

# Lecture Notes in Artificial Intelligence 3559

Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Peter Auer Ron Meir (Eds.)

# Learning Theory

18th Annual Conference on Learning Theory, COLT 2005  
Bertinoro, Italy, June 27-30, 2005  
Proceedings

## Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA  
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

## Volume Editors

Peter Auer  
University of Leoben  
Department of Mathematics and Information Technologies  
Franz-Josef-Strasse 18, 8700 Leoben, Austria  
E-mail: auer@unileoben.ac.at

Ron Meir  
Technion, Israel Institute of Technology  
Department of Electrical Engineering  
Haifa 3200, P.O. Box, Israel  
E-mail: rmeir@ee.technion.ac.il

Library of Congress Control Number: 2005927736

CR Subject Classification (1998): I.2.6, I.2.3, I.2, F.4.1, F.2, F.1.1

ISSN	0302-9743
ISBN-10	3-540-26556-2 Springer Berlin Heidelberg New York
ISBN-13	978-3-540-26556-6 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media  
springeronline.com

© Springer-Verlag Berlin Heidelberg 2005  
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India  
Printed on acid-free paper      SPIN: 11503415      06/3142      5 4 3 2 1 0

# Preface

This volume contains papers presented at the Eighteenth Annual Conference on Learning Theory (previously known as the Conference on Computational Learning Theory) held in Bertinoro, Italy from June 27 to 30, 2005.

The technical program contained 45 papers selected from 120 submissions, 3 open problems selected from among 5 contributed, and 2 invited lectures. The invited lectures were given by Sergiu Hart on “Uncoupled Dynamics and Nash Equilibrium”, and by Satinder Singh on “Rethinking State, Action, and Reward in Reinforcement Learning”. These papers were not included in this volume.

The Mark Fulk Award is presented annually for the best paper co-authored by a student. The student selected this year was Hadi Salmasian for the paper titled “The Spectral Method for General Mixture Models” co-authored with Ravindran Kannan and Santosh Vempala.

The number of papers submitted to COLT this year was exceptionally high. In addition to the classical COLT topics, we found an increase in the number of submissions related to novel classification scenarios such as ranking. This increase reflects a healthy shift towards more structured classification problems, which are becoming increasingly relevant to practitioners. The large number of quality submissions placed a heavy burden on the program committee of the conference: Shai Ben David (University of Waterloo), Avrim Blum (Carnegie Mellon University), Peter Bartlett (University of California, Berkeley), Nader Bshouty (Technion), Ran El-Yaniv (Technion), Yoav Freund (Columbia University), Ralf Herbrich (Microsoft Research, Cambridge), Marcus Hutter (IDSIA, Switzerland), Tony Jebara (Columbia University), Balazs Kegl (University of Montreal), Vladimir Koltchinskii (University New Mexico), Phil Long (Columbia University), Gábor Lugosi (Pompeu Fabra University), Shie Mannor (McGill University), Shahar Mendelson (Australian National University), Massimiliano Pontil (University College London), Daniel Reidenbach (University of Kaiserslautern), Dan Roth (University Illinois Urbana-Champaign), Michael Schmitt (Ruhr University Bochum), Rocco Servedio (Columbia University), Hans Ulrich Simon (Ruhr University Bochum), Volodya Vovk (Royal Holloway), Manfred Warmuth (University of California, Santa Cruz), Tong Zhang (IBM Research, Yorktown). We take this opportunity to thank all reviewers for the excellent job performed over a relatively short period of time. Some of them were even willing to review additional papers beyond those initially assigned. Their efforts have led to the selection of an exceptional set of papers, which ensured an outstanding conference. We would like to have mentioned the sub-reviewers who assisted the program committee in reaching their assessments, but unfortunately space constraints do not permit us to include this long list of names and we must simply ask them to accept our thanks anonymously.

We are particularly grateful to Nicolò Cesa-Bianchi and Claudio Gentile, the conference local chairs. Together they handled the conference publicity and all the local arrangements to ensure a successful event. We would also like to thank Microsoft for providing the software used in the program committee deliberations and Dori Peleg for creating the conference web site. Jyrki Kivinen assisted the organization of the conference in his role as head of the COLT Steering Committee.

This work was also supported in part by the IST Programme of the European Community, under the PASCAL Network of Excellence, IST-2002-506778.

Finally, we would like to thank the Machine Learning Journal, Google Inc., the Bertinoro International Center for Informatics, and the Università degli Studi di Milano for their sponsorship of the conference.

April, 2005

Peter Auer,  
Ron Meir  
Program Co-chairs COLT 2005

Sponsored by:



# Table of Contents

## Learning to Rank

Ranking and Scoring Using Empirical Risk Minimization <i>Stéphan Cléménçon, Gábor Lugosi, Nicolas Vayatis</i> .....	1
Learnability of Bipartite Ranking Functions <i>Shivani Agarwal, Dan Roth</i> .....	16
Stability and Generalization of Bipartite Ranking Algorithms <i>Shivani Agarwal, Partha Niyogi</i> .....	32
Loss Bounds for Online Category Ranking <i>Koby Crammer, Yoram Singer</i> .....	48

## Boosting

Margin-Based Ranking Meets Boosting in the Middle <i>Cynthia Rudin, Corinna Cortes, Mehryar Mohri, Robert E. Schapire</i> .....	63
Martingale Boosting <i>Philip M. Long, Rocco A. Servedio</i> .....	79
The Value of Agreement, a New Boosting Algorithm <i>Boaz Leskes</i> .....	95

## Unlabeled Data, Multiclass Classification

A PAC-Style Model for Learning from Labeled and Unlabeled Data <i>Maria-Florina Balcan, Avrim Blum</i> .....	111
Generalization Error Bounds Using Unlabeled Data <i>Matti Kääriäinen</i> .....	127
On the Consistency of Multiclass Classification Methods <i>Ambuj Tewari, Peter L. Bartlett</i> .....	143
Sensitive Error Correcting Output Codes <i>John Langford, Alina Beygelzimer</i> .....	158

## Online Learning I

### Data Dependent Concentration Bounds for Sequential Prediction Algorithms

*Tong Zhang* ..... 173

### The Weak Aggregating Algorithm and Weak Mixability

*Yuri Kalnishkan, Michael V. Vyugin* ..... 188

### Tracking the Best of Many Experts

*András György, Tamás Linder, Gábor Lugosi* ..... 204

### Improved Second-Order Bounds for Prediction with Expert Advice

*Nicolò Cesa-Bianchi, Yishay Mansour, Gilles Stoltz* ..... 217

## Online Learning II

### Competitive Collaborative Learning

*Baruch Awerbuch, Robert D. Kleinberg* ..... 233

### Analysis of Perceptron-Based Active Learning

*Sanjoy Dasgupta, Adam Tauman Kalai, Claire Monteleoni* ..... 249

### A New Perspective on an Old Perceptron Algorithm

*Shai Shalev-Shwartz, Yoram Singer* ..... 264

## Support Vector Machines

### Fast Rates for Support Vector Machines

*Ingo Steinwart, Clint Scovel* ..... 279

### Exponential Convergence Rates in Classification

*Vladimir Koltchinskii, Olexandra Beznosova* ..... 295

### General Polynomial Time Decomposition Algorithms

*Nikolas List, Hans Ulrich Simon* ..... 308

## Kernels and Embeddings

### Approximating a Gram Matrix for Improved Kernel-Based Learning

*Petros Drineas, Michael W. Mahoney* ..... 323

Learning Convex Combinations of Continuously Parameterized Basic Kernels <i>Andreas Argyriou, Charles A. Micchelli, Massimiliano Pontil</i> .....	338
On the Limitations of Embedding Methods <i>Shahar Mendelson</i> .....	353
Leaving the Span <i>Manfred K. Warmuth, S.V.N. Vishwanathan</i> .....	366
<b>Inductive Inference</b>	
Variations on U-Shaped Learning <i>Lorenzo Carlucci, Sanjay Jain, Efim Kinber, Frank Stephan</i> .....	382
Mind Change Efficient Learning <i>Wei Luo, Oliver Schulte</i> .....	398
On a Syntactic Characterization of Classification with a Mind Change Bound <i>Eric Martin, Arun Sharma</i> .....	413
<b>Unsupervised Learning</b>	
Ellipsoid Approximation Using Random Vectors <i>Shahar Mendelson, Alain Pajor</i> .....	429
The Spectral Method for General Mixture Models <i>Ravindran Kannan, Hadi Salmasian, Santosh Vempala</i> .....	444
On Spectral Learning of Mixtures of Distributions <i>Dimitris Achlioptas, Frank McSherry</i> .....	458
From Graphs to Manifolds – Weak and Strong Pointwise Consistency of Graph Laplacians <i>Matthias Hein, Jean-Yves Audibert, Ulrike von Luxburg</i> .....	470
Towards a Theoretical Foundation for Laplacian-Based Manifold Methods <i>Mikhail Belkin, Partha Niyogi</i> .....	486



## Generalization Bounds

Permutation Tests for Classification <i>Polina Golland, Feng Liang, Sayan Mukherjee, Dmitry Panchenko</i> . . .	501
Localized Upper and Lower Bounds for Some Estimation Problems <i>Tong Zhang</i> . . . . .	516
Improved Minimax Bounds on the Test and Training Distortion of Empirically Designed Vector Quantizers <i>András Antos</i> . . . . .	531
Rank, Trace-Norm and Max-Norm <i>Nathan Srebro, Adi Shraibman</i> . . . . .	545

## Query Learning, Attribute Efficiency, Compression Schemes

Learning a Hidden Hypergraph <i>Dana Angluin, Jiang Chen</i> . . . . .	561
On Attribute Efficient and Non-adaptive Learning of Parities and DNF Expressions <i>Vitaly Feldman</i> . . . . .	576
Unlabeled Compression Schemes for Maximum Classes <i>Dima Kuzmin, Manfred K. Warmuth</i> . . . . .	591

## Economics and Game Theory

Trading in Markovian Price Models <i>Sham M. Kakade, Michael Kearns</i> . . . . .	606
From External to Internal Regret <i>Avrim Blum, Yishay Mansour</i> . . . . .	621

## Separation Results for Learning Models

Separating Models of Learning from Correlated and Uncorrelated Data <i>Ariel Elbaz, Homin K. Lee, Rocco A. Servedio, Andrew Wan</i> . . . . .	637
Asymptotic Log-Loss of Prequential Maximum Likelihood Codes <i>Peter Grünwald, Steven de Rooij</i> . . . . .	652

Teaching Classes with High Teaching Dimension Using Few Examples <i>Frank J. Balbach</i> .....	668
---------------------------------------------------------------------------------------------------	-----

## Open Problems

Optimum Follow the Leader Algorithm <i>Dima Kuzmin, Manfred K. Warmuth</i> .....	684
-------------------------------------------------------------------------------------	-----

The Cross Validation Problem <i>John Langford</i> .....	687
------------------------------------------------------------	-----

Compute Inclusion Depth of a Pattern <i>Wei Luo</i> .....	689
--------------------------------------------------------------	-----

<b>Author Index</b> .....	691
---------------------------	-----