1. Record Nr. UNINA9910972287003321

Autore	Nisbet Robert
Titolo	Handbook of statistical analysis and data mining applications / / Robert Nisbet, John Elder, Gary Miner
Pubbl/distr/stampa	Amsterdam ; ; Boston, : Academic Press/Elsevier, c2009
ISBN	9786612168314 9781282168312 1282168312 9780080912035 0080912036
Edizione	[First edition.]
Descrizione fisica	1 online resource (859 pages)
Classificazione	31.73
Altri autori (Persone)	ElderJohn F (John Fletcher) MinerGary
Disciplina	006.3/12 22 519.5 006.312015195
Soggetti	Data mining - Statistical methods Multivariate analysis
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 and index.
Nota di contenuto	Front Cover; Handbook of Statistical Analysis and Data Mining Applications; Copyright Page; Table of Contents; Foreword 1; Foreword 2; Preface; Introduction; List of Tutorials by Guest Authors; Part 1: History of Phases of Data Analysis, Basic Theory, and the Data Mining Process; Chapter 1: The Background for Data Mining Practice; Assumptions of the Parametric Model; Two Views of Reality; Aristotle; Plato; The Rise of Modern Statistical Analysis: The Second Generation; Machine Learning Methods: The Third Generation; Statistical Learning Theory: The Fourth Generation Chapter 2: Theoretical Considerations for Data MiningMajor Issues in Data Mining; General Requirements for Success in a Data Mining Project; The Importance of Domain Knowledge; Postscript; Some Caveats with Data Mining Solutions; Chapter 3: The Data Mining Process; CRISP-DM; Assess the Business Environment for Data Mining; Data Understanding (Mostly Science); References; Preamble; Chapter 4:

Data Understanding and Preparation; Preamble; Issues That Should be Resolved; Splitting Data
Part 1: Using a Wrapper Approach in Weka to Determine the Most Appropriate Variables for Your Neural Network ModelExample 4; Data Extraction; Data Weighting and Balancing; Data Filtering and Smoothing; Data Abstraction; Data Reduction; Data Sampling; Data Discretization; Data Derivation; Postscript; Chapter 5: Feature Selection; Inductive Database Approach; Bi-variate Methods; Multivariate Methods; Postscript; Complex Methods; The Other Two Ways of Using Feature Selection in STATISTICA: Interactive Workspace; Preamble; Chapter 6: Accessory Tools for Doing Data Mining; Preamble; Introduction
Basic Descriptive StatisticsCombining Groups (Classes) for Predictive Data Mining; Generalized Linear Models (GLMs); Data Miner Workspace Templates; Comparison of Models with and Without Time-Based Features; Example: The IDP Facility of STATISTICA Data Miner; Ensembles in General; Part 2: The Algorithms in Data Mining and Text Mining, the Organization of the Three most common Data Mining Tools, and Selected Speci...; Chapter 7: Basic Algorithms for Data Mining: A Brief Overview; Preamble; STATISTICA Data Miner Recipe (DMRecipe); Automated Neural Nets; Generalized Additive Models (GAMs)
Outputs of GAMsRecursive Partitioning; Pruning Trees; Bibliography; Chapter 8: Advanced Algorithms for Data Mining; The Physical Data Mart; Summary; Micro-Target the Profitable Customers; Quality Control Data Mining and Root Cause Analysis; Chapter 9: Text Mining and Natural Language Processing; The Development of Text Mining; Chapter 10: The Three Most Common Data Mining Software Tools; Preamble; SPSS Clementine Overview; Preamble; Setting the Default Directory; Visual Data Preparation for Data Mining: Taking Photos, Moving Pictures, and Objects into Spreadsheets Representing the Photos...
Preamble

Sommario/riassunto

The Handbook of Statistical Analysis and Data Mining Applications is a comprehensive professional reference book that guides business analysts, scientists, engineers and researchers (both academic and industrial) through all stages of data analysis, model building and implementation. The Handbook helps one discern the technical and business problem, understand the strengths and weaknesses of modern data mining algorithms, and employ the right statistical methods for practical application. Use this book to address massive and complex datasets with novel statistical approaches and be a