KDD 98

The Fourth International Conference on Knowledge Discovery and Data Mining

New York City
August 27-31
1998

Sponsored by The American Association for Artificial Intelligence

In cooperation with the Very Large Data Bases Conference (VLDB '98)
With support from SAS Institute, Inc.

Registration Brochure
www.aaai.org

KDD-98: A Preview

Business and scientific endeavors now routinely gather large quantities of data that can help with understanding critical problems and with making strategic and tactical decisions. However, manual analysis becomes difficult as databases grow, and techniques for discovering and extracting knowledge automatically become necessary.

The Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98) addresses the science and technology of automated discovery, drawing from the fields of statistics, databases, machine learning, data visualization, high performance computing, knowledge acquisition, and knowledge-based systems. Besides scientific and technical contributions from the latest research and applications, this year's program will also feature free tutorials given by leading experts, demonstrations by leading tool vendors, and several invited talks summarizing the state of the art and looking forward to the future. The field's interdisciplinary nature facilitates the cross-fertilization of ideas, and this summer KDD-98 will be collocated with the 24th International Conference on Very Large Databases (VLDB-98). The KDD-98 program promises many opportunities for exciting interactions—but to fulfill that promise we need you!

Please join us this August in New York City!

KDD-98 Organization

General Conference Chair
Gregory Piatetsky-Shapiro, *Knowledge Stream Partners*

Program Cochairs
Rakesh Agrawal, *IBM Almaden Research Center*
Paul Stolorz, *Jet Propulsion Laboratory*

Publicity Chair
Foster Provost, *Bell Atlantic Science and Technology*

Tutorial Chair
Padhraic Smyth, *University of California, Irvine*

Panel Chair
Willi Kloesgen, *GMD, Germany*

Workshops Chair
Ronny Kohavi, *Silicon Graphics*

Exhibits Chair
Ismail Parsa, *Epsilon*

Poster Sessions Chair
David Jensen, *University of Massachusetts, Amherst*

Local Arrangements Chair
Kyusoek Shim, *Bell Laboratories*

Sponsorship Chair
Ramasamy Uthurusamy, *General Motors Corporation*

Program Committee Members:

Rakesh Agrawal, *IBM Almaden Research Center*
Tej Anand, *Golden Books Family Entertainment*
Chid Apte, *IBM TJ Watson Research Center*
Andreas Arning, *IBM, Germany*
Roberto Bayardo, *IBM Almaden Research Center*
Carla Brodley, *Purdue University*
Wray Buntine, *Ultimode Systems*
Michael Burl, *Jet Propulsion Laboratory*
Soumen Chakrabarti, *IBM Almaden Research Center*
Ernest Chan, *Credit Suisse First Boston*
Surajit Chaudhuri, *Microsoft Research*
Corinna Cortes, *AT&T Laboratories*
Bruce Croft, *University of Massachusetts, Amherst*
Umeshwar Dayal, *Hewlett Packard Laboratories*
Pedro Domingos, *Instituto Superior Tecnico, Portugal*
Sue Dumais, *Microsoft Research*
William Eddy, *Carnegie Mellon University*
Charles Elkan, *University of California, San Diego*
Christos Faloutsos, *Carnegie Mellon University*
Tom Fawcett, *Bell Atlantic Science and Technology*
Usama M. Fayyad, *Microsoft Research*
Ronen Feldman, *Bar-Ilan University, Israel*
Stephen Gallant, *Knowledge Stream Partners*
Clark Glymour, *University of California, San Diego*
Moises Goldszmidt, *SRI International*
Georges Grinstein, *University of Massachusetts, Lowell*
Dimitrios Gunopulos, *IBM Almaden Research Center*
Jiawei Han, *Simon Fraser University, Canada*
David Hand, *Open University, UK*
David Heckerman, *Microsoft Research*
Tomas Imielinski, *Rutgers University*
Yannis Ioannidis, *University of Athens, Greece and University of Wisconsin, Madison*
H.V. Jagadish, *AT&T Laboratories*
David Jensen, *University of Massachusetts, Amherst*
George H. John, *Epiphany*
Pete Johnson, *Mellon Bank Strategic Technology*
Michael Jordan, *Massachusetts Institute of Technology*
Daniel Keim, *University of Halle-Wittenberg, Germany*
Willi Kloesgen, *GMD, Germany*
Ronny Kohavi, *Silicon Graphics*
Hans-Peter Kriegel, *University of Munich, Germany*
T.Y. Lin, *San Jose State University*
Peter Lockemann, *Universitaet Karlsruhe, Germany*
Hongjun Lu, *National University of Singapore, Singapore*
Neil Mackin, *WhiteCross Data Exploration*
David Madigan, *University of Washington*
Heikki Mannila, *University of Helsinki, Finland*
Brij Masand, *GTE Laboratories*
Gary McDonald, *General Motors Global Research and Development Operations*
Eric Mjolsness, *Jet Propulsion Laboratory*
Sally Morton, *Rand Corporation*
Richard Muntz, *University of California, Los Angeles*
Raymond Ng, *University of British Columbia, Canada*
Shojiro Nishio, *Osaka University, Japan*
Ismail Parsa, *Epsilon*
Gregory Piatetsky-Shapiro, *Knowledge Stream Partners*
Daryl Pregibon, *AT&T Laboratories*
Foster Provost, *Bell Atlantic Science and Technology*
Raghu Ramakrishnan, *University of Wisconsin, Madison*
Patricia Riddle, *University of Auckland, New Zealand*
Ted Senator, *National Association of Securities Dealers*
Jude Shavlik, *University of Wisconsin, Madison*
Wei-Min Shen, *University of Southern California*
Kyusoek Shim, *Bell Laboratories*
Arno Siebes, *CWI, The Netherlands*
Evangelos Simoudis, *IBM*
Padhraic Smyth, *University of California, Irvine*
Ramakrishnan Srikant, *IBM Almaden Research Center*
Salvatore J. Stolfo, *Columbia University*
Paul Stolorz, *Jet Propulsion Laboratory*
Kurt Thearling, *Exchange Applications*
Hannu Toivonen, *University of Helsinki, Finland*
Shalom Tsur, *Hitachi America*
Michael Turmon, *Jet Propulsion Laboratory*
Alexander Tuzhilin, *New York University*
Jeffrey D. Ullman, *Stanford University*
Ramasamy Uthurusamy, *General Motors Corporation*
Graham Williams, *CSIRO Mathematical and Information Sciences, Australia*
David Wolpert, *NASA Ames Research Center*
Jan M. Zytkow, *University of North Carolina*

Invited Speakers, Panels, and Tutorials

Keynote Address and Invited Speakers

The keynote address and invited speaker program will be announced on the KDD-98 web site (www-aig.jpl.nasa.gov/public/kdd98/)

Panels

Behind-the-Scenes Data Mining

Moderator: George H. John, Epiphany
Panelists: Graham Spencer, Excite; Gerald Fahner, Fair, Isaac & Co.; and Paul DuBose Analytika, Inc.

Truly successful technologies become invisible. Data mining has a long way to go—or does it? Most stories of data mining applications focus on an expert's use of raw technology to solve one particular problem, but millions of people use or are affected by data mining technology every day, without even being aware of it. The panelists will discuss examples and characteristics of this "behind-the-scenes" variety of data mining.

Database-Data Mining Coupling

Moderator: Rakesh Agrawal, IBM Almaden Research Center
Please see www-aig.jpl.nasa.gov/public/kdd98/ for details.

Privacy and Data Mining

Moderator: Ellen Spertus, Mills College and Computer Professionals for Social Responsibility
Panelists: Jason Catlett, Junkbusters Corp.; Dan Jaye, Engage Technologies; and Daryl Pregibon, AT&T Labs

Data mining allows unprecedented opportunities for targeted marketing, which can be seen either as a boon for advertisers and consumers or as an enormous invasion of privacy. This panel considers ownership of personal information and explores how maximum benefits can be obtained from data mining while respecting individuals' privacy.

KDD-98 Tutorial Program

Tutorials will be presented on Thursday, August 27 and Friday, August 28, 1998. Admission to four tutorials and their accompany syllabi are included in the conference registration fee.

Thursday, 8/27

Session 1

2:00 – 4:00 PM	T1—Mannila (single session)
4:30 – 6:30 PM	T2—Jagadish and Faloutsos (single session)

Friday, 8/28

Session 2

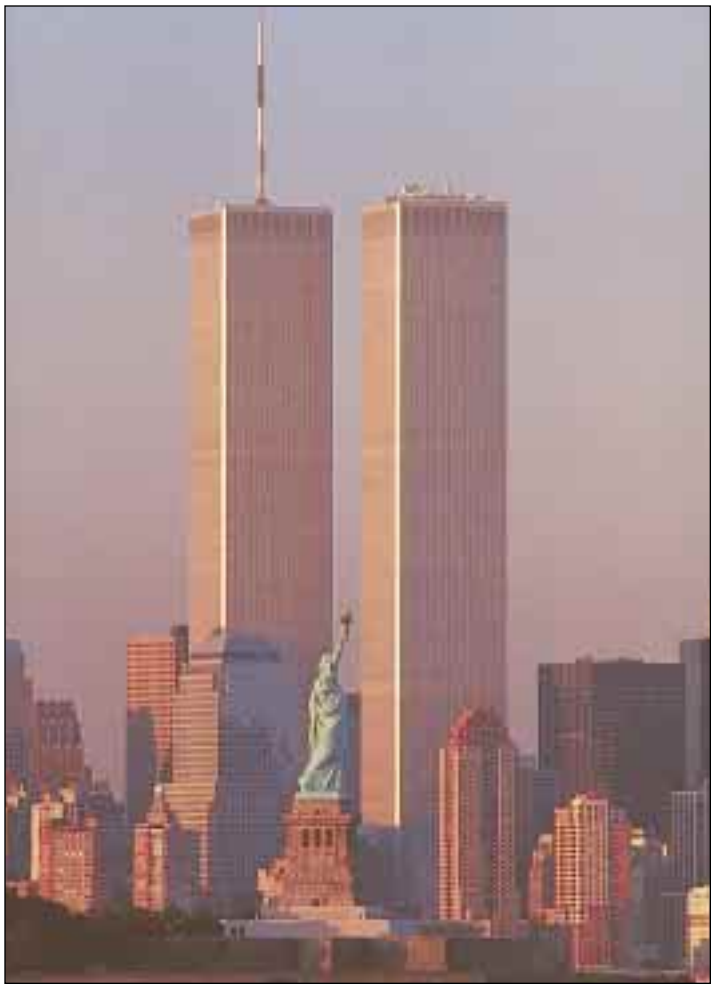
8:00 – 10:00 AM	T3 – Grossman and Bailey
10:30 AM – 12:30 PM	T6 – Marron

Session 3

8:00 – 10:00 AM	T4 – Donoho and Bennett
10:30 AM – 12:30 PM	T7 – Provost and Jensen

Session 4

8:00 – 10:00 AM	T5 – Banks and Levenson
10:30 AM – 12:30 PM	T8 – Elder and Abbott



Tutorial 1

Database Methods for Data Mining

Heikki Mannila, University of Helsinki, Finland

Thursday, August 27, 2:00 – 4:00 PM,

The tutorial explains the basic database techniques that can be used in data mining. We start by discussing the basics of database methods and the ideas behind data warehousing. Then we describe OLAP (on-line analytical processing) and the similarities and differences between OLAP and data mining.

The main part of the tutorial shows how simple descriptive patterns such as association rules, sequential patterns, and integrity constraints can efficiently be found from large databases using relatively straightforward methods. After that we discuss methods and data structures that can be used to store large amounts of multidimensional data, and how learning algorithms can be scaled to handle large datasets. The final part of the tutorial concerns emerging trends such as discovery from semistructured data such as web pages.

Heikki Mannila is a professor of computer science at the University of Helsinki in Finland. He currently works on data mining on event sequences and semistructured data, foundations of data mining and problems of fitting large statistical models on big datasets. He is Editor-in-Chief of the *Data Mining and Knowledge Discovery* journal.

Tutorial 2

Data Reduction

H. V. Jagadish, AT&T Laboratories and Christos Faloutsos, Carnegie Mellon University and University of Maryland, College Park

Thursday, August 27, 4:30 – 6:30 PM

Given a warehouse with a very large data set, we describe how to reduce the data set size used for analysis, through appropriate approximate representations. In the data mining process, there is a critical data representation step, typically including data reduction, after data acquisition and cleaning, but before data analysis. In this tutorial we will learn about the choices one can make in this step, and the tradeoffs involved.

The central issue in data reduction is to control the error introduced by the approximation and to trade this off against the savings in storage and processing cost. Given the rich diversity of data analysis techniques, a matching diversity of corresponding data reduction techniques is required. We will describe several of these techniques, including some recent database tools for data mining in massive datasets: feature extraction for multimedia indexing by content, singular value decomposition for lossy compression, and methods for information reconstruction.

Prerequisites: Familiarity with B-trees, and with basic concepts of relational database systems (tables, attributes, joins)

Christos Faloutsos is working on physical data base design, text and spatial access methods, multimedia databases and data mining. He has received two "best paper" awards (SIGMOD-94, VLDB-97). He has filed for three patents, and he has published over 70 refereed articles and one monograph. He is currently on leave at Carnegie Mellon University.

H. V. Jagadish obtained his Ph.D from Stanford University in 1985, and has since been with AT&T, where he currently heads the Database Research Department. He has published more than seventy-five articles, and has filed more than thirty patents. He was the co-organizer of an expert workshop that resulted in the "New Jersey Data Reduction Report."

Tutorial 3

A Tutorial Introduction to High Performance Data Mining

Robert Grossman, Magnify, Inc. and National Center for Data Mining, University of Illinois at Chicago and Stuart Bailey, National Center for Data Mining, University of Illinois at Chicago

Friday, August 28, 8:00 – 10:00 AM

Data mining is automatic discovery of patterns, associations, changes, anomalies, and statistically significant structures and events in data. Scaling data mining to large data sets is a fundamental problem, with important practical applications. The goal of the tutorial is to provide researchers, practitioners, and advanced students with an introduction to mining large data sets by exploiting techniques from high performance computing and high performance data management. We will describe several architectural frameworks for high performance data mining systems and discuss their advantages and disadvantages. We will use several case studies involving mining large data sets, from 10 – 1000 gigabytes in size, as running examples.

Robert Grossman is the President of Magnify, Inc. and Director of the National Center for Data Mining at the University of Illinois at Chicago. He has been a leader in the development of high performance and wide area data mining systems for over ten years. He has published widely on data mining and related areas and speaks frequently on the subject.

Stuart Bailey is a member of the technical staff at the National Center for Data Mining at the University of Illinois at Chicago. He has led the software development effort for the Terabyte Challenge Data Mining Demonstrations during the last three Supercomputing Conferences.

Tutorial 4

Fraud Detection and Discovery

Steven K. Donoho and Scott W. Bennett, SRA International, Inc.

Friday, August 28, 8:00 – 10:00 AM

This tutorial covers automated techniques for detecting fraud in areas such as health care, insurance, banking, telecommunications, and finance. Particular emphasis is given to how knowledge discovery techniques can be used to discover new fraud scenarios. Fraud detection is ripe for the application of KDD techniques because of the large volume of data involved and the amount of money lost each year to fraud. Of the over $1 trillion spent on health care each year, the U.S. General Accounting Office estimates that fraud accounts for 3 to 10 percent. Cellular phone fraud in the US is estimated at $5 billion per year.

We cover typical fraud detection workflow issues such as realtime vs. offline detection, alert investigation, gathering further information, alert explainability, and actions taken when fraud is still suspected after investigation. KDD issues covered include profiling, supervised and unsupervised approaches, analyzing sequences of events, practices vs. isolated incidents, and concept drift. Participants will go away with an understanding of fraud detection work to date, how discovery fits into the larger scheme of fraud detection, and challenges, pitfalls, and open research issues related to fraud detection.

Steve Donoho is a research scientist at SRA International. He received his Ph.D in 1996 from the University of Illinois focusing on automated change of representation techniques. His work at SRA has included customizing KDD techniques to detect fraud on the NASDAQ Market, devising scalable techniques for very large data sets, and developing data visualization tools for analyzing suspected fraud cases.

Scott Bennett is Technology Director of the Intelligent Information Systems Division at SRA International. His data mining experience includes work with very large structured and unstructured data sets, parallel implementations, discovery and detection algorithms, and graphical tools for analysis of mining results. He received his Ph.D in 1993 from the University of Illinois in machine learning and planning.

Tutorial 5

New-Wave Nonparametric Regression Methods for KDD

David Banks and Mark S. Levenson,
National Institute of Standards and Technology

Friday, August 28, 8:00 – 10:00 AM

In the last decade, statisticians have developed a group of modeling procedures designed to handle large, high-dimensional datasets. These procedures, which we call new-wave nonparametric regression methods, tend to be very flexible and interpretable. They offer competitive alternatives both in applicability and feasibility to traditional statistical and machine learning methods. The first half of this tutorial will review the current front-running methods, including multivariate adaptive regression splines (MARS), projection pursuit regression (PPR), additive models (AM), recursive partitioning regression (known commercially as CART), alternating conditional expectations (ACE), an approach based on variance stabilization (AVAS), neural nets, and locally weighted regression (LOESS), as well as classical multiple linear regression and stepwise regression. During the second half of the tutorial, we will address practical considerations in the choice and use of the methods. This will be based on a large-scale simulation experiment and examples of the methods applied to various datasets. Freeware software will be used to demonstrate the procedures.

David Banks is a mathematical statistician, at the National Institute of Standards and Technology (and formerly an associate professor of statistics at Carnegie Mellon University). He is coeditor of the *Encyclopedia of Statistical Sciences*, and chair-elect of the Classification Society of North America. An applied statistician, he pursues forays into mathematics and information technology.

Mark Levenson is a mathematical statistician at the National Institute of Standards and Technology. He received a Ph.D in 1993 from the Department of Statistics at The University of Chicago. His research interests are in the areas of image processing, data mining, and statistical problems in the physical and engineering sciences.

Tutorial 6

Smoothing Methods for Learning from Data

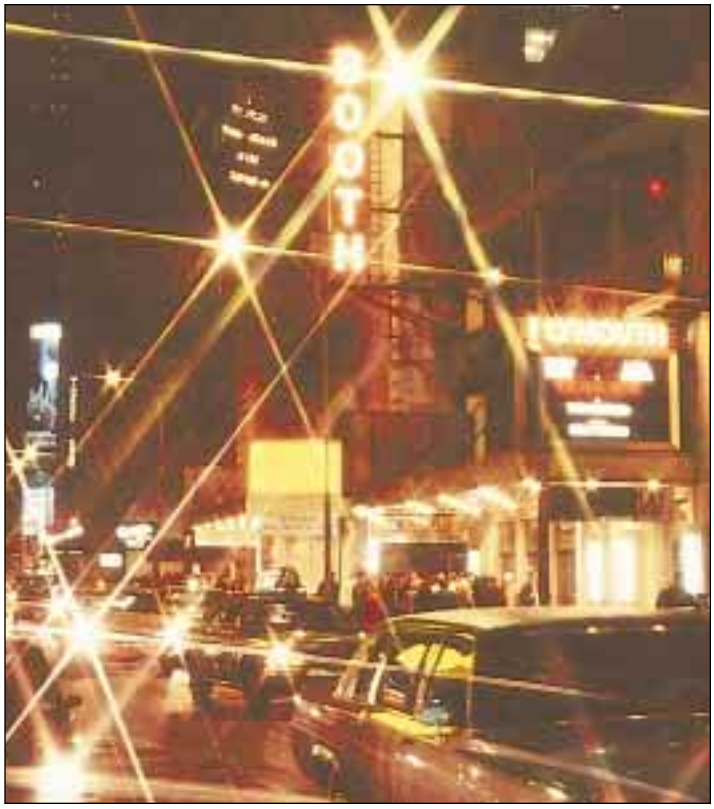
J. S. Marron, University of North Carolina

Friday, August 28, 10:30 AM – 12:30 PM

Real data examples are used to illustrate how smoothing methods provide a powerful tool for gaining insights from data. The three crucial issues for the application of smoothing methods in massive data contexts are, first, computational methods that scale appropriately; second, choice of window width; and third, assessing which observed features (e.g. "bumps") are "really there" (instead of being sampling artifacts). Nonobvious, recent approaches to each of these issues are described and assessed.

The fast smoothing method that scales up well is called "binning" or "WARPing," and will be explained together with some enhancements. The large literature on data based choice of window width for smoothing methods will not be reviewed in detail since the "family approach," seems more useful for data mining applications. Determination of statistical significance of features will be addressed from the viewpoints of formal mode testing, and SiZer (based on scale space ideas).

J. S. Marron is a professor of statistics at the University of North Carolina, Chapel Hill. In 1982, he earned a Ph.D in mathematics from the University of California, Los Angeles. His research interests include: statistical smoothing methods, smoothing parameter selection, fast implementations of smoothers and curves and surfaces as data.



Tutorial 7

Evaluating Knowledge Discovery and Data Mining

Foster Provost, Bell Atlantic Science and Technology and David Jensen, University of Massachusetts, Amherst

Friday, August 28, 10:30 AM – 12:30 PM

Both the science and practice of KDD stand to benefit from a common understanding of the strengths and limitations of the many frameworks for evaluating results. We will explain and criticize a wide variety of evaluation techniques, illustrating the similarities, but focusing on the important small differences. We first discuss the difference between evaluating models and evaluating model-building algorithms, which leads into a description of the traditional scientific frameworks for comparing KDD results. We then show where these frameworks are weak statistically and recommend techniques for strengthening them. Next, we discuss weaknesses of these frameworks when it comes to the practical application of data mining results. We show how to make evaluations more robust for a wide variety of real-world data mining scenarios, comparing and contrasting metrics such as sensitivity, specificity, positive predictive value, precision, and recall, and frameworks such as lift and ROC curves. Finally, expanding our view, we consider the general problem of searching for interesting patterns. We describe a diverse collection of techniques, including Bayesian and Bonferroni adjustments, blindfold trials, interestingness criteria, and the use of prior domain knowledge.

Foster Provost's research concentrates on weakening the simplifying assumptions that prevent inductive algorithms from being applied successfully. He received his Ph.D in computer science from the University of Pittsburgh in 1992. He has worked on automated knowledge discovery in science, and is currently with Bell Atlantic Science and Technology.

David Jensen is research assistant professor of computer science at the University of Massachusetts, Amherst. His research focuses on learning and KDD, particularly the statistical properties of KDD algorithms. He is managing editor of *Evaluation of Intelligent Systems*, a web-accessible resource about empirical methods for studying AI systems.

Tutorial 8

A Comparison of Leading Data Mining Tools

John F. Elder IV and Dean W. Abbott, Elder Research

Friday, August 28, 10:30 AM – 12:30 PM

Several high-performance, but costly, software products for knowledge discovery and data mining have recently been introduced. Most feature multiple modeling and classification algorithms and/or increased support for key data-handling and interpretation stages of the KDD process. Still, they compete with a healthy (and growing) lineup of desktop products vying for survival—some of which are focused on particular vertical markets. Natural questions to ask include "Which product is best?" "Will a general-purpose tool suffice for my application?" and "Are the high-end ones worth it?" This tutorial will address such questions by providing an overview of the current field of data mining software tools. The instructors will highlight the distinctive properties and relative strengths of several major products, and share practical insights and observations from their use.

Though technical in parts, the tutorial should benefit both "enterprise technologists" and "line-of-business executives." All participants having the focus of a practical application to solve should gain insight into tools likely to add near-term value. A brief outline of the tutorial includes descriptions of algorithms (classical statistical, neural network, decision tree, polynomial network, density-based, rule induction); survey of high-end products (ease of use, comparative strengths, distinctive properties, cost); and lower-end products with complementary abilities (as time allows).

John Elder heads a small, but growing, data mining research and consulting firm in Charlottesville, Virginia. He chairs the Adaptive and Learning Systems Group of the IEEE-SMC, is an adjunct at the University of Virginia, has created influential DM algorithms and short courses, and writes and speaks often on KDD.

Dean Abbott is a senior research scientist at Elder Research, in San Diego, California. He has engineering and mathematics degrees from the University of Virginia and Rensselaer Polytechnic Institute, and experience at three consulting firms. An expert in pattern discovery, Abbott has designed and implemented algorithms for commercial DM software.

Workshop Program

Workshop KW1

Data Mining in Finance

Workshop Chairs

Tae Horn Hann
Institute for Statistics and Mathematical Economics
University of Karlsruhe
Rachenzentrum, Zirkel 2
76128 Karlsruhe, Germany
E-mail: thh@vwl3sun1.wiwi.uni-karlsruhe.de
Telephone: 49 721 608 3383
Fax: 49 721 608 3491

Gholamreza Nakhaeizadeh
Daimler-Benz AG, Research and Technology
Postfach 2360
89013 Ulm, Germany
E-mail: nakhaeizadeh@dbag.ulm.DailmerBenz.com

Workshop KW2

Distributed Data Mining

Workshop Organizers

Hillol Kargupta
Faculty of Computer Science
School of Electrical Engineering and Computer Science
Washington State University
Pullman, WA 99164-2752
E-mail: hillol@eecs.wsu.edu
Telephone: (509) 335-6602
Fax: (509) 335-3818

Philip Chan
Computer Science
Florida Institute of Technology
150 West University Boulevard
Melbourne, FL 32901
Email: pkc@cs.fit.edu

Workshop KW3

Keys to the Commercial Success of Data Mining

Workshop Organizers

Kurt Thearling
Director of Advanced Analytics
Exchange Applications, Boston

Roger M. Stein
Vice President, Senior Credit Officer
Quantitative Analytics and Knowledge Based Systems
Moody's Investors Service, New York
E-mail: kdd-workshop@exapps.com

Exhibits and Demonstrations

Following the success of the demonstration sessions in previous KDD conferences, the KDD-98 program will also include demonstrations of knowledge discovery products, knowledge discovery applications and research prototypes. We will clearly differentiate between commercial product demonstrations and research demonstrations.

Exhibits

We invite all commercial knowledge discovery and data mining tool vendors, consultants providing related services, academics with research prototypes, publishers and corporations with significant applications to exhibit at KDD-98. The exhibitor fee for KDD-98 will be a nominal $500.00. Exhibitors will be provided with a 6 foot table top. In this space, exhibitors will be allowed to distribute product/service/company literature, display product/service demonstrations and set up signage. Exhibitors are responsible for bringing all hardware and software required for their demonstrations.

The exhibit area will be open Friday, August 28 from 2:00 PM – 6:00 PM and Saturday, August 29 from 10:00 AM – 6:00 PM. Total attendance at KDD-97 was 618. Of these 24 percent were affiliated with universities and 76 percent were affiliated with industry. If you would like to exhibit at KDD-98, please contact AAAI at kdd@aaai.org, 650-328-3123 (telephone), or 650-321-4457 (fax) or write to:

AAAI
KDD-98 Exhibits
445 Burgess Drive
Menlo Park, CA 94025-3442 USA

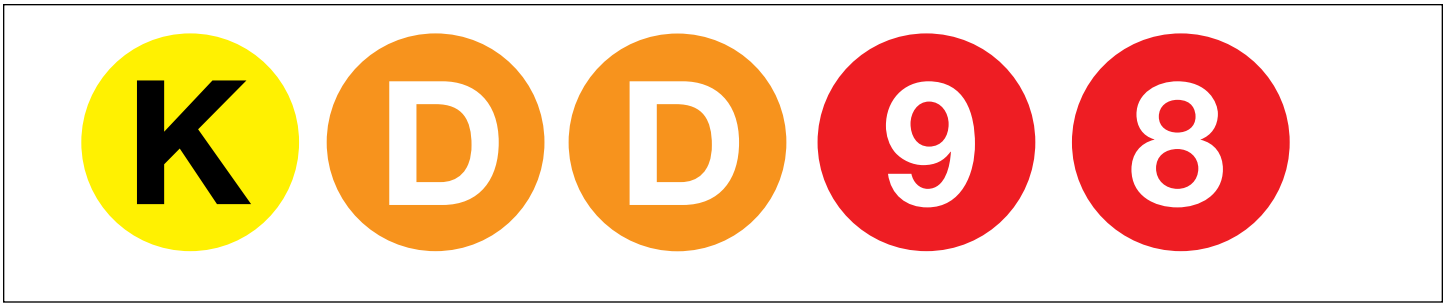
Research Prototype Demonstrations

We are also soliciting demonstrations of research prototypes at KDD-98. This demonstration session will be held on Friday, August 28 from 2:00 PM – 6:00 PM and Saturday, August 29 from 10:00 AM – 6:00 PM. This year we will give priority to demonstrations in conjunction with accepted papers at KDD-98. Within budget and space constraints, we will make every effort to accommodate as many demonstrations as possible. If you would like your demonstration to be considered for KDD-98, please provide the following information to Ismail Parsa at iparsa@epsilon.com, fax (781) 272-8604 or mail to:

Ismail Parsa
Epsilon
50 Cambridge Street
Burlington, MA 01803 USA

Please provide the following information:

- Name of demonstration
- Title of paper (if this demonstration is in conjunction with a paper/poster at KDD-98)
- Development team
- Affiliations of development team members
- Contact telephone number
- Contact address (email & postal)
- Description of demonstration (200 words maximum)
- What is unique about your system or application? (No more than 50 words)
- Status: Is the system a research prototype, a commercially available product, or a fielded application?
- Will you bring your own hardware? If no, please specify hardware requirements
- Operating system
- WAN connection required? (If yes, please state any special modem requirements)
- Any other requirements?



Registration Fees

The KDD-98 program registration includes admission to all technical paper, poster, and exhibit/demo sessions, four tutorials and the accompanying four tutorial syllabi, the *KDD-98 Conference Proceedings*, Thursday and Saturday evening receptions, and coffee breaks. Onsite registration will be located in the foyer outside the Westside Ballroom, New York Marriott Marquis, fifth floor. Individuals who are also registering for the 24th International Conference on Very Large Databases (VLDB'98) may deduct $20 from the appropriate fee below. For information about how to register for VLDB'98, please see www.research.att.com/conf/vldb98.

Early Registration

(Postmarked by July 15)

AAAI Members			
Regular	$350	Students	$100
Nonmembers			
Regular	$430	Students	$160

Late Registration

(Postmarked by August 5)

AAAI Members			
Regular	$410	Students	$130
Nonmembers			
Regular	$490	Students	$190

On-Site Registration

(Postmarked after August 5 or onsite.)

AAAI Members			
Regular	$470	Students	$160
Nonmembers			
Regular	$550	Students	$220

Workshop Registration

Workshop registration is limited to those active participants determined by the organizer prior to the conference. All workshop participants must preregister. Workshop registration materials will be sent directly to invited participants. The workshop registration fee is $100.

Payment Information

Prepayment of registration fees is required. Checks, international money orders, bank transfers and travelers' checks must be in US dollars. American Express, MasterCard, VISA, and government purchase orders are also accepted. Registration applications postmarked after the early registration deadline will be subject to the late registration fees. Registration applications postmarked after the late registration deadline will be subject to on-site registration fees. Student registrations must be accompanied by proof of full-time student status.

Refund Requests

The deadline for refund requests is August 12, 1998. All refund requests must be made in writing. A $75.00 processing fee will be assessed for all refunds granted.

Registration Hours

Registration hours will be as follows:

Thursday, August 27	11:00 AM – 6:00 PM
Friday–Saturday, August 28–29	7:30 AM – 6:00 PM
Sunday, August 30	8:00 AM – 3:00 PM
Monday, August 31	8:00 AM – 12:00 PM.

All attendees must pick up their registration packets for admittance to programs.

General Information

Housing

AAAI has reserved a block of rooms at the New York Marriott Marquis Hotel at reduced conference rates. Conference attendees must contact the hotel directly and identify themselves as KDD-98 registrants to qualify for the reduced rates. Hotel rooms are priced as singles (1 person, 1 bed), and doubles (2 persons, 2 beds). Rooms will be assigned on a first-come, first-served basis. Hotel rooms are subject to applicable state and local taxes (currently 8.25% New York State Tax, 5% New York City Tax and a $2.00 per room per day Occupancy Tax) in effect at the time of check in.

Headquarters Hotel
New York Marriott Marquis
1535 Broadway
New York, New York 10036
Telephone: 212-398-1900
Fax: 212-704-8930
Single room: $147.00
Double room: $147.00
Additional Person: $15.00
Check-in time: 3:00 PM
Check-out time: 12:00 noon
Cut-off date: August 5, 1998.

All reservation requests must be accompanied by a first night room deposit, or guaranteed with a major credit card. The hotel will not hold any reservations unless guaranteed by one of the above methods. All reservation requests will require a one night's advance deposit that will be refundable only if the reservation is canceled at least twenty-four hours prior to the arrival date.

Transporation

The following information provided is the best available at press time. Please confirm fares when making reservations.

Air Transportation and Car Rental

New York City, New York – Get there for less! Discounted fares have been negotiated for this event. Call Conventions in America at 800-929-4242 and ask for Group #428. You will receive 5 – 10 percent off the lowest applicable fares on American Airlines and United Airlines, or the guaranteed lowest available fare on any carrier. Take an additional 5 percent off if you purchase at least 60 days prior to departure. Travel between August 24 – September 3, 1998. All attendees booking through CIA will receive free flight insurance. Avis Rent A Car is also offering special low rates, with unlimited free mileage. Call Conventions in America at 1-800-929-4242, ask for Group #428. Reservation hours: Monday – Friday, 6:30 AM – 5:00 PM Pacific time. Outside US and Canada, call 619-453-3686 or fax 619-453-7976. E-mail: address flycia@scitravel.com If you call direct: American 1-800-433-1790, ask for index #10309, United 800-521-4041, ask for tour code #512QW, Avis 800-331-1600 ask for AWD #J947822.

Airport Connections

Gray Line Air Shuttle: John F. Kennedy, LaGuardia, and Newark International Airports: 800-451-0455. The fare from JFK, LaGuardia, or Newark Airports to the New York Marriott Marquis Hotel is $14.00 per person. Traveler's checks and cash are accepted.

Taxi

Taxis are available at JFK, LaGuardia, and Newark Airports to the New York Marriott Marquis Hotel . The approximate fare from JFK to the New York Marriott Marquis is $30, from LaGuardia, $25, and from Newark, $45.

Bus

Port Authority Bus Terminal: The depot is located at Eighth Avenue at 40th-42nd Streets, Manhattan. For information on fares and scheduling, call 212-564-8484 or 201-659-8823.

Rail

Amtrak—National Railroad Passenger Corporation is located at Penn Station, Seventh Avenue (31st–33rd Streets), New York City. For general information and ticketing, call 800-USA-RAIL.

Metro North Railroad operates from the Grand Central Terminal to 119 stations in New York, Connecticut and New Jersey. The Grand Central Terminal is located at 42nd Street and Park Avenue. For general information and ticketing, call 800-METRO-INFO or 212-532-4900.

City Transit System

Public Buses and Subways: MTA — New York City Transit Fare is $1.50 regardless of distance traveled. (Exact change, subway token or MetroCard required.) For general information call 718-330-1234.

Parking

Parking is available at the New York Marriott Marquis Hotel. The rate for valet parking for the first six hours is $20.00 and for 24 hour parking $30.00. There is no self parking.

KDD-98 Preregistration Form

Name _____ Company/Univ. _____

☐ Home ☐ Work

Address _____ Dept./MS _____

City _____ State _____ Zip _____

Country _____ Daytime Telephone & Fax _____

Membership No. _____ E-mail Address _____

Circle courses and fees that apply. Students must submit registration receipt or letter from faculty advisor.	**Early Registration Rates** *Postmarked by July 15*				**Late Registration Rates** *Postmarked by August 5*				**TOTAL**
	AAAI Member		**Nonmember**		**AAAI Member**		**Nonmember**		
	Regular	**Student**	**Regular**	**Student**	**Regular**	**Student**	**Regular**	**Student**	
KDD-98	$350	$100	$430	$160	$410	$130	$490	$190	_____
KDD-98 Fees for Persons also preregistering for VLDB '98	$330	$80	$410	$140	$390	$110	$470	$170	_____
Tutorials *(Circle choices—only 1 per line)* Thursday T1 – – / T2 – – / *Friday:* T3 T4 T5 / T6 T7 T8	*Your conference registration includes admission to four consecutive tutorials and the accompanying syllabi. The other four syllabi are available for purchase at $60.00 per set.*								
	Additional sets at $60.00 per set: no. sets: _____								_____
Welcoming Reception *(Included in registration fee)*	*Spouse, child, or guest @ $15.00 per person* no. persons: _____								_____
Discovery Reception *(Included in registration fee)*	*Spouse, child, or guest @ $25.00 per person* no. persons: _____								_____
AAAI Membership	*(continued from reverse)*								_____
AAAI Sponsored Journals	*(continued from reverse)*								_____
AAAI Press Books	*(continued from reverse)*								_____
***Data Mining & Knowledge Discovery* Journal** *KDD-98 attendees are eligible to purchase a one-year subscription to Kluwer's Data Mining and Knowledge Discovery Journal at the discounted rate of $55.00*									_____

Method of Payment *(Circle One):*

MC Visa Amex Check (payable to KDD–98 and drawn on a US Bank) US Govt. PO

Card Number _____ Exp. Date _____

Name (as it appears on card) _____ Signature _____

Total Enclosed

_____

The refund request deadline is August 12, 1998. A $75.00 processing fee will be assessed for all refunds.
Registrations postmarked after August 5 are subject to on-site rates.

On-site registration will be located in the foyer outside the Westside Ballroom.
New York Marriott Marquis Hotel,
1535 Broadway, New York, New York 10036

Send with payment to KDD-98
445 Burgess Drive, Menlo Park, CA 94025-3442. 650 / 328-3123; Fax 650 / 321-4457

AAAI Membership Application / Renewal

Now it's even easier to become a member of the AAAI. Just fill out and mail both sides of this form, and we'll ensure that you receive all the benefits that thousands of members worldwide enjoy each year.

Here are just a few of the benefits you'll receive:

• *AI Magazine* • AAAI Electronic Library Access • Reduced rates on selected AI-related journals and publications

Information about all of AAAI's events and programs, including:

• Spring and Fall Symposium Series • AAAI Press Publications • Conference on Innovative Applications
• Tutorial Program • Exhibit Program • KDD Conference
• Technical Program of the National Conference on Artificial Intelligence • AAAI–Sponsored Workshops

Take the initiative to join the association that will keep you informed about the latest developments in your exciting field.

Renew your membership or become a member of the AAAI today.

Application Type

☐ New Application

☐ Change of Address

☐ Renewal

(Please include your membership number on the reverse side of this form)

☐ Do not include me in the *Membership Directory*

☐ Do not release my name to outside groups

I am interested in the following subgroup:

☐ AI in Medicine n AI and the Law

☐ AI in Manufacturing n AI in Business

Journals

(Offer limited to individuals only).

☐ Send me the 1998 *AI Journal*. I enclose an additional $116.00.

☐ Send me the 1998 *Machine Learning Journal*. I enclose an additional $170.00

Books

(Offer limited to individuals only).

☐ Send me the *1995 KDD Proceedings*. I enclose an additional $54.50 (US destinations) or $56.00 (foreign destinations). *(California residents must add their local sales tax.)*

☐ Send me the *1996 KDD Proceedings*. I enclose an additional $59.50 (US destinations) or $61.00 (foreign destinations). *(California residents must add their local sales tax.)*

☐ Send me the 1997 KDD Proceedings. I enclose an additional $64.50 or $66.00 (foreign destinations). *(California residents must add their local sales tax.)*

☐ Send me *Advances in Knowledge Discovery and Data Mining*, edited by Usama M. Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth, and Ramasamy Uthurusamy. I enclose an additional $55.00 (US destinations) or $56.50 (foreign destinations). *(California residents must add their local sales tax.)*

Membership Categories

Please circle desired term and amount

Individual US / Canadian Member

One Year	*Three Year*	*Five Year*	*Life*
$50	$150	$250	$700

Individual Foreign Member

One Year	*Three Year*	*Five Year*	*Life*
$75	$225	$375	$1000

Institution / Library—US / Canadian

One Year	*Three Year*	*Five Year*	*Life*
$75	$225	$375	n/a

Institution / Library—Foreign

One Year	*Three Year*	*Five Year*	*Life*
$100	$300	$500	n/a

Full-Time US/Canadian Student

One Year	*Three Year*	*Five Year*	*Life*
$20	n/a	n/a	n/a

Full-Time Foreign Student

One Year	*Three Year*	*Five Year*	*Life*
$45	n/a	n/a	n/a

Please Note

Order cannot be processed if information is incomplete or illegible. Student applicants must send legible proof of student status, i.e., a letter from your faculty advisor verifying full-time enrollment in a degree-bearing program, or a copy of your current registration receipt. Prepayment is required for all orders. Memberships begin with the next published issue of AI Magazine.

Be sure to enter your complete name and address on the reverse side of this form!

Membership subtotal

_____

(Enter here and on reverse)

Books subtotal

_____

(Enter here and on reverse)

Journals subtotal

_____

(Enter here and on reverse)

New York, New York!

New York City, New York!

New York City offers incomparable attractions, world-renowned restaurants, hotels, theaters, entertainment and shopping. Whether you are visiting the Empire State Building, the United Nations, the World Trade Center, The Metropolitan Museum of Art , Times Square or Broadway you are assured memories to last a lifetime. Come visit the Big Apple and see why last year alone 32 million people visited the city.

Photographs In This Brochure

Page 2—Statute of Liberty.
Page 4—World Trade Center.
Page 5—Rockefeller Center.
Page 6—Metropolitan Museum of Art.
Page 7—Broadway.
Page 9—Times Square.
Page 10—Brooklyn Bridge.
Page 15—Empire State and Chrysler Buildings.

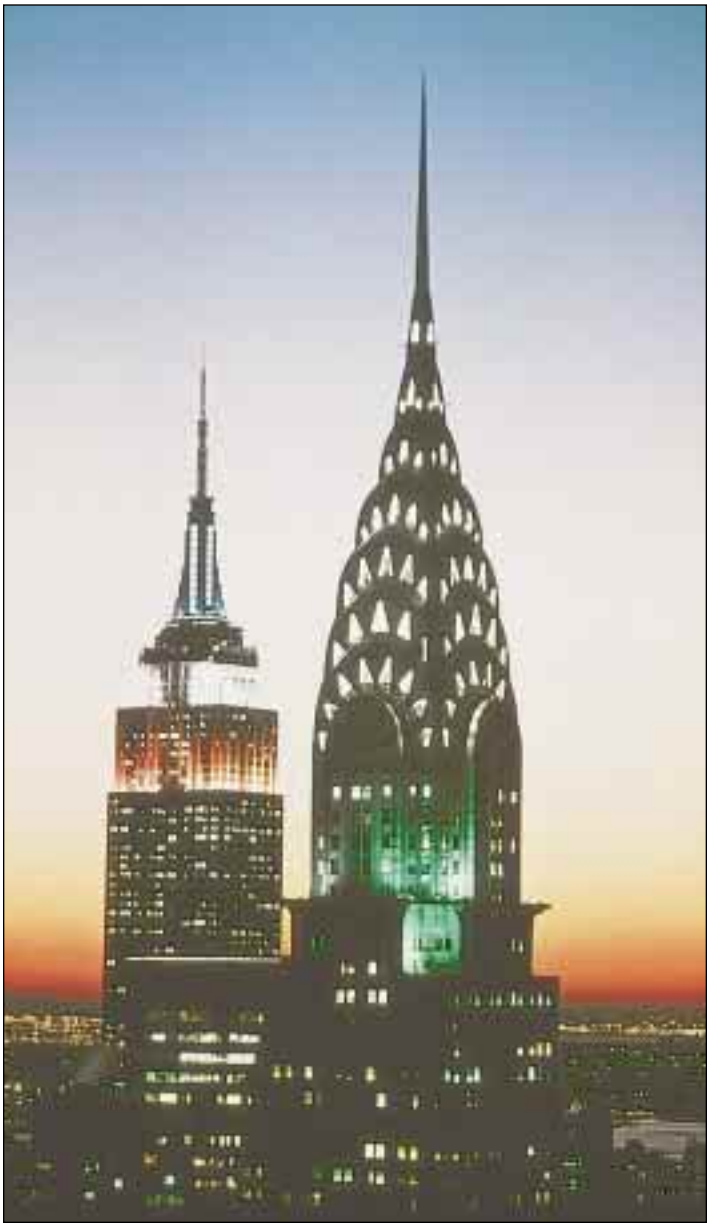
Photographs in this brochure are courtesy, New York Convention and Visitors Bureau.

Disclaimer

In offering American Airlines, Avis Rent A Car, the New York Marriott Marquis Hotel, United Airlines and all other service providers, (hereinafter referred to as "Supplier[s]") for the Fourth International Conference on Knowledge Discovery and Data Mining, AAAI acts only in the capacity of agent for the Suppliers which are the providers of the service. Because AAAI has no control over the personnel, equipment or operations of providers of accommodations or other services included as part of the KDD-98 program, AAAI assumes no responsibility for and will not be liable for any personal delay, inconveniences or other damage suffered by conference participants which may arise by reason of (1) any wrongful or negligent acts or omissions on the part of any Supplier or its employees, (2) any defect in or failure of any vehicle, equipment or instrumentality owned, operated or otherwise used by any Supplier, or (3) any wrongful or negligent acts or omissions on the part of any other party not under the control, direct or otherwise, of AAAI.

New York City Visitor Information

A concierge desk is available in the New York Marriott Marquis Hotel. They can assist with dining reservations, directions, tour bookings, entertainment suggestions, and transportation information. Maps and brochures are also available. Information about New York City is also available on the world wide web: www.nycvisit.com.



Contents