

1. Record Nr.	UNINA9910798428703321
Autore	Kozlov Alex
Titolo	Mastering Scala machine learning : advance your skills in efficient data analysis and data processing using the powerful tools of Scala, Spark, and Hadoop // Alex Kozlov
Pubbl/distr/stampa	Birmingham : , : Packt Publishing, , 2016
ISBN	1-78588-526-X
Edizione	[1st edition]
Descrizione fisica	1 online resource (310 pages) : color illustrations
Collana	Community experience distilled
Disciplina	006.31
Soggetti	Scala (Computer program language) Machine learning Electronic data processing
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Includes index.
Sommario/riassunto	Advance your skills in efficient data analysis and data processing using the powerful tools of Scala, Spark, and Hadoop About This Book This is a primer on functional-programming-style techniques to help you efficiently process and analyze all of your data Get acquainted with the best and newest tools available such as Scala, Spark, Parquet and MLlib for machine learning Learn the best practices to incorporate new Big Data machine learning in your data-driven enterprise to gain future scalability and maintainability Who This Book Is For Mastering Scala Machine Learning is intended for enthusiasts who want to plunge into the new pool of emerging techniques for machine learning. Some familiarity with standard statistical techniques is required. What You Will Learn Sharpen your functional programming skills in Scala using REPL Apply standard and advanced machine learning techniques using Scala Get acquainted with Big Data technologies and grasp why we need a functional approach to Big Data Discover new data structures, algorithms, approaches, and habits that will allow you to work effectively with large amounts of data Understand the principles of supervised and unsupervised learning in machine learning Work with unstructured data and serialize it using Kryo, Protobuf, Avro, and

AvroParquet Construct reliable and robust data pipelines and manage data in a data-driven enterprise Implement scalable model monitoring and alerts with Scala In Detail Since the advent of object-oriented programming, new technologies related to Big Data are constantly popping up on the market. One such technology is Scala, which is considered to be a successor to Java in the area of Big Data by many, like Java was to C/C++ in the area of distributed programming. This book aims to take your knowledge to next level and help you impart that knowledge to build advanced applications such as social media mining, intelligent news portals, and more. After a quick refresher on functional programming concepts using REPL, you will see some practical examples of setting up the development environment and tinkering with data. We will then explore working with Spark and MLlib using k-means and decision trees. Most of the data that we produce today is unstructured and raw, and you will learn to tackle this type of data with advanced topics such as regression, classification, integration, and working with graph algorithms. Finally, you will discover at how to use Scala to perform complex...
