Advances in Data Mining

Applications in Image Mining, Medicine and Biotechnology, Management and Environmental Control, and Telecommunications

4th Industrial Conference on Data Mining, ICDM 2004
Leipzig, Germany, July 4-7, 2004
Revised Selected Papers
Preface

The Industrial Conference on Data Mining ICDM-Leipzig was the fourth meeting in a series of annual events which started in 2000, organized by the Institute of Computer Vision and Applied Computer Sciences (IBaI) in Leipzig.

The mission of the conference is to bring together researchers and people from industry in order to discuss together new trends and applications in data mining. This year a broad spectrum of work of different applications was presented ranging from image mining, medicine and biotechnology, management and environmental control, to telecommunications. Besides that an industrial exhibition showed the successful application of data mining methods by industries in different areas such as medical devices, mass data management systems, data mining tools, etc.

During the discussion many projects were inspired leading to new and joint work. The fruitful discussions, the exchange of ideas and the spirit of the conference made it a remarkable event for both sides, industry and research.

We would like to express our appreciation to the reviewers for their precise and highly professional work. We appreciate the help and understanding of the editorial staff at Springer and in particular Alfred Hofmann, who supported the publication of these proceedings in the LNAI series.

Last, but not least, we wish to thank all speakers, participants and industrial exhibitors who contributed to the success of the conference.

We are looking forward to welcoming you to ICDM 2005 (www.data-mining-forum.de) and to the new work you will present there.

July 2004

Petra Perner
4th Industrial Conference on Data Mining ICDM'2004

The Aim of the Conference

This conference is the fourth conference in a series of industrial conferences on Data Mining that will be held on a yearly basis. Experts from different fields will present their applications and the results obtained by applying data mining. Besides that, newcomers in the field can get a fast introduction to Data Mining by taking the tutorial running in connection with the conference. In a problem/solution hour you will have the opportunity to present your application and ask for support by others or for cooperation in solving the problem.

Paper submission

Paper submissions should be related but not limited to any of the following topics: Applications of Data Mining in...

- Marketing
- Medicine
- E-Commerce
- Biotechnology
- Quality Management
- Knowledge extraction from text, video, signals and images.

Potentials speakers should submit an abstract of their talk (not more than 2 pages long) to the e-mail address: ibai@aol.com at the:

Institut für Bildverarbeitung
und angewandte Informatik
Koernerstraße 10
04107 Leipzig/Germany

Telephone: +49 (0) 341 86 12 273
FAX: +49 (0) 341 86 12 275
Homepage: www.ibai-institut.de

For the latest news as well as for the program of the previous ICDM and the publications within ICDM look at:

http://www.data-mining-forum.de

The abstracts will be reviewed by the program committee. The abstract of accepted talks will be published in a special edition of an international journal by Springer Verlag. Extended versions of selected papers will be published after the workshop.

Former proceedings from ICDM have been published by Springer Verlag in Advances in Data Mining, P. Perner (Ed.).
# Table of Contents

## Case-Based Reasoning

- Neuro-symbolic System for Business Internal Control  
  *Juan M. Corchado, M. Lourdes Borrajo, María A. Pellicer, J. Carlos Yáñez* ………………………………………………………………………... 1

- Applying Case Based Reasoning Approach in Analyzing Organization Change Management Data  
  *Orit Raphaeli, Jacob Zahavi, Ron Kenett* ………………………………………. 11

- Improving the K-NN Classification with the Euclidean Distance Through Linear Data Transformations  
  *Leon Bobrowski, Magdalena Topczewska* ……………………………………… 23

- An IBR System to Quantify the Ocean’s Carbon Dioxide Budget  
  *Juan M. Corchado, Emilio S. Corchado, Jim Aiken* …………………………………… 33

- A Beta-Cooperative CBR System for Constructing a Business Management Model  
  *Emilio S. Corchado, Juan M. Corchado, Lourdes Sáiz, Ana Lara* ………………… 42

## Image Mining

- Braving the Semantic Gap: Mapping Visual Concepts from Images and Videos  
  *Da Deng* ……………………………………………………………………………………………………… 50

- Mining Images to Find General Forms of Biological Objects  
  *Petra Perner, Horst Perner, Angela Bühring, Silke Jänichen* ………………… 60

## Applications in Process Control and Insurance

- The Main Steps to Data Quality  
  *Joachim Schmid* ……………………………………………………………………………………….. 69

- Cost-Sensitive Design of Claim Fraud Screens  
  *Stijn Viaene, Dirk Van Gheel, Mercedes Ayuso, Montserrat Guillén* ………………… 78

- An Early Warning System for Vehicle Related Quality Data  
  *Matthias Grabert, Markus Prechtel, Tomas Hrycej, Winfried Günther* ………………… 88
Clustering and Association Rules

Shape-Invariant Cluster Validity Indices
   Greet Frederix, Eric J. Pauwels.................................................. 96

Mining Indirect Association Rules
   Shinichi Hamano, Masako Sato.................................................. 106

An Association Mining Method for Time Series and Its Application in the
Stock Prices of TFT-LCD Industry
   Chiung-Fen Huang, Yen-Chu Chen, An-Pin Chen............................ 117

Clustering of Web Sessions Using Levenstein Metric
   Andrei Scherbina, Sergey Kuznetsov.......................................... 127

Telecommunication

A Data Mining Approach for Call Admission Control and Resource
Reservation in Wireless Mobile Networks
   Sherif Rashad, Mehmed Kantardzic, Anup Kumar............................ 134

Mining of an Alarm Log to Improve the Discovery of Frequent Patterns
   Françoise Fessant, Fabrice Clérot, Christophe Dousson.................... 144

Medicine and Biotechnology

Feature Selection and Classification Model Construction on Type 2
Diabetic Patient's Data
   Yue Huang, Paul McCullagh, Norman Black, Roy Harper.................... 153

Knowledge Based Phylogenetic Classification Mining
   Isabelle Bichindaritz, Stephen Potter, Société Française de Systématique.... 163

Author Index......................................................................................... 173