

Takao Terano Huan Liu
Arbee L.P. Chen (Eds.)



© 2008 AGI-Information Management Consultants
May be used for personal purposes only or by
libraries associated to dandelion.com network.

Knowledge Discovery and Data Mining

Current Issues and New Applications

4th Pacific-Asia Conference, PAKDD 2000
Kyoto, Japan, April 18-20, 2000
Proceedings



Springer

Table of Contents

Keynote Speeches and Invited Talk

Perspective on Data Mining from Statistical Viewpoints	1
<i>Yoshiharu Sato</i>	
Inductive Databases and Knowledge Scouts	2
<i>Ryszard S. Michalski</i>	
Hyperlink-Aware Mining and Analysis of the Web	4
<i>Andrew Tomkins</i>	

Data Mining Theory

Polynomial Time Matching Algorithms for Tree-Like Structured Patterns in Knowledge Discovery	5
<i>Tetsuhiro Miyahara, Takayoshi Shoudai, Tomoyuki Uchida, Kenichi Takahashi, Hiroaki Ueda</i>	
Fast Discovery of Interesting Rules	17
<i>Nobuhiro Yugami, Yuiko Ohta, Seishi Okamoto</i>	
Performance Controlled Data Reduction for Knowledge Discovery in Distributed Databases	29
<i>Slobodan Vucetic, Zoran Obradovic</i>	
Minimum Message Length Criterion for Second-Order Polynomial Model Discovery	40
<i>Grace W. Rumantir</i>	
Frequent Itemset Counting Across Multiple Tables	49
<i>Viviane Crestana Jensen, Nandit Soparkar</i>	
Frequent Closures as a Concise Representation for Binary Data Mining ...	62
<i>Jean-François Boulicaut, Artur Bykowski</i>	
An Optimization Problem in Data Cube System Design	74
<i>Edward Hung, David W. Cheung, Ben Kao, Yilong Liang</i>	
Exception Rule Mining with a Relative Interestingness Measure	86
<i>Farhad Hussain, Huan Liu, Eimoshin Suzuki, Hongjun Lu</i>	

Feature Selection and Transformation

Consistency Based Feature Selection	98
<i>Manoranjan Dash, Huan Liu, Hiroshi Motoda</i>	

Feature Selection for Clustering 110
Manoranjan Dash, Huan Liu

A Simple Dimensionality Reduction Technique for Fast Similarity Search
in Large Time Series Databases 122
Eamonn J. Keogh, Michael J. Pazzani

Missing Value Estimation Based on Dynamic Attribute Selection 134
K. C. Lee, J. S. Park, Y. S. Kim, Y. T. Byun

On Association, Similarity and Dependency of Attributes 138
Yi Yu Yao, Ning Zhong

Clustering

Prototype Generation Based on Instance Filtering and Averaging 142
Chi-Kin Keung, Wai Lam

A Visual Method of Cluster Validation with Fastmap 153
Zherue Huang, Tao Lin

COE: Clustering with Obstacles Entities. A Preliminary Study 165
Anthony K.H. Tung, Jean Hou, Jiawei Han

Combining Sampling Technique with DBSCAN Algorithm for Clustering
Large Spatial Databases 169
Shuigeng Zhou, Aoying Zhou, Jing Cao, Jin Wen, Ye Fan, Yunfa Hu

Predictive Adaptive Resonance Theory and Knowledge Discovery
in Databases 173
Ah-Hwee Tan, Hui-Shin Vivien Soon

Improving Generalization Ability of Self-Generating Neural Networks
Through Ensemble Averaging 177
Hiroataka Inoue, Hiroyuki Narihisa

Application of Data Mining

Attribute Transformations on Numerical Databases 181
Tsau Young Lin, Joseph Tremba

Efficient Detection of Local Interactions in the Cascade Model 193
Takashi Okada

Extracting Predictors of Corporate Bankruptcy: Empirical Study on Data
Mining Methods 204
Cindy Yoshiko Shirata, Takao Terano

Evaluating Hypothesis-Driven Exception-Rule Discovery with Medical Data Sets	208
<i>Einoshin Suzuki, Shusaku Tsumoto</i>	
Discovering Protein Functional Models Using Inductive Logic Programming.....	212
<i>Takashi Ishikawa, Masayuki Numao, Takao Terano</i>	
Mining Web Transaction Patterns in an Electronic Commerce Environment	216
<i>Ching-Huang Yun, Ming-Syan Chen</i>	

Association Rules and Related Topics

Making Use of the Most Expressive Jumping Emerging Patterns for Classification	220
<i>Jinyan Li, Guozhu Dong, Ramamohanarao Kotagiri</i>	
Mining Structured Association Patterns from Databases	233
<i>Hirofumi Matsuzawa, Takeshi Fukuda</i>	
Association Rules	245
<i>Tao Zhang</i>	
Density-Based Mining of Quantitative Association Rules	257
<i>David W. Cheung, Lian Wang, S.M. Yiu, Bo Zhou</i>	
AViz: A Visualization System for Discovering Numeric Association Rules..	269
<i>Jiancho Han, Nick Cercone</i>	
Discovering Unordered and Ordered Phrase Association Patterns for Text Mining	281
<i>Ryoichi Fujino, Hiroki Arimura, Setsuo Arikawa</i>	
Using <i>Random Walks</i> for Mining Web Document Associations	294
<i>K. Selçuk Candan, Wen-Syan Li</i>	

Induction

A Concurrent Approach to the Key-Preserving Attribute-Oriented Induction Method.....	306
<i>Maybin K. Mueyba, John A. Keane</i>	
Scaling Up a Boosting-Based Learner via Adaptive Sampling	317
<i>Carlos Domingo, Osamu Watanabe</i>	
Adaptive Boosting for Spatial Functions with Unstable Driving Attributes	329
<i>Aleksandar Lazarevic, Tim Fiez, Zoran Obradovic</i>	

Robust Ensemble Learning for Data Mining	341
<i>Gunnar Rätsch, Bernhard Schölkopf, Alexander Johannes Smola, Sebastian Mika, Takashi Onoda, Klaus-Robert Müller</i>	
Interactive Visualization in Mining Large Decision Trees	345
<i>Trong Dung Nguyen, Tu Bao Ho, Hiroshi Shimodaira</i>	
VQTree: Vector Quantization for Decision Tree Induction	349
<i>Shlomo Geva, Lawrence Buckingham</i>	
Making Knowledge Extraction and Reasoning Closer	360
<i>Fosca Giannotti, Giuseppe Manco</i>	
Discovery of Relevant Weights by Minimizing Cross-Validation Error	372
<i>Kazumi Saito, Ryohei Nakano</i>	
Efficient and Comprehensible Local Regression	376
<i>Luís Torgo</i>	
Information Granules for Spatial Reasoning	380
<i>Andrzej Skowron, Jaroslaw Stepaniuk, Shusaku Tsumoto</i>	
Text, Web, and Graph Mining	
Uncovering the Hierarchical Structure of Text Archives by Using an Unsupervised Neural Network with Adaptive Architecture	384
<i>Dieter Merkl, Andreas Rauber</i>	
Mining Access Patterns Efficiently from Web Logs	396
<i>Jian Pei, Jiawei Han, Behzad Mortazavi-asl, Hua Zhu</i>	
A Comparative Study of Classification Based Personal E-mail Filtering ...	408
<i>Yanlei Diao, Hongjun Lu, Dekai Wu</i>	
Extension of Graph-Based Induction for General Graph Structured Data ..	420
<i>Takashi Matsuda, Tadashi Horiuchi, Hiroshi Motoda, Takashi Washio</i>	
Text-Source Discovery and GLOSS Update in a Dynamic Web	432
<i>Chi-Yuen Ng, Ben Kao, David Cheung</i>	
Extraction of Fuzzy Clusters from Weighted Graphs	442
<i>Seiji Hotta, Kohei Inoue, Kiichi Urahama</i>	
Text Summarization by Sentence Segment Extraction Using Machine Learning Algorithms	454
<i>Wesley T. Chuang, Jihoon Yang</i>	
Author Index	459