural Language Learning ...................................................... 641
Claire Cardie

A Guided Tour of Finite Mixture Models: From Pearson to the Web ............................................................... 642
Padhraic Smyth

Learning Structure from Sequences, with Applications in a Digital Library ..................................................... 643
Ian H. Witten

Author Index ...................................................................................................................................................... 644