

1. Record Nr.	UNINA9910298991203321
Autore	Qin Zengchang
Titolo	Uncertainty Modeling for Data Mining : A Label Semantics Approach // by Zengchang Qin, Yongchuan Tang
Pubbl/distr/stampa	Berlin, Heidelberg : , : Springer Berlin Heidelberg : , : Imprint : Springer, , 2014
ISBN	3-642-41251-3
Edizione	[1st ed. 2014.]
Descrizione fisica	1 online resource (303 p.)
Collana	Advanced Topics in Science and Technology in China, , 1995-6827
Disciplina	004 004.0151 005.7 006.3 006.312
Soggetti	Data mining Artificial intelligence Computer networks Computer science - Mathematics Data Mining and Knowledge Discovery Artificial Intelligence Computer Communication Networks Mathematical Applications in Computer Science
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Description based upon print version of record.
Nota di bibliografia	Includes bibliographical references.
Nota di contenuto	Cover; Title Page; Copyright Page; Dedication Page; Preface; Acknowledgements; Table of Contents; Acronyms; Notations; 1 Introduction; 1.1 Types of Uncertainty; 1.2 Uncertainty Modeling and Data Mining; 1.3 Related Works; References; 2 Induction and Learning; 2.1 Introduction; 2.2 Machine Learning; 2.2.1 Searching in Hypothesis Space; 2.2.2 Supervised Learning; 2.2.3 Unsupervised Learning; 2.2.4 Instance-Based Learning; 2.3 Data Mining and Algorithms; 2.3.1 Why Do We Need Data Mining?; 2.3.2 How Do We do Data Mining?; 2.3.3 Artificial Neural Networks; 2.3.4 Support Vector Machines 2.4 Measurement of Classifiers 2.4.1 ROC Analysis for Classification;

2.4.2 Area Under the ROC Curve; 2.5 Summary; References; 3 Label Semantics Theory; 3.1 Uncertainty Modeling with Labels; 3.1.1 Fuzzy Logic; 3.1.2 Computing with Words; 3.1.3 Mass Assignment Theory; 3.2 Label Semantics; 3.2.1 Epistemic View of Label Semantics; 3.2.2 Random Set Framework; 3.2.3 Appropriateness Degrees; 3.2.4 Assumptions for Data Analysis; 3.2.5 Linguistic Translation; 3.3 Fuzzy Discretization; 3.3.1 Percentile-Based Discretization; 3.3.2 Entropy-Based Discretization; 3.4 Reasoning with Fuzzy Labels 3.4.1 Conditional Distribution Given Mass Assignments 3.4.2 Logical Expressions of Fuzzy Labels; 3.4.3 Linguistic Interpretation of Appropriate Labels; 3.4.4 Evidence Theory and Mass Assignment; 3.5 Label Relations; 3.6 Summary; References; 4 Linguistic Decision Trees for Classification; 4.1 Introduction; 4.2 Tree Induction; 4.2.1 Entropy; 4.2.2 Soft Decision Trees; 4.3 Linguistic Decision for Classification; 4.3.1 Branch Probability; 4.3.2 Classification by LDT; 4.3.3 Linguistic ID3 Algorithm; 4.4 Experimental Studies; 4.4.1 Influence of the Threshold; 4.4.2 Overlapping Between Fuzzy Labels 4.5 Comparison Studies 4.6 Merging of Branches; 4.6.1 Forward Merging Algorithm; 4.6.2 Dual-Branch LDTs; 4.6.3 Experimental Studies for Forward Merging; 4.6.4 ROC Analysis for Forward Merging; 4.7 Linguistic Reasoning; 4.7.1 Linguistic Interpretation of an LDT; 4.7.2 Linguistic Constraints; 4.7.3 Classification of Fuzzy Data; 4.8 Summary; References; 5 Linguistic Decision Trees for Prediction; 5.1 Prediction Trees; 5.2 Linguistic Prediction Trees; 5.2.1 Branch Evaluation; 5.2.2 Defuzzification; 5.2.3 Linguistic ID3 Algorithm for Prediction; 5.2.4 Forward Branch Merging for Prediction 5.3 Experimental Studies 5.3.1 3D Surface Regression; 5.3.2 Abalone and Boston Housing Problem; 5.3.3 Prediction of Sunspots; 5.3.4 Flood Forecasting; 5.4 Query Evaluation; 5.4.1 Single Queries; 5.4.2 Compound Queries; 5.5 ROC Analysis for Prediction; 5.5.1 Predictors and Probabilistic Classifiers; 5.5.2 AUC Value for Prediction; 5.6 Summary; References; 6 Bayesian Methods Based on Label Semantics; 6.1 Introduction; 6.2 Naive Bayes; 6.2.1 Bayes Theorem; 6.2.2 Fuzzy Naive Bayes; 6.3 Fuzzy Semi-Naive Bayes; 6.4 Online Fuzzy Bayesian Prediction; 6.4.1 Bayesian Methods; 6.4.2 Online Learning 6.5 Bayesian Estimation Trees

Sommario/riassunto

Machine learning and data mining are inseparably connected with uncertainty. The observable data for learning is usually imprecise, incomplete or noisy. Uncertainty Modeling for Data Mining: A Label Semantics Approach introduces 'label semantics', a fuzzy-logic-based theory for modeling uncertainty. Several new data mining algorithms based on label semantics are proposed and tested on real-world datasets. A prototype interpretation of label semantics and new prototype-based data mining algorithms are also discussed. This book offers a valuable resource for postgraduates, researchers and other professionals in the fields of data mining, fuzzy computing and uncertainty reasoning. Zengchang Qin is an associate professor at the School of Automation Science and Electrical Engineering, Beihang University, China; Yongchuan Tang is an associate professor at the College of Computer Science, Zhejiang University, China.