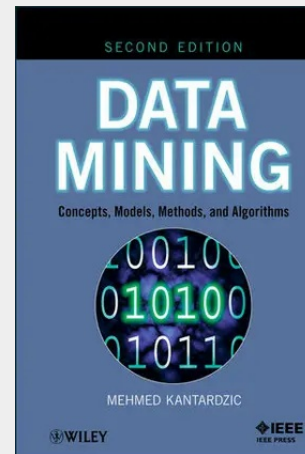


Data Mining

This book reviews state-of-the-art methodologies and techniques for analyzing enormous quantities of raw data in high-dimensional data spaces, to extract new information for decision making. The goal of this book is to provide a single introductory source, organized in a systematic way, in which we could direct the readers in analysis of large data sets, through the explanation of basic concepts, models and methodologies developed in recent decades. To obtain instructor's materials, please send an email to: pressbooks@ieee.org

Now updated--the systematic introductory guide to modern analysis of large data sets As data sets continue to grow in size and complexity, there has been an inevitable move towards indirect, automatic, and intelligent data analysis in which the analyst works via more complex and sophisticated software tools. This book reviews state-of-the-art methodologies and techniques for analyzing enormous quantities of raw data in high-dimensional data spaces to extract new information for decision-making. This Second Edition of Data Mining: Concepts, Models, Methods, and Algorithms discusses data mining principles and then describes representative state-of-the-art methods and algorithms originating from different disciplines such as statistics, machine learning, neural networks, fuzzy logic, and evolutionary computation. Detailed algorithms are provided with necessary explanations and illustrative examples, and questions and exercises for practice at the end of each chapter. This new edition features the following new techniques/methodologies: * Support Vector Machines (SVM)--developed based on statistical learning theory, they have a large potential for applications in predictive data mining * Kohonen Maps (Self-Organizing Maps - SOM)--one of very applicative neural-networks-based methodologies for descriptive data mining and multi-dimensional data visualizations * DBSCAN, BIRCH, and distributed DBSCAN clustering algorithms--representatives of an important class of density-based clustering methodologies * Bayesian Networks (BN) methodology often used for causality modeling * Algorithms for measuring Betweenness and Centrality parameters in graphs, important for applications in mining large social networks * CART algorithm and Gini index in building decision trees * Bagging & Boosting approaches to ensemble-learning methodologies, with details of AdaBoost algorithm * Relief algorithm, one of the core feature selection algorithms inspired by instance-based learning * PageRank algorithm for mining and authority ranking of web pages * Latent Semantic Analysis (LSA) for text mining and measuring semantic similarities between text-based documents * New sections on temporal, spatial, web, text, parallel, and distributed data mining * More emphasis on business, privacy, security, and legal aspects of data mining technology This text offers guidance on how and when to use a particular software tool (with the companion data sets) from among the hundreds offered when faced with a data set to mine. This allows analysts to create and perform their own data mining experiments using their knowledge of the methodologies and techniques provided. The book emphasizes the selection of appropriate methodologies and data analysis software, as well as parameter tuning. These critically important, qualitative decisions can only be made with the deeper understanding of parameter meaning and its role in the technique that is offered here. This volume is primarily intended as a data-mining textbook for computer science, computer engineering, and computer information systems majors at the graduate level. Senior students at the undergraduate level and with the appropriate background can also successfully comprehend all topics presented here.



119,50 €

111,68 € (zzgl. MwSt.)

Nicht mehr lieferbar

Artikelnummer: 9780470890455

Medium: Buch

ISBN: 978-0-470-89045-5

Verlag: Wiley-Blackwell

Erscheinungstermin: 23.09.2011

Sprache(n): Englisch

Auflage: 2. Auflage 2011

Produktform: Gebunden

Gewicht: 911 g

Seiten: 552

Format (B x H): 167 x 245 mm

