Table of Contents

Closed Frequent Itemset Mining over Fast Data Stream Based on Hadoop 1
Shan Jicheng and Liu Qingbao

Robustness of Bisecting k-means Clustering-based Collaborative Filtering Algorithm 7
Alper Bilge and Huseyin Polat

Improving Relevance Effectiveness in Data Leakage Detection Using Feature Selection 14
Adrienn Skrop

Analyzing and Improving Educational Process Models using Process Mining Techniques 17
Awatef Hicheur Cairns, Billel Gueni, Joseph Assu, Christian Joubert, and Nasser Khelifa

Streamlining the Detection of Accounting Fraud through Web Mining and Interpretable Internal Representations 23
Duarte Trigueiros and Carolina Sam

Exponential Moving Maximum Filter for Predictive Analytics in Network Reporting 27
Bin Yu, Les Smith, and Mark E. Threefoot

Automatic KDD Data Preparation Using Multi-criteria Features 33
Youssef Hmamouche, Christian Ersnt, and Alain Casali

Towards Predictive Policing: Knowledge-based Monitoring of Social Networks 39
Michael Spranger, Florian Heinke, Steffen Grunert, and Dirk Labudde

Real-time Partition of Streamed Graphs for Data Mining Over Large Scale Data 41
Victor Medel and Unai Arronategui

Sketch of Big Data Real-Time Analytics Model 48
Bakhtiar Amen and Joan Lu

Augmenting Data Files with Semantics for Coherency, Extensibility, and Reproducibility 54
John McCloud and Subhasish Mazumdar

Automatically Triggering Activity and Product Predictions in Mobile Phone Based on Individual's Activity 61
Kalpana Algotar and Sanjay Addicam

Recommender Systems for Museums: Evaluation on a Real Dataset 65
Ivan Keller and Emmanuel Viennet

An Extensible Conceptual Model for Tabular Scientific Datasets 72

Javad Chamanara, Michael Owonibi, Alsayed Algergawy, and Roman Gerlach

Context-aware Healthcare Dataset- A Case Study from Pakistan 77
Shahid Mahmud, Rahat Iqbal, and Faiyaz Doctor

Query Acceleration in Multimedia Database Systems 83
Ramzi Haraty and Rawa Karaki