Hujun Yin Peter Tino Emilio Corchado Will Byrne Xin Yao (Eds.)

Intelligent Data Engineering and Automated Learning – IDEAL 2007

8th International Conference Birmingham, UK, December 16-19, 2007 Proceedings



Volume Editors

Hujun Yin The University of Manchester Manchester, M60 1QD, UK E-mail: hujun.yin@manchester.ac.uk

Peter Tino
Will Byrne
Xin Yao
University of Birmingham
Birmingham B15 2TT, UK
E-mail: {p.tino, w.f.byrne, x.yao}@cs.bham.ac.uk

Emilio Corchado University of Burgos 09001 Burgos, Spain

E-mail: escorchado@ubu.es

Library of Congress Control Number: 2007941157

CR Subject Classification (1998): H.2.8, F.2.2, I.2, F.4, K.4.4, H.3, H.4

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

ISSN 0302-9743

ISBN-10 3-540-77225-1 Springer Berlin Heidelberg New York ISBN-13 978-3-540-77225-5 Springer Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer. Violations are liable to prosecution under the German Copyright Law.

Springer is a part of Springer Science+Business Media

springer.com

© Springer-Verlag Berlin Heidelberg 2007 Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India Printed on acid-free paper SPIN: 12202126 06/3180 5 4 3 2 1 0

Preface

After a vibrant and successful event in Burgos, Spain, last year, this year's international conference on Intelligent Data Engineering and Automated Learning—IDEAL 2007 (http://events.cs.bham.ac.uk/ideal07/)—was held in the second largest city of the UK, Birmingham. The IDEAL conference has become a unique multidisciplinary forum for researchers in both theoretical and practical aspects of learning and information processing, data mining, retrieval and management, bioinformatics and bio-inspired models, agents and hybrid systems, and financial engineering. A special feature of the IDEAL conferences is the cross-disciplinary exchange of ideas in emerging techniques and applications in these areas. Data engineering and associated learning paradigms are playing increasingly important roles in an increasing number of disciplines and fields. The multidisciplinary nature of contemporary research and modern technology is pushing boundaries and one of the principal aims of the IDEAL conference is to promote interactions and collaborations across disciplines.

This volume of Lecture Notes in Computer Science contains accepted papers presented at IDEAL 2007 held at the University of Birmingham, UK, during December 16–19, 2007. This year, the conference received over 270 submissions from around the world, which were subsequently peer-referred by the Programme Committee comprising leading scholars in the field. Each paper was rigorously reviewed by two reviewers and only papers that had received positive comments from both reviewers were accepted and included in the proceedings in order to maintain the highest quality of the conference. This resulted in about 110 top quality papers for the conference and the proceedings. The acceptance rate was about 40%. The buoyant numbers of submissions in recent years are a clear indication of the importance of the fields related to IDEAL and the popularity of the IDEAL conference and community. This year's conference had five regular themes: Learning and Information Processing, Data Mining and Information Management, Bioinformatics and Neuroinformatics, Agents and Distributed Systems, and Financial Engineering and Modelling. Two special sessions, Agentbased Approach to Service Sciences and Neural-evolutionary Fusion Algorithms and Their Application, were also organized by Akira Namatame and Sancho Salcedo-Sanz, respectively. Many new ideas, novel algorithms and emerging techniques were reported and discussed at the conference.

This 8th IDEAL conference also enjoyed outstanding keynote speeches by distinguished guest speakers: Piero Bonissone, GE Global Research, Kevin Kelly, Carnegie Mellon University, Toby Gibson, European Molecular Biology Laboratory, and Soo-Young Lee of the Korea Advanced Institute of Science and Technology. Their in-depth coverage and work on various challenging topics served as both exceptional examples and inspirations for others.

This year IDEAL also teamed up with two international journals, namely, the *International Journal of Neural Systems* and the *Journal of Mathematical Modelling and Algorithm* for two special issues. The extended papers, together with contributed articles received in response to subsequent open calls, will go through further rounds of peer refereeing in the remits of these two journals.

We would like to thank the International Advisory Committee and the Steering Committee for the guidance and advice. We would particularly like to acknowledge the diligent work of our Programme Committee members, who performed review tasks admirably under tight deadline pressures.

Particular thanks go to CERCIA (The Centre of Excellence for Research in Computational Intelligence and Applications) and the School of Computer Science (especially Ceinwen Cushway) of the University of Birmingham, for their support and organization of IDEAL 2007.

We are also grateful to the publisher, Springer, especially Alfred Hofmann and Anna Kramer at the LNCS Editorial Office, for their continued support and collaboration in this demanding publication project.

Last but not the least we thank all the authors and participants for their contributions that made this conference such a successful and enjoyable event.

October 2007

Hujun Yin Peter Tino Emilio Corchado Will Byrne Xin Yao

Organization

General Co-chairs

Xin Yao University of Birmingham, UK

Benjamin Wah University of Illinois, Urbana-Champaign, USA

International Advisory Committee

Lei Xu (Chair) Chinese University of Hong Kong

Yaser Abu-Mostafa CALTECH, USA Shun-ichi Amari RIKEN, Japan

Michael Dempster University of Cambridge, UK Sun-Yung Kung Princeton University, USA

Erkki Oja Helsinki University of Technology, Finland

Latit M. Patnaik Indian Institute of Science, India

Steering Committee

Hujun Yin (Co-chair) University of Manchester, UK Lai-Wan Chan (Co-chair) Chinese University of Hong Kong

Nigel Allinson University of Sheffield, UK

Yiu-ming Cheung Hong Kong Baptist University, Hong Kong

Emilio Corchado University of Burgos, Spain

Marcus Gallagher University of Queensland, Australia

Marc van Hulle K.U. Leuven, Belgium

John Keane University of Manchester, UK

Jimmy Lee Chinese University of Hong Kong, Hong Kong

Malik Magdon-Ismail Rensselaer Polytechnic Institute, USA

Zheng Rong Yang University of Exeter, UK

Ron Sun Rensselaer Polytechnic Institute, USA Ning Zhong Maebashi Institute of Technology, Japan

Programme Committee Chairs

Hujun Yin (Chair)

Peter Tino (Technical Chair)

Emilio Corchado (Technical Co-chair)

Malik Magdon-Ismail (Technical Co-chair)

University of Birmingham, UK

University of Burgos, Spain

Rensselaer Polytechnic Institute,

USA

Zheng Rong Yang (Technical Co-chair) University of Exeter, UK

Will Byrne (Organizing Chair)

University of Birmingham, UK

Programme Committee

Ajith Abraham José Adserias Khurshid Ahmad Nigel Allinson Ángel Alonso Luis Alonso Martyn Amos Davide Anguita Bruno Apolloni Jiyuan An Javier Bajo Federico Barber Bruno Baruque Lubica Benuskova Michael Biehl Alan Blair Mikael Boden Lourdes Borrajo Juan Botía Vicente Botti Teodoro Calonge Carlos Carrascosa Andre de Carvalho Matthew Casev Sheng Chen Songcan Chen Sung-Bae Cho Sungzoon Cho

Juan Manuel Corchado Rafael Corchuelo

David Corne
Ernesto Costa
Robert Dale
Bernard De Baets
Yanira De Paz
Ricardo Del Olmo
Miguel Delgado
Fernando Díaz
Zhao Yang Dong
José Dorronsoro
Richard Everson

Fernando Fernández

Florentino Fernández

Marcus Frean

Igor Farkas

Richard Freeman Toshio Fukuda Colin Fyfe Bogdan Gabrys Marcus Gallagher John Qiang Gan Francisco Garijo Mark Girolami Antonio F. Gómez Ana González Angélica González Manuel González Daniel González Francisco Herrera Alvaro Herrero James Hogan Tony Holden Jaakko Hollmen Robert J. Howlett David Hoyle Hisao Ishibuchi

Ata Kaban Hoon Kang Juha Karhunen Samuel Kaski Dong Hwa Kim Irwin King Aldebaro Klautau

Paul Jackway

Gareth Jones

Vicente Julián

Mario Köppen
Kostadin Korutchev
Rudolf Kruse
Rosalía Laza
Kwong S. Leung
Carlos Linares
Paulo Lisboa
Eva Lorenzo
Frederic Maire
Roque Marín
José F. Martínez

Aitor Mata José Ramón Méndez

Simon Miles

José Mira

José Manuel Molina Carla Möller-Levet

Joaquín Pacheco

Juan Pavón David Pelta David Powers José Principe

José Ramirez Omer Rana

Vic Rayward-Smith Perfecto Reguera

Bernadete Ribeiro José Riquelme

Ramón Rizo Roman Rosipal Dymitr Ruta

Shazia Sadiq José Santos Michael Small P.N. Suganthan David Taniar Dante Israel Tapia

Miguel Toro Marcos Valiño Marc Van Hulle Alfredo Vellido

José Ramón Villar

Thomas Villmann Lipo Wang Dong-Qing Wei Ian Wood

Gordon Wyeth Yong Xu

Nobuyoshi Yabuki Ronald R.Yager Du Zhang Yanqing Zhang Ning Zhong

Rodolfo Zunino

Table of Contents

Learning and Information Processing	
Support Function Machines	1
Different Bayesian Network Models in the Classification of Remote Sensing Images	10
Group Decision Making with Triangular Fuzzy Linguistic Variables Zeshui Xu	17
Sparse Kernel Modelling: A Unified Approach	27
Advanced Forecasting and Classification Technique for Condition Monitoring of Rotating Machinery	37
Out of Bootstrap Estimation of Generalization Error Curves in Bagging Ensembles	47
An Edit Distance Approach to Shallow Semantic Labeling	57
A Comparison of One-Class Classifiers for Novelty Detection in Forensic Case Data	67
Variational GTM	77
Skill Combination for Reinforcement Learning	87
A New Recurring Multistage Evolutionary Algorithm for Solving Problems Efficiently	97
The Effect of Missing Wind Speed Data on Wind Power Estimation Fatih Onur Hocaoğlu and Mehmet Kurban	107

Exploration of a Text Collection and Identification of Topics by Clustering	115
Asynchronous BCI Control of a Robot Simulator with Supervised Online Training	125
Fuzzy Ridge Regression with Non Symmetric Membership Functions and Quadratic Models	135
A Subjective and Objective Integrated Method for MAGDM Problems with Multiple Types of Exact Preference Formats	145
Energy Saving by Means of Fuzzy Systems	155
A Comparative Study of Local Classifiers Based on Clustering Techniques and One-Layer Neural Networks	168
Filter Methods for Feature Selection – A Comparative Study	178
FPGA-Based Architecture for Computing Testors	188
Minimal BSDT Abstract Selectional Machines and Their Selectional and Computational Performance	198
Active Learning for Regression Based on Query by Committee	209
Influence of Wavelet Frequency and Orientation in an SVM-Based Parallel Gabor PCA Face Verification System	219
Wrapping the Naive Bayes Classifier to Relax the Effect of Dependences	229

Vincenzo Cannella, Riccardo Rizzo, and Roberto Pirrone

356

A Linear Learning Method for Multilayer Perceptrons Using	
Least-Squares	365
Deutiz I etez-Sanchez, ana I auta Fragueta	
A Discriminative Model Corresponding to Hierarchical HMMs	375
Finding Unsatisfiable Subformulas with Stochastic Method	385
A New Efficient Approach in Clustering Ensembles	395
An Evolutionary Hyperheuristic to Solve Strip-Packing Problems Pablo Garrido and María-Cristina Riff	406
Statistical Analysis of Sample-Size Effects in ICA	416
HPGP: An Abstraction-Based Framework for Decision-Theoretic	
Planning	426
Correction of Medical Handwriting OCR Based on Semantic	
Similarity	437
Multiple Classifier Fusion Using k -Nearest Localized Templates Jun-Ki Min and Sung-Bae Cho	447
Data Mining and Information Management	
Color Image Segmentation Applied to Medical Domain Liana Stanescu, Dan Dumitru Burdescu, and Cosmin Stoica	457
Hierarchical Program Representation for Program Element Matching Fernando Berzal, Juan-Carlos Cubero, and Aída Jiménez	467
A Combination-of-Tools Method for Learning Interpretable Fuzzy Rule-Based Classifiers from Support Vector Machines	477
An Effective Content-Based Image Retrieval System by Hierachical Segmentation	487
Knowledge Extraction from Unstructured Surface Meshes	497

Table of Contents	XV
Clustering with Reinforcement Learning	507
Mining Frequent Itemsets in Large Data Warehouses: A Novel Approach Proposed for Sparse Data Sets	517
A Sparse Bayesian Position Weighted Bio-Kernel Network	527
Square Penalty Support Vector Regression	537
Constructing Accurate Fuzzy Rule-Based Classification Systems Using Apriori Principles and Rule-Weighting	547
Visualization of Topology Representing Networks	557
The Outer Impartation Information Content of Rules and Rule Sets Dan Hu and Yuanfu Feng	567
An Engineering Approach to Data Mining Projects	578
Classifying Polyphonic Melodies by Chord Estimation Based on Hidden Markov Model	589
Elastic Non-contiguous Sequence Pattern Detection for Data Stream Monitoring	599
Joint Cutoff Probabilistic Estimation Using Simulation: A Mailing Campaign Application	609
Segmentation and Annotation of Audiovisual Recordings Based on Automated Speech Recognition	620
Mining Disjunctive Sequential Patterns from News Stream	630

A New Dissimilarity Measure Between Trees by Decomposition of Unit-Cost Edit Distance	643
Takanori Yokoyama Optimizing Web Structures Using Web Mining Techniques	653
Jonathan Jeffrey, Peter Karski, Björn Lohrmann, Keivan Kianmehr, and Reda Alhajj	000
A Collaborative Recommender System Based on Asymmetric User Similarity	663
Stop Wasting Time: On Predicting the Success or Failure of Learning for Industrial Applications	673
Parallel Wavelet Transform for Spatio-temporal Outlier Detection in Large Meteorological Data	684
A Tool for Web Usage Mining	695
An Algorithm to Mine General Association Rules from Tabular Data Siyamand Ayubi, Maybin Muyeba, and John Keane	705
Intrusion Detection at Packet Level by Unsupervised Architectures Álvaro Herrero, Emilio Corchado, Paolo Gastaldo, Davide Leoncini, Francesco Picasso, and Rodolfo Zunino	718
Quality of Adaptation of Fusion ViSOM	728
Classification Based on the Trace of Variables over Time	739
Extracting Meaningful Contexts from Mobile Life Log Youngseol Lee and Sung-Bae Cho	750
Topological Tree Clustering of Social Network Search Results	760
Bioinformatics and Neuroinformatics	
A Framework to Analyze Biclustering Results on Microarray Experiments	770
Rodrigo Santamaría, Roberto Therón, and Luis Quintales	110

Methods to Bicluster Validation and Comparison in Microarray Data Rodrigo Santamaría, Luis Quintales, and Roberto Therón	780
Capturing Heuristics and Intelligent Methods for Improving Micro-array Data Classification	790
Classification of Microarrays with kNN: Comparison of Dimensionality Reduction Methods	800
Protein Data Condensation for Effective Quaternary Structure Classification	810
PINCoC: A Co-clustering Based Approach to Analyze Protein-Protein Interaction Networks	821
Discovering α-Patterns from Gene Expression Data	831
Biclusters Evaluation Based on Shifting and Scaling Patterns	840
A Deterministic Model to Infer Gene Networks from Microarray Data	850
Profiling of High-Throughput Mass Spectrometry Data for Ovarian Cancer Detection	860
Adapting Machine Learning Technique for Periodicity Detection in Nucleosomal Locations in Sequences	870
Analysis of Tiling Microarray Data by Learning Vector Quantization and Relevance Learning	880
Discriminating Microbial Species Using Protein Sequence Properties and Machine Learning	890

Automatic Prognostic Determination and Evolution of Cognitive Decline Using Artificial Neural Networks	898
Agents and Distributed Systems	
SCSTallocator: Sized and Call-Site Tracing-Based Shared Memory Allocator for False Sharing Reduction in Page-Based DSM Systems Jongwoo Lee, Youngho Park, and Yongik Yoon	908
A Hybrid Social Model for Simulating the Effects of Policies on Residential Power Consumption	919
On Intelligent Interface Agents for Human Based Computation F. Aznar, M. Sempere, M. Pujol, and R. Rizo	930
Reverse Engineering an Agent-Based Hidden Markov Model for Complex Social Systems	940
Effects of Neighbourhood Structure on Evolution of Cooperation in N-Player Iterated Prisoner's Dilemma	950
Interface Agents' Design for a DRT Transportation System Using PASSI	960
A Multi-agent System Approach to Power System Topology Verification	970
Financial Engineering and Modelling	
A System for Efficient Portfolio Management	980
Partitioning-Clustering Techniques Applied to the Electricity Price Time Series	990
Time-Series Prediction Using Self-Organising Mixture Autoregressive Network	1000

Adjusting the Generalized Pareto Distribution with Evolution Strategies – An application to a Spanish Motor Liability Insurance Database	1010
Independent Factor Reinforcement Learning for Portfolio Management	1020
Discrete Time Portfolio Selection with Lévy Processes	1032
Agent-Based Approach to Service Sciences	
Analyzing the Influence of Overconfident Investors on Financial Markets Through Agent-Based Model	1042
Modularity, Product Innovation, and Consumer Satisfaction: An Agent-Based Approach	1053
An Agent-Based Model of Interactions in the Payment Card Market Biliana Alexandrova-Kabadjova, Andreas Krause, and Edward Tsang	1063
The Possibility of an Epidemic Meme Analogy for Web Community Population Analysis	1073
The Econometric Analysis of Agent-Based Models in Finance: An Application	1081
Short Run Dynamics in an Artificial Futures Market with Human Subjects	1092
Video-Based Conjoint Analysis and Agent Based Simulation for Estimating Customer's Behavior	1102
Effect of the Number of Users and Bias of Users' Preference on Recommender Systems	1112

Exploring Quantitative Evaluation Criteria for Service and Potentials of New Service in Transportation: Analyzing Transport Networks of Railway, Subway, and Waterbus	1122
Neural-evolutionary Fusion Algorithms and Their Applications	
Saw-Tooth Algorithm Guided by the Variance of Best Individual Distributions for Designing Evolutionary Neural Networks	1131
Using a Genetic Algorithm for Editing k-Nearest Neighbor Classifiers \dots $R.$ Gil-Pita and $X.$ Yao	1141
An Evolution of Geometric Structures Algorithm for the Automatic Classification of HRR Radar Targets	1151
Hybrid Cross-Entropy Method/Hopfield Neural Network for Combinatorial Optimization Problems	1160
Author Index	1171