1. Record Nr. UNISA996466246103316

Titolo	Discovery Science [[electronic resource]] : 20th International Conference, DS 2017, Kyoto, Japan, October 15–17, 2017, Proceedings / / edited by Akihiro Yamamoto, Takuya Kida, Takeaki Uno, Tetsuji Kuboyama
Pubbl/distr/stampa	Cham : , : Springer International Publishing : , : Imprint : Springer, , 2017
ISBN	3-319-67786-1
Edizione	[1st ed. 2017.]
Descrizione fisica	1 online resource (XV, 357 p. 90 illus.)
Collana	Lecture Notes in Artificial Intelligence ; ; 10558
Disciplina	501
Soggetti	Artificial intelligence Data mining Application software Artificial Intelligence Data Mining and Knowledge Discovery Information Systems Applications (incl. Internet) Computer Appl. in Social and Behavioral Sciences
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Preface -- Organization -- Abstracts of Invited Talks -- Machine Learning from Weak Supervision - Towards Accurate Classification with Low Labeling Costs -- Automatic Design of Functional Molecules and Materials -- Contents -- Online Learning -- Context-Based Abrupt Change Detection and Adaptation for Categorical Data Streams -- 1 Introduction -- 2 Background -- 2.1 DILCA Context-Based Similarity Measure -- 2.2 CDCStream Categorical Drift Detector -- 3 FG-CDCStream Algorithm -- 4 Experimentation -- 4.1 Results -- 4.2 Discussion -- 5 Conclusion -- References -- A New Adaptive Learning Algorithm and Its Application to Online Malware Detection -- 1 Introduction -- 2 Related Work -- 2.1 Batch Learning in Malware Detection -- 2.2 Online Learning in Malware Detection -- 3 Problem Statement -- 4 Methodology -- 4.1 Batch Learning -- LR -- 4.2 Online Learning -- FTRL-DP -- 5 Data Collection -- 5.1 Malware

Collection -- 5.2 Malware Execution -- 5.3 Feature Extraction -- 6 Evaluation -- 6.1 Experiment with LR -- 6.2 Experiment with FTRL Algorithms -- 7 Discussion -- 7.1 Prediction Accuracy -- 7.2 Running Time -- 8 Conclusions and Future Work -- A Proof of Theorem1 -- References -- Real-Time Validation of Retail Gasoline Prices -- 1 Introduction -- 2 Related Research -- 2.1 Factors Affecting Gasoline Prices -- 2.2 The Most Common Action Model -- 3 The PCR Real-Time Prediction Model -- 3.1 Price Change Rules -- 3.2 Prediction Using Price Change Rules -- 3.3 Description of the Method -- 3.4 Evaluation of the PCR Method -- 4 Experimental Evaluation -- 5 Conclusion -- References -- Regression -- General Meta-Model Framework for Surrogate-Based Numerical Optimization -- 1 Introduction -- 2 Background and Related Work -- 2.1 Numerical Optimization -- 2.2 Estimating Parameters of Ordinary Differential Equations -- 2.3 Surrogate-Based Numerical Optimization.
3 Meta Model for Surrogate-Based Optimization -- 3.1 Meta-Model Framework -- 3.2 Uninformed Meta Model -- 3.3 Relevator Meta Model -- 4 Empirical Evaluation of the Meta-Model Variants -- 4.1 Parameter Estimation Problems -- 4.2 Experimental Setup -- 4.3 Results -- 5 Conclusion -- References -- Evaluation of Different Heuristics for Accommodating Asymmetric Loss Functions in Regression -- 1 Introduction -- 2 Problem Formulation -- 3 Related Work -- 4 Empirical Approaches Based on Prediction Shift -- 4.1 Constant Shift -- 4.2 Pointwise Shift -- 4.3 Learned Model-Based Shift -- 4.4 Assumed Error Model-Based Shift -- 4.5 k-Nearest Neighbors Based Methods -- 5 Experimental Comparison of Methods -- 5.1 Regression Methods Used -- 5.2 Loss Functions Used -- 5.3 Results -- 6 Conclusion -- References -- Differentially Private Empirical Risk Minimization with Input Perturbation -- 1 Introduction -- 2 Problem Definition and Preliminary -- 3 Input Perturbation -- 3.1 Loss Function for Input Perturbation -- 3.2 Input Perturbation Method -- 3.3 Privacy of Input Perturbation -- 3.4 Utility Analysis -- 4 Conclusion -- References -- Label Classification -- On a New Competence Measure Applied to the Dynamic Selection of Classifiers Ensemble -- 1 Introduction -- 2 Multiclassifier System -- 2.1 Preliminaries -- 2.2 Measure of Competence -- 2.3 DES Systems -- 3 Experiments -- 3.1 Experimental Setup -- 3.2 Results and Discussion -- 4 Conclusion -- References -- Multi-label Classification Using Random Label Subset Selections -- 1 Introduction -- 2 MLC Using Random Label Subset Selections -- 3 Experimental Design -- 4 Results -- 5 Conclusions and Future Work -- References -- Option Predictive Clustering Trees for Hierarchical Multi-label Classification -- 1 Introduction -- 2 Option Predictive Clustering Trees -- 3 Experimental Design -- 4 Results and Discussion.
5 Conclusions -- References -- Deep Learning -- Re-training Deep Neural Networks to Facilitate Boolean Concept Extraction -- 1 Introduction -- 2 Knowledge Distillation from Neural Networks -- 2.1 Rule Extraction -- 2.2 Connection Pruning -- 3 The DEEPRED Algorithm -- 3.1 Overview -- 3.2 Extraction of DNF Formulas from Trees -- 3.3 Simplification and Post-pruning of Expressions -- 4 Retraining DNNs to Extract Better Representations -- 4.1 Weight Sparseness Pruning -- 4.2 Activation Polarization -- 5 Experiments -- 5.1 Experimental Setup -- 5.2 Characteristics of the Trained Networks -- 5.3 Reconstruction Using the Entire Dataset -- 5.4 Reconstruction Using Part of the Dataset -- 6 Conclusion -- References -- An In-Depth Experimental Comparison of RNTNs and CNNs for Sentence Modeling -- 1 Introduction -- 2 Method -- 3 Experiments -- 3.1 Experimental Settings -- 3.2 Task 1: Sentiment Analysis -- 3.3 Task 2: Sentence Categorization -- 3.4 Comparison of CNN Architectures -- 4

Conclusions -- References -- Feature Selection -- Improving Classification Accuracy by Means of the Sliding Window Method in Consistency-Based Feature Selection -- 1 Introduction -- 2 Feature Selection Methods -- 2.1 Feature Ranking Methods -- 2.2 Pairwise Evaluation Methods -- 2.3 Consistency-Based Algorithms -- 3 Our Proposal -- 3.1 Defieciencies of steepest-descent Search -- 4 Experiments -- 4.1 Numbers of Features Selected and Auc-Roc Scores -- 4.2 Efficiency -- 5 Conclusion and Future Works -- References -- Feature Ranking for Multi-target Regression with Tree Ensemble Methods -- 1 Introduction -- 2 Predictive Clustering Trees for Multi-target Regression -- 3 Feature Ranking via Ensembles of PCTs -- 3.1 Ensembles of PCTs -- 3.2 Ensemble Feature Ranking Methods -- 4 Experimental Design -- 4.1 Experimental Questions -- 4.2 Data Description -- 4.3 Evaluation Methodology.
4.4 Statistical Analysis of the Results -- 4.5 Parameter Instantiation -- 5 Results and Discussion -- 5.1 Are the Obtained Feature Rankings Relevant? -- 5.2 Comparison of the Different Ranking Methods -- 5.3 Comparison of the Different Ensemble Methods -- 5.4 Selecting the Best Ensemble-Ranking Pair -- 6 Conclusions -- References -- Recommendation System -- Recommending Collaborative Filtering Algorithms Using Subsampling Landmarkers -- 1 Introduction -- 2 Related Work -- 2.1 Collaborative Filtering -- 2.2 Metalearning -- 2.3 Algorithm Selection for CF -- 3 Subsampling Landmarkers for Collaborative Filtering -- 3.1 Subsampling Landmarkers -- 3.2 Experimental Procedure -- 4 Results and Discussion -- 4.1 Metalevel Evaluation -- 4.2 Baselevel Performance Analysis -- 4.3 Metaknowledge -- 5 Conclusions and Future Work -- References -- Community Detection -- Recursive Extraction of Modular Structure from Layered Neural Networks Using Variational Bayes Method -- 1 Introduction -- 2 Layered Neural Networks -- 3 Recursive Extraction of Modular Structure from Layered Neural Networks Using Variational Bayes Method -- 3.1 Community Detection Using Variational Bayes Method -- 3.2 Modular Representation of Layered Neural Networks -- 3.3 Recursive Extraction of Modular Structure and Pruning -- 4 Experiment -- 4.1 Discovery of a Hidden Structure in a Layered Neural Network -- 4.2 Analysis of STUDENT ALCOHOL CONSUMPTION Data Set -- 5 Discussion -- 6 Conclusion -- References -- Discovering Hidden Knowledge in Carbon Emissions Data: A Multilayer Network Approach -- 1 Introduction -- 2 Data Collection and Preparation -- 3 Proposed Multilayer Network -- 3.1 Building a Multilayer Network -- 3.2 Community Detection on Multilayer Networks -- 4 Results and Discussion -- 4.1 Knowledge Discovery in Carbon Emissions Networks -- 4.2 Community Detection on CETA-MLN.
4.3 Impact of Varying Coupling and Resolution Parameters -- 5 Conclusion and Future Work -- References -- Topic Extraction on Twitter Considering Author's Role Based on Bipartite Networks -- 1 Introduction -- 2 Related Work -- 3 Community Detection Technique from Bipartite Networks -- 4 Proposed Method -- 4.1 Step 1: Creation of Author-Word Count Matrices -- 4.2 Step 2: Clustering and Step 3: Feature Selection -- 4.3 Step 4: Topic Detection -- 5 Experimental Result -- 5.1 Target Data -- 5.2 Step 1: Creation of Author-Word Count Matrices -- 5.3 Step 2: Clustering and Step 3: Feature Selection -- 5.4 Step 4: Topic Detection -- 6 Conclusion -- References -- Pattern Mining -- Mining Strongly Closed Itemsets from Data Streams -- 1 Introduction -- 2 Related Work -- 3 The Problem Setting -- 4 The Mining Algorithm -- 4.1 Sampling -- 4.2 Incremental Update -- 5 Empirical Evaluation -- 5.1 Speedup -- 5.2 Approximation Quality -- 6 Concluding Remarks -- References -- Extracting Mutually Dependent

Multisets -- 1 Introduction -- 2 Mutually Dependent Multisets -- 3 Tail Occurrence of Multisets -- 4 Algorithm to Extract Mutually Dependent Multisets -- 5 Experimental Results -- 5.1 Antibiogram -- 5.2 Artificial Data -- 6 Conclusion and Future Works -- References -- Bioinformatics -- LOCANDA: Exploiting Causality in the Reconstruction of Gene Regulatory Networks -- 1 Introduction -- 2 Related Work -- 3 The Method LOCANDA -- 4 Experiments -- 5 Conclusions and Future Work -- References -- Discovery of Salivary Gland Tumors' Biomarkers via Co-Regularized Sparse-Group Lasso -- 1 Background and Motivation -- 2 Previous Work -- 3 Methodology -- 3.1 Co-Regularized Sparse-Group Lasso -- 3.2 Stability Selection and Randomization Test -- 4 Experiments -- 5 Results and Discussion -- 5.1 Model Performance Evaluation -- 5.2 Biological Interpretation -- 6 Conclusions -- References.
Knowledge Discovery.

Sommario/riassunto

This book constitutes the proceedings of the 20th International Conference on Discovery Science, DS 2017, held in Kyoto, Japan, in October 2017, co-located with the International Conference on Algorithmic Learning Theory, ALT 2017. The 18 revised full papers presented together with 6 short papers and 2 invited talks in this volume were carefully reviewed and selected from 42 submissions. The scope of the conference includes the development and analysis of methods for discovering scientific knowledge, coming from machine learning, data mining, intelligent data analysis, big data analysis as well as their application in various scientific domains. The papers are organized in topical sections on machine learning: online learning, regression, label classification, deep learning, feature selection, recommendation system; and knowledge discovery: recommendation system, community detection, pattern mining, misc.