1.	Record Nr.	UNINA9910460169903321
	Autore	Larose Daniel T.
	Titolo	Data mining and predictive analytics / / Daniel T. Larose, Chantal D. Larose
	Pubbl/distr/stampa	Hoboken, New Jersey : , : John Wiley & Sons, , 2015 ©2015
	ISBN	1-118-86870-6 1-118-86867-6
	Edizione	[Second edition.]
	Descrizione fisica	1 online resource (827 p.)
	Collana	Wiley Series on Methods and Applications in Data Mining
	Disciplina	006.3/12
	Soggetti	Data mining Prediction theory Electronic books.
	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	Cover; Contents; Preface; Acknowledgments; Part I Data Preparation; Chapter 1 An Introduction to Data Mining and Predictive Analytics; 1.1 What is Data Mining? What is Predictive Analytics?; 1.2 Wanted: Data Miners; 1.3 The Need for Human Direction of Data Mining; 1.4 The Cross-Industry Standard Process for Data Mining: CRISP-DM; 1.4.1 CRISP-DM: The Six Phases; 1.5 Fallacies of Data Mining; 1.6 What Tasks Can Data Mining Accomplish; 1.6.1 Description; 1.6.2 Estimation; 1.6.3 Prediction; 1.6.4 Classification; 1.6.5 Clustering; 1.6.6 Association; The R Zone; R References; Exercises Chapter 2 Data Preprocessing2.1 Why do We Need to Preprocess the Data?; 2.2 Data Cleaning; 2.3 Handling Missing Data; 2.4 Identifying Misclassifications; 2.5 Graphical Methods for Identifying Outliers; 2.6 Measures of Center and Spread; 2.7 Data Transformation; 2.8 Min-Max Normalization; 2.9 Z-Score Standardization; 2.10 Decimal Scaling; 2.11 Transformations to Achieve Normality; 2.12 Numerical Methods for Identifying Outliers; 2.13 Flag Variables; 2.14 Transforming Categorical Variables into Numerical Variables; 2.15 Binning Numerical Variables; 2.16 Reclassifying Categorical Variables 2.17 Adding an Index Field2.18 Removing Variables that are not Useful;

Sommario/riassunto	Learn methods of data analysis and their application to real-world data sets This updated second edition serves as an introduction to data mining methods and models, including association rules, clustering, neural networks, logistic regression, and multivariate analysis. The authors apply a unified "white box" approach to data mining methods and models. This approach is designed to walk readers through the operations and nuances of the various methods, using small data sets, so readers can gain an insight into the inner workings of the method under review. Chapters provide readers with hands