



Practical Data Mining: Applications Experiences and Challenges

Business, government and science organizations are increasingly moving towards decision-making processes that are based on information. In parallel, the amount of data representing the activities of organizations that is stored in databases is also growing. Therefore, the pressure to extract as much useful information as possible from this data is very strong. Data and text mining (DTM) is essentially a technological field enabling the improvement of the aforementioned processes. DTM research is, thus, essentially geared towards practical problems and requires permanent feedback from practical applications. Feedback from people with a business-oriented perspective on DTM is therefore useful to assess current research work and directions, to assess how theoretical results and methods developed in the lab stand the test of real world problems. And, not the least, to provide researchers with new challenges to work on.

Besides two excellent invited talks, you will learn about state-of-the-art applications in areas including marketing, medicine and justice, presented by leading researchers and practitioners from universities and companies, such as DeutschePost, Unilever and Capgemini. A detailed program is given in the next page.

Chairs

- Markus Ackermann (Institute for Computer Science, University of Leipzig)
- Carlos Soares (Faculty of Economics, University of Porto)
- Bettina Guidemann (SAS Deutschland)

Objectives

Contribute to the interaction between theory and practice by attracting people with a business, banking or industrial perspective on data and text mining to the ECML-PKDD conference, and to actually bring together the developers of new data mining methods with the practitioners applying them to business problems. We encourage the submission of:

- case studies and experience reports of the application of data mining (DM), text mining (TM) and business intelligence (BI) technologies, emphasizing the benefits as well as the lessons learned
- challenges and innovative applications of DM, TM & BI
- integration of DM, TM & BI with other technologies

Invited Speakers

Stefan Wrobel
(Director of the Fraunhofer Institute for Autonomous Intelligent Systems)

Geo Intelligence – New Business Opportunities and Research Challenges in Spatial Data Mining and Business Intelligence

Target Audience

- business managers and executives (marketing, finance, customer relationship, sales, product development, etc.)
- researchers from areas which can benefit from better data analysis tools and methodologies (bioinformatics, pharmaceuticals, economics, engineering, etc.)
- researchers and professionals on DM, BI and IT in general
- anyone who makes decisions based on data

Ulrich Reincke
(Executive Expert Analytics, SAS Institute)

Directions of Analytics, Data and Text Mining – A Software Vendor's View

More Information

<http://wortschatz.uni-leipzig.de/~macker/dmbizo6>

Contact

dmbiz@liacc.up.pt

The workshop on Practical Data Mining will be held in conjunction with the 17th European Conference on Machine Learning (ECML) and the 10th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD). Participants of the workshop may attend Friday morning's invited talk:

ECML/PKDD Invited Talk

Jonathan Schaeffer (University of Alberta)
Learning to Have Fun

Workshop Program

Oral Presentations

Customer churn prediction - a case study in retail banking
by Teemu Mutanen, Jussi Ahola and Sami Nousiainen
(VTT Technical Research Center of Finland)

Sequence Mining for Customer Behaviour Predictions in Telecommunications
by Frank Eichinger, Detlef D. Nauck and Frank Klawonn
(Univ. Karlsruhe, BT Group and Univ. Applied Sciences Braunschweig/Wolfenbüttel)

Onto Clustering of Criminal Careers
by Jeroen S. de Bruin, Tim K. Cox, Walter A. Kusters, Jeroen F.J. Laros and Joost N. Kok
(Leiden Univ.)

Bayesian Network Classifiers for Forecasting Country Risk
by Ernesto Coutinho Colla, Jaime Shinsuke Ide and Fabio Gagliardi Cozman
(Univ. of São Paulo)

Experiences from Automatic Summarization of IMF Staff Reports
by Shuhua Liu and Johnny Lindroos
(Abo Academy)

Online Ensembles for Financial Trading
by Jorge Barbosa and Luís Torgo
(Univ. of Porto)

Mining Medical Administrative Data - The PKB System
by Aaron Ceglar, Richard Morrall and John F. Roddick
(Flinders Univ. and PowerHealth Solutions)

Clustering of Psychological Personality Tests of Criminal Offenders
by Markus Breitenbach, Tim Brennan, William Dieterich and Gregory Z. Grudic
(Univ. of Colorado at Boulder, Univ. of Denver and Northpointe Institute for Public Management)

Predictive Marketing
by Jochen Werner
(SPSS)

Deployment of Association Analysis for Cross Selling Purposes at the Deutsche Post Philatelic
by Johann Domogalla
(DeutschPost)

Combination of different text mining strategies to classify requests about involuntary childlessness to an internet medical expert forum
by Wolfgang Himmel, Ulrich Reincke and Hans Wilhelm Michelmann
(Univ. of Göttingen and SAS Institute)

Poster Presentations

Carancho - A Decision Support System for Customs
by Norton Trevisan Roman, Everton Rufino Constantino, Helder Ribeiro, Jorge Jambeiro Filho, Antonella Lanna, Siome Klein Goldenstein and Jacques Wainer
(Univ. of Campinas)

Mining in Health Data by GUHA method
by Jan Rauch
(Univ. of Economics of Prague and Center of Biomedical Informatics)

Bridging the gap between commercial and open-source data mining tools: a case study
by Matjaz Kukar and Andrej Zega
(Univ. of Ljubljana)

Customer churn prediction using knowledge extraction from emergent structure maps
by Alfred Ultsch and Lutz Herrmann
(Phillipps Univ. of Marburg)

Sequential patterns extraction in multitemporal satellite images
by Andreea Julea, Nicolas Méger and Emmanuel Trouv
(Univ. Politechnica Bucuresti, Univ. of Savoie and Inst. de Stunte Spatiale Bucuresti)

Contrasting Large-Scale Document Collections: Applications, Experiences and Challenges
by Moustafa Ghanem, Cecilia Eyre and George Milis
(InforSense, Unilever and Imperial College of London)

Towards Better Understanding of CFB Boilers: Getting Domain Experts to Look at the Data Mining Perspective
by Mykola Pechenizkiy, Tommi Karkkainen, Andriy Ivannikov, Antti Tourunen and Heidi Hasa
(Univ. of Jyväskylä and VTT Processes)

An Investigation into a Beta-Carotene/Retinol Dataset Using Rough Sets
by Kenneth Revett, Florin Gorunescu and Marina Gorunescu
(Univ. of Westminster and Univ. of Medicine and Pharmacology of Cracovia)

Taping into the European Energy Exchange (www.eex.de) to feed the stochastic Portfolio Optimization Tool "SpOt" for Electric Utilities
by Ulrich Reincke, Heiner Lake, Michael Wigbels, Andre Rothig and Michael Lucht
(SAS Institute, Capgemini and Fraunhofer Institute)

Registration

Registration for the ECML/PKDD conferences includes participation in the workshop: <http://www.ecmlpkdd2006.org>

Registration for the **workshop only** is possible **until August 5** for EUR 290 and includes:

- printed proceedings of DMBiz workshop
- online access to all workshop and tutorial materials
- **conference banquet at Zitadelle Spandau** on Thursday evening
- coffee breaks on Friday

To register for the workshop only, please contact Prof. Tobias Scheffer (scheffer@informatik.hu-berlin.de).

