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
[6.63 MB]

Reviews

It in one of the most popular ebook. It usually fails to price an excessive amount of. Its been printed in an extremely basic way in fact it is merely right after i finished reading through this book in which really altered me, change the way i believe.
-- Sigrid Brown

Absolutely one of the best pdf We have ever read. I really could comprehended every little thing using this written e book. I am easily could get a satisfaction of reading a written publication.
-- Dr. Odie Hamill

DMCA Notice | Terms