

Table of Contents

Session 1A – Towards Broader Foundations	
Multi-relational Data Mining, Using UML for ILP	1
<i>Arno J. Knobbe, Arno Siebes, Hendrik Blockeel, and Daniël Van Der Wallen</i>	
An Apriori-Based Algorithm for Mining Frequent Substructures from Graph Data	13
<i>Akihiro Inokuchi, Takashi Washio, and Hiroshi Motoda</i>	
Basis of a Fuzzy Knowledge Discovery System	24
<i>Maurice Bernadet</i>	
Session 1B – Rules and Trees	
Confirmation Rule Sets	34
<i>Dragan Gamberger and Nada Lavrač</i>	
Contribution of Dataset Reduction Techniques to Tree-Simplification and Knowledge Discovery	44
<i>Marc Sebban and Richard Nock</i>	
Combining Multiple Models with Meta Decision Trees	54
<i>Ljupčo Todorovski and Sašo Džeroski</i>	
Session 2A – Databases and Reward-Based Learning ...	
Materialized Data Mining Views	65
<i>Tadeusz Morzy, Marek Wojciechowski, and Maciej Zakrzewicz</i>	
Approximation of Frequency Queries by Means of Free-Sets	75
<i>Jean-François Boulicaut, Arthur Bykowski, and Christophe Rigotti</i>	
Application of Reinforcement Learning to Electrical Power System Closed-Loop Emergency Control	86
<i>Christophe Druet, Damien Ernst, and Louis Wehenkel</i>	
Efficient Score-Based Learning of Equivalence Classes of Bayesian Networks	96
<i>Paul Munteanu and Denis Cau</i>	

Session 2B – Classification

Quantifying the Resilience of Inductive Classification Algorithms	106
<i>Melanie Hilario and Alexandros Kalousis</i>	
Bagging and Boosting with Dynamic Integration of Classifiers	116
<i>Alexey Tsymbal and Seppo Puuronen</i>	
Zoomed Ranking: Selection of Classification Algorithms Based on Relevant Performance Information	126
<i>Carlos Soares and Pavel B. Brazdil</i>	
Some Enhancements of Decision Tree Bagging	136
<i>Pierre Geurts</i>	

Session 3A – Association Rules and Exceptions

Relative Unsupervised Discretization for Association Rule Mining	148
<i>Marcus-Christopher Ludl and Gerhard Widmer</i>	
Mining Association Rules: Deriving a Superior Algorithm by Analyzing Today's Approaches	159
<i>Jochen Hipp, Ulrich Güntzer, and Gholamreza Nakhaeizadeh</i>	
Unified Algorithm for Undirected Discovery of Exception Rules	169
<i>Einoshin Suzuki and Jan Żytkow</i>	
Sampling Strategies for Targeting Rare Groups from a Bank Customer Database	181
<i>Jean-Hugues Chauchat, Ricco Rakotomalala, and Didier Robert</i>	

Session 3B – Instance-Based Discovery

Instance-Based Classification by Emerging Patterns	191
<i>Jinyan Li, Guozhu Dong, and Kotagiri Ramamohanarao</i>	
Context-Based Similarity Measures for Categorical Databases	201
<i>Gautam Das and Heikki Mannila</i>	
A Mixed Similarity Measure in Near-Linear Computational Complexity for Distance-Based Methods	211
<i>Ngoc Binh Nguyen and Tu Bao Ho</i>	
Fast Feature Selection Using Partial Correlation for Multi-valued Attributes	221
<i>Stéphane Lallich and Ricco Rakotomalala</i>	

Session 4A – Clustering and Classification

Fast Hierarchical Clustering Based on Compressed Data and OPTICS	232
<i>Markus M. Breunig, Hans-Peter Kriegel, and Jörg Sander</i>	

Accurate Recasting of Parameter Estimation Algorithms Using Sufficient Statistics for Efficient Parallel Speed-Up: Demonstrated for Center-Based Data Clustering Algorithms	243
<i>Bin Zhang, Meichun Hsu, and George Forman</i>	

Predictive Performance of Weighted Relative Accuracy	255
<i>Ljupčo Todorovski, Peter Flach, and Nada Lavrač</i>	

Quality Scheme Assessment in the Clustering Process	265
<i>Maria Halkidi, M. Vazirgiannis, and Y. Batistakis</i>	

Session 5A – Time Series

Algorithm for Matching Sets of Time Series	277
<i>Iztok Savnik, Georg Lausen, Hans-Peter Kahle, Heinrich Spiecker, and Sebastian Hein</i>	

MSTS: A System for Mining Sets of Time Series	289
<i>Georg Lausen, Iztok Savnik, and Aldar Dougarjapov</i>	

Learning First Order Logic Time Series Classifiers: Rules and Boosting . . .	299
<i>Juan J. Rodríguez, Carlos J. Alonso, and Henrik Boström</i>	

Posters

Learning Right Sized Belief Networks by Means of a Hybrid Methodology .	309
<i>Sylvia Acid and Luis M. De Campos</i>	

Algorithms for Mining Share Frequent Itemsets Containing Infrequent Subsets	316
<i>Brock Barber and Howard J. Hamilton</i>	

Discovering Task Neighbourhoods through Landmark Learning Performances	325
<i>Hilan Bensusan and Christophe Giraud-Carrier</i>	

Induction of Multivariate Decision Trees by Using Dipolar Criteria	331
<i>Leon Bobrowski and Marek Krętowski</i>	

Inductive Logic Programming in Clementine	337
<i>Sam Brewer and Tom Khabaza</i>	

A Genetic Algorithm-Based Solution for the Problem of Small Disjuncts . .	345
<i>Deborah R. Carvalho and Alex A. Freitas</i>	

Clustering Large, Multi-level Data Sets: An Approach Based on Kohonen Self Organizing Maps	353
<i>Antonio Ciampi and Yves Lechevallier</i>	
Trees and Induction Graphs for Multivariate Response	359
<i>Antonio Ciampi, Djamel A. Zighed, and J��r��my Clech</i>	
CEM - Visualisation and Discovery in Email	367
<i>Richard Cole, Peter Eklund and Gerd Stumme</i>	
Image Access and Data Mining: An Approach	375
<i>Chabane Djeraba</i>	
Decision Tree Toolkit: A Component-Based Library of Decision Tree Algorithms	381
<i>Nikos Drossos, Athanasios Papagelis, and Dimitris Kalles</i>	
Determination of Screening Descriptors for Chemical Reaction Databases .	388
<i>Laurent Dury, Laurence Leherter, and Daniel P. Vercauteren</i>	
Prior Knowledge in Economic Applications of Data Mining	395
<i>A.J. Feelders</i>	
Temporal Machine Learning for Switching Control	401
<i>Pierre Geurts and Louis Wehenkel</i>	
Improving Dissimilarity Functions with Domain Knowledge, Applications with IKBS System	409
<i>David Grosser, Jean Diatta, and No��l Conr��yt</i>	
Mining Weighted Association Rules for Fuzzy Quantitative Items	416
<i>Attila Gyenesei</i>	
Centroid-Based Document Classification: Analysis and Experimental Results	424
<i>Eui-Hong (Sam) Han and George Karypis</i>	
Applying Objective Interestingness Measures in Data Mining Systems	432
<i>Robert J. Hilderman and Howard J. Hamilton</i>	
Observational Logic Integrates Data Mining Based on Statistics and Neural Networks	440
<i>Martin Hole��a</i>	
Supporting Discovery in Medicine by Association Rule Mining of Bibliographic Databases	446
<i>Dimitar Hristovski, Sa��o D��zeroski, Borut Peterlin, and Anamarija Rozic-Hristovski</i>	

Collective Principal Component Analysis from Distributed, Heterogeneous Data	452
<i>Hillol Kargupta, Weiyun Huang, Krishnamoorthy Sivakumar, Byung-Hoon Park, and Shuren Wang</i>	
Hierarchical Document Clustering Based on Tolerance Rough Set Model ..	458
<i>Saori Kawasaki, Ngoc Binh Nguyen, and Tu Bao Ho</i>	
Application of Data-Mining and Knowledge Discovery in Automotive Data Engineering	464
<i>Jörg Keller, Valerij Bauer, and Wojciech Kwedlo</i>	
Towards Knowledge Discovery from cDNA Microarray Gene Expression Data	470
<i>Jan Komorowski, Torgeir R. Hvidsten, Tor-Kristian Jenssen, Dyre Tjeldvoll, Eivind Hovig, Arne K. Sanvik, and Astrid Lægreid</i>	
Mining with Cover and Extension Operators	476
<i>Marzena Kryszkiewicz</i>	
A User-Driven Process for Mining Association Rules	483
<i>Pascale Kuntz, Fabrice Guillet, Rémi Lehn, and Henri Briand</i>	
Learning from Labeled and Unlabeled Documents: A Comparative Study on Semi-Supervised Text Classification	490
<i>Carsten Lanquillon</i>	
Schema Mining: Finding Structural Regularity among Semistructured Data	498
<i>P.A. Laur, F. Massegli, and P. Poncelet</i>	
Improving an Association Rule Based Classifier	504
<i>Bing Liu, Yiming Ma, and Ching Kian Wong</i>	
Discovery of Generalized Association Rules with Multiple Minimum Supports	510
<i>Chung-Leung Lui and Fu-Lai Chung</i>	
Learning Dynamic Bayesian Belief Networks Using Conditional Phase-Type Distributions	516
<i>Adele Marshall, Sally McClean, Mary Shapcott, and Peter Millard</i>	
Discovering Differences in Patients with Uveitis through Typical Testors by Class	524
<i>José F. Martínez-Trinidad, Miriam Velasco-Sánchez, and Edgar E. Contreras-Aravelo</i>	

Web Usage Mining: How to Efficiently Manage New Transactions and New Clients	530
<i>F. Masseglia, P. Poncelet, and M. Teisseire</i>	
Mining Relational Databases	536
<i>Frédéric Moal, Teddy Turmeaux, and Christel Vrain</i>	
Interestingness in Attribute-Oriented Induction (AOI): Multiple-Level Rule Generation	542
<i>Maybin K. Muyebe and John A. Keane</i>	
Discovery of Characteristic Subgraph Patterns Using Relative Indexing and the Cascade Model	550
<i>Takashi Okada and Mayumi Oyama</i>	
Transparency and Predictive Power: Explaining Complex Classification Models	558
<i>Gerhard Paass and Jörg Kindermann</i>	
Clustering Distributed Homogeneous Datasets	566
<i>Srinivasan Parthasarathy and Mitsunori Ogihara</i>	
Empirical Evaluation of Feature Subset Selection Based on a Real-World Data Set	575
<i>Petra Perner and Chid Apte</i>	
Discovery of Ambiguous Patterns in Sequences: Application to Bioinformatics	581
<i>Gerard Ramstein, Pascal Bunelle, and Yannick Jacques</i>	
Action-Rules: How to Increase Profit of a Company	587
<i>Zbigniew W. Ras and Alicja Wiecezorkowska</i>	
Aggregation and Association in Cross Tables	593
<i>Gilbert Ritschard and Nicolas Nicoloyannis</i>	
An Experimental Study of Partition Quality Indices in Clustering	599
<i>Céline Robardet, Fabien Feschet, and Nicolas Nicoloyannis</i>	
Expert Constrained Clustering: A Symbolic Approach	605
<i>Fabrice Rossi and Frédérick Vautrain</i>	
An Application of Association Rules Discovery to Geographic Information Systems	613
<i>Ansaf Salleb and Christel Vrain</i>	
Generalized Entropy and Projection Clustering of Categorical Data	619
<i>Dan A. Simovici, Dana Cristofor, and Laurentiu Cristofor</i>	

Supporting Case Acquisition and Labelling in the Context of Web Mining	626
<i>Vojtěch Svátek and Martin Kavalec</i>	
Indirect Association: Mining Higher Order Dependencies in Data	632
<i>Pang-Ning Tan, Vipin Kumar, and Jaideep Srivastava</i>	
Discovering Association Rules in Large, Dense Databases	638
<i>Tudor Teusan, Gilles Nachouki, Henri Briand, and Jacques Philippe</i>	
Providing Advice to Website Designers Towards Effective Websites	
Re-organization	646
<i>Peter Tselios, Agapios Platis, and George Vouros</i>	
Clinical Knowledge Discovery in Hospital Information Systems: Two Case Studies	652
<i>Shusaku Tsumoto</i>	
Knowledge Discovery Using Least Squares Support Vector Machine	
Classifiers: A Direct Marketing Case	657
<i>Stijn Viaene, B. Baesens, T. Van Gestel, J.A.K. Suykens, D. Van Den Poel, J. Vanthienen, D. De Moor, and G. Dedene</i>	
Lightweight Document Clustering	665
<i>Sholom M. Weiss, Brian F. White, and Chidanand V. Apte</i>	
Automatic Category Structure Generation and Categorization of Chinese Text Documents	673
<i>Hsin-Chang Yang and Chung-Hong Lee</i>	
Mining Generalized Multiple-Level Association Rules	679
<i>Show-Jane Yen</i>	
An Efficient Approach to Discovering Sequential Patterns in Large Databases	685
<i>Show-Jane Yen and Chung-Wen Cho</i>	
Using Background Knowledge as a Bias to Control the Rule Discovery Process	691
<i>Ning Zhong, Juzhen Dong, and Setsuo Ohsuga</i>	
Author Index	699

Principles of Data Mining and Knowledge Discovery
4th European Conference, PKDD, 2000, Lyon, France,
September 13-16, 2000 Proceedings
Zighed, D.A.; Komorowski, J.; Zytkow, J. (Eds.)
2000, XV, 701 p., Softcover
ISBN: 978-3-540-41066-9