Contents

Regular Papers

Self Similar Layered Hidden Markov Models ................................. 1
	Jafar Adibi and Wei-Min Shen

Automatic Text Summarization Using Unsupervised and Semi-supervised
Learning................................................................. 16
	Massih-Reza Amini and Patrick Gallinari

Detecting Temporal Changes in Event Sequences: An Application to
Demographic Data......................................................... 29
	Hendrik Blockeel, Johannes Fürnkranz, Alexia Prskawetz, and
	Francesco C. Billari

Knowledge Discovery in Multi-label Phenotype Data....................... 42
	Amanda Clare and Ross D. King

Computing Association Rules Using Partial Totals .......................... 54
	Frans Coenen, Graham Goulbourne, and Paul Leng

Gaphyl: A Genetic Algorithms Approach to Cladistics ..................... 67
	Clare Bates Congdon

Parametric Approximation Algorithms for High-Dimensional Euclidean
Similarity........................................................................ 79
	Ömer Eğecioğlu

Data Structures for Minimization of Total Within-Group Distance for
Spatio-Temporal Clustering.................................................. 91
	Vladimir Estivill-Castro and Michael E. Houle

Non-crisp Clustering by Fast, Convergent, and Robust Algorithms ...... 103
	Vladimir Estivill-Castro and Jianhua Yang

Pattern Extraction for Time Series Classification.......................... 115
	Pierre Geurts

Specifying Mining Algorithms with Iterative User-Defined Aggregates:
A Case Study ............................................................. 128
	Fosca Giannotti, Giuseppe Manco, and Franco Turini

Interesting Fuzzy Association Rules in Quantitative Databases......... 140
	Jeannette M. de Graaf, Walter A. Kosters, and
	Jeroen J.W. Witteman
Interestingness Measures for Fuzzy Association Rules .................. 152
Attila Gyenesesi and Jukka Teuhola

A Data Set Oriented Approach for Clustering Algorithm Selection ...... 165
Maria Halkidi and Michalis Vazirgiannis

Fusion of Meta-knowledge and Meta-data for Case-Based Model Selection. 180
Melanie Hilario and Alexandros Kalousis

Discovery of Temporal Patterns. Learning Rules about the Qualitative
Behaviour of Time Series ................................................... 192
Frank Höppner

Temporal Rule Discovery for Time-Series Satellite Images and Integration
with RDB ................................................................. 204
Rie Honda and Osamu Komishi

Using Grammatical Inference to Automate Information Extraction
from the Web .................................................................. 216
Theodore W. Hong and Keith L. Clark

Biological Sequence Data Mining ......................................... 228
Yuh-Jyh Hu

Implication-Based Fuzzy Association Rules ................................. 241
Eyke Hüllermeier

A General Measure of Rule Interestingness ............................... 253
Szymon Jaroszewicz and Dan A. Simovici

Error Correcting Codes with Optimized Kullback-Leibler Distances for Text
Categorization .............................................................. 266
Jörg Kindermann, Gerhard Paass, and Edda Leopold

Propositionalisation and Aggregates ...................................... 277
Arno J. Knobbe, Marc de Haas, and Arno Siebes

Algorithms for the Construction of Concept Lattices and Their Diagram
Graphs ........................................................................ 289
Sergei Kuznetsov and Sergei Obiedkov

Data Reduction Using Multiple Models Integration ..................... 301
Aleksandar O. Lazarevic and Zoran A. Obradovic

Discovering Fuzzy Classification Rules with Genetic Programming and
Co-evolution ............................................................. 314
Roberto R.F. Mendes, Fabricio de B. Voznika, Alex A. Freitas, and
Julio C. Nievola
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence Filtering for Information Extraction in Genomics, a Classification</td>
<td>326</td>
</tr>
<tr>
<td>Claire Nédellec, Mohamed Ould Abdel Vetah, and Philippe Bessières</td>
<td></td>
</tr>
<tr>
<td>Text Categorization and Semantic Browsing with Self-Organizing Maps on Non-euclidean Spaces</td>
<td>338</td>
</tr>
<tr>
<td>Jörg Ontrup and Helge Ritter</td>
<td></td>
</tr>
<tr>
<td>A Study on the Hierarchical Data Clustering Algorithm Based on Gravity Theory</td>
<td>350</td>
</tr>
<tr>
<td>Yen-Jen Oyang, Chien-Yu Chen, and Tsai-Wei Yang</td>
<td></td>
</tr>
<tr>
<td>Internet Document Filtering Using Fourier Domain Scoring</td>
<td>362</td>
</tr>
<tr>
<td>Laurence A.F. Park, Marimuthu Palaniswami, and Ramamohanarao Koitagiri</td>
<td></td>
</tr>
<tr>
<td>Distinguishing Natural Language Processes on the Basis of fMRI-Measured Brain Activation</td>
<td>374</td>
</tr>
<tr>
<td>Francisco Pereira, Marcel Just, and Tom Mitchell</td>
<td></td>
</tr>
<tr>
<td>Automatic Construction and Refinement of a Class Hierarchy over Multivalued Data</td>
<td>386</td>
</tr>
<tr>
<td>Nathalie Pernelle, Marie-Christine Rousset, and Véronique Ventos</td>
<td></td>
</tr>
<tr>
<td>Comparison of Three Objective Functions for Conceptual Clustering</td>
<td>399</td>
</tr>
<tr>
<td>Céline Robardet and Fabien Feschet</td>
<td></td>
</tr>
<tr>
<td>Identification of ECG Arrhythmias Using Phase Space Reconstruction</td>
<td>411</td>
</tr>
<tr>
<td>Felice M. Roberts, Richard J. Povinelli, and Kristina M. Ropella</td>
<td></td>
</tr>
<tr>
<td>Finding Association Rules That Trade Support Optimally against Confidence</td>
<td>424</td>
</tr>
<tr>
<td>Tobias Scheffer</td>
<td></td>
</tr>
<tr>
<td>Bloomy Decision Tree for Multi-objective Classification</td>
<td>436</td>
</tr>
<tr>
<td>Einoshin Suzuki, Masafumi Gotoh, and Yuta Choki</td>
<td></td>
</tr>
<tr>
<td>Discovery of Temporal Knowledge in Medical Time-Series Databases Using Moving Average, Multiscale Matching, and Rule Induction</td>
<td>448</td>
</tr>
<tr>
<td>Shusaku Tsumoto</td>
<td></td>
</tr>
<tr>
<td>Mining Positive and Negative Knowledge in Clinical Databases Based on Rough Set Model</td>
<td>460</td>
</tr>
<tr>
<td>Shusaku Tsumoto</td>
<td></td>
</tr>
<tr>
<td>The TwoKey Plot for Multiple Association Rules Control</td>
<td>472</td>
</tr>
<tr>
<td>Antony Unwin, Heike Hofmann, and Klaus Bernt</td>
<td></td>
</tr>
</tbody>
</table>
Lightweight Collaborative Filtering Method for Binary-Encoded Data
Sholom M. Weiss and Nitin Indurkhya

Invited Papers

Support Vectors for Reinforcement Learning
Thomas G. Dietterich and Xin Wang

Combining Discrete Algorithmic and Probabilistic Approaches
in Data Mining
Heikki Mannila

Statistification or Mystification? The Need for Statistical Thought
in Visual Data Mining
Antony Unwin

The Musical Expression Project: A Challenge for Machine Learning and
Knowledge Discovery
Gerhard Widmer

Scalability, Search, and Sampling: From Smart Algorithms
to Active Discovery
Stefan Wrobel

Author Index
Principles of Data Mining and Knowledge Discovery
5th European Conference, PKDD 2001, Freiburg, Germany, September 3-5, 2001 Proceedings
Raedt, L. de; Siebes, A. (Eds.)
2001, DXXII, 514 p., Softcover
ISBN: 978-3-540-42534-2