Table of Contents

Data Mining Applications in Industry and Government

Using Scan-Statistical Correlations for Network Change Analysis 1
Adriel Cheng and Peter Dickinson

Predicting High Impact Academic Papers Using Citation Network Features 14
Daniel McNamara, Paul Wong, Peter Christen, and Kee Siong Ng

An OLAP Server for Sensor Networks Using Augmented Statistics Trees 26
Neil Dunstan

Indirect Information Linkage for OSINT through Authorship Analysis of Aliases 36
Robert Layton, Charles Perez, Babiga Birregah, Paul Watters, and Marc Lemercier

Dynamic Similarity-Aware Inverted Indexing for Real-Time Entity Resolution 47
Banda Ramadan, Peter Christen, Huizhi Liang, Ross W. Gayler, and David Hawking

Identifying Dominant Economic Sectors and Stock Markets: A Social Network Mining Approach 59
Ram Babu Roy and Uttam Kumar Sarkar

Ensemble Learning Model for Petroleum Reservoir Characterization: A Case of Feed-Forward Back-Propagation Neural Networks 71
Fatai Anifowose, Jane Labadin, and Abdulazeez Abdulraheem

Visual Data Mining Methods for Kernel Smoothed Estimates of Cox Processes 83
David Rohde, Ruth Huang, Jonathan Corcoran, and Gentry White

Real-Time Television ROI Tracking Using Mirrored Experimental Designs 95
Brendan Kitts, Dyng Au, and Brian Burdick

On the Evaluation of the Homogeneous Ensembles with CV-Passports 109
Vladimir Nikulin, Aneesha Bakharia, and Tian-Hsiang Huang

Bibliografische Informationen
http://d-nb.info/103812042X

digitalisiert durch DEUTSCHE NATIONAL BIBLIOTHEK

Parallel Sentiment Polarity Classification Method with Substring Feature Reduction 121
Yaowen Zhang, Xiaojun Xiang, Cunyan Yin, and Lin Shang

Identifying Authoritative and Reliable Contents in Community Question Answering with Domain Knowledge 133
Lifan Guo and Xiaohua Hu

Data Analytics for Targeted Healthcare

On the Application of Multi-class Classification in Physical Therapy Recommendation 143
Jing Zhang, Douglas Gross, and Osmar R. Zaïane

EEG-MINE: Mining and Understanding Epilepsy Data 155
SunHee Kim, Christos Faloutsos, and Hyung-Jeong Yang

A Constraint and Rule in an Enhancement of Binary Particle Swarm Optimization to Select Informative Genes for Cancer Classification 168
Mohd Saberi Mohamad, Sigeru Omatu, Safaai Deris, and Michifumi Yoshioka

Parameter Estimation Using Improved Differential Evolution (IDE) and Bacterial Foraging Algorithm to Model Tyrosine Production in Mus Musculus (Mouse) 179
Jia Xing Yeoh, Chuii Khim Chong, Yee Wen Choon, Lian En Chai, Safaai Deris, Rosli Md. Illias, and Mohd Saberi Mohamad

Threonine Biosynthesis Pathway Simulation Using IBMDE with Parameter Estimation 191
Chuii Khim Chong, Mohd Saberi Mohamad, Safaai Deris, Mohd Shahir Shamsir, Yee Wen Choon, and Lian En Chai

A Depression Detection Model Based on Sentiment Analysis in Micro-blog Social Network 201
Xinyu Wang, Chunhong Zhang, Yang Ji, Li Sun, Leijia Wu, and Zhana Bao

Modelling Gene Networks by a Dynamic Bayesian Network-Based Model with Time Lag Estimation 214
Lian En Chai, Mohd Saberi Mohamad, Safaai Deris, Chuii Khim Chong, and Yee Wen Choon

Identifying Gene Knockout Strategy Using Bees Hill Flux Balance Analysis (BHFBA) for Improving the Production of Succinic Acid and Glycerol in *Saccharomyces cerevisiae* 223
Yee Wen Choon, Mohd Saberi Mohamad, Safaai Deris, Rosli Md. Illias, Lian En Chai, and Chuii Khim Chong

Mining Clinical Process in Order Histories Using Sequential Pattern Mining Approach 234
Shusaku Tsumoto and Hidenao Abe

Multiclass Prediction for Cancer Microarray Data Using Various Variables Range Selection Based on Random Forest 247
Kohbalan Moorthy, Mohd Saberi Mohamad, and Safaai Deris

A Hybrid of SVM and SCAD with Group-Specific Tuning Parameters in Identification of Informative Genes and Biological Pathways 258
Muhammad Faiz Misman, Weng Howe Chan, Mohd Saberi Mohamad, and Safaai Deris

Structured Feature Extraction Using Association Rules 270
Nan Tian, Yue Xu, Yuefeng Li, and Gabriella Pasi

Quality Issues, Measures of Interestingness and Evaluation of Data Mining Models

Evaluation of Error-Sensitive Attributes 283
William Wu and Shichao Zhang

Mining Correlated Patterns with Multiple Minimum All-Confidence Thresholds 295
R. Uday Kiran and Masaru Kitsuregawa

A Novel Proposal for Outlier Detection in High Dimensional Space 307
Zhana Bao and Wataru Kameyama

CPPG: Efficient Mining of Coverage Patterns Using Projected Pattern Growth Technique 319
P. Gowtham Srinivas, P. Krishna Reddy, and A.V. Trinath

A Two-Stage Dual Space Reduction Framework for Multi-label Classification 330
Eakasit Pacharawongsakda and Thanaruk Theeramunkong

Effective Evaluation Measures for Subspace Clustering of Data Streams 342
Marwan Hassani, Yunsu Kim, Seungjin Choi, and Thomas Seidl

Objectively Evaluating Interestingness Measures for Frequent Itemset Mining 354
Albrecht Zimmermann

A New Feature Selection and Feature Contrasting Approach Based on Quality Metric: Application to Efficient Classification of Complex Textual Data 367
Jean-Charles Lamirel, Pascal Cuxac, Aneesh Sreevallabh Chivukula, and Kafil Hajlaoui

Evaluation of Position-Constrained Association-Rule-Based Classification for Tree-Structured Data 379
Dang Bach Bui, Fedja Hadzic, and Michael Hecker

Enhancing Textual Data Quality in Data Mining: Case Study and Experiences 392
Yi Feng and Chunhua Ju

Cost-Based Quality Measures in Subgroup Discovery 404
Rob M. Konijn, Wouter Duivesteijn, Marvin Meeng, and Arno Knobbe

Biological Inspired Techniques for Data Mining

Applying Migrating Birds Optimization to Credit Card Fraud Detection 416
Ekrem Duman and Ilker Elikucuk

Clustering in Conjunction with Quantum Genetic Algorithm for Relevant Genes Selection for Cancer Microarray Data 428
Manju Sardana, R.K. Agrawal, and Baljeet Kaur

On the Optimality of Subsets of Features Selected by Heuristic and Hyper-heuristic Approaches 440
Kourosh Neshatian and Lucianne Varn

A PSO-Based Cost-Sensitive Neural Network for Imbalanced Data Classification 452
Peng Cao, Dazhe Zhao, and Osmar R. Zaïane

Binary Classification Using Genetic Programming: Evolving Discriminant Functions with Dynamic Thresholds 464
Jill de Jong and Kourosh Neshatian

Constraint Discovery and Cloud Service Discovery

Incremental Constrained Clustering: A Decision Theoretic Approach ... 475
Swapna Raj Prabakara Raj and Balaraman Ravindran

Querying Compressed XML Data 487
Olfa Arfaoui and Minyar Sassi-Hidri

Mining Approximate Keys Based on Reasoning from XML Data 499
Liu Yijun, Ye Feiyue, and He Sheng

A Semantic-Based Dual Caching System for Nomadic Web Service 511
Panpan Han, Liang Chen, and Jian Wu

FTCRank: Ranking Components for Building Highly Reliable Cloud Applications 522
Hanze Xu, Yanan Xie, Dinglong Duan, Liang Chen, and Jian Wu

Research on SaaS Resource Management Method Oriented to Periodic User Behavior 533
Jun Guo, Hongle Wu, Hao Huang, Fang Liu, and Bin Zhang

Weight Based Live Migration of Virtual Machines 543
Baiyou Qiao, Kai Zhang, Yanpeng Guo, Yutong Li, Yuhai Zhao, and Guoren Wang

Author Index 555