Knowledge Discovery and Data Mining: Challenges and Realities
with Real-word Data

Table of Content

Foreword
Philip S. Yu
IBM TJ Watson Research Center

Introduction

Data Mining in Software Quality Modeling

1. Software Quality Modeling with Limited Apriori Defect Data
   Naeem Seliya¹, Taghi M. Khoshgoftaar²
   ¹ Computer and Information Science, University of Michigan
   ² Computer Science and Engineering, Florida Atlantic University

Knowledge Discovery from Genetic and Medical Data

2. Genome-Wide Analysis of Epistasis Using Multifactor Dimensionality
   Reduction: Feature Selection and Construction in the Domain of Human
   Genetics
   Jason H. Moore
   Departments of Genetics and Community and Family Medicine, Dartmouth Medical School

3. Mining Clinical Trial Data
   Jose Ma. J. Alvir¹, Javier Cabrera², Frank Caridi¹, Ha Nguyen¹
   ¹ Clinical Research Development/Biostatistics, Pfizer Inc, USA
   ² Biostatistics Institute, Dept of Statistics, Rutgers University

Data Mining in Mixed Media Data

4. Cross-Modal Correlation Mining using Graph Algorithms
   Jia-Yu Pan¹, Hyung-Jeong Yang², Christos Faloutsos¹, Pinar Duygulu³
   ¹ Carnegie Mellon University
   ² Chonnam National University, Chonnam National University
   ³ Bilkent University, Turkey

Mining Image Data Repository

5. Image Mining for the Construction of Semantic-Inference Rules and for the
   Development of Automatic Image Diagnosis Systems
   Petra Perner
   Institute of Computer Vision and Applied Computer Sciences, Ibai-Institute, Germany
6. **A Successive Decision Tree Approach to Mining Remotely Sensed Image Data**  
Jianting Zhang, Wieguo Liu, Le Gruenwald  
*the University of New Mexico, USA*

**Data Mining and Business Intelligence**

7. **The Business Impact of Predictive Analytics**  
Tilmann Bruckhaus  
*Numetrics Management Systems, California, USA*

8. **Beyond Classification: Challenges of Data Mining For Credit Scoring**  
Anna Olecka  
*Credit Risk Department, Barclaycard US*

**Data Mining and Ontology Engineering**

9. **Semantics Enhancing Knowledge Discovery and Ontology Engineering Using Mining Techniques: A Crossover Review**  
Irina Neaga  
*University of Amsterdam, The Netherlands*

10. **Knowledge Discovery in Biomedical Data facilitated by Domain Ontologies**  
Amandeep S. Sidhu, Paul J. Kennedy, Simeon Simoff, Tharam S. Dillon  
*University of Technology Sydney, Australia*

**Traditional Data Mining Algorithms**

11. **Effective Intelligent Data Mining Using Dempster-Shafer Theory**  
Malcolm J. Beynon  
*Cardiff University, UK*

12. **Outlier detection strategy using the Self-Organizing Map**  
Fedja Hadzic, Tharam S. Dillon, Henry Tan  
*University of Technology Sydney, Australia*

13. **Re-sampling Based Data Mining Using Rough Set Theory**  
Benjamin Griffiths, Malcolm J. Beynon  
*Cardiff University, UK*