Preface

This volume contains the papers accepted to the *27th International Conference on Machine Learning* (ICML 2010). ICML is the annual conference of the International Machine Learning Society (IMLS) and provides a venue for the presentation and discussion of current research in the field of machine learning. These proceedings can also be found online at `http://www.machinelearning.org`.

ICML 2010 was held June 21 – June 25 in Haifa, Israel, making it the first ICML to be located in Israel. The conference was immediately followed by the 23rd Annual Conference on Computational Learning Theory (COLT 2010), both conferences sharing a joint workshop day and a leisure day with an organized tour to Jerusalem.

We received 594 manuscripts for review, which roughly matches the number of submissions in the previous year. Of those submissions, eventually 152 papers were accepted using a thorough two-phase reviewing process with a program committee consisting of 46 Area Chairs and 517 Reviewers. The reviewing process was double blind, meaning that Reviewers could not see the author names. Area Chairs and Program Chairs, however, were aware of the author identities. All accepted papers received both an oral and a poster presentation at the conference. These proceedings are evidence of the resulting high-quality and exciting program.

In the call for papers, we specified 64 area keywords. These were used for selecting 42 Area Chairs in a strictly area-driven process. We tried to account not only for area popularity but also for breadth of coverage, so that we had suitable chairs even for areas that had low submission rates in the past. Four additional Area Chairs were nominated after the submission deadline to better match the actual area distribution of the submissions.

Each of the 46 Area Chairs was responsible for 2 – 4 areas, and most areas were covered by more than one chair. Authors could mark their submissions with up to six areas (one primary and five secondary), and Area Chairs were asked to bid on papers in their areas. This information was then used to automatically generate candidate assignments of papers to Area Chairs. The final assignment was made by the Program Chairs. Deliberately loosening the constraint of a balanced work-load throughout the reviewing process, together with the above-mentioned redundancy in the area-to-area-chair mapping gave us more flexibility to assign papers to appropriate chairs and diminished the risk of the proverbial "odd paper" in the batch. Eventually, each Area Chair was responsible for 10 – 15 papers.

Initially, more than 600 Reviewers were nominated by the Area Chairs, and about 80% of them accepted our invitations. More Reviewers were later added by the Area Chairs throughout the reviewing process.

Papers were reviewed in two phases. For Phase I, Reviewers bid on papers and each Reviewer was assigned 2–3 papers based on these bids. At the end of Phase I, every paper had 2 reviews. These two reviews were the basis of the Author Rebuttal. Authors could respond to

the Reviewers, but could also make confidential comments to the Area Chair. In several cases, the confidential comments identified shortcomings and problems with the review assignments of Phase I, which the Area Chair could act upon for the assignments in Phase II.

If, on the basis of the two reviews from Phase I and the Author Rebuttal, the Area Chair identified substantial problems that left no hope for eventually accepting the paper, the Area Chair could reject a paper at this point. This early reject option was used for 78 submissions (including a few submissions that were withdrawn by the authors after seeing the first two reviews).

In Phase II, the Area Chairs were able to manually assign additional Reviewers to each of their papers. All papers that were not early rejects received at least one further review, but the Area Chair could assign as many additional Reviewers as necessary to make an confident decision. All papers that made it into Phase II received at least 3 reviews in total, some receiving up to 6 reviews. Because of time constraints, there was no Author Rebuttal for Phase II reviews.

During and after Phase II, Reviewers who had already submitted their review were able to discuss the paper. Based on all reviews, the Author Rebuttal of Phase-I reviews, the discussion, and possibly reading the paper themselves, the Area Chair made a recommendation for acceptance/rejection as argued in their meta-review. Each Area Chair also had a "Buddy" Area Chair, who had read-access to all of their reviewing data and could be asked for a second opinion on difficult papers.

All recommendations were reviewed by the Program Chairs. In several cases the Program Chairs assigned additional Reviewers or Area Chairs to come to confident conclusions on all papers.

One guiding principle in designing this two-phase review process was to give all involved parties some influence in the most crucial decisions within the reviewing process, namely the assignment of Reviewers and Area Chairs to papers.

- Authors were able to provide keywords, through which they could target particular Area Chairs responsible for these areas. Furthermore, they were able to point out shortcomings of the first two reviews in their Rebuttal, influencing the Area Chair's choice of Reviewers for Phase II.
- Reviewers were able to pick their assignments for Phase I through bidding. In this phase, virtually all Reviewers received only papers they had bid on.
- Area Chairs were also able to bid on papers, but were restricted to their stated and published subject areas. Furthermore, Area Chairs were free in their choice of Reviewers in Phase II, up to whatever number of reviews they deemed necessary.

We believe that the joint expertise of all parties provided good assignments of Reviewers to papers, and hope that nobody felt that their voice was not heard.

The *Invited Application Track* was a new introduction this year. To bring expertise and interesting research questions from related fields to ICML, we invited 7 papers from interesting application areas to the conference. These papers were not reviewed for ICML, but have already been reviewed and published at recent conferences in these application areas. The

papers to invite were identified by a committee, particularly focusing on papers that opened interesting new questions and applications for machine learning research. The committee was asked to select one paper in each of the following subject areas: Robotics, Computer Systems, Computer Vision, Natural Language Processing, Computer Music and Entertainment, Games and Planning, and Computational Biology. The Invited Application Papers received a regular slot in a parallel session, as well as the opportunity to publish a write-up of their work in the proceedings. We hope that this Invited Applications Track leads to an interesting exchange of ideas between fields and further extends the importance of machine learning approaches.

We are excited to have won Tom Mitchell, Nir Friedman, and Duncan Watts for the plenary invited talks, as well as Nobel Prize winner Robert Aumann for a special guest lecture.

The conference program was flanked by an exciting tutorial and workshop program, starting off with 8 tutorials on the first day of the conference, and finishing with 9 workshops on its last day.

The Best Paper Committee selected the "Best Paper", the "10-Year Best Paper Award", the "Best Student Paper", the "Best Application Paper", and their runner-ups. We thank Springer, publisher of *Machine Learning*, for sponsoring our student paper awards.

Many people worked together in organizing ICML 2010 and ensuring its success. All of them are named in the following sections of the proceedings, including the members of the Invited Applications Committee, the members of the Best Paper Awards Committee, the officers and board members of the IMLS, the Area Chairs, and the Reviewers. A big thank you goes to Stefan Wrobel for providing his experience and sound judgment in difficult decisions, to John Langford for organizing an exciting workshop program, and to Ben Taskar for an excellent selection of tutorials. To Alan Fern we owe a great list of sponsors, whom we thank as well. Jennifer Dy made sure travel scholarships were fairly awarded among the student applicants, and Hal Daumé III expertly produced these proceedings and the program booklet. Last but not least we would like to thank all authors who submitted their work to this conference.

Our special thanks go to the Local Chairs, Shai Fine and Dan Pelleg, for handling the budget and organizing everything from conference rooms, to hotels, to meals, to pins for putting up the posters. Noam Slonim organized the many volunteers, Oded Margalit handled visa issues, Efrat Maimon and Ettie Gilead took care of publicity and the Web-Site, and Vered Aharon provided local and technical assistance. We are grateful to IBM Research in Haifa for providing the many resources that went into the local organization of ICML 2010.

Many more people spent many hours on ICML 2010 – many more than we could possibly acknowledge in this short preface. But we hope that they found their involvement just as interesting and fun as we did.

Thorsten Joachims

Johannes Fürnkranz

ICML 2010 Program Chairs

Dedicated to Sam Roweis
(27.4.1972 - 12.01.2010)

ICML 2010 Organization

General Chair

Stefan Wrobel, *Fraunhofer IAIS & Universität Bonn, Germany*

Program Co-Chairs

Johannes Fürnkranz, *Technische Universität Darmstadt, Germany*
Thorsten Joachims, *Cornell University, U.S.A.*

Local Arrangements Co-Chairs

Shai Fine, *IBM Research Haifa, Israel*
Dan Pelleg, *IBM Research Haifa, Israel*

Workshop Chair

John Langford, *Yahoo! Research, U.S.A.*

Tutorials Chair

Ben Taskar, *University of Pennsylvania, U.S.A.*

Sponsorship Chair

Alan Fern, *Oregon State University, U.S.A.*

Scholarship Chair

Jennifer Dy, *Northeastern University, U.S.A.*

Publications Chair

Hal Daumé III, *University of Maryland, U.S.A.*

Volunteers Chair

Noam Slonim, *IBM Research Haifa, Israel*

Publicity and Production Chair

Ettie Gilead, *IBM Research Haifa, Israel*

Technical Assistance and Local Chair

Vered Aharaon, *IBM Research Haifa, Israel*

Webmaster

Efrat Maimon, *IBM Research Haifa, Israel*

Logo and Cover Design

Tamar Dekel, *IBM Research Haifa, Israel*

Governing Body

The International Machine Learning Society, http://www.machinelearning.org

Board of the International Machine Learning Society

President, Raymond J. Mooney, *University of Texas at Austin, U.S.A.*
Secretary, Jeff Schneider, *Carnegie Mellon University, U.S.A.*
Treasurer, Jennifer Neville, *Purdue University, U.S.A.*

Member, David Blei
Member, Corinna Cortes
Member, Luc De Raedt
Member, Tom Dietterich
Member, Lise Getoor
Member, Zoubin Ghahramani
Member, Rob Holte
Member, Thorsten Joachims
Member, Andrew McCallum
Member, Tom Mitchell
Member, Foster Provost
Member, Sam Roweis
Member, Bernhard Schölkopf
Member, Dale Schuurmans
Member, John Shawe-Tayler
Member, Stefan Wrobel

Area Chairs

Francis Bach	Mikhail Belkin	Mikhail Bilenko	David Blei
Hendrik Blockeel	Henrik Boström	Léon Bottou	Rich Caruana
Corinna Cortes	Padraig Cunningham	Sanjoy Dasgupta	Luc De Raedt
Inderjit Dhillon	Kurt Driessens	Pierre Dupont	Alan Fern
Peter Flach	Eibe Frank	Thomas Gärtner	Joao Gama
Katherine Heller	Thomas Hofmann	Eyke Hüllermeier	Sathiya Keerthi
Eamonn Keogh	Kristian Kersting	John Langford	Jure Leskovec
Phil Long	Shie Mannor	Elena Marchiori	Katharina Morik
Filip Radlinski	Dan Roth	Volker Roth	Ashutosh Saxena
Tobias Scheffer	Alex Smola	Nathan Srebro	Amos Storkey
Csaba Szepesvàri	Naftali Tishby	Kai Yu	Tong Zhang
Xiaojin Zhu	Martin Zinkevich		

Best Paper Committee

Peter Auer	Carla Brodley	Johannes Fürnkranz
Zoubin Ghahramani	Tony Jebara	Thorsten Joachims
Kristian Kersting	Tobias Scheffer	Rich Sutton

Invited Applications Track Committee

Wolfram Burgard	Pedro Felzenszwalb	Johannes Fürnkranz
Lillian Lee	Christina Leslie	Thorsten Joachims
Jonathan Schaeffer	Yoav Tock	Gerhard Widmer

Reviewers

Pieter Abbeel
Sreangsu Acharyya
Ryan Adams
Pieter Adriaans
Alekh Agarwal
Shivani Agarwal
David Aha
Amr Ahmed
Fabio Aiolli
Edo Airoldi
Klaus-Dieter Althoff
Yasemin Altun
Massih-Reza Amini
David Andrzejewski
Annalisa Appice
Cedric Archambeau
Andreas Argyriou
Ira Assent
Jean-Yves Audibert
Anne Auger
Jose Avila
Suhrid Balakrishnan
Arindam Banerjee
Wolfgang Banzhaf
Aharon Bar Hillel
Ziv Bar-Joseph
Sugato Basu
Roberto Battiti
Ron Bekkerman
Kedar Bellare
Samy Bengio
Paul Bennett
Tanya Berger-Wolf
Ralph Bergmann
Alina Beygelzimer
Chiranjib Bhattacharyya
Jinbo Bi
Isabelle Bichindaritz
Albert Bifet
Mario Boley
Edwin Bonilla
Gianluca Bontempi
Antoine Bordes
Christian Borgelt
Karsten Borgwardt
Guillaume Bouchard
Y-Lan Boureau
Jordan Boyd-Graber
Mikio Braun
Pavel Brazdil
Ulf Brefeld
Emma Brunskill
Joachim Buhmann
Florentina Bunea
Chris Burges
Alberto Busetto
Charles Cadieu
Tiberio Caetano
Stephane Canu
Lawrence Carin
Xavier Carreras
Ben Carterette
Andre Carvalho
Francisco Casacuberta
Kian Ming Chai
Deepay Chakrabarti
Jonathan Chang
Kai-Wei Chang
Ming-Wei Chang
Kamalika Chaudhuri
Nitesh Chawla
Beechung Chen
Weiwei Cheng
Haibin Cheng
Yun Chi
Silvia Chiappa
Sumit Chopra
Wei Chu
Alexander Clark
Stephan Clemencon
Ira Cohen
Kevyn Collins-Thompson
Ronan Collobert
Fabrizio Costa
Timothee Cour
Koby Crammer
W. Bruce Croft
Tom Croonenborghs
Lehel Csato
John Cunningham
James Cussens
Marco Cuturi
Alfredo Cuzzocrea
Arnak Dalalyan
Hal Daumé III
Mark Davenport
Jesse Davis
Nando De Freitas
Ramon De Mantaras
Boris Defourny
Olivier Dellaleau
Krzysztof Dembczyński
Francois Denis
Marie desJardins
Dotan Di Castro
Fernando Diaz
Laura Dietz
Federico Divina
Santosh Divvala
Chuong Do
Finale Doshi-Velez
Doug Downey
Philip Drewe
Chris Drummond
Miroslav Dudik
Alain Dutech
Jennifer Dy
Sašo Džeroski
Jacob Eisenstein
Pradheep Elango
Tina Eliassi-Rad
Charles Elkan
Tapio Elomaa
Barbara Engelhardt
Damien Ernst
Amir Farahmand
Ad Feelders
Carlos Ferreira
Cesar Ferri
Mario Figueiredo
Thomas Finley
George Forman
Emily Fox
Vojtech Franc
Damien Francois
Jordan Frank
Peter Frazier
Kenji Fukumizu
Mohamed Gaber
Gemma Garriga
Jan Gasthaus
Ricard Gavalda
Claudio Gentile
Pierre Geurts
Rayid Ghani
Mohammad Ghavamzadeh
Joydeep Ghosh
Ran Gilad-Bachrach
Mark Girolami
Andrew Goldberg
Sreenivas Gollapudi
Joseph Gonzalez
Stephen Gould
David Grangier
Derek Greene
Arthur Gretton
Perry Groot
Vincent Guigue
Yunsong Guo
Gholamreza Haffari
Patrick Haffner
Peter Haider
Barbara Hammer
Steve Hanneke
Zaid Harchaoui
Elad Hazan
Tamir Hazan
Xiaofei He
Matthias Hein
Philipp Hennig
Mark Herbster
Daniel Hernandez Lobato
Jose Hernandez-Orallo
Tom Heskes
Melanie Hilario
Vasant Honavar
Tamas Horvath
Daniel Hsu
Rebecca Hutchinson
Christian Igel
Elena Ikonomovska
Laurent Jacob
Manfred Jaeger
Prateek Jain
Rodolphe Jenatton
Chad Jenkins
David Jensen
Shuiwang Ji
Rong Jin
Michael Johanson
Ulf Johansson
Rosie Jones
Alipio Jorge
Felix Jungermann
Ata Kaban
Leslie Kaelbling
Satyen Kale
Alexandros Kalousis
Nikolaos Karampatziakis
Alexandros Karatzoglou
Hisashi Kashima
Samuel Kaski
Roni Khardon
Jörg-Uwe Kietz
Seyoung Kim
Angelika Kimmig
Gunnar Klau
Alexandre Klementiev

David Knowles
Igor Kononenko
Jussi Kujala
Sanjiv Kumar
Simon Lacoste-Julien
Hugo Larochelle
Quoc Le
Ann Lee
Christina Leslie
Ping Li
Yuanqing Lin
Heitor Lopes
Rich Maclin
Gideon Mann
Shaul Markovitch
Julian McAuley
Francisco Melo
Lilyana Mihalkova
Mirjam Minor
Stefania Montani
Joao Moreira
Tsuyoshi Murata
Amir Navot
David Newman
Ann Nowe
Manfred Opper
Bo Pang
Dmitry Pechyony
Bernhard Pfahringer
Christian Plagemann
Philippe Preux
Emmanuel Rachelson
Alain Rakotomamonjy
Vinayak Rao
Soumya Ray
Martin Riedmiller
Nan Rong
Afshin Rostamizadeh
Cynthia Rudin
Yvan Saeys
Rajhans Samdani
Purnamrita Sarkar
Marc Schoenauer
Marc Sebban
Cosma Shalizi
Jianbo Shi
David Silver
Aarti Singh
Josef Sivic
Paris Smaragdis
Peter Sollich
Elaine Sousa
Ida Sprinkhuizen-Kuyper
Mark Steyvers

Alek Kolcz
Stefan Kramer
Alex Kulesza
Dima Kuzmin
Michail Lagoudakis
Pedro Larranaga
David Leake
Daniel Lee
Guy Lever
Wei Li
Bing Liu
Daniel Lowd
Malik Magdon-Ismail
Vikash Mansinghka
Andre Martins
Jon McAuliffe
Prem Melville
Kurt Miller
Andriy Mnih
Claire Monteleoni
Stephen Muggleton
Kevin Murphy
Claire Nédellec
Regis-Ghislain Newo Kenmogne
Sebastian Nowozin
Ronald Ortner
Shivaswamy Pannagadatta
Elżbieta Pękalska
Joelle Pineau
Enric Plaza
Yanjun Qi
Ali Rahimi
Liva Ralaivola
Gunnar Rätsch
Ben Recht
Stephen Roberts
Michal Rosen-Zvi
Dmitri Roussinov
Ulrich Rückert
Yucel Saigin
Adam Sanborn
Issei Sato
Martin Scholz
Matthias Seeger
Amnon Shashua
Lavi Shpigelman
Tomas Simon
Ajit Singh
Tomas Skopal
Noah Snavely
Le Song
Alessandro Sperduti
Suvrit Sra
Peter Stone

Zico Kolter
Andreas Krause
Brian Kulis
James Kwok
Guanghui Lan
Mark Last
Guy Lebanon
Honglak Lee
Congcong Li
Percy Liang
Jun Liu
Eneldo Loza Mencia
Julien Mairal
Yi Mao
Florent Masseglia
Qiaozhu Mei
Aditya Menon
David Mimno
Shakir Mohamed
Joris Mooij
Sayan Mukherjee
Boaz Nadler
Gergely Neu
Yael Niv
Tim Oates
Aline Paes
Andrea Passerini
Francisco Pereira
Cristiano Pitangui
Pascal Poupart
Yuan Qi
Tapani Raiko
Jan Ramon
Magnus Rattray
Lev Reyzin
Marko Robnik-Sikonja
Saharon Rosset
Daniel Roy
Bryan Russell
Lorenza Saitta
Scott Sanner
Bruno Scherrer
Benjamin Schowe
Yevgeny Seldin
Christian Shelton
Sajid Siddiqi
Özgür Şimşek
Kaushik Sinha
Noam Slonim
Matthijs Snel
Sören Sonnenburg
Dan Spielman
Nathan Srebro
David Stracuzzi

Imre Risi Kondor
Eyal Krupka
Pawan Kumar
Nicolas Lachiche
Niels Landwehr
Alessandro Lazaric
Yann LeCun
Wee Sun Lee
Lihong Li
Chih-Jen Lin
Gaelle Loosli
Zhengdong Lu
Donato Malerba
Steve Marcus
Michael May
Thorsten Meinl
Donald Metzler
Einat Minkov
Mehryar Mohri
Taesup Moon
Remi Munos
Sriraam Natarajan
Jennifer Neville
Blaž Novak
Guillaume Obozinski
John Paisley
Vladimir Pavlovic
Jan Peters
Clara Pizzuti
Doina Precup
Novi Quadrianto
Shyamsundar Rajaram
Marc'Aurelio Ranzato
Pradeep Ravikumar
Sebastian Riedel
Pedro Rodrigues
Fabrice Rossi
Benjamin Rubinstein
Sivan Sabato
Ruslan Salakhutdinov
Vitor Santos Costa
Mikkel Schmidt
Sebastian Schultheiss
Shai Shalev-Shwartz
Shirish Shevade
Ricardo Silva
Vikas Sindhwani
Janne Sinkkonen
Kevin Small
Carlos Soares
David Sontag
Eduardo Spinosa
David Stern
Matt Streeter

Nathan Sturtevant
Charles Sutton
Eiji Takimoto
Michael Taylor
Ivan Titov
Lisa Torrey
Evgeni Tsivtsivadze
Rich Turner
Raquel Urtasun
Martijn Van Otterlo
Michel Verleysen
S.V.N. Vishwanathan
Kiri Wagstaff
Chong Wang
Markus Weimer
Christian Widmer
Ran Wolff
Ying Yang
Huizhen Yu
Yaoliang Yu
Filip Zelezny
Min-Ling Zhang
Zhi-Hua Zhou
Andrew Zisserman

Masashi Sugiyama
Umar Syed
El-Ghazali Talbi
Choon-Hui Teo
Michalis Titsias
Marc Toussaint
Grigorios Tsoumakas
Karl Tuyls
Nicolas Usunier
Maarten Van Someren
Jean-Philippe Vert
Fabio Vitale
Aaron Wilson
Liwei Wang
Kilian Weinberger
Marco Wiering
Frank Wood
Jieping Ye
Jia Yuan Yu
Yisong Yue
Luke Zettlemoyer
Nevin Zhang
Jun Zhu

S Sundararajan
Itsvan Szita
Ben Taskar
Ambuj Tewari
Hannu Toivonen
Mikhail Traskin
Koji Tsuda
Lyle Ungar
Jurgen Van Gael
Sergei Vassilivitski
Enrique Vidal
Michalis Vlachos
Christian Walder
Takashi Washio
Jason Weston
Sinead Williamson
Mingrui Wu
Emine Yilmaz
Philip Yu
Mikhail Zaslavskiy
Chenxiang Zhai
Alice Zheng
Shenghuo Zhu

Ilya Sutskever
Prasad Tadepalli
Matthew Taylor
Jo-Anne Ting
Luis Torgo
Volker Tresp
Douglas Turnbull
Takeaki Uno
Marcel van Gerven
Shankar Vembu
Ricardo Vilalta
Willem Waegeman
Hanna Wallach
Louis Wehenkel
Shimon Whiteson
David Wilson
Zhao Xu
Elad Yom-Tov
Shipeng Yu
Gerson Zaverucha
Jian Zhang
Dengyong Zhou
Albrecht Zimmermann

Student Volunteers

Barak Almog
Yahel David
Anat Hashavit
Ester Koilis
Lester Mackey
Kira Radinsky
Ben Zion Vatashsky
Michael Wunder

Ossant Bar-shira
Kirill Dyagilev
Hananel Hazan
Gal Lavee
Yariv Maron
Ariel Raviv
Thomas Walsh
David Yanay

Ofer Bartal
Maayan Gal-on
Erez Karpas
Kfir Yehuda Levy
Avihai Mejer
Han-Hsing Tu
Wei Wang

Nicholas Bartlett
Sean Gerrish
Noam Koenigstein
Omer Levy
Uri Patish
Daniel Vainsencher
Ari Weinstein

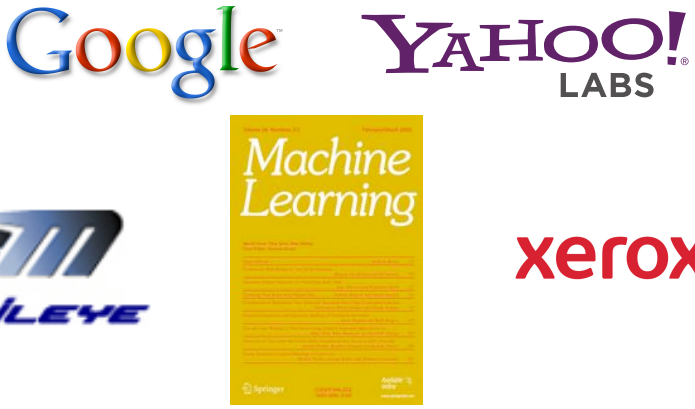
ICML 2010 Sponsors

The organizers of ICML 2010 would like to thank the following sponsors:

Platinum

Gold

Silver



Supporting

ICML 2010 Invited Talks

Invited Talk:

Never-Ending Learning

Tom M. Mitchell, Carnegie Mellon University, U.S.A.

Abstract:

What would it take to develop machine learners that run forever, each day improving their performance and also the accuracy with which they learn? This talk will describe our attempt to build a never-ending language learner, NELL, that runs 24 hours per day, forever, and that each day has two goals: (1) extract more structured information from the web to populate its growing knowledge base, and (2) learn to read better than yesterday, by using previously acquired knowledge to better constrain its subsequent learning. The approach implemented by NELL is based on two key ideas: coupling the semi-supervised training of hundreds of diffent functions that extract different types of information from different web sources, and automatically discovering new constraints that more tightly couple the training of these functions over time. NELL has been running nonstop since January 2010 (follow it at `http://rtw.ml.cmu.edu/readtheweb.html`), and had extracted a knowledge based containing hundreds of thousands of beliefs as of May 2010. This talk will describe NELL, its successes and its failures, and use it as a case study to explore the question of how to design never-ending learners.

Biography:

Tom M. Mitchell is the E. Fredkin University Professor and head of the Machine Learning Department at Carnegie Mellon University. His research interests lie in machine learning, natural language processing, artificial intelligence, and cognitive neuroscience. Mitchell is a member of the U.S. National Academy of Engineering, a Fellow of the American Association for the Advancement of Science (AAAS), and a Fellow and Past President of the Association for the Advancement of Artificial Intelligence (AAAI). Mitchell believes the field of machine learning will be the fastest growing branch of computer science during the 21st century.

Invited Talk:

Using the Web to do Social Science

Duncan Watts, Yahoo! Research, U.S.A.

Abstract:

Social science is often concerned with the emergence of collective behavior out of the interactions of large numbers of individuals, but in this regard it has long suffered from a severe measurement problemnamely that interactions between people are hard to observe, especially at scale, over time, and at the same time as observing behavior. In this talk, I will argue that the technological revolution of the Internet is beginning to lift this constraint. To illustrate, I will describe several examples of internet-based research that would have been impractical to perform until recently, and that shed light on some longstanding sociological questions. Although internet-based research still faces serious methodological and procedural obstacles, I propose that the ability to study truly social dynamics at individual-level resolution will have dramatic consequences for social science.

Biography:

Duncan Watts is a principal research scientist at Yahoo! Research, where he directs the Human Social Dynamics group. He is also an adjunct senior research fellow at Columbia University, and an external faculty member of the Santa Fe Institute and Nuffield College, Oxford. His research on social networks and collective dynamics has appeared in a wide range of journals, from Nature, Science, and Physical Review Letters to the American Journal of Sociology. He is also the author of Six Degrees: The Science of a Connected Age (W.W. Norton, 2003) and Small Worlds: The Dynamics of Networks between Order and Randomness (Princeton University Press, 1999). He holds a B.Sc. in Physics from the University of New South Wales, and Ph.D. in Theoretical and Applied Mechanics from Cornell University.

Invited Guest Lecture:

IF

Robert (Yisrael) Aumann, Hebrew University of Jerusalem, Israel

Abstract:

An exploration of the conceptual foundations of the backward induction algorithm – the algorithm that lies at the basis of all chess-playing programs, including Deep Blue. These foundations are not nearly as clear as they may at first seem. One of the central issues is that of "counterfactual conditionals": sentences like "If I had pushed my pawn, he could have trapped my queen."

Biography:

Robert Aumann was born in Frankfurt am Main, Germany, in 1930, to a well-to-do orthodox Jewish family. Fleeing Nazi persecution, he emigrated to the United States with his family in 1938, settling in New York. In the process, his parents lost everything, but nevertheless gave their two children an excellent Jewish and general education. Aumann attended Yeshiva elementary and high schools, got a bachelor's degree from the City College of New York in 1950, and a Ph.D. in mathematics from MIT in 1955. He joined the mathematics department at the Hebrew University of Jerusalem in 1956, and has been there ever since. In 1990, he was among the founders of the Center for Rationality at the Hebrew University, an interdisciplinary research center, centered on Game Theory, with members from over a dozen different departments, including Business, Economics, Psychology, Computer Science, Law, Mathematics, Ecology, Philosophy, and others. Aumann is the author of over ninety scientific papers and six books, and has held visiting positions at Princeton, Yale, Berkeley, Louvain, Stanford, Stony Brook, and NYU. He is a member of the American Academy of Arts and Sciences, the National Academy of Sciences (USA), the British Academy, and the Israel Academy of Sciences; holds honorary doctorates from the Universities of Chicago, Bonn, Louvain, City University of New York, and Bar Ilan University; and has received numerous prizes, including the Nobel Memorial Prize in Economic Sciences for 2005. Aumann is married and had five children (the oldest was killed in Lebanon in 1982). Also, he has twenty-one grandchildren, and seven great-grandchildren. When not working, he likes to hike, ski, cook, and study the Talmud.

Invited Talk:

Genes, Chromatin, and Transcription

Nir Friedman, Hebrew University of Jerusalem, Israel

Abstract:

A central question in molecular biology is understanding how cells process information and react accordingly. A crucial level of response is by regulating gene transcription, the first step in producing the protein that the gene encodes. Transcriptional regulation is crucial for defining the cells identity and its ability to function. The main dogma is that regulatory instructions are part of the genetic blueprint encoded in the genome, the sequence of DNA base pairs. In recent years there is growing evidence for additional layers of epigenetic information that are passed from a cell to its daughter cells not through the DNA sequence. One of these layers is chromatin, the protein-DNA complex that forms chromosomes. The basic unit of chromatin is a nucleosome, around which about 150 base pairs of DNA are wound. Each nucleosome can be modified by addition of multiple discrete marks, which in turn can be recognized by specific regulatory proteins that modify nucleosomes or impact transcription. As such nucleosomes serve as a substrate for recording information by some regulatory mechanisms and reading it by others, and for passing information to daughter cells following cell division. These new discoveries raise basic questions of what does chromatin state encodes, how it is maintained, updated, passed to next generations, and how it interact with transcription. The research to answer these questions relies on new methodologies that collect massive amount of data about chromatin state in each location along the genome. In this talk I will provide an overview of the field and describe ongoing investigations that attempt to answer these questions.

Biography:

Nir Friedman is a Professor at the Hebrew University of Jerusalem, holding a joint appointment at the School of Computer Science and Engineering and the Institute of Life Sciences. He received his doctorate from Stanford in 1997, and after a two year post-doctoral fellowship at U.C. Berkeley, joined the Hebrew University. He started his research career in Artificial Intelligence, mainly specializing in learning Graphical Probabilistic Models. In the last decade he became more involved in research in Molecular and Cellular Biology and Systems Biology. His current research interests are in understanding regulatory mechanisms involved in gene expression, and in particular how the mechanisms combine to achieve specific expression of different genes. He is a co-founder of the B.Sc. in Computational Biology program at the Hebrew University, one of the first of its kind. Together with Daphne Koller he is a co-author of "Probabilistic Graphical Models: Principles and Techniques."

ICML 2010 Best Paper Awards

Every year, ICML honors its best contributions with best paper awards. This year, the selection committee consisting of the program chairs and seven senior members of the community, chosen to guarantee a sufficient breadth of expertise, selected the contributions listed below. All best paper award winners will receive a certificate and a check over $1000, all runner-ups will receive a check over $500. We are grateful to a generous contribution from Springer, the publishers of *Machine Learning*, for helping to fund these awards.

Best Paper Award

For the Best Paper Award, we prepared a list of 19 papers that received excellent reviews from the reviewers and the area chair. In a two-step selection process, the following paper emerged as the clear winner of the ICML 2010 Best Paper Award.

- *Le Song, Byron Boots, Sajid Saddiqi, Geoffrey Gordon, Alex Smola*:
  Hilbert Space Embeddings of Hidden Markov Models
  Wed, 23.06.2010, 8.30, Oren Hall (Plenary Session)

The runner-up in this category is the paper

- *Julia Vogt, Sandhya Prabhakaran, Thomas Fuchs, Volker Roth*:
  The Translation-invariant Wishart-Dirichlet Process for Clustering Distance Data
  Wed, 23.06.2010, 15.40, Tamar (Clustering 1)

Best Student Paper Award

The best student papers were selected from the same 19 papers as the best paper awards, again in a two-step selection process. Eligible were only papers which had a student as the primary author. The winner is

- *John Duchi, Lester Mackey, Michael Jordan*:
  On the Consistency of Ranking Algorithms
  Thu, 24.06.2010, 8.30, Oren Hall (Plenary Session)

The close runner-up was:

- *Brian Ziebart, Drew Bagnell, Anind Dey*:
  Modeling Interaction via the Principle of Maximum Causal Entropy
  Wed, 23.06.2010, 15.40, Rimon (Causal Inference)

Best Application Paper Award

For the best application, we formed a short-list of six candidates with a strong focus on applications. We are excited to give the award to a paper that brings together basic research in transfer learning and cognitive science:

- *Kevin Canini, Mikhail Shashkov, Tom Griffiths*:
  Modeling Transfer Learning in Human Categorization with the Hierarchical Dirichlet Process
  Wed, 23.06.2010, 14.30, Oren Hall (Plenary Session)

The runner-up in this category is

- *Ali Shoeb, John Guttag*:
  Application of Machine Learning To Epileptic Seizure Detection
  Thu, 24.06.2010, 15.40, Rimon (Time-Series Analysis)

Best 10-Year Paper

Last but not least we have the privilege of honoring the most influential paper of ICML 2000 (Stanford). In a first phase, we selected a short-list of six papers based on suggestions from the best paper committee and several senior members of the community. We then invited all area chairs of ICML 2010 and all members of the IMLS board to name and rank their three favorite papers among these six. From this vote, the following paper emerged as a clear winner:

- *Erin L. Allwein, Robert E. Schapire, Yoram Singer*:
  Reducing Multiclass to Binary: A Unifying Approach for Margin Classifiers
  Tue, 22.06.2010, 17.30, Oren Hall (Plenary Session)

We are glad to have Rob Schapire take a look back upon this great piece of work and discuss its influence on the community.

ICML 2010
Tutorial and Workshop Summaries

Overview of Tutorials and Workshops

As in previous years we were pleased to have a strong program of tutorials for ICML 2010. These were held on June 21, immediately preceding the main conference. The program featured eight tutorials covering a wide range of methods for, and applications of, machine learning. There were tutorials on: domain adaptation (Daumé & Blitzer), geometric tools for identifying structure in networks (Mahoney), kernel adaptive filters (Príncipe & Liu), learning through exploration (Beygelzimer and Langford), metric learning (Kulis), privacy-preserving data mining (Matwin), sparse modeling (Rish & Grabarnik), stochastic optimization (Srebro & Tewari). We would like to thank the community for the high-quality tutorial proposals that were received, the presenters for their extensive efforts in preparing and delivering the selected tutorials, and the local arrangements, program, and general chairs of ICML for their hard work in organizing such a stimulating conference.

Ben Taskar
ICML 2010 Tutorial Chair

ICML solicited workshops on topics related to machine learning, in coordination with the COLT program chairs (Adam Kalai and Mehryar Mohri). We have a set of 9 workshops, each covering a topic of substantial interest in the areas of machine learning and learning theory.

All workshops are held on June 25th, immediately after ICML and two days before COLT. The workshop organizers are the key to a broad set of interesting workshops, so we should thank them all for their effort putting together these programs. We would also like to thank local organizers Dan Pelleg and Shai Fine who have substantially contributed to the organization of facilities and refreshment.

John Langford
ICML 2010 Workshop Chair

Tutorials

Domain Adaptation

Hal Daumé III, John Blitzer

Almost anyone who has deployed machine learning systems in the real world has encountered the task of domain adaptation: We build our models from some fixed source domain, but we wish to deploy them across one or more different target domains. For example, large-scale speech recognition systems need to work well across arbitrary speech, regardless of background noise or accents. Text processing systems trained on news often need to be applied to blogs or forum posts. Gene finders are trained on a particular organism, but often we wish to identify the genes of another organism or even group of organisms. Face recognition systems might be trained under certain pose, lighting, and occlusion settings, but applied to arbitrary sets of pose, lighting, and occlusion. The purpose of this tutorial is to introduce participants to the problem of domain adaptation, the variety of forms it takes, the techniques that have been used to solve it, and our current understanding of when these techniques can and cannot work. We'll also cover interesting open problems in domain adaptation.

Geometric Tools for Identifying Structure in Large Social and Information Networks

Michael W. Mahoney

The tutorial will cover recent algorithmic and statistical work on identifying and exploiting "geometric" structure in large informatics graphs such as large social and information networks. Such tools (e.g., Principal Component Analysis and related non-linear dimensionality reduction methods) are popular in many areas of machine learning and data analysis due to their relatively-nice algorithmic properties and their connections with regularization and statistical inference. These tools are not, however, immediately-applicable in many large informatics graphs applications since graphs are more combinatorial objects; due to the noise and sparsity patterns of many real-world networks, etc. Recent theoretical and empirical work has begun to remedy this, and in doing so it has already elucidated several surprising and counterintuitive properties of very large networks. Topics include: underlying theoretical ideas; tips to bridge the theory-practice gap; empirical observations; and the usefulness of these tools for such diverse applications as community detection, routing, inference, and visualization.

Kernel Adaptive Filters

José Príncipe, Weifeng Liu

This tutorial introduces a family of sequential active learning algorithms in reproducing kernel Hilbert spaces (RKHS), collectively called kernel adaptive filters. A filter is the equivalent of a regressor for time series. Kernel adaptive filters are linear adaptive filters in RKHS and correspond to universal learning systems in the input space for a majority of kernels. Moreover, by exploiting the structure of the RKHS, inner product operations with infinite dimensional vectors can be easily computed in the input space by kernel evaluations. Last but not the least, these nonlinear adaptive filters do not suffer from the local minima problem unlike traditional neural networks. These three characteristics make kernel adaptive filters an intriguing and potentially very useful class of nonlinear adaptive filters and learning machines.

Learning through Exploration

Alina Beygelzimer, John Langford

This tutorial is about learning through exploration. The goal is to learn how to make decisions when the payoff of only a chosen action is observed rather than all choices. The setting we address is simpler than general reinforcement learning, because we consider situations where future rewards are not affected by past decisions, although the algorithms we discuss do have applications in this more general setting. A canonical example of this problem is the ad display problem, where based on a query and user history, an ad is chosen, and then the ad is either clicked on or not. In general, feedback for one ad provides no information about the quality of a different ad. This problem is much more general, of course, because it is really about the process of learning from interactive feedback where there is no "reset" button on history. We will cover algorithms that can learn from logged data where exploration decisions have already been made. This tutorial draws on material from several different communities including Reinforcement Learning, Bandit Online Learning, Supervised Learning and Active Learning.

Metric Learning

Brian Kulis

Metric learning aims to learn an appropriate distance/similarity function for a given problem. Because many learning problems involve some notion of distance or similarity, metric learning can be applied in a variety of settings and has become a popular problem for many learning tasks. This tutorial will cover several approaches to metric learning and discuss recent applications. A majority of research in metric learning has focused on learning Mahalanobis distances, which learn global linear transformations of the data. We will discuss a general framework for several of these methods, as well as online variants and extensions to non-linear and high-dimensional transformations via kernelization. We will also cover some non-Mahalanobis metric learning algorithms, such as local (per-example) methods, and compare and contrast metric learning work with research in kernel learning and dimensionality reduction. Recent progress has made metric learning practical for a host of applications, especially high-dimensional ones such as those in vision, text, and multimedia; we will provide a overview of such applications, with the hope that practitioners can begin to apply metric learning methods to novel domains.

Privacy-preserving Data Mining

Stan Matwin

There is little doubt that data mining technologies create new challenges in the area of data privacy. In this tutorial, we will review some of the new developments in Privacy-preserving Data Mining. In particular, we will look at the existing definitions of privacy, discuss techniques in which data mining results can reveal personal data, and how this can be prevented. We will discus some of the existing anonymization techniques which protect the data against disclosing private information about individuals, and the limitations of these methods. We will look at the practically interesting situations where data to be mined is distributed among several parties, and how cryptographic techniques can assist in this task. We will argue that methods that effectively protect personal data, while at the same time preserve the quality of the data from the data analysis perspective, are some of the principal new challenges before the field.

Sparse Modeling: Theory, Algorithms and Applications

Irina Rish, Genady Grabarnik

Sparse modeling is a rapidly developing area on the intersection of statistics, machine-learning and signal processing. Indeed, discovering sparse structure in the data, such as a subset of highly predictive variables, is essential in many practical applications, such as, for example, biomedical ones, where interpretability of a statistical model is as important as its predictive power. Moreover, efficient recovery of high-dimensional sparse signals from a relatively low number of observations is the main focus of compressed sensing, a rapidly growing and extremely popular area of signal processing. Recent years have witnessed a flurry of research on algorithms and theory for sparse modeling and sparse signal recovery, mainly focused on L1-optimization, a convex relaxation of the (NP-hard) smallest subset selection problem. Examples include sparse regression, such as Lasso and its various extensions (Elastic Net, fused Lasso, group Lasso, simultaneous Lasso and multi-task learning, adaptive Lasso, boosted Lasso, and so on), as well sparse graphical model selection and sparse dimensionality reductions. Theoretical work has provided some conditions when various relaxation methods are capable of recovering an underlying sparse signal/model, provided bounds on sample complexity, and investigated tradeoffs between different choices of design matrix properties that guarantee good performance. The aim of this tutorial is to provide a survey of key recent developments in the above fields.

Stochastic Optimization for Machine Learning

Nathan Srebro, Ambuj Tewari

We discuss the role of Stochastic Optimization in Statistical Learning, the relationship between them, survey recent results in Stochastic Optimization and discuss their relevance for machine learning. Our goal is both to explore the historical, conceptual and theoretical connections (and differences) between Stochastic Optimization and Statistical Learning, and to introduce the machine learning community to the latest research and techniques coming out of the Stochastic Optimization community, exploring how these can be used in learning problems.

Workshops

Learning in Non-(geo)metric Spaces

Joachim Buhmann, Robert Duin, Mario Figueiredo,
Edwin Hancock, Vittorio Murino, Marcello Pelillo

In the last few years, interest around purely similarity-based techniques has grown considerably. For example, within the supervised learning paradigm (where expert-labeled training data is assumed to be available) the now famous "kernel trick" shifts the focus from the choice of an appropriate set of features to the choice of a suitable kernel, which is related to object similarities. However, this shift of focus is only partial, as the classical interpretation of the notion of a kernel is that it provides an implicit transformation of the feature space rather than a purely similarity-based representation. Similarly, in the unsupervised domain, there has been an increasing interest around pairwise or even multiway algorithms, such as spectral and graph-theoretic clustering methods, which avoid the use of features altogether. By departing from vector-space representations one is confronted with the challenging problem of dealing with (dis)similarities that do not necessarily possess the Euclidean behavior or not even obey the requirements of a metric. The lack of the Euclidean and/or metric properties undermines the very foundations of traditional pattern recognition theories and algorithms, and poses totally new theoretical/computational questions and challenges.

Budgeted Learning

Dragos Margineantu, Russell Greiner, Tomas Singliar, Prem Melville

One of the major challenges emerging from applying learning algorithms to practical tasks is the need for principled approaches for training, prediction, and human-in-the-loop learning on a fixed budget. The ICML-2010 Budgeted Learning workshop exposes the machine learning community to the challenges of this problem and will discuss recent progress in machine learning research towards addressing it. Classifier training on a fixed budget for feature value acquisition, active learning of resource-bounded predictors, PAC-learnability under budget constraints, learning and inference primarily limited by expensive human involvement, and anytime learning algorithms are some of the topics that will be discussed. The workshop will also analyze how well-known learning performance metrics relate to budgeted learning problems. The workshop will start with an introductory overview followed by contributed and invited talks, a poster session, and will conclude with an open panel discussion.

Learning to Rank Challenge

Tie-Yan Liu, Olivier Chapelle, Yi Chang

Though over 100 papers have been published in the learning to rank field (LTR), most of the large-scale, real-world datasets are not publicly available. This makes drawing comparisons between algorithms difficult. In the spirit of changing this, Yahoo! has hosted the Learning to Rank Challenge and released two datasets used for learning Yahoo's search ranking function. This exciting machine learning challenge consisted of two tracks: the first is a standard LTR track and the second is a transfer-learning track. Top competitors in both tracks will come and present their algorithms. In addition, the workshop will also feature presentations on ranking which are not directly related to the challenge. It will be an opportunity to discuss future research directions regarding learning to rank.

Learning from Multi-Label Data

Min-Ling Zhang, Grigorios Tsoumakas, Zhi-Hua Zhou

The 2nd International Workshop on Learning from Multi-Label Data (MLD'10) is devoted to the theme of learning from multi-label data, a.k.a. multi-label learning. Multi-label learning deals with data where each object is associated with multiple labels simultaneously instead of a single label. Many real-world applications such as text categorization naturally fall into this learning scenario. During the past decade, significant amount of progress has been witnessed on learning from multi-label data, including design of novel multi-label learning algorithms, real-world applications of multi-label learning techniques, and also extensions of the main multi-label learning tasks. The program of MLD'10 will feature an invited talk given by Prof. Eyke Huellermeier on "Multi-Label Classification: Challenges, Pitfalls and Perspectives". In addition, three oral presentation sessions (1 morning + 2 afternoon) are arranged for 10 accepted regular papers, which address various aspects of multi-label learning such as new strategies toward learning from multi-label data, case-studies of multi-label learning techniques, applications in bioinformatics and document classification, and multi-instance multi-label learning (MIML).

Machine Learning Open Source Software 2010

Sören Sonnenburg, Mikio Braun, Cheng Soon Ong, Patrik Hoyer

We believe that the wide-spread adoption of open source software policies will have a tremendous impact on the field of machine learning. The goal of this workshop is to further support the current developments in this area and give new impulses to it. Following the success of the inaugural NIPS-MLOSS workshop held at NIPS 2006, the Journal of Machine Learning Research (JMLR) has started a new track for machine learning open source software initiated by the workshop's organizers. Many prominent machine learning researchers have co-authored a position paper advocating the need for open source software in machine learning. To date 11 machine learning open source software projects have been published in JMLR. Furthermore, the workshop's organizers have set up a community website mloss.org where people can register their software projects, rate existing projects and initiate discussions about projects and related topics. This website currently lists 221 such projects including many prominent projects in the area of machine learning.

The main goal of this workshop is to bring the main practitioners in the area of machine learning open source software together in order to initiate processes which will help to further improve the development of this area. In particular, we have to move beyond a mere collection of more or less unrelated software projects and provide a common foundation to stimulate cooperation and interoperability between different projects. An important step in this direction will be a common data exchange format such that different methods can exchange their results more easily.

Machine Learning and Games

Kurt Driessens, Olana Missura, Chistian Thurau

In recent years there has been a growing interest in games as a source of new and challenging problems for the machine learning community. Games evolved from single player games to complex world simulations with hundreds or even millions of players simultaneously. This does not only impose new challenges to game design, it also requires more and more sophisticated methods for automation. Fraud detection, story generation, adapting game AI, or matchmaking are only a few of the novel challenges that are being targeted by the industry - and that offer exciting research topics. While in this workshop we want to reflect the ongoing changes in the games industry, the primary goal is to provide an opportunity for the researchers working on these problems to meet and interchange their results and ideas. We want to provide a venue for discussing future directions for machine learning in games, both for academia and the industry.

Social analytics: Learning from Human Interactions

Elad Yom-Tov, Shie Mannor, Yossi Richter

Our social interactions are often made by electronic means, and are thus recorded in accessible formats. This opens a range of possibilities for studying human interactions from data such as social network sites on the web and cell phone communications. These tasks usually involve massive amount of data (billions of records) that is often quite noisy and even corrupted. Social and artificial networks suggest new challenges in modelling machine learning problems as there are strong spatial and temporal correlations. In this workshop we will hear researchers from academia and industry share their insights about the theory and applications of social analytics. Our focus will be on graph mining and predictive tools that can be used in this area. We will also consider different applications from recommender systems, marketing, search, network optimization and other emerging areas.

Our goal is to bring people from different fields, ranging from computer science to marketing, from both academia and industry, to foster a discussion on the different approaches people have taken to the study of social phenomena. The workshop will introduce people to the tools and data available for social research in various communities. A primary goal of the workshop is to create a set of common theoretical problems to work on and to solicit data from companies to facilitate common research challenges.

Reinforcement Learning and Search in Very Large Spaces

Peter Auer, Samuel Kaski, Csaba Szepesvàri

This workshop is about reinforcement learning in large state/action spaces, learning to optimize search, and the relation of these two. It addresses in particular the following two questions:

- Identify cases when realistically large problems with delayed feedback can be solved successfully, possibly but not necessarily by reinforcement learning algorithms. Such algorithm may need to exploit the special structure of the learning problem. As an example we see content-based information retrieval.
- Application of learning techniques to develop powerful interactive search algorithms: optimizing a single search or learning across searches, with or without probabilistic assumptions.

Topic Models: Structure, Applications, Evaluation and Extensions

Michal Rosen-Zvi, Amit Gruber, Richard Zemel

Latent topic models serve as an unsupervised machine learning technique for mining data in large corpora. These models have been successfully applied to various types of data, including text documents, images, biological data, and more. Various methods for approximate inference have been suggested and many variations and extensions of the basic models have been published over the past decade in the machine learning and data mining communities. This workshop aims at initiating further discussion on improving topic models and exploring new directions of research.

ICML 2010 Invited Applications Papers