

Lecture Notes in Artificial Intelligence 4005

Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Gabor Lugosi Hans Ulrich Simon (Eds.)

Learning Theory

19th Annual Conference on Learning Theory, COLT 2006
Pittsburgh, PA, USA, June 22-25, 2006
Proceedings



Springer

Series Editors

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

Volume Editors

Gabor Lugosi
Pompeu Fabra University
Ramon Trias Fargas 25-27, 08005, Barcelona, Spain
E-mail: lugosi@upf.es

Hans Ulrich Simon
Ruhr University Bochum, Department of Mathematics
Building NA 1/73, 44780 Bochum, Germany
E-mail: simon@lmi.rub.de

Library of Congress Control Number: 2006927286

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

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN 0302-9743
ISBN-10 3-540-35294-5 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-35294-5 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

springer.com

© Springer-Verlag Berlin Heidelberg 2006
Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 11776420 06/3142 5 4 3 2 1 0

Preface

This volume contains papers presented at the 19th Annual Conference on Learning Theory (previously known as the Conference on Computational Learning Theory) held at the Carnegie Mellon University in Pittsburgh, USA, June 22–25, 2006.

The technical program contained 43 papers selected from 102 submissions, 2 open problems selected from among 4 contributed, and 3 invited lectures. The invited lectures were given by Luc Devroye on “Random Multivariate Search Trees,” by György Turán on “Learning and Logic,” and by Vladimir Vovk on “Predictions as Statements and Decisions.” The abstracts of these papers are included in this volume.

The Mark Fulk Award is presented annually for the best paper co-authored by a student. This year the Mark Fulk award was supplemented with three further awards funded by the *Machine Learning Journal*. We were therefore able to select four student papers for prizes. The students selected were Guillaume Lécué for the single-author paper “Optimal Oracle Inequality for Aggregation of Classifiers Under Low Noise Condition,” Homin K. Lee and Andrew Wan for the paper “DNF are Teachable in the Average Case” (co-authored by Rocco A. Servedio), Alexander A. Sherstov for the paper “Improved Lower Bounds for Learning the Intersections of Halfspaces” (co-authored by Adam R. Klivans), and Dávid Pál for the paper “A Sober Look at Clustering Stability” (co-authored by Ulrike von Luxburg and Shai Ben-David).

The trend of the previous two years of receiving more than 100 submissions continued. The selected papers cover a wide range of topics (including clustering, un- and semisupervised learning, statistical learning theory, regularized learning and kernel methods, query learning and teaching, inductive inference, learning algorithms and limitations on learning, online aggregation, online prediction and reinforcement learning). Online Prediction with 11 selected papers is particularly well represented. The large number of quality submissions placed a heavy burden on the Program Committee of the conference: Peter Auer (University of Leoben), Peter Bartlett (UC Berkeley), Léon Bottou (NEC Laboratories America), Nicolò Cesa-Bianchi (Università degli Studi di Milano), Koby Crammer (University of Pennsylvania), Yoav Freund (UC, San Diego), Claudio Gentile (Università dell’Insubria, Varese), Lisa Hellerstein (Polytechnic University, Brooklyn, NY), Ralf Herbrich (Microsoft Research Cambridge), Sham M. Kakade (Toyota Technology Institute), Ravi Kannan (Yale University), Jyrki Kivinen (University of Helsinki), Shie Mannor (McGill University), Shahar Mendelson (The Australian National University and Technion, I.I.T.), Massimiliano Pontil (University College London), Dan Roth (University of Illinois at Urbana-Champaign), Alex Smola (National ICT Australia), Ingo Steinwart (Los Alamos National Laboratory), Christino Tamon (Clarkson University), Santosh Vempala (MIT),

Ulrike von Luxburg (Fraunhofer IPSI), Vladimir Vovk (Royal Holloway), Thomas Zeugmann (Hokkaido University), and Tong Zhang (Yahoo). We are extremely grateful for their careful and thorough reviewing and for the detailed discussions that ensured the very high quality of the final program. We would like to have mentioned the sub-reviewers who assisted the Program Committee, 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 Avrim Blum, the conference Local Chair, and his administrative assistant Nicole Stenger. 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, Niko List for creating the conference website, and Sanjoy Dasgupta for maintaining the learningtheory.org website. Jyrki Kivinen assisted the organization of the conference in his role as head of the COLT Steering Committee. We would also like to thank the ICML organizers for ensuring a smooth co-location of the two conferences including an “overlap day,” June 25.

Finally, we would like to thank Google, IBM, the *Machine Learning Journal*, and National ICT Australia (Statistical Machine Learning Program) for their sponsorship of the conference.

April 2006

Gábor Lugosi
Hans Ulrich Simon
Program Co-chairs
COLT 2006

Sponsored by:



Table of Contents

Invited Presentations

Random Multivariate Search Trees <i>Luc Devroye</i>	1
On Learning and Logic <i>György Turán</i>	2
Predictions as Statements and Decisions <i>Vladimir Vovk</i>	4

Clustering, Un-, and Semisupervised Learning

A Sober Look at Clustering Stability <i>Shai Ben-David, Ulrike von Luxburg, Dávid Pál</i>	5
PAC Learning Axis-Aligned Mixtures of Gaussians with No Separation Assumption <i>Jon Feldman, Rocco A. Servedio, Ryan O'Donnell</i>	20
Stable Transductive Learning <i>Ran El-Yaniv, Dmitry Pechyony</i>	35
Uniform Convergence of Adaptive Graph-Based Regularization <i>Matthias Hein</i>	50

Statistical Learning Theory

The Rademacher Complexity of Linear Transformation Classes <i>Andreas Maurer</i>	65
Function Classes That Approximate the Bayes Risk <i>Ingo Steinwart, Don Hush, Clint Scovel</i>	79
Functional Classification with Margin Conditions <i>Magalie Fromont, Christine Tuleau</i>	94
Significance and Recovery of Block Structures in Binary Matrices with Noise <i>Xing Sun, Andrew Nobel</i>	109

Regularized Learning and Kernel Methods

Maximum Entropy Distribution Estimation with Generalized Regularization <i>Miroslav Dudík, Robert E. Schapire</i>	123
Unifying Divergence Minimization and Statistical Inference Via Convex Duality <i>Yasemin Altun, Alex Smola</i>	139
Mercer's Theorem, Feature Maps, and Smoothing <i>Ha Quang Minh, Partha Niyogi, Yuan Yao</i>	154
Learning Bounds for Support Vector Machines with Learned Kernels <i>Nathan Srebro, Shai Ben-David</i>	169

Query Learning and Teaching

On Optimal Learning Algorithms for Multiplicity Automata <i>Laurence Bisht, Nader H. Bshouty, Hanna Mazzawi</i>	184
Exact Learning Composed Classes with a Small Number of Mistakes <i>Nader H. Bshouty, Hanna Mazzawi</i>	199
DNF Are Teachable in the Average Case <i>Homin K. Lee, Rocco A. Servedio, Andrew Wan</i>	214
Teaching Randomized Learners <i>Frank J. Balbach, Thomas Zeugmann</i>	229

Inductive Inference

Memory-Limited U-Shaped Learning <i>Lorenzo Carlucci, John Case, Sanjay Jain, Frank Stephan</i>	244
On Learning Languages from Positive Data and a Limited Number of Short Counterexamples <i>Sanjay Jain, Efim Kinber</i>	259
Learning Rational Stochastic Languages <i>François Denis, Yann Esposito, Amaury Habrard</i>	274
Parent Assignment Is Hard for the MDL, AIC, and NML Costs <i>Mikko Koivisto</i>	289

Learning Algorithms and Limitations on Learning

Uniform-Distribution Learnability of Noisy Linear Threshold Functions
with Restricted Focus of Attention

Jeffrey C. Jackson 304

Discriminative Learning Can Succeed Where Generative Learning
Fails

Philip M. Long, Rocco A. Servedio 319

Improved Lower Bounds for Learning Intersections of Halfspaces

Adam R. Klivans, Alexander A. Sherstov 335

Efficient Learning Algorithms Yield Circuit Lower Bounds

Lance Fortnow, Adam R. Klivans 350

Online Aggregation

Optimal Oracle Inequality for Aggregation of Classifiers Under Low
Noise Condition

Guillaume Lecué 364

Aggregation and Sparsity Via ℓ_1 Penalized Least Squares

Florentina Bunea, Alexandre B. Tsybakov,
Marten H. Wegkamp 379

A Randomized Online Learning Algorithm for Better Variance
Control

Jean-Yves Audibert 392

Online Prediction and Reinforcement Learning I

Online Learning with Variable Stage Duration

Shie Mannor, Nahum Shimkin 408

Online Learning Meets Optimization in the Dual

Shai Shalev-Shwartz, Yoram Singer 423

Online Tracking of Linear Subspaces

Koby Crammer 438

Online Multitask Learning

Ofer Dekel, Philip M. Long, Yoram Singer 453

Online Prediction and Reinforcement Learning II

The Shortest Path Problem Under Partial Monitoring <i>András György, Tamás Linder, György Ottucsák</i>	468
Tracking the Best Hyperplane with a Simple Budget Perceptron <i>Nicolò Cesa-Bianchi, Claudio Gentile</i>	483
Logarithmic Regret Algorithms for Online Convex Optimization <i>Elad Hazan, Adam Kalai, Satyen Kale, Amit Agarwal</i>	499
Online Variance Minimization <i>Manfred K. Warmuth, Dima Kuzmin</i>	514

Online Prediction and Reinforcement Learning III

Online Learning with Constraints <i>Shie Mannor, John N. Tsitsiklis</i>	529
Continuous Experts and the Binning Algorithm <i>Jacob Abernethy, John Langford, Manfred K. Warmuth</i>	544
Competing with Wild Prediction Rules <i>Vladimir Vovk</i>	559
Learning Near-Optimal Policies with Bellman-Residual Minimization Based Fitted Policy Iteration and a Single Sample Path <i>András Antos, Csaba Szepesvári, Rémi Munos</i>	574

Other Approaches

Ranking with a P-Norm Push <i>Cynthia Rudin</i>	589
Subset Ranking Using Regression <i>David Cossock, Tong Zhang</i>	605
Active Sampling for Multiple Output Identification <i>Shai Fine, Yishay Mansour</i>	620
Improving Random Projections Using Marginal Information <i>Ping Li, Trevor J. Hastie, Kenneth W. Church</i>	635

Open Problems

Efficient Algorithms for General Active Learning <i>Claire Monteleoni</i>	650
Can Entropic Regularization Be Replaced by Squared Euclidean Distance Plus Additional Linear Constraints <i>Manfred K. Warmuth</i>	653
Author Index	655