

1. Record Nr.	UNINA9910254940303321
Autore	Suthaharan Shan
Titolo	Machine Learning Models and Algorithms for Big Data Classification : Thinking with Examples for Effective Learning // by Shan Suthaharan
Pubbl/distr/stampa	New York, NY : , : Springer US : , : Imprint : Springer, , 2016
ISBN	1-4899-7641-8
Edizione	[1st ed. 2016.]
Descrizione fisica	1 online resource
Collana	Integrated Series in Information Systems, , 1571-0270 ; ; 36
Disciplina	658.4033
Soggetti	Management Database management Artificial intelligence Database Management Artificial Intelligence
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	Science of Information -- Part I Understanding Big Data -- Big Data Essentials -- Big Data Analytics -- Part II Understanding Big Data Systems -- Distributed File System -- MapReduce Programming Platform -- Part III Understanding Machine Learning -- Modeling and Algorithms -- Supervised Learning Models -- Supervised Learning Algorithms -- Support Vector Machine -- Decision Tree Learning -- Part IV Understanding Scaling-Up Machine Learning -- Random Forest Learning -- Deep Learning Models -- Chandelier Decision Tree -- Dimensionality Reduction.
Sommario/riassunto	This book presents machine learning models and algorithms to address big data classification problems. Existing machine learning techniques like the decision tree (a hierarchical approach), random forest (an ensemble hierarchical approach), and deep learning (a layered approach) are highly suitable for the system that can handle such problems. This book helps readers, especially students and newcomers to the field of big data and machine learning, to gain a quick understanding of the techniques and technologies; therefore, the theory, examples, and programs (Matlab and R) presented in this book have been simplified, hardcoded, repeated, or spaced for

improvements. They provide vehicles to test and understand the complicated concepts of various topics in the field. It is expected that the readers adopt these programs to experiment with the examples, and then modify or write their own programs toward advancing their knowledge for solving more complex and challenging problems. The presentation format of this book focuses on simplicity, readability, and dependability so that both undergraduate and graduate students as well as new researchers, developers, and practitioners in this field can easily trust and grasp the concepts, and learn them effectively. It has been written to reduce the mathematical complexity and help the vast majority of readers to understand the topics and get interested in the field. This book consists of four parts, with the total of 14 chapters. The first part mainly focuses on the topics that are needed to help analyze and understand data and big data. The second part covers the topics that can explain the systems required for processing big data. The third part presents the topics required to understand and select machine learning techniques to classify big data. Finally, the fourth part concentrates on the topics that explain the scaling-up machine learning, an important solution for modern big data problems.
