

Changjie Tang Charles X. Ling
Xiaofang Zhou Nick J. Cercone Xue Li (Eds.)

Advanced Data Mining and Applications

4th International Conference, ADM A 2008
Chengdu, China, October 8-10, 2008
Proceedings

^{A ^}
Springer

Table of Contents

Keynotes

An Introduction, to Transfer Learning	1
<i>Qiang Yang</i>	
Autonomy-Oriented Computing (AOC), Self-organized Computability, and Complex Data, Mining	2
<i>Jiming Liu</i>	

Regular Papers

Improving Angle Based Mappings.	3
<i>Frank Rehm and Frank Klawonn</i>	
Mining Natural Language Programming Directives with Class-Oriented Bayesian Networks.	15
<i>Manolis Maragoudakis, Nikolaos Co'srnas, and Amstogiannis Garbis</i>	
Boosting over Groups and Its Application to Acronym-Expansion Extraction.	27
<i>Weijian Ni, Yalou Huang, Dong Li, and Yang Wang</i>	
A Genetic-Based Feature Construction Method for Data Summarisation.	39
<i>B.ayner Alfred</i>	
Suicidal Risk Evaluation Using a Similarity-Based Classifier.	51
<i>S. Chattopadhyay, P. Ray. H.S. Chen, M.B. Lee, and H. C. Chiang</i>	
Gene Selection for Cancer Classification Using DCA.	62
<i>Hoax An Le Thi, Van Vinh Nguyen, and, Samir Ouchani</i>	
FARS: A Multi-relational Feature and Relation Selection Approach for Efficient Classification.	73
<i>Bo Iiu. Hongyan Liu, Jun He, and Xiaoyong Du</i>	
Enhancing Text Categorization Using Sentence Semantics.	87
<i>Shady Shehata, Fakhri Karray, and Mohamed Kamel</i>	
Mining Evolving Web Sessions and Clustering Dynamic Web Documents for Similarity-Aware Web Content Management	99
<i>Jitian Xiao</i>	

Data Quality in Privacy Preservation for Associative Classification.111
Nattapon Hamsamut, Jiigapong Natwichai, Xingzhi Sun, and Xue Li

Timeline Analysis of Web News Events. 123
Jiangtao Qiu, Chuan Li, Shaojie Qiao, Taiyong Li, and Jim, Zhu

Analysis of Alarm Sequences in a Chemical Plant135
Savo Kordic, Peng Lam, Jitian Xiao, and H'uaizhong Li,

Speed Up SVM Algorithm for Massive Classification Tasks. 147
Thanh-Nghi Do. Van-Hoa Nguyen, and Francois Poulet

Mining Supplemental Frequent Patterns.158
Yintian Liu, Yingming Liu, Tao Zeng, Kaikuo Xu, and R.onq Tang

A Distributed Privacy-Preserving Association Rules Mining Scheme
Using Frequent-Pattern Tree.170
Chunhua Su and, Kouichi Sakurai •

Dichotomy Method toward Interactive Testing-Based Fault
Localization182
Ji-Rong Sun, Zhi-Shu Li. and Jian-Cheng Ni

Maintaining the Maximum Normalized Mean and Applications in Data
Stream Mining. 194
Jan Peter Patist

Identification of Interface Residues Involved in Protein-Protein
Interactions Using Naive Bayes Classifier. 207
Chishe Wang, Jiaxing Cheng, Shoubao Su, and Dongzhe Xu

Negative Generator Border for Effective Pattern Maintenance. 217
Mengling Feng, Jinyan Li, Limsoon Wong, and Yap-Peng Tan

CommTracker: A Core-Based Algorithm of Tracking Community
Evolution. 229
Yi Wang, Bin Wu. and, Xih Pei

Face Recognition Using Clustering Based Optimal Linear Discriminant
Analysis.241
Wenxin Yang, Shuqin Rao, Jina Wang, Jian Yin, and Man Chen

A Novel Immune Based Approach for Detection of Windows PE
Virus. 250
Yu Zhang, Tao Li. Jia Sun. and, Renchao Qin

Using Genetic Algorithms for Parameter Optimization in Building
Predictive Data Mining Models. 260
Ashish S'ureka and, Kishore Varrna, Indukuri.

Using Data Mining Methods to Predict Personally Identifiable Information in Emails	272
<i>Liqiang Geng, Larry Korba, Xin Wang, Yunli Wang, Hongyu Liu, and Yonghua You</i>	
Iterative Reinforcement Cross-Domain Text Classification.	282
<i>Di Zhang, Gui-Rong Xue, and, Yong Yu</i>	
Extracting Decision Rules from Sigmoid Kernel.	294
<i>Quanzhong Liu, Yang Zhang, and, Zhengguo Hu</i>	
DMGrid: A Data, Mining System Based on Grid Computing.	305
<i>Yi Wang, Liutong Xu, Guanhuai Geng, Xiangang Zhao, and Nan Du</i>	
S-SimRank: Combining Content and Link Information to Cluster Papers Effectively and Efficiently.	317
<i>Yuanzhe Cai, Pei Li, Hongyan Liu, Jun He, and Xiaoyong Du</i>	
Open Domain Recommendation: Social Networks and Collaborative Filtering.	330
<i>Sarah K. Tyler and Yi Zhang</i>	
An Effective Approach for Identifying Evolving Three-Dimensional Structural Motifs in Protein Folding Data	342
<i>Hui Yang and Lin Han</i>	
Texture Image Retrieval Based on Contourlet Transform and Active Perceptual Similarity Learning.	355
<i>Huaijing Qu, Yuhuo, Peng, Honglin Wan, and Min Han</i>	
A Temporal Dominant Relationship Analysis Method	367
<i>Jing Yang, Yuanxi Wu, Cuiping Li, Hong Chen, and, Bo Qu</i>	
Leakage-Aware Energy Efficient Scheduling for Fixed-Priority Tasks with Preemption Thresholds.	379
<i>XiaoChuan Lie and Yan Jia</i>	

Short Papers

Learning and Inferences of the Bayesian Network with Maximum Likelihood Parameters.	391
<i>JiaDong Zhang, Kun Yue, and WeiYi Liu</i>	
TARtool: A Temporal Dataset Generator for Market Basket Analysis.	400
<i>Asem Omari, Regina Langer, and Stefan, Conrad</i>	
Dimensionality Reduction for Classification: Comparison of Techniques and Dimension Choice.	411
<i>Frank Plastria, Steven De Bruyne, and Emilio Carrizosa</i>	

Trajectories Mining for Traffic Condition Renewing. 419
Danhuai Guo

Mining Bug Classifier and Debug Strategy Association Rules for
Web-Based Applications. 427
Linn Yu, Changzhu Kong, Lei Xu, Jingtao Zhao, and Huihui Zhang

lest the Overall Significance of p-Values by Using Joint Tail Probability
of Ordered p-Values as Test Statistic. 435
Y'ongxiang Fang and Ernst Wit

Mining Interesting Infrequent and Frequent Iternsets Based on MLMS
Model. 444
*Xiongjun Dong, Zhendong Niu, Donghua Zhu, Zhiyun Zheng, and,
Qiuting Ma*

Text Learning and Hierarchical Feature; Selection in Webpage
Classification. 452
Xiaogang Peng, Zhong Ming, and, Haitao Wang

The RSO Algorithm for Reducing Number of Set Operations in
Association Rule Mining. 460
Muhammad Sarwar and Onaiza Maqbool

Predictive Performance of Clustered Feature-Weighting Case-Based
Reasoning. 469
Sung Ho Ha, Jong Sik Jin, and Jeong Won Yang

Selecting the Right Features for Bipartite-Based Text Clustering 477
Chao Qu, Yong Li, Me Zhang, Tianmmg Hu, and, Qian Chen

Image Emotional Classification Based on Color Semantic Description 485
Kaiping Wei, Bin He, Tao Zhang, and Wenya He

A Semi-supervised Clustering Algorithm Based on Must-Link Set 492
Haichao Huang, Yong Cheng, and Ruilian Zhao

T-rotation: Multiple Publications of Privacy Preserving Da.ta
Sequence. 500
*Youdong Tao, Yunhai Tong, Shaohua Tan, Shiwei Tang, and,
Dongqing Yang*

The Integrated Methodology of KPCA and Wavelet Support Vector
Machine for Predicting Financial Distress. 508
Jian-guo Zhou, Tao Ba,i, and, Ji-rning Tian

Outlier Detection Based on Voronoi Diagram. 516
Jilin Qu

AWSum - Data Mining for Insight	524
<i>Anthony Quinn, Andrew Stranieri, John Yearwood. and Gaudenz Hafen</i>	
Integrative Neural Network Approach for Protein Interaction Prediction from Heterogeneous Data	532
<i>Xue-wen Chen, Mei Liu, and Yong Hu</i>	
Rules Extraction Based on Data Summarisation Approach Using DARA	540
<i>R.ayner Alfred</i>	
A R.ough-Apriori Technique in Mining Linguistic Association Rules....'.	548
<i>Yun-Huoy Choo, Azuraliza Abu Bakar, and, Abdul Razak Harndan</i>	
Mining Causal Knowledge from Diagnostic Knowledge.	556
<i>Xiangdong An and Nick Cercone</i>	
Modified Particle Swarm Optimizer with Adaptive Dynamic Weights for Cancer Combinational Chemotherapy.	563
<i>Harish Chandra, Soundararajan, Jagannathan Raman, and R., Muthucuma.raswamy</i>	
MPSQAR: Mining Quantitative Association Rules Preserving Semantics.	572
<i>Chunqiu Zeng, Jie Zuo, Chuan Li, Kaikuo Xu, Shengqiao Ni, Liang Tang, Yue Zhang, and Shaojie Qiao</i>	
Using Support Vector Regression for Classification.	581
<i>Bo Huang, Zhihua Cai, Qiong Gu, and Changjun Chen</i>	
Dynamic Growing Self-organizing Neural Network for Clustering	589
<i>Daxin Tian, Yueou Ren, and, Qiuju Li</i>	
A Design of Reward Function Based on Knowledge in Multi-agent Learning.	596
<i>Bo Fan and Jiexin Pu</i>	
A Learning Method of Detecting Anomalous Pedestrian.	604
<i>Yue Liu, Jun Zhang, and Zhijing Liu</i>	
Moment-h: Mining Closed Frequent Itemsets over Data Stream	612
<i>Haifeng Li and, Hong Chen</i>	
CDPM: Finding and Evaluating Community Structure in Social Networks.	620
<i>Li Wan, Jianxin Liao, and Xiaom.in Zhu</i>	
Using Matrix Model to Find Association Rule Core for Diverse Compound Critiques.	628
<i>Li Yu</i>	

Link-Contexts for Ranking 636
Jessica, Gronski

DC-Tree: An Algorithm for Skyline Query on Data Streams. 644
Jing Yang, Bo Qu, Cui-Ping Li, and, Hong Chen

Sequential Pattern Mining for Protein Function Prediction. 652
Mia.o Wang, Xue-qun Sha.ng, and Zhan-huai Li

Improving Web Search by Categorization, Clustering, and
Personalization 659
Dengya Zhu and Heinz Dreher

JSNVA: A Ja-va Straight-Line Drawing Framework for Network Visual
Analysis. 667
Qi, Ye, Bin Wu, and Bai Wang

Recognition of Data Records in Semi-structured Web-Pages Using
Ontology and χ^2 Statistical Distribution. 675
*Amin Keshavarzi, Amir Masoud Rahmani,
Mehran Mohsenzadeh, and Reza Keshavarzi*

Organizing Structured Deep Web by ClusteringQuery Interfaces Link
Graph. 683
Pengpeng Zhao, Li Huang, Wei Fang, and Zhirning Cui

CB.P: A New Efficient Method for Mining Multilevel and Generalized
Frequent Itemsets. 691
Yu Xing Mao and Bai Le Shi

Supporting Customer Retention through Real-Time-Monitoring of
Individual Web Usage. 699
Peter' I. Hofgesang and Jan Peter Patist

A Comparative Study of Correlation Measurements for Searching
Similar Tags. 709
*Kaikuo Xu, Yu Chen, Yexi Jiang, Rong Tang, Yintian Liu, and
Jie Gong*

Structure of Query Modification Process: Branchings. 717
Nikolai Buzikashvili

Mining Top-n Local Outliers in Constrained Spatial Networks. 725
Chongsheng Zhang, Zhongbo Wu, Bo Qu. and, Hong Chen

Mining Concept-Drifting Data Streams-with Multiple Semi-Random
Decision Trees. 733
Peipei Li, Xuegang Hu, and, Xindong Wu

Automatic Web Tagging and Person Tagging Using Language
Models. 741
Qia.ozhu Mei and Yi Zhang

Real-Time Peison Tracking Based on Data Field. 749
Shuliang Wang, Juebo Wu, Feng Cheng, and Hong Jin

Author Index 757