

Table of Contents

Invited Talks

Mathematics Based on Learning	1
<i>Susumu Hayashi</i>	
Data Mining with Graphical Models	2
<i>Rudolf Kruse, Christian Borgelt</i>	
On the Eigenspectrum of the Gram Matrix and Its Relationship to the Operator Eigenspectrum	12
<i>John Shawe-Taylor, Chris Williams, Nello Cristianini, Jaz Kandola</i>	
In Search of the Horowitz Factor: Interim Report on a Musical Discovery Project	13
<i>Gerhard Widmer</i>	
Learning Structure from Sequences, with Applications in a Digital Library	33
<i>Ian H. Witten</i>	

Regular Papers

Application of Discovery to Natural Science

Discovering Frequent Structured Patterns from String Databases: An Application to Biological Sequences	34
<i>Luigi Palopoli, Giorgio Terracina</i>	
Discovery in Hydrating Plaster Using Machine Learning Methods	47
<i>Judith E. Devaney, John G. Hagedorn</i>	
Revising Qualitative Models of Gene Regulation	59
<i>Kazumi Saito, Stephen Bay, Pat Langley</i>	

Knowledge Discovery from Unstructured and Semi-structured Data

SEuS: Structure Extraction Using Summaries	71
<i>Shayan Ghazizadeh, Sudarshan S. Chawathe</i>	
Discovering Best Variable-Length-Don't-Care Patterns	86
<i>Shunsuke Inenaga, Hideo Bannai, Ayumi Shinohara, Masayuki Takeda, Setsuo Arikawa</i>	

Meta-learning and Analysis of Machine Learning Algorithms

A Study on the Effect of Class Distribution Using Cost-Sensitive Learning	98
<i>Kai Ming Ting</i>	
Model Complexity and Algorithm Selection in Classification	113
<i>Melanie Hilario</i>	
Experiments with Projection Learning	127
<i>Tapio Elomaa, J.T. Lindgren</i>	
Improved Dataset Characterisation for Meta-learning	141
<i>Yonghong Peng, Peter A. Flach, Carlos Soares, Pavel Brazdil</i>	

Combining Machine Learning Algorithms

Racing Committees for Large Datasets	153
<i>Eibe Frank, Geoffrey Holmes, Richard Kirkby, Mark Hall</i>	
From Ensemble Methods to Comprehensible Models	165
<i>C. Ferri, J. Hernández-Orallo, M.J. Ramírez-Quintana</i>	

Neural Networks and Statistical Methods

Learning the Causal Structure of Overlapping Variable Sets	178
<i>David Danks</i>	
Extraction of Logical Rules from Data by Means of Piecewise-Linear Neural Networks	192
<i>Martin Holeňa</i>	
Structuring Neural Networks through Bidirectional Clustering of Weights	206
<i>Kazumi Saito, Ryohei Nakano</i>	

New Approaches to Knowledge Discovery

Toward Drawing an Atlas of Hypothesis Classes: Approximating a Hypothesis via Another Hypothesis Model	220
<i>Osamu Maruyama, Takayoshi Shoudai, Satoru Miyano</i>	
Datascape Survey Using the Cascade Model	233
<i>Takashi Okada</i>	
Learning Hierarchical Skills from Observation	247
<i>Ryutaro Ichise, Daniel Shapiro, Pat Langley</i>	

Poster Papers

Applications of Knowledge Discovery to Natural Science

Image Analysis for Detecting Faulty Spots from Microarray Images	259
<i>Salla Ruosaari, Jaakko Hollmén</i>	
Inferring Gene Regulatory Networks from Time-Ordered Gene Expression Data Using Differential Equations	267
<i>Michiel de Hoon, Seiya Imoto, Satoru Miyano</i>	
DNA-Tract Curvature Profile Reconstruction: A Fragment Flipping Algorithm	275
<i>Daniele Masotti</i>	
Evolution Map: Modeling State Transition of Typhoon Image Sequences by Spatio-Temporal Clustering	283
<i>Asanobu Kitamoto</i>	
Structure-Sweetness Relationships of Aspartame Derivatives by GUHA . . .	291
<i>Jaroslava Halova, Premysl Zak, Pavel Stopka, Tomoaki Yuzuri, Yukino Abe, Kazuhisa Sakakibara, Hiroko Suezawa, Minoru Hirota</i>	

Knowledge Discovery from Texts

A Hybrid Approach for Chinese Named Entity Recognition	297
<i>Xiaoshan Fang, Huanye Sheng</i>	
Extraction of Word Senses from Human Factors in Knowledge Discovery	302
<i>Yoo-Jin Moon, Minkoo Kim, Youngho Hwang, Pankoo Kim, Ki-joon Choi</i>	
Event Pattern Discovery from the Stock Market Bulletin	310
<i>Fang Li, Huanye Sheng, Dongmo Zhang</i>	
Email Categorization Using Fast Machine Learning Algorithms	316
<i>Jihoon Yang, Sung-Yong Park</i>	
Discovery of Maximal Analogies between Stories	324
<i>Makoto Haraguchi, Shigetora Nakano, Masaharu Yoshioka</i>	
Automatic Wrapper Generation for Multilingual Web Resources	332
<i>Yasuhiro Yamada, Daisuke Ikeda, Sachio Hirokawa</i>	
Combining Multiple K-Nearest Neighbor Classifiers for Text Classification by Reducts	340
<i>Yongguang Bao, Naohiro Ishii</i>	

ARISTA Causal Knowledge Discovery from Texts 348
John Kontos, Areti Elmaoglou, Ioanna Malagardi

Applications of Knowledge Discovery to Social Science

Knowledge Discovery as Applied to Music: Will Music Web Retrieval
Revolutionize Musicology? 356
Francis Rousseaux, Alain Bonardi

Process Mining: Discovering Direct Successors in Process Logs 364
*Laura Maruster, A.J.M.M. Weijters,
W.M.P. van der Aalst, Antal van den Bosch*

The Emergence of Artificial Creole by the EM Algorithm 374
Makoto Nakamura, Satoshi Tojo

Generalized Musical Pattern Discovery by Analogy from
Local Viewpoints 382
Olivier Lartillot

Machine Learning Approaches

Using Genetic Algorithms-Based Approach for Better Decision
Trees: A Computational Study 390
Zhiwei Fu

Handling Feature Ambiguity in Knowledge Discovery from Time Series . . . 398
Frank Höppner

A Compositional Framework for Mining Longest Ranges 406
Haiyan Zhao, Zhenjiang Hu, Masato Takeichi

Post-processing Operators for Browsing Large Sets of
Association Rules 414
Alipio Jorge, João Poças, Paulo Azevedo

Mining Patterns from Structured Data by Beam-Wise
Graph-Based Induction 422
Takashi Matsuda, Hiroshi Motoda, Tetsuya Yoshida, Takashi Washio

Feature Selection for Propositionalization 430
Mark-A. Krogel, Stefan Wrobel

Subspace Clustering Based on Compressibility 435
Masaki Narahashi, Einoshin Suzuki

New Approaches to Knowledge Discovery

The Extra-Theoretical Dimension of Discovery. Extracting Knowledge by Abduction	441
<i>Lorenzo Magnani, Matteo Piazza, Riccardo Dossena</i>	
Discovery Process on the WWW: Analysis Based on a Theory of Scientific Discovery	449
<i>Hitomi Saito, Kazuhisa Miwa</i>	
Invention vs. Discovery (A Critical Discussion)	457
<i>Carlotta Piscopo, Mauro Birattari</i>	
Author Index	463

Discovery Science

5th International Conference, DS 2002, Lubeck,
Germany, November 24-26, 2002, Proceedings

Lange, S.; Satoh, K.; Smith, C.H. (Eds.)

2002, XIV, 470 p., Softcover

ISBN: 978-3-540-00188-1