

Table of Contents

Long Papers

Pattern Mining

- Predictive Graph Mining 1
Andreas Karwath and Luc De Raedt
- An Efficient Algorithm for Enumerating Closed Patterns
in Transaction Databases 16
Takeaki Uno, Tatsuya Asai, Yuzo Uchida, and Hiroki Arimura
- Finding Optimal Pairs of Cooperative and Competing Patterns
with Bounded Distance 32
*Shunsuke Inenaga, Hideo Bannai, Heikki Hyvrö, Ayumi Shinohara,
Masayuki Takeda, Kenta Nakai, and Satoru Miyano*

Classification

- Mining Noisy Data Streams via a Discriminative Model 47
Fang Chu, Yizhou Wang, and Carlo Zaniolo
- CorClass: Correlated Association Rule Mining for Classification 60
Albrecht Zimmermann and Luc De Raedt
- Maximum a Posteriori Tree Augmented Naive Bayes Classifiers 73
Jesús Cerquides and Ramon López de Màntaras

Outlier Detection

- Improving Prediction of Distance-Based Outliers 89
Fabrizio Angiulli, Stefano Basta, and Clara Pizzuti
- Detecting Outliers via Logical Theories and Its Data Complexity 101
Fabrizio Angiulli, Gianluigi Greco, and Luigi Palopoli

Clustering

- Fast Hierarchical Clustering Algorithm Using Locality-Sensitive Hashing . . 114
Hisashi Koga, Tetsuo Ishibashi, and Toshinori Watanabe

Measuring the Similarity for Heterogenous Data:
An Ordered Probability-Based Approach 129
SiQuang Le and TuBao Ho

Feature Construction and Generation

Constructive Inductive Learning Based on Meta-attributes 142
Kouzou Ohara, Yukio Onishi, Noboru Babaguchi, and Hiroshi Motoda

Resemblance Coefficient and a Quantum Genetic Algorithm
for Feature Selection 155
Gexiang Zhang, Laizhao Hu, and Weidong Jin

Knowledge Acquisition

Extracting Positive Attributions from Scientific Papers 169
Son Bao Pham and Achim Hoffmann

Enhancing SVM with Visualization 183
Thanh-Nghi Do and François Poulet

An Associative Information Retrieval
Based on the Dependency of Term Co-occurrence 195
Mina Akaishi, Ken Satoh, and Yuzuru Tanaka

On the Convergence of Incremental Knowledge Base Construction 207
Tri M. Cao, Eric Martin, and Paul Compton

Discovery Science in Reality

Privacy Problems with Anonymized Transaction Databases 219
Taneli Mielikäinen

A Methodology for Biologically Relevant Pattern Discovery
from Gene Expression Data 230
Ruggero G. Pensa, Jérémy Besson, and Jean-François Boulicaut

Using the Computer
to Study the Dynamics of the Handwriting Processes 242
Gilles Caporossi, Denis Alamargot, and David Chesnet

Product Recommendation in e-Commerce
Using Direct and Indirect Confidence for Historical User Sessions 255
Przemysław Kazienko

Regular Papers

Pattern Mining

- Optimal Discovery of Subword Associations in Strings 270
Alberto Apostolico, Cinzia Pizzi, and Giorgio Satta
- Tiling Databases 278
Floris Geerts, Bart Goethals, and Taneli Mielikäinen
- A Clustering of Interestingness Measures 290
Benoît Vaillant, Philippe Lenca, and Stéphane Lallich
- Extracting Minimal and Closed Monotone DNF Formulas 298
*Yoshikazu Shima, Shinji Mitsuishi, Kouichi Hirata,
and Masateru Harao*
- Characterizations of Multivalued Dependencies and Related Expressions . . 306
José Luis Balcázar and Jaume Baixeries

Machine Learning Algorithms

- Outlier Handling in the Neighbourhood-Based Learning
of a Continuous Class 314
Fabrice Muhlenbach, Stéphane Lallich, and Djamel A. Zighed
- A New Clustering Algorithm Based On Cluster Validity Indices 322
Minho Kim and R.S. Ramakrishna
- An Efficient Rules Induction Algorithm for Rough Set Classification 330
Songbo Tan and Jun Gu
- Analysing the Trade-Off Between Comprehensibility
and Accuracy in Mimetic Models 338
*Ricardo Blanco-Vega, José Hernández-Orallo,
and María José Ramírez-Quintana*
- Generating AVTs Using GA for Learning Decision Tree Classifiers
with Missing Data 347
Jinu Joo, Jun Zhang, Jihoon Yang, and Vasant Honavar

Web Mining

- Using WWW-Distribution of Words in Detecting Peculiar Web Pages 355
Masayuki Hirose and Eimoshin Suzuki

DHT Facilitated Web Service Discovery
Incorporating Semantic Annotation 363
Shoujian Yu, Jianwei Liu, and Jiajin Le

Discovering Relationships Among Catalogs 371
Ryutaro Ichise, Masahiro Hamasaki, and Hideaki Takeda

Applications of Predictive Methods

Reasoning-Based Knowledge Extraction for Text Classification 380
Chiara Cumbo, Salvatore Iritano, and Pasquale Rullo

A Useful System Prototype
for Intrusion Detection – Architecture and Experiments 388
Ye Du, Huiqiang Wang, and Yonggang Pang

Discovery of Hidden Similarity on Collaborative Filtering
to Overcome Sparsity Problem 396
Sanghack Lee, Jihoon Yang, and Sung-Yong Park

Interdisciplinary Approaches

Seamlessly Supporting Combined Knowledge Discovery
and Query Answering: A Case Study 403
Marcelo A.T. Aragão and Alvaro A.A. Fernandes

A Structuralist Approach Towards Computational Scientific Discovery 412
*Charles-David Wajnberg, Vincent Corruble, Jean-Gabriel Ganascia,
and C. Ulises Moulines*

Extracting Modal Implications and Equivalences from Cognitive Minds ... 420
Radoslaw Piotr Katarzyniak

Author Index 429