Learning Spark: Lightning-Fast Big Data Analysis

By Holden Karau, Andy Kowinski, Mark Hamstra, Matei Zaharia

O'Reilly Media, Inc, USA. Paperback. Book Condition: new. BRAND NEW, Learning Spark: Lightning-Fast Big Data Analysis, Holden Karau, Andy Kowinski, Mark Hamstra, Matei Zaharia, 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
[9.44 MB]

Reviews

Merely no phrases to spell out. I actually have read through and i am certain that i will gonna study once again again later on. You wont truly feel monotony at at any time of your time (that's what catalogues are for about should you check with me).
-- *Jaiden Konopelski*

Most of these publication is the ideal ebook readily available. it was actually writtern very flawlessly and beneficial. I discovered this book from my i and dad suggested this book to find out.
-- *Prof. Lavern Brakus*

DMCA Notice | Terms