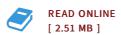# Learning Spark: Lightning-Fast Big Data Analysis (Paperback)

By Holden Karau, Andy Kowinski, Mark Hamstra

O Reilly Media, Inc, USA, United States, 2015. Paperback. Condition: New. Language: English . Brand New Book. Data in all domains is getting bigger. How can you work with it efficiently? This book introduces Apache Spark, the open source cluster computing system that makes data analytics fast to write and fast to run. With Spark, you can tackle big datasets quickly through simple APIs in Python, Java, and Scala. Written by the developers of Spark, this book will have data scientists and engineers up and running in no time. You ll learn how to express parallel jobs with just a few lines of code, and cover applications from simple batch jobs to stream processing and machine learning. Quickly dive into Spark capabilities such as distributed datasets, in-memory caching, and the interactive shell Leverage Spark s powerful built-in libraries, including Spark SQL, Spark Streaming, and MLlib Use one programming paradigm instead of mixing and matching tools like Hive, Hadoop, Mahout, and Storm Learn how to deploy interactive, batch, and streaming applications Connect to data sources including HDFS, Hive, JSON, and S3 Master advanced topics like data partitioning and shared variables.

READ ONLINE
[2.51 MB]

Reviews

I actually started out reading this article ebook. This is for those who statte that there had not been a worth reading. Its been developed in an extremely easy way and it is just after i finished reading this book in which in fact modified me, change the way i really believe.
-- Antonetta Ritchie IV

A brand new eBook with a brand new standpoint. It can be rally fascinating throgh reading through time. I am happy to let you know that this is the greatest ebook i have go through within my very own daily life and can be he best book for at any time.
-- Leanne Cremin