

AISTATS 2005

Proceedings of the

Tenth International Workshop on
Artificial Intelligence and Statistics

Edited by

Robert Cowell

City University, London

Zoubin Ghahramani

University College London

January 6-8, 2005

The Savannah Hotel, Barbados

Published by *The Society for Artificial Intelligence and Statistics*

The web site for the AISTATS 2005 workshop may be found at

<http://www.gatsby.ucl.ac.uk/aistats/>

from which individual papers in these proceedings may be downloaded, together with a bibtex file with details all of the workshop papers. Citations of articles that appear in the proceedings should adhere to the format of the following example:

References

- [1] Shivani Agarwal, Sarel Har-Peled, and Dan Roth. A uniform convergence bound for the area under the ROC curve. In Robert G. Cowell and Zoubin Ghahramani, editors, *Proceedings of the Tenth International Workshop on Artificial Intelligence and Statistics, Jan 6-8, 2005, Savannah Hotel, Barbados*, pages 1–8. Society for Artificial Intelligence and Statistics, 2005. (Available electronically at <http://www.gatsby.ucl.ac.uk/aistats/>)

Copyright Notice:

©2005 by The Society for Artificial Intelligence and Statistics

This volume is published electronically by The Society for Artificial Intelligence and Statistics. The copyright of each paper in this volume resides with its authors. These papers appear in these electronic conference proceedings by the authors' permission being implicitly granted by their knowledge that they would be published electronically by The Society for Artificial Intelligence and Statistics on the AISTATS 2005 workshop web site according to the instructions on the AISTATS 2005 workshop *Call For Papers* web page.

ISBN 0-9727358-1-X

Contents

<i>Preface</i>	v
<i>Acknowledgements</i>	vi
A Uniform Convergence Bound for the Area Under the ROC Curve <i>Shivani Agarwal, Sariel Har-Peled and Dan Roth</i>	1
On the Path to an Ideal ROC Curve: Considering Cost Asymmetry in Learning Classifiers <i>Francis Bach, David Heckerman and Eric Horvitz</i>	9
On Manifold Regularization <i>Misha Belkin, Partha Niyogi and Vikas Sindhwani</i>	17
Distributed Latent Variable Models of Lexical Co-occurrences <i>John Blitzer, Amir Globerson and Fernando Pereira</i>	25
On Contrastive Divergence Learning <i>Miguel Á. Carreira-Perpiñán and Geoffrey Hinton</i>	33
OoBN for Forensic Identification through Searching a DNA profiles' Database <i>David Cavallini and Fabio Corradi</i>	41
Active Learning for Parzen Window Classifier <i>Olivier Chapelle</i>	49
Semi-Supervised Classification by Low Density Separation <i>Olivier Chapelle and Alexander Zien</i>	57
Learning spectral graph segmentation <i>Timothée Cour, Nicolas Gogin and Jianbo Shi</i>	65
A Graphical Model for Simultaneous Partitioning and Labeling <i>Philip J. Cowans and Martin Szummer</i>	73
Restructuring Dynamic Causal Systems in Equilibrium <i>Denver Dash</i>	81
Probability and Statistics in the Law <i>Philip Dawid</i>	89
Efficient Non-Parametric Function Induction in Semi-Supervised Learning <i>Olivier Delalleau, Yoshua Bengio and Nicolas Le Roux</i>	96

Structured Variational Inference Procedures and their Realizations	104
<i>Dan Geiger and Chris Meek</i>	
Kernel Constrained Covariance for Dependence Measurement	112
<i>Arthur Gretton, Alexander Smola, Olivier Bousquet, Ralf Herbrich, Andrei Belytski, Mark Augath, Yusuke Murayama, Jon Pauls, Bernhard Schölkopf and Nikos Logothetis</i>	
Semisupervised alignment of manifolds	120
<i>Jihun Ham, Daniel Lee and Lawrence Saul</i>	
Learning Causally Linked Markov Random Fields	128
<i>Geoffrey Hinton, Simon Osindero and Kejie Bao</i>	
Hilbertian Metrics and Positive Definite Kernels on Probability Measures	136
<i>Matthias Hein and Olivier Bousquet</i>	
Fast Non-Parametric Bayesian Inference on Infinite Trees	144
<i>Marcus Hutter</i>	
Restricted concentration models – graphical Gaussian models with concentration parameters restricted to being equal	152
<i>Søren Højsgaard and Steffen Lauritzen</i>	
Fast maximum a-posteriori inference on Monte Carlo state spaces	158
<i>Mike Klaas, Dustin Lang and Nando de Freitas</i>	
Generative Model for Layers of Appearance and Deformation	166
<i>Anitha Kannan, Nebojsa Jojic and Brendan Frey</i>	
Toward Question-Asking Machines: The Logic of Questions and the Inquiry Calculus	174
<i>Kevin Knuth</i>	
Convergent tree-reweighted message passing for energy minimization	182
<i>Vladimir Kolmogorov</i>	
Instrumental variable tests for Directed Acyclic Graph Models	190
<i>Manabu Kuroki and Zhihong Cai</i>	
Estimating Class Membership Probabilities using Classifier Learners	198
<i>John Langford and Bianca Zadrozny</i>	
Loss Functions for Discriminative Training of Energy-Based Models	206
<i>Yann LeCun and Fu Jie Huang</i>	
Probabilistic Soft Interventions in Conditional Gaussian Networks	214
<i>Florian Markowetz, Steffen Grossmann, and Rainer Spang</i>	
Unsupervised Learning with Non-Ignorable Missing Data	222
<i>Benjamin M. Marlin, Sam T. Roweis and Richard S. Zemel</i>	

Regularized spectral learning	230
<i>Marina Meilă, Susan Shortreed and Liang Xu</i>	
Approximate Inference for Infinite Contingent Bayesian Networks	238
<i>Brian Milch, Bhaskara Marthi, David Sontag, Stuart Russell, Daniel L. Ong and Andrey Kolobov</i>	
Hierarchical Probabilistic Neural Network Language Model	246
<i>Frederic Morin and Yoshua Bengio</i>	
Greedy Spectral Embedding	253
<i>Marie Ouimet and Yoshua Bengio</i>	
FastMap, MetricMap, and Landmark MDS are all Nystrom Algorithms	261
<i>John Platt</i>	
Bayesian Conditional Random Fields	269
<i>Yuan Qi, Martin Szummer and Tom Minka</i>	
Poisson-Networks: A Model for Structured Poisson Processes	277
<i>Shyamsundar Rajaram, Graepel Thore and Ralf Herbrich</i>	
Deformable Spectrograms	285
<i>Manuel Reyes-Gomez, Nebojsa Jojic and Daniel Ellis</i>	
Variational Speech Separation of More Sources than Mixtures	293
<i>Steven J. Rennie, Kannan Achan, Brendan J. Frey and Parham Aarabi</i>	
Learning Bayesian Network Models from Incomplete Data using Importance Sampling	301
<i>Carsten Riggelsen and Ad Feelders</i>	
On the Behavior of MDL Denoising	309
<i>Teemu Roos, Petri Myllymäki and Henry Tirri</i>	
Focused Inference	317
<i>Romer Rosales and Tommi Jaakkola</i>	
Kernel Methods for Missing Variables	325
<i>Alex J. Smola, S. V. N. Vishwanathan and Thomas Hofmann</i>	
Semiparametric latent factor models	333
<i>Yee Whye Teh, Matthias Seeger and Michael I. Jordan</i>	
Efficient Gradient Computation for Conditional Gaussian Models	341
<i>Bo Thiesson and Chris Meek</i>	
Very Large SVM Training using Core Vector Machines	349
<i>Ivor Tsang, James Kwok and Pak-Ming Cheung</i>	

Streaming Feature Selection using IIC	357
<i>Lyle H. Ungar, Jing Zhou, Dean P. Foster and Bob A. Stine</i>	
Defensive Forecasting	365
<i>Vladimir Vovk, Akimichi Takemura and Glenn Shafer</i>	
Inadequacy of interval estimates corresponding to variational Bayesian approximations	373
<i>Bo Wang and D. M. Titterton</i>	
Nonlinear Dimensionality Reduction by Semidefinite Programming and Kernel Matrix Factorization	381
<i>Kilian Weinberger, Benjamin Packer, and Lawrence K. Saul</i>	
An Expectation Maximization Algorithm for Inferring Offset-Normal Shape Distributions	389
<i>Max Welling</i>	
Learning in Markov Random Fields with Contrastive Free Energies	397
<i>Max Welling and Charles Sutton</i>	
Robust Higher Order Statistics	405
<i>Max Welling</i>	
Online (and Offline) on an Even Tighter Budget	413
<i>Jason Weston, Antoine Bordes and Leon Bottou</i>	
Approximations with Reweighted Generalized Belief Propagation	421
<i>Wim Wiegand</i>	
Recursive Autonomy Identification for Bayesian Network Structure Learning	429
<i>Raanan Yehezkel and Boaz Lerner</i>	
Dirichlet Enhanced Latent Semantic Analysis	437
<i>Kai Yu, Shipeng Yu and Volker Tresp</i>	
Gaussian Quadrature Based Expectation Propagation	445
<i>Onno Zoeter and Tom Heskes</i>	

Preface

The Society for Artificial Intelligence and Statistics (SAIAS) is dedicated to facilitating interactions between researchers in AI and Statistics. The primary responsibility of the society is to organize the biennial International Workshops on Artificial Intelligence and Statistics, as well as maintain the AI-Stats home page and mailing list on the Internet. The tenth such meeting took place in January 2005 in Barbados, a new venue for this conference and the first time it had been held outside of the United States of America. Details about the conference may be found at <http://www.gatsby.ucl.ac.uk/aistats/>.

Papers from a large number of different areas at the interface of statistics and AI were presented at the workshop. In addition to traditional areas of strength at AISTATS, such as probabilistic graphical models and approximate inference algorithms, the workshop also benefitted from high quality presentations in a broader set of new topics, such as semi-supervised learning, kernel methods, spectral learning, dimensionality reduction, and learning theory, to name a few. This diversity contributed to a strong and stimulating programme.

A novel feature of this workshop were prizes awarded to students for Best Student Papers. Three awards were made to Francis Bach, Philip J. Cowans and Kilian Weinberger. This was made possible by a donation from the NITCA at the Australian National University.

There were approximately 150 submissions. Almost every paper was assigned three reviewers, other than the program chairs. On the basis of these reviews, 21 papers were selected for presentation in the plenary session and 36 were selected for the poster sessions, based on their interest and relevance to the conference and on their originality and clarity of exposition. In deciding on which papers to accept we drew heavily on the reviews of the program committee. The standard was rigorous and impartial, and we note in passing that several members of the program committee had papers rejected.

The United States was the country with the most submissions (57) with Canada (17) and the United Kingdom (14) being the next largest contributing countries. Most of the remaining submissions came from Europe, with submissions from Belgium, Denmark, Finland, France, Germany, Ireland, Italy, Slovakia, Slovenia, Spain, Switzerland and The Netherlands. Papers submitted from outside of Europe and North America originated from Algeria, Australia, Brazil, Chile, Hong Kong, Iran, Israel, Japan, Malaysia and Russia. This range of countries emphasizes the truly international character of the conference. It is worth noting that several papers had their co-authors based in different countries.

Equal acceptance criteria were used for all submissions, and our decision of how each paper was presented was aimed at creating a varied programme rather than drawing a distinction between poster papers and plenary papers. Accordingly, in these proceedings we have ordered all of the papers alphabetically by the lead author's name. The conference web page may be consulted to see how individual papers were presented.

Robert Cowell and Zoubin Ghahramani, Program Chairs

Acknowledgements

The volume would not exist without the help of many people: the authors, the reviewers, the participants, the sponsors, and the Society for Artificial Intelligence and Statistics.

We should like to thank our invited speakers: Craig Boutilier, of the Department of Computer Science University of Toronto, who talked about “Regret-based Methods for Decision Making and Preference Elicitation”; Nir Friedman, of the School of Computer Science and Engineering, Hebrew University, who described “Probabilistic Models for Identifying Regulation Networks: From Qualitative to Quantitative Models”; Tom Minka, of Microsoft Research (Cambridge, UK), who spoke about “Some Intuitions About Message Passing”, and Steffen Lauritzen, of the Department of Statistics at the University of Oxford, who talked about “Identification and Separation of DNA Mixtures using Peak Area Information”. Regrettably, our fifth invited speaker, Tommi Jaakkola of the MIT Computer Science and Artificial Intelligence Laboratory was unable to attend the meeting.

We are also very grateful to all authors: authors of plenary papers, poster papers and those authors whose papers, though submitted, were not included in the conference program, which resulted in work of such a high standard.

We would like to thank our sponsors, NITCA at the Australian National University for the prize money for the Best Student Paper awards, and also Microsoft Research and the European PASCAL initiative for donations towards the running costs of the conference.

We would also like to thank Mrs. Faye Wharton-Parris of Premier Events, who took care among other things of local arrangements such as collecting registration fees, organizing accommodation and transportation between hotels, multimedia hire and wireless networking within the conference room. Her company’s professional services ensured that the day-to-day running of the conference went smoothly.

Special thanks also to Katherine Heller at the Gatsby Unit, who designed and maintained the website, and helped with numerous aspects of the conference organization; to Daryl Pregibon at Google who handled finances for the deposits; and to Chani Johnson and others of Microsoft Research, who maintained the Conference Management Toolkit with which the reviewing process was coordinated.

Finally, we would especially like to thank the members of the program committee and a few external reviewers for agreeing to give up their time to review papers. The thorough and conscientious reviews of their allocated papers, which they carried out in a short period of time, ensured a high standard of presentations for the conference. The members of the program committee are listed on the following page.

Program Committee

Yasemin Altun, Brown University
Hagai Attias, Golden Metallic, Inc.
Francis Bach, University of California, Berkeley
Matthew Beal, SUNY Buffalo
Yoshua Bengio, University of Montreal
Christopher Bishop, Microsoft Research
David Blei, University of California, Berkeley
Olivier Bousquet, Pertinence
Wray Buntine, HIIT
Olivier Chapelle, Max Planck Institute
Guido Consonni, University of Pavia
Greg Cooper, University of Pittsburgh
Adrian Corduneanu, MIT
Denver Dash, Intel Research
Phil Dawid, University College London
Nando de Freitas, University of British Columbia
Vanessa Didelez, University College London
Michael Duff, Trinity College
Nir Friedman, Hebrew University of Jerusalem
Dan Geiger, Technion - Israel Inst. of Technology
Lise Getoor, University of Maryland
Paolo Giudici, University of Pavia
Tom Griffiths, Stanford University
Peter Grunwald, CWI, Netherlands
Carlos Guestrin, Carnegie Mellon University
David Heckerman, Microsoft Research
Ralf Herbrich, Microsoft Research
Tom Heskes, University of Nijmegen
Thomas Hofmann, Brown University
Chris Holmes, University of Oxford
Tommi Jaakkola, MIT
Nebojsa Jojic, Microsoft Research
Anitha Kannan, University of Toronto
Uffe Kjærulff, Aalborg University
John Lafferty, Carnegie Mellon University
John Langford, TTI - Chicago
Neil Lawrence, University of Sheffield
Guy Lebanon, Carnegie Mellon University
Juan Lin, Rutgers University
David Madigan, Rutgers University
Chris Meek, Microsoft Research
Marina Meilă, University of Washington
Tom Minka, Microsoft Research
Quaid Morris, University of Toronto
Iain Murray, Gatsby Unit
Kevin Murphy, MIT
Petri Myllymäki, University of Helsinki
Nuria Oliver, Microsoft Research
Manfred Opper, University of Southampton
Fernando Pérez-Cruz, Gatsby Unit
John Platt, Microsoft Research
Thomas Richardson, University of Washington
Imre Risi Kondor, Columbia University
Michal Rosen-Zvi, University of California, Irvine
Sam Roweis, University of Toronto
Lawrence Saul, University of Pennsylvania
Dale Schuurmans, University of Alberta
Paola Sebastiani, Boston University
Matthias Seeger, University of California, Berkeley
Prakash Shenoy, University of Kansas
Yoram Singer, Hebrew University of Jerusalem
Jim Smith, University of Warwick
Alex Smola, Australian National University
Bo Thiesson, Microsoft Research
Henry Tirri, University of Helsinki
Volker Tresp, Siemens Research
Linda van der Gaag, University of Utrecht
Vladimir Vovk, Royal Holloway University of London
Martin Wainwright, University of California, Berkeley
Max Welling, University of California, Irvine
Jason Weston, NEC Labs America
Joe Whittaker, Lancaster University
Yee Whye Teh, University of California, Berkeley
Chris Williams, University of Edinburgh
John Winn, Microsoft Research
Eric Xing, Carnegie Mellon University
Xiaojin Zhu, Carnegie Mellon University