

1. Record Nr.	UNISA996465681503316
Titolo	New Frontiers in Mining Complex Patterns [[electronic resource]] : 4th International Workshop, NFMCP 2015, Held in Conjunction with ECML-PKDD 2015, Porto, Portugal, September 7, 2015, Revised Selected Papers // edited by Michelangelo Ceci, Corrado Loglisci, Giuseppe Manco, Elio Masciari, Zbigniew W. Ras
Pubbl/distr/stampa	Cham : , : Springer International Publishing : , : Imprint : Springer, , 2016
ISBN	3-319-39315-4
Edizione	[1st ed. 2016.]
Descrizione fisica	1 online resource (X, 239 p. 57 illus.)
Collana	Lecture Notes in Artificial Intelligence ; ; 9607
Disciplina	006.3
Soggetti	Data mining Database management Information storage and retrieval Artificial intelligence Data Mining and Knowledge Discovery Database Management Information Storage and Retrieval Artificial Intelligence
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Includes index.
Nota di contenuto	Intro -- New Frontiers in Mining Complex Patterns (NFMCP 2015) -- Organization -- Contents -- Data Stream Mining -- Adaptive Ensembles for Evolving Data Streams -- Combining Block-Based and Online Solutions -- 1 Introduction -- 2 Concept Drift in Data Streams -- 3 Ensembles for Evolving Data Streams -- 4 AUE and OAUE Ensembles -- 4.1 Accuracy Updated Ensemble -- 4.2 Online Accuracy Updated Ensemble -- 5 Final Remarks and Open Issues -- References -- Comparison of Tree-Based Methods for Multi-target Regression on Data Streams -- 1 Introduction -- 2 Background and Related Work -- 2.1 Multi-target Regression -- 2.2 Data Streams -- 2.3 Multi-target Regression on Data Streams -- 3 Tree-Based Approaches for Multi-target Regression on Data Streams -- 3.1 A Local Approach to MTR --

3.2 A Global Approach to MTR -- 3.3 Ensemble of Trees for MTR on Data Streams -- 3.4 Baseline Method -- 4 Experimental Setup -- 4.1 Experimental Questions -- 4.2 Evaluation Measures and Experimental Methodology -- 4.3 Datasets -- 4.4 Compared Methods -- 5 Results -- 5.1 Predictive Performance (RMAE) -- 5.2 Time Consumption -- 5.3 Memory Consumption -- 6 Conclusions and Further Work -- References -- Frequent Itemsets Mining in Data Streams Using Reconfigurable Hardware -- 1 Introduction -- 2 Theoretical Basis -- 2.1 Reconfigurable Computing -- 3 Related Works -- 4 A Method for Frequent Itemsets Mining Using Reconfigurable Hardware -- 4.1 Frequent 1-Itemsets Detection -- 4.2 Proposed Method -- 5 Results -- 5.1 Discussion -- 6 Conclusions -- References -- Discovering and Tracking Organizational Structures in Event Logs -- 1 Introduction -- 2 Basics -- 3 Time-Evolving Organization Structure Tracker -- 3.1 Resource Community Detection -- 3.2 Tracking Evolutions of Resource Communities -- 4 Case Studies -- 5 Conclusions -- References. Intelligent Adaptive Ensembles for Data Stream Mining: A High Return on Investment Approach -- 1 Introduction -- 2 Background -- 2.1 Size of Ensemble Object in Memory -- 2.2 Ensemble Size versus Utility -- 3 Adaptive Ensemble Size Algorithm -- 4 Experimentation -- 4.1 Results -- 5 Conclusion -- References -- Mining Periodic Changes in Complex Dynamic Data Through Relational Pattern Discovery -- 1 Introduction -- 2 Basics and Definitions -- 3 The Method -- 3.1 Relational Frequent Pattern Discovery -- 3.2 Emerging Pattern Extraction -- 3.3 Periodic Change Detection -- 4 Experiments -- 5 Related Works -- 6 Conclusions -- References -- Classification -- The Usefulness of Roughly Balanced Bagging for Complex and High-Dimensional Imbalanced Data -- 1 Introduction -- 2 Related Works -- 3 Studying the Role of Components in Roughly Balanced Bagging -- 3.1 Choosing Algorithms to Learn Component Classifiers -- 3.2 The Influence of the Number of Component Classifiers -- 3.3 Diversity of Component Classifiers -- 4 Influence of the Type of Examples -- 5 Applying a Random Selection of Attributes -- 6 Discussion and Final Remarks -- References -- Classifying Traces of Event Logs on the Basis of Security Risks -- 1 Introduction -- 2 Preliminaries -- 3 The Classification Problem and Our Approach for Solving It -- 3.1 The Challenges of Evaluating a Classification and Our Solution -- 4 The Monte Carlo Classification Algorithm -- 5 Experimental Validation -- 6 Related Work -- 7 Conclusions and Future Work -- References -- Redescription Mining with Multi-target Predictive Clustering Trees -- 1 Introduction -- 2 Notation and Definitions -- 3 The CLUS-RM Algorithm -- 3.1 The Procedure for Creating Redescriptions -- 3.2 Rule Size Minimization -- 3.3 Algorithm Time Complexity -- 4 Mining Redescriptions on Data Describing Countries -- 5 Algorithm Evaluation and Comparison. 6 Conclusion -- A Appendix -- References -- Mining Complex Data -- Generalizing Patterns for Cross-Domain Analogy -- 1 Introduction -- 2 Preliminaries -- 3 Analogy and Inference -- 3.1 Representation Formalism -- 3.2 Analogy -- 3.3 Inference and Re-representation -- 4 Analogical Pattern Generalization -- 4.1 Formal Definition -- 4.2 Evaluation -- 4.3 Addition and Union -- 4.4 Military and Medical Strategy -- 4.5 Patterns Assessment -- 5 Conclusions -- References -- Spectral Features for Audio Based Vehicle Identification -- 1 Introduction -- 1.1 Related Work -- 1.2 Vehicle Classes -- 2 Data Collection and Description -- 3 Feature Set -- 3.1 Feature Selection -- 4 Experiments -- 4.1 Classifiers -- 4.2 Data -- 4.3 Classification Results -- 4.4 Hierarchical Classification -- 5 Summary and Conclusions -- References -- Probabilistic Frequent Subtree Kernels -- 1 Introduction -- 2 The Probabilistic Frequent Subtree Kernel -- 2.1

Probabilistic Frequent Subtrees -- 2.2 Implementation Issues and Runtime Analysis -- 3 Experiments -- 3.1 Datasets -- 3.2 Runtime -- 3.3 Recall -- 3.4 Stability of Probabilistic Subtree Patterns -- 3.5 Predictive Performance -- 4 Conclusion and Future Work -- References -- Heterogeneous Network Decomposition and Weighting with Text Mining Heuristics -- 1 Introduction -- 2 Related Work -- 3 Methodology -- 3.1 Network Decomposition -- 3.2 Classification -- 4 Methodology Improvement -- 4.1 Imbalanced Data Sets and Label Propagation -- 4.2 Text Mining Inspired Weights Calculation -- 5 Experimental Setting and Results -- 5.1 Data Set Description -- 5.2 Experiment Description -- 5.3 Experimental Results -- 6 Conclusions and Further Work -- References -- Sequences -- Semi-supervised Multivariate Sequential Pattern Mining -- 1 Introduction -- 2 Related Work -- 3 The Semi-supervised Learning Framework -- 3.1 Graph Construction. 3.2 Label Propagation -- 3.3 Extension to Out-of-Samples -- 4 Experimental Analysis -- 5 Conclusion -- References -- Evaluating a Simple String Representation for Intra-day Foreign Exchange Prediction -- 1 Introduction -- 2 Previous Work -- 3 Simple Strategy -- 4 String Subsequences Strategy -- 4.1 n-Grams -- 4.2 Time Decay n-Grams -- 5 Experiments -- 5.1 Simple String Strategy: Word vs. Alphabet Length -- 5.2 Parzen Window Strategy: Regression vs. Classification -- 5.3 Simple Strings Strategy vs. SVM Classification -- 5.4 Time Decay n-Grams: Decay Analysis -- 5.5 Simple String Strategy vs. Time Decay n-Grams -- 6 Conclusions -- References -- Author Index.

Sommario/riassunto

This book constitutes the thoroughly refereed post-conference proceedings of the 4th International Workshop on New Frontiers in Mining Complex Patterns, NFMCP 2015, held in conjunction with ECML-PKDD 2015 in Porto, Portugal, in September 2015. The 15 revised full papers presented together with one invited talk were carefully reviewed and selected from 19 submissions. They illustrate advanced data mining techniques which preserve the informative richness of complex data and allow for efficient and effective identification of complex information units present in such data. The papers are organized in the following sections: data stream mining, classification, mining complex data, and sequences.
