Conference Schedule

Twentieth International Conference on Machine Learning

Washington, D.C. USA **August 21–24, 2003**

General Notes:

- **Registration** is in the Regency Gallery every day. Hours are:
  - Thursday 8-5
  - Friday 8-5
  - Saturday 8-12:30
- All paper presentations are **22 minutes** long: 20 minutes for the talk and 2 minutes for questions.
- **Continental breakfasts** are available every morning from 8 to 9 AM. They will be in the **Palladian Foyer** on Thursday morning and in the **Regency Ballroom** on subsequent mornings. These breakfasts will have coffee, tea, juices, pastries, breads, fruit preserves, jam and butter.
- There will be two 30-minute **breaks** every day. They will be catered with coffee, tea, light snacks, etc.
- There will be two **receptions/poster sessions**, one on Thursday and one on Saturday. There should be enough food for a light dinner. There will be vegetarian options at all meals.

  The researchers presenting posters will be determined by the last name of the first author (or student). People with names beginning with A-K will present posters in the Thursday evening reception, and people with names beginning with L-Z will present in the Saturday evening reception.

Foster Provost

The Role of Applications in the Science of Machine Learning

Friday, August 22, 2:00–3:00

Abstract

Applications play a variety of roles for the science of machine learning. Obviously, high-profile applications create excitement about the field, helping to attract research funding and high-caliber students.

Less obviously, applications provide important stimulus to the field and help to keep it vital and relevant. They highlight deficiencies in the state of the art and thereby steer research in new directions. I will examine a case study involving machine learning for fraud detection, and will show how this application revealed several topics requiring research attention. Applying machine learning to fraud detection reveals:

1. multirelational data
2. networked examples
3. models that must monitor activity over time
4. various kinds of costs
5. unbalanced and changing class distributions.

I will point to recent advances in some of these areas.

Unfortunately, despite repeated calls for applications papers, few get submitted and almost none get published. I also will discuss why this is and what we as authors and reviewers might do about it.

Biography

Professor Foster Provost, of NYU's Stern School of Business, teaches and speaks about advances in machine learning, data mining, and knowledge systems and about their alignment with business problems. His research focuses on issues involved with aligning machine learning techniques with real-world problems. These issues include taking costs into account when learning, robust learning in the face of imprecision, and learning profiles in order to monitor activity. He currently is studying learning from relational and network-structured data.

Professor Provost recently was elected as a founding board member of the International Machine Learning Society, is an editor of the journal Machine Learning and a member of the editorial boards of the Journal of Machine Learning Research and the Journal of Artificial Intelligence Research. He was Program Chair of KDD-2003 (with R. Srikant).

He received his Ph.D. from the University of Pittsburgh in 1992, under the supervision of Prof. Bruce Buchanan. Prior to joining NYU, he worked in the research labs of NYNEX/Bell Atlantic, studying the application of machine learning methods to problems of fraud detection, network diagnosis, network monitoring, and others.

Michael Kearns

Computational Learning Theory: A Retrospective

Saturday, August 23, 2:00–3:00

Abstract

It has been nearly two decades since the publication of Valiant's "A Theory of the Learnable", and 15 years since the first COLT conference. During this period, computational learning theory has matured dramatically, growing in mathematical depth and finding its natural connections to many other disciplines. The models and algorithms of the field have had widespread impact on the practice of machine learning.

This talk will be an eclectic history of the ideas, results and people of computational learning theory. I will survey the algorithms that have influenced applications, the models that have shaped the language of machine learning, and ideas that ultimately did neither but were still interesting. Connections with the simultaneous lines of thought in the practice of machine learning will be given, along with some amusing anecdotes.

Biography

Michael Kearns is a professor in the Computer and Information Science Department at the University of Pennsylvania, and the co-director of Penn's interdisciplinary Institute for Research in Cognitive Science. He also has a joint appointment in the Operations and Information Management (OPIM) department of the Wharton School.

Prof. Kearns did his undergraduate studies at the University of California at Berkeley in math and computer science, graduating in 1985. He received a Ph.D. in computer science from Harvard University in 1989; the title of his dissertation was The Computational Complexity of Machine Learning, under Prof. L.G. Valiant. Following postdoctoral positions at the Laboratory for Computer Science at M.I.T. and at the International Computer Science Institute in Berkeley, in 1991 he joined the research staff of AT&T Bell Labs.

His primary research interests are in artificial intelligence and machine learning, including computational learning theory, reinforcement learning, probabilistic inference and graphical models, and computational game theory. Prof. Kearns has worked on a variety of applications of AI to human-computer interaction, including spoken dialogue systems and software agents in MUDs. He also has interests in cryptography and network security, theoretical computer science, and computational finance.

Daphne Koller
Statistical Learning from Relational Data

Sunday, August 24, 2:00–3:00

Abstract

Much of the data in the world is relational in nature, involving multiple objects, related to each other in a variety of ways. Examples include both structured databases such as customer transaction data, semi-structured data such as hyperlinked pages on the world-wide web or networks of interacting genes, and unstructured data such as text. In this talk, I will describe a statistical framework for learning from relational data. The approach is based on probabilistic models, which have been applied with great success to a variety of machine learning tasks. Generally, this framework has been applied to data represented as fixed-length attribute-value vectors, or to sequenceq data. I will describe the language of probabilistic relational models (PRMs), which extend probabilistic graphical models with the expressive power of object-relational languages. PRMs model the uncertainty over the attributes of objects in the domain as well as uncertainty over the existence of relations between objects. I will present techniques for automatically learning PRMs directly from a relational data set, and applications of these techniques to various tasks, such as: collective classification of an entire set of related entities; clustering a set of linked entities into coherent groups; and even predicting the existence of links between entities. The talk will demonstrate the applicability of the techniques on several domains, such as web data and biological data.

Biography

Daphne Koller received her PhD from Stanford University in 1994. After a two-year postdoc at Berkeley, she returned to Stanford, where she is now an Associate Professor in the Computer Science Department. Her main research interest is in creating large-scale systems that reason and act under uncertainty, using techniques fro decision theory and economics. Daphne Koller is the author of over 70 refereed publications, which have appeared in AI, theoretical computer science, and economics venues. She was the co-chair of the recent UAI 2001 conference, has served on numerous program committes, and as associate editor of the Journal of Artificial Intelligence Research and of the Machine Learning Journal. She was awarded the Arthur Samuel Thesis Award in 1994, the Sloan Foundation Faculty Fellowship in 1996, the ONR Young Investigator Award in 1998, the Presidential Early Career Award for Scientists and Engineers (PECASE) in 1999, and the IJCAI Computers and Thought Award at the IJCAI 2001 conference.

Thursday morning (August 21)

08:00–9:00 **Continental Breakfast**
Palladian Foyer

Morning tutorials:

9:00–10:30 **MT1 Mining Time Series Data**
Palladian Ballroom
Christos Faloutsos

9:00–10:30 **MT2 The State of the Art in Language Modeling**
Empire Ballroom
Joshua Goodman

Workshops:

9:00–5:30 **W1 Machine Learning Technologies for Autonomous Space Applications**
Forum Room
Kiri Wagstaff • Amy McGovern • Terran Lane

9:00–5:30 **W2 Machine Learning in Bioinformatics**
Executive Room
Vladimir Pavlovic • Ashutosh Garg • Simon Kasif

9:00–5:30 **W3 Learning from Imbalanced Data Sets II**
Congressional A
Nitesh Chawla • Nathalie Japkowicz • Aleksandr Kolcz

9:00–5:30 **W2 The Continuum from Labeled to Unlabeled Data in Machine Learning and Data Mining**
Congressional B
Rayid Ghani • Rosie Jones • Chuck Rosenberg

10:30–11:00 **Break**
Palladian Foyer

12:30–2:00 **Lunch Break**

Thursday afternoon (August 21)

Afternoon tutorials:

2:00–5:30 **AT1 Machine Learning and Genetic Microarrays**
Palladian Ballroom
Jude Shavlik • David Page

2:00–5:30 **AT2 Practical Sample Complexity**
Empire Ballroom
John Langford

Workshops (continuing):

9:00–5:30 **W1 Machine Learning Technologies for Autonomous Space Applications**
Forum Room
Kiri Wagstaff • Amy McGovern • Terran Lane

9:00–5:30 **W2 Machine Learning in Bioinformatics**
Executive Room
Vladimir Pavlovic • Ashutosh Garg • Simon Kasif

9:00–5:30 **W3 Learning from Imbalanced Data Sets II**
Congressional A
Nitesh Chawla • Nathalie Japkowicz • Aleksandr Kolcz

9:00–5:30 **W2 The Continuum from Labeled to Unlabeled Data in Machine Learning and Data Mining**
Congressional B
Rayid Ghani • Rosie Jones • Chuck Rosenberg

3:30–4:00 **Break**
Palladian Foyer

5:30pm–7:00pm **Welcome Reception**
Poster session: A–K
Ambassador Ballroom

Friday morning (August 22)

08:00–9:00 **Continental Breakfast**
Regency Ballroom

9:00–10:30 **1 Reinforcement learning I** *(Chair: Peter Stone)*
Ambassador Ballroom

Model-based Policy Gradient Reinforcement Learning
Xin Wang • Thomas Dietterich

The Cross Entropy method for Fast Policy Search
Shie Mannor • Reuven Rubinstein • Yohai Gat

Hierarchical Policy Gradient Algorithms
Mohammad Ghavamzadeh • Sridhar Mahadevan

Relativized Options: Choosing the Right Transformation
Balaraman Ravindran • Andrew Barto

9:00–10:30 **2 ROC graphs & extensions** *(Chair: David Jensen)*
Executive Room

Optimizing Classifier Performance via an Approximation to the Wilcoxon-Mann-Whitney Statistic
Lian Yan • Robert Dodier • Michael Mozer • Richard Wolniewicz

Improving accuracy and cost of two-class and multi-class probabilistic classifiers using ROC curves
Nicolas Lachiche • Peter Flach

The Geometry of ROC Space: Understanding Machine Learning Metrics through ROC Isometrics
Peter Flach

Regression Error Characteristic Curves
Jinbo Bi • Kristin Bennett

9:00–10:30 **3 Text I** *(Chair: Marko Grobelnik)*
Congressional Room

Text Classification Using Stochastic Keyword Generation
Cong Li • Ji-Rong Wen • Hang Li

Classification of Text Documents Based on Minimum System Entropy
Raghu Krishnapuram • Krishna Chitrapura • Sachindra Joshi

Tackling the Poor Assumptions of Naive Bayes Text Classifiers
Jason Rennie • Lawrence Shih • Jaime Teevan • David R. Karger

Text Bundling: Statistics Based Data-Reduction
Lawrence Shih • Jason Rennie • Yu-Han Chang • David R. Karger

---

10:30–11:00 **Break**
Regency Ballroom

11:00–12:30 **4 Reinforcement Learning II** *(Chair: Satinder Singh)*

AMBASSADOR BALLROOM

Bayes Meets Bellman: The Gaussian Process Approach to Temporal Difference Learning

⟹ ICML STUDENT PAPER AWARD WINNER ⟸

Yaakov Engel • Shie Mannor • Ron Meir

TD(0) Converges Provably Faster than the Residual Gradient Algorithm

Ralf Schoknecht • Artur Merke

The Significance of Temporal-Difference Learning in Self-Play Training TD-Rummy versus EVO-rummy

Jugal Kalita • Cliff Kotnik

Combining TD-learning with Cascade-correlation Networks

François Rivest • Doina Precup

11:00–12:30 **5 Dimension reduction** *(Chair: Carla Brodley)*

EXECUTIVE ROOM

Weighted Low-Rank Approximations

Nathan Srebro • Tommi Jaakkola

Informative Discriminant Analysis

Samuel Kaski • Jaakko Peltonen

Isometric Embedding and Continuum ISOMAP

Hongyuan Zha • Zhenyue Zhang

Cross-Entropy Directed Embedding of Network Data

Takeshi Yamada • Kazumi Saito • Naonori Ueda

11:00–12:30 **6 Ensembles I** *(Chair: Rob Schapire)*

CONGRESSIONAL ROOM

On the Convergence of Boosting Procedures

Tong Zhang • Bin Yu

The Use of the Ambiguity Decomposition in Neural Network Ensemble Learning Methods

Gavin Brown • Jeremy Wyatt

Low Bias Bagged Support Vector Machines

Giorgio Valentini • Thomas Dietterich

Linear Programming Boosting for Uneven Datasets

Jurij Leskovec • John Shawe-Taylor

12:30–2:00 **Lunch Break**

2:00–3:00 **Invited Talk: Foster Provost**
The Role of Applications in the Science of Machine Learning
Regency Ballroom

03:00–03:45 **7 Text II** *(Chair: Rayid Ghani)*
Ambassador Ballroom
A Loss Function Analysis for Classification Methods in Text Categorization
Fan Li • Yiming Yang
Modified Logistic Regression: An Approximation to SVM and Its Applications in Large-Scale Text Categorization
Jian Zhang • Rong Jin • Yiming Yang • Alex Hauptmann

03:00–03:45 **8 Mixture modeling I** *(Chair: John Langford)*
Executive Room
Mixtures of Conditional Maximum Entropy Models
Dmitry Pavlov • Alexandrin Popescul • David Pennock • Lyle Ungar
Learning Mixture Models with the Latent Maximum Entropy Principle
Shaojun Wang • Dale Schuurmans • Fuchun Peng • Yunxin Zhao

03:00–03:45 **9 Planning and problem solving** *(Chair: Prasad Tadepalli)*
Congressional Room
DISTILL: Learning Domain-Specific Planners by Example
Elly Winner • Manuela Veloso
Characteristics of Long-term Learning in Soar and its Application to the Utility Problem
William Kennedy • Kenneth De Jong

3:45–4:15 **Break**
Regency Ballroom

Friday afternoon (August 22)

4:15–5:00 **10A Learning with skew/error costs** *(Chair: Michael Pazzani)*
Ambassador Ballroom

Perceptron Based Learning with Example Dependent and Noisy Costs
Peter Geibel • Fritz Wysotzki

Adaptive Feature-Space Conformal Transformation for Imbalanced-Data Learning
Gang Wu • Edward Chang

5:00–6:05 **10B Information filtering** *(Chair: Thomas Hofmann)*
Ambassador Ballroom

Online Ranking/Collaborative filtering using the Perceptron Algorithm
Edward Harrington

Flexible Mixture Model for Collaborative Filtering
Luo Si • Rong Jin

Exploration and Exploitation in Adaptive Filtering Based on Bayesian Active Learning
Yi Zhang • Wei Xu • Jamie Callan

4:15–6:05 **11 Mixture modeling II** *(Chair: Yoshua Bengio)*
Executive Room

A Faster Iterative Scaling Algorithm For Conditional Exponential Model
Rong Jin • Rong Yan • Jian Zhang • Alex Hauptmann

Adaptive Overrelaxed Bound Optimization Methods
Ruslan Salakhutdinov • Sam Roweis

Optimization with EM and Expectation-Conjugate-Gradient
Ruslan Salakhutdinov • Sam Roweis • Zoubin Ghahramani

Unsupervised Learning with Permuted Data
Sergey Kirshner • Sridevi Parise • Padhraic Smyth

Discriminative Gaussian Mixture Models: A Comparison with Kernel Classifiers
Aldebaro Klautau • Nikola Jevtic • Alon Orlitsky

4:15–6:05 **12 Reinforcement learning III** *(Chair: Rich Sutton)*
Congressional Room

Learning Predictive State Representations
Satinder Singh • Michael Littman • Nicholas Jong • David Pardoe • Peter Stone

Reinforcement Learning as Classification: Leveraging Modern Classifiers
Michail Lagoudakis • Ronald Parr

Correlated Q-Learning
Amy Greenwald • Keith Hall

Q-Decomposition for Reinforcement Learning Agents
Stuart Russell • Andrew Zimdars

Action Elimination and Stopping Conditions for Reinforcement Learning
Eyal Even-Dar • Shie Mannor • Yishay Mansour

07:00pm–10:00pm **Conference Banquet**
Palladian Ballroom

08:00–9:00 **Continental Breakfast**
Regency Ballroom

9:00–10:30 **13 Relational learning** *(Chair: Stan Matwin)*
Ambassador Ballroom

Relational Instance Based Regression for Relational Reinforcement Learning
Kurt Driessens • Jan Ramon

Identifying Predictive Structures in Relational Data Using Multiple Instance Learning
Amy McGovern • David Jensen

Avoiding Bias when Aggregating Relational Data with Degree Disparity
David Jensen • Jennifer Neville • Michael Hay

On Kernel Methods for Relational Learning
Chad Cumby • Dan Roth

9:00–10:30 **14 Decision tree and rule learning** *(Chair: David Page)*
Executive Room

Decision-tree Induction from Time-series Data Based on a Standard-example Split Test
Yuu Yamada • Einoshin Suzuki • Hideto Yokoi • Katsuhiko Takabayashi

Decision Tree with Better Ranking
Charles Ling • Robert (Jun) Yan

Boosting Lazy Decision Trees
Xiaoli Fern • Carla Brodley

An Analysis of Rule Evaluation Metrics
Johannes Fürnkranz • Peter Flach

9:00–10:30 **15 SVM I** *(Chair: Tony Jebara)*
Congressional Room

The Set Covering Machine with Data-Dependent Half-Spaces
Mario Marchand • Mohak Shah • John Shawe-Taylor • Marina Sokolova

Kernel PLS-SVC for Linear and Nonlinear Classification
Roman Rosipal • Leonard Trejo • Bryan Matthews

Multi-Objective Programming in SVMs
Jinbo Bi

SimpleSVM
S V N Vishwanathan • Alex Smola • Narashima Murty

10:30–11:00 **Break**
Regency Ballroom

Saturday morning (August 23)

11:00–12:30 16 Semi-supervised learning *(Chair: Dan Roth)*

Ambassador Ballroom

Semi-Supervised Learning of Mixture Models
Fabio Cozman • Ira Cohen • Marcelo Cirelo

Learning with Positive and Unlabeled Examples Using Weighted Logistic Regression
Wee Sun Lee • Bing Liu

Semi-Supervised Learning Using Gaussian Fields and Harmonic Functions
Xiaojin Zhu • Zoubin Ghahramani • John Lafferty

Transductive Learning via Spectral Graph Partitioning
Thorsten Joachims

11:00–12:30 17 Bayes nets *(Chair: Marie desJardins)*

Executive Room

Probabilistic Classifiers and the Concepts they Recognize
Manfred Jaeger

Tractable Bayesian Learning of Tree Augmented Naive Bayes Models
Jesús Cerquides • Ramon López de Màntaras

Optimal Reinsertion: A new search operator for accelerated and more accurate Bayesian network structure learning
Andrew Moore • Weng-Keen Wong

Learning with Knowledge from Multiple Experts
Matthew Richardson • Pedro Domingos

11:00–12:30 18 Applications *(Chair: Mark Schwabacher)*

Congressional Room

Hierarchical Latent Knowledge Analysis for Co-occurrence Data
Hiroshi Mamitsuka

Learning Logic Programs for Layout Analysis Correction
Margherita Berardi • Michelangelo Ceci • Floriana Esposito • Donato Malerba

Robust Induction of Process Models from Time-Series Data
Pat Langley • Dileep George • Stephen Bay • Kazumi Saito

Goal-directed Learning to Fly
Andrew Isaac • Claude Sammut

12:30–2:00 **Lunch Break**

12:30–2:00 **Tutorial for Students: The Three R's of Publishing Machine Learning Papers: Research, 'Riting, and Reviews**
Congressional Room
Marie desJardins • Rob Holte • Rob Schapire

2:00–3:00 **Invited Talk: Michael Kearns**
Computational Learning Theory: A Retrospective
Regency Ballroom

03:00–03:45 **19 Evolutionary Methods** *(Chair: Shie Mannor)*
Ambassador Ballroom
Utilizing Domain Knowledge in Neuroevolution
James Fan • Raymond Lau • Risto Miikkulainen
Evolutionary MCMC sampling and optimization in discrete spaces
Malcolm Strens

03:00–03:45 **20 Active learning** *(Chair: Sally Goldman)*
Executive Room
Incorporating Diversity in Active Learning with Support Vector Machines
Klaus Brinker
Online Choice of Active Learning Algorithms
Yoram Baram • Ran El-Yaniv • Kobi Luz

03:00–03:45 **21 Clustering** *(Chair: Rich Caruana)*
Congressional Room
Using the Triangle Inequality to Accelerate k-Means
Charles Elkan
Random Projection for High Dimensional Data Clustering: A Cluster Ensemble Approach
Xiaoli Fern • Carla Brodley

3:45–4:15 **Break**
Regency Ballroom

4:15–6:05 **22 COLT and online learning** *(Chair: Shai Ben-David)*

AMBASSADOR BALLROOM

BL-WoLF: A Framework For Loss-Bounded Learnability In Zero-Sum Games
Vincent Conitzer • Tuomas Sandholm

Stochastic Local Search in k-term DNF Learning
Ulrich Rueckert • Stefan Kramer

Using Linear-threshold Algorithms to Combine Multi-class Sub-experts
Chris Mesterharm

Testing Exchangeability On-Line
Vladimir Vovk • Ilia Nouretdinov • Alex Gammerman

Online Convex Programming and Generalized Infinitesimal Gradient Ascent
Martin Zinkevich

4:15–5:21 **23A Link analysis** *(Chair: Thorsten Joachims)*

EXECUTIVE ROOM

Finding Underlying Connections: A Fast Graph-Based Method for Link Analysis and Collaboration Queries
Jeremy Kubica • Andrew Moore • David Cohn • Jeff Schneider

Evolving Strategies for Focused Web Crawling
Judy Johnson • Kostas Tsioutsiouliklis • C. Lee Giles

Link-based Classification
Qing Lu • Lise Getoor

5:21–6:05 **23B Learning distance functions** *(Chair: Thorsten Joachims)*

EXECUTIVE ROOM

Learning Distance Functions using Equivalence Relations
Aharon Bar Hillel • Tomer Hertz • Noam Shental • Daphna Weinshall

Learning Metrics via Discriminant Kernels and Multidimensional Scaling: Toward Expected Euclidean Representation
Zhihua Zhang

4:15–5:21 **24A Multi-agent learning** *(Chair: Ronald Parr)*

CONGRESSIONAL ROOM

Justification-based Multiagent Learning
Santi Ontan • Enric Plaza

Learning To Cooperate in a Social Dilemma: A Satisficing Approach to Bargaining
Jeffrey Stimpson • Michael Goodrich

AWESOME: A General Multiagent Learning Algorithm that Converges in Self-Play and Learns a Best Response Against Stationary Opponents
Vincent Conitzer • Tuomas Sandholm

5:21–6:05 **24B** **Reinforcement learning IV** *(Chair: Ronald Parr)*

CONGRESSIONAL ROOM

The Influence of Reward on the Speed of Reinforcement Learning: An Analysis of Shaping

Adam Laud • Gerald DeJong

Principled Methods for Advising Reinforcement Learning Agents

Eric Wiewiora • Garrison Cottrell • Charles Elkan

6:15pm–7:00pm **ICML Business Meeting**

All are welcome

EXECUTIVE ROOM

07:00pm–10:00pm **Reception**

Poster session: L–Z

AMBASSADOR BALLROOM

Sunday morning (August 24)

08:00–9:00 **Continental Breakfast**
Regency Ballroom

9:00–10:30 **25 SVM II** *(Chair: Thore Graepel)*
Ambassador Ballroom

New ν-Support Vector Machines and their sequential minimal optimization
Xiaoyun Wu • Rohini Srihari

Hidden Markov Support Vector Machines
Yasemin Altun • Ioannis Tsochantaridis • Thomas Hofmann

Fast Query-Optimized Kernel Machine Classification Via Incremental Approximate Nearest Support Vectors
Dennis DeCoste • Dominic Mazzoni

Margin Distribution and Learning
Ashutosh Garg • Dan Roth

9:00–10:30 **26 Reinforcement learning V** *(Chair: Michail Lagoudakis)*
Executive Room

Exploration in Metric State Spaces
Sham Kakade • Michael Kearns • John Langford

Error Bounds for Approximate Policy Iteration
Remi Munos

Planning in the Presence of Cost Functions Controlled by an Adversary
H. Brendan McMahan • Avrim Blum • Geoffrey Gordon

Design for an Optimal Probe

Diffusion Approximation for Bayesian Markov Chains
Michael Duff

9:00–10:30 **27 Feature Selection & Generation** *(Chair: Claire Cardie)*
Congressional Room

Online Feature Selection using Grafting
Simon Perkins • James Theiler

An Evaluation on Feature Selection for Text Clustering
Tao Liu • Shengping Liu • Zheng Chen • Wei-Ying Ma

Visual Learning by Evolutionary Feature Synthesis
Krzysztof Krawiec • Bir Bhanu

Feature Selection for High-Dimensional Data: A Fast Correlation-Based Filter Solution
Lei Yu • Huan Liu

10:30–11:00 **Break**
Regency Ballroom

11:00–12:30 **28 Noise/imperfect data** *(Chair: Terran Lane)*
Ambassador Ballroom

On State Merging in Grammatical Inference: A Statistical Approach for Dealing with Noisy Data
Marc Sebban • Jean-Christophe Janodet

Solving Noisy Linear Operator Equations by Gaussian Processes: Application to Ordinary and Partial Differential Equations
Thore Graepel

Eliminating Class Noise in Large Datasets
Xingquan Zhu • Xindong Wu • Qijun Chen

Learning from Attribute Value Taxonomies and Partially Specified Instances
Jun Zhang • Vasant Honavar

11:00–12:30 **29 Kernel Methods/SVM** *(Chair: Jose Luis Balcazar)*
Executive Room

The Pre-Image Problem in Kernel Methods
James Kwok • Ivor Tsang

Marginalized Kernels Between Labeled Graphs
Hisashi Kashima • Koji Tsuda • Akihiro Inokuchi

Machine Learning with Hyperkernels
Cheng Soon Ong • Alex Smola

Learning with Idealized Kernels
James Kwok • Ivor Tsang

11:00–11:45 **30A Meta-learning** *(Chair: Dan Oblinger)*
Congressional Room

Representational Issues in Meta-Learning
Alexandros Kalousis • Melanie Hilario

Choosing between two learning algorithms based on calibrated tests
Remco Bouckaert

11:45–12:30 **30B Ensembles II** *(Chair: Dan Oblinger)*
Congressional Room

A New Boosting Algorithm Using Input-Dependent Regularizer
Rong Jin • Yan Liu • Luo Si • Jaime Carbonell • Alex Hauptmann

Weighted Order Statistic Classifiers with Large Rank-Order Margin
Reid Porter • Damian Eads • Don Hush • James Theiler

12:30–2:00 **Lunch Break**

12:30–2:00 **MLJ Editorial Board Meeting**
KDD HOTEL

Note: The remaining Sunday events are held at the KDD site.
Please see the colocation map for details. The KDD site is a short (1/2 mile) walk from the ICML hotel. All ICML attendees are welcome.

2:00–3:00 **Joint Invited Talk: Daphne Koller**
Statistical Learning from Relational Data
KDD HOTEL

3:00–4:00 **Joint ICML/KDD Session I** *(Chair: Pedro Domingos)*
KDD INTERNATIONAL BALLROOM – CENTER

Maximizing the Spread of Influence through a Social Network
⟹ KDD BEST RESEARCH PAPER AWARD ⟸
David Kempe • Jon Kleinberg • Eva Tardos

Bayesian Network Anomaly Pattern Detection for Disease Outbreaks
Weng-Keen Wong • Andrew Moore • Gregory Cooper • Michael Wagner

4:00–4:30 **Break**

4:30–6:30 **Joint ICML/KDD Session II** *(Chair: Tom Fawcett)*
KDD INTERNATIONAL BALLROOM – CENTER

XRules: An Effective Structural Classifier for XML Data
Mohammed Zaki • Charu Aggarwal

Learning on the Test Data: Leveraging "Unseen" Features
Ben Taskar • Ming Fai Wong • Daphne Koller

Information-Theoretic Co-clustering
Inderjit Dhillon • Subramanyam Mallela • Dharmendra Modha

A Kernel between Sets of Vectors
⟹ ICML STUDENT PAPER AWARD ⟸
Risi Kondor • Tony Jebara