

Table of Contents

Keynote Presentations

Incompleteness in Data Mining	1
<i>Hosagrahar Visvesvaraya Jagadish</i>	
Mining E-Commerce Data: The Good, the Bad, and the Ugly	2
<i>Ronny Kohavi</i>	
Seamless Integration of Data Mining with DBMS and Applications	3
<i>Hongjun Lu</i>	

Web Mining

Applying Pattern Mining to Web Information Extraction	4
<i>Chia-Hui Chang, Shao-Chen Lui, Yen-Chin Wu</i>	
Empirical Study of Recommender Systems Using Linear Classifiers	16
<i>Vijay S. Iyengar, Tong Zhang</i>	
iJADE eMiner—A Web-Based Mining Agent Based on Intelligent Java Agent Development Environment (iJADE) on Internet Shopping	28
<i>Raymond S.T. Lee, James N.K. Liu</i>	
A Characterized Rating Recommend System	41
<i>Yao-Tsung Lin, Shian-Shyong Tseng</i>	
Discovery of Frequent Tree Structured Patterns in Semistructured Web Documents	47
<i>Tetsuhiro Miyahara, Takayoshi Shoudai, Tomoyuki Uchida, Kenichi Takahashi, Hiroaki Ueda</i>	

Text Mining

Text Categorization Using Weight Adjusted k -Nearest Neighbor Classification	53
<i>Eui-Hong Han, George Karypis, Vipin Kumar</i>	
Predictive Self-Organizing Networks for Text Categorization	66
<i>Ah-Hwee Tan</i>	
Meta-learning Models for Automatic Textual Document Categorization ...	78
<i>Kwok-Yin Lai, Wai Lam</i>	
Efficient Algorithms for Concept Space Construction	90
<i>Chi Yuen Ng, Joseph Lee, Felix Cheung, Ben Kao, David Cheung</i>	

Topic Detection, Tracking, and Trend Analysis Using Self-Organizing Neural Networks	102
<i>Kanagasabi Rajaraman, Ah-Hwee Tan</i>	

Automatic Hypertext Construction through a Text Mining Approach by Self-Organizing Maps	108
<i>Hsin-Chang Yang, Chung-Hong Lee</i>	

Applications and Tools

Semantic Expectation-Based Causation Knowledge Extraction: A Study on Hong Kong Stock Movement Analysis	114
<i>Boon-Toh Low, Ki Chan, Lei-Lei Choi, Man-Yee Chin, Sin-Ling Lay</i>	

A Toolbox Approach to Flexible and Efficient Data Mining	124
<i>Ole M. Nielsen, Peter Christen, Markus Hegland, Tatiana Semenova, Timothy Hancock</i>	

Determining Progression in Glaucoma Using Visual Fields	136
<i>Andrew Turpin, Eibe Frank, Mark Hall, Ian H. Witten, Chris A. Johnson</i>	

Seabreeze Prediction Using Bayesian Networks	148
<i>Russell J. Kennett, Kevin B. Korb, Ann E. Nicholson</i>	

Semi-supervised Learning in Medical Image Database	154
<i>Chun Hung Li, Pong Chi Yuen</i>	

On Application of Rough Data Mining Methods to Automatic Construction of Student Models	161
<i>Feng-Hsu Wang, Shiou-Wen Hung</i>	

Concept Hierarchies

Concept Approximation in Concept Lattice	167
<i>Keyun Hu, Yuefei Sui, Yuchang Lu, Ju Wang, Chunyi Shi</i>	

Generating Concept Hierarchies/Networks: Mining Additional Semantics in Relational Data	174
<i>T. Y. Lin</i>	

Representing Large Concept Hierarchies Using Lattice Data Structure	186
<i>Yanee Kachai, Kitsana Waiyamai</i>	

Feature Selection

Feature Selection for Temporal Health Records	198
<i>Rohan A. Baxter, Graham J. Williams, Hongxing He</i>	

Boosting the Performance of Nearest Neighbour Methods with Feature Selection	210
<i>Shlomo Geva</i>	

Feature Selection for Meta-learning	222
<i>Alexandros Kalousis, Melanie Hilario</i>	

Interestingness

Efficient Mining of Niches and Set Routines	234
<i>Guozhu Dong, Kaustubh Deshpande</i>	

Evaluation of Interestingness Measures for Ranking Discovered Knowledge	247
<i>Robert J. Hilderman, Howard J. Hamilton</i>	

Peculiarity Oriented Mining and Its Application for Knowledge Discovery in Amino-Acid Data	260
<i>Ning Zhong, Muneaki Ohshima, Setsuo Ohsuga</i>	

Sequence Mining

Mining Sequence Patterns from Wind Tunnel Experimental Data for Flight Control	270
<i>Zhenyu Liu, Wesley W. Chu, Adam Huang, Chris Folk, Chih-Ming Ho</i>	

Scalable Hierarchical Clustering Method for Sequences of Categorical Values	282
<i>Tadeusz Morzy, Marek Wojciechowski, Maciej Zakrzewicz</i>	

FFS—An I/O-Efficient Algorithm for Mining Frequent Sequences	294
<i>Minghua Zhang, Ben Kao, Chi-Lap Yip, David Cheung</i>	

Sequential Index Structure for Content-Based Retrieval	306
<i>Maciej Zakrzewicz</i>	

Spatial and Temporal Mining

The S^2 -Tree: An Index Structure for Subsequence Matching of Spatial Objects	312
<i>Haiyun Wang, Chang-Shin Perng</i>	

Temporal Data Mining Using Hidden Markov-Local Polynomial Models . . .	324
<i>Weiqiang Lin, Mehmet A. Orgun, Graham J. Williams</i>	

Patterns Discovery Based on Time-Series Decomposition	336
<i>Jeffrey Xu Yu, Michael K. Ng, Joshua Zhexue Huang</i>	

Criteria on Proximity Graphs for Boundary Extraction and Spatial Clustering	348
<i>Vladimir Estivill-Castro, Ickjai Lee, Alan T. Murray</i>	

Micro Similarity Queries in Time Series Database	358
<i>Xiao-ming Jin, Yuchang Lu, Chunyi Shi</i>	

Association Mining

Mining Optimal Class Association Rule Set	364
<i>Jiuyong Li, Hong Shen, Rodney Topor</i>	

Generating Frequent Patterns with the Frequent Pattern List	376
<i>Fan-Chen Tseng, Ching-Chi Hsu</i>	

User-Defined Association Mining	387
<i>Ke Wang, Yu He</i>	

Direct and Incremental Computing of Maximal Covering Rules	400
<i>Marzena Kryszkiewicz</i>	

Towards Efficient Data Re-mining (DRM)	406
<i>Jiming Liu, Jian Yin</i>	

Data Allocation Algorithm for Parallel Association Rule Discovery	413
<i>Anna M. Manning, John A. Keane</i>	

Classification and Rule Induction

Direct Domain Knowledge Inclusion in the PA3 Rule Induction Algorithm	421
<i>Pedro de Almeida</i>	

Hierarchical Classification of Documents with Error Control	433
<i>Chun-hung Cheng, Jian Tang, Ada Wai-chee Fu, Irwin King</i>	

An Efficient Data Compression Approach to the Classification Task	444
<i>Claudia Diamantini, Maurizio Panti</i>	

Combining the Strength of Pattern Frequency and Distance for Classification	455
<i>Jinyan Li, Kotagiri Ramamohanarao, Guozhu Dong</i>	

A Scalable Algorithm for Rule Post-pruning of Large Decision Trees	467
<i>Trong Dung Nguyen, Tu Bao Ho, Hiroshi Shimodaira</i>	

Optimizing the Induction of Alternating Decision Trees	477
<i>Bernhard Pfahringer, Geoffrey Holmes, Richard Kirkby</i>	

Building Behaviour Knowledge Space to Make Classification Decision	488
<i>Xiuzhen Zhang, Guozhu Dong, Kotagiri Ramamohanarao</i>	

Clustering

Efficient Hierarchical Clustering Algorithms Using Partially Overlapping Partitions	495
<i>Manoranjan Dash, Huan Liu</i>	
A Rough Set-Based Clustering Method with Modification of Equivalence Relations	507
<i>Shoji Hirano, Tomohiro Okuzaki, Yutaka Hata, Shusaku Tsumoto, Kouhei Tsumoto</i>	
Importance of Individual Variables in the k -Means Algorithm	513
<i>Juha Vesanto</i>	
A Hybrid Approach to Clustering in Very Large Databases	519
<i>Aoying Zhou, Weining Qian, Hailei Qian, Jin Wen, Shuigeng Zhou, Ye Fan</i>	

Advanced Topics and New Methods

A Similarity Indexing Method for the Data Warehousing—Bit-Wise Indexing Method	525
<i>Wei-Chou Chen, Shian-Shyong Tseng, Lu-Ping Chang, Mon-Fong Jiang</i>	
Rule Reduction over Numerical Attributes in Decision Trees Using Multilayer Perceptron	538
<i>DaeEun Kim, Jaeho Lee</i>	
Knowledge Acquisition from Both Human Expert and Data	550
<i>Takuya Wada, Hiroshi Motoda, Takashi Washio</i>	
Neighborhood Dependencies for Prediction	562
<i>Renaud Bassée, Jef Wijsen</i>	
Learning Bayesian Networks with Hidden Variables Using the Combination of EM and Evolutionary Algorithms	568
<i>Fengzhan Tian, Yuchang Lu, Chunyi Shi</i>	
Interactive Construction of Decision Trees	575
<i>Jianchao Han, Nick Cercone</i>	
An Improved Learning Algorithm for Augmented Naive Bayes	581
<i>Huajie Zhang, Charles X. Ling</i>	

Generalised RBF Networks Trained Using an IBL Algorithm for
Mining Symbolic Data 587
 Liviu Vladutu, Stergios Papadimitriou, Severina Mavroudi,
 Anastasios Bezerianos

Author Index 595

Advances in Knowledge Discovery and Data Mining
5th Pacific-Asia Conference, PAKDD 2001 Hong Kong,
China, April 16-18, 2001. Proceedings
Cheung, D.; Williams, G.J.; Li, Q. (Eds.)
2001, XVII, 599 p., Softcover
ISBN: 978-3-540-41910-5