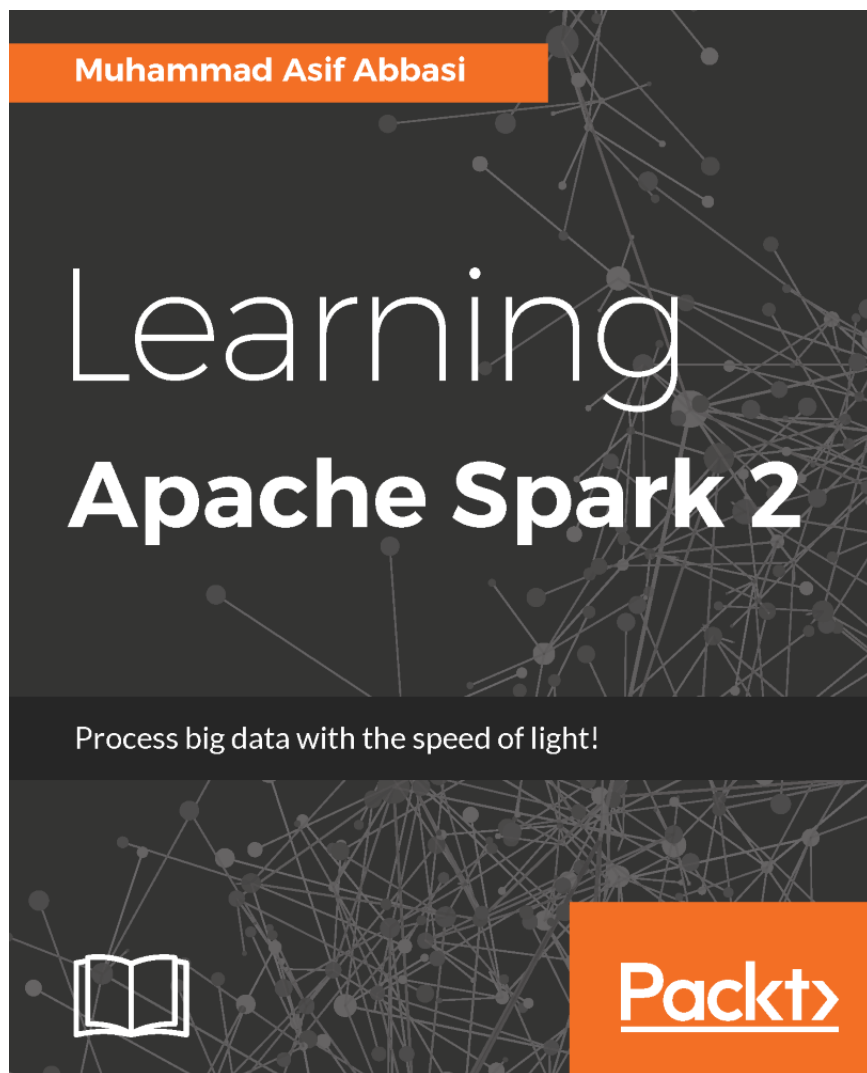


[\[电子书\]Learning Apache Spark 2 PDF下载](#)

本书于2017-03由Packt Publishing出版，作者Muhammad Asif Abbasi，全书356页。

通过本书你将学到以下知识：

- Get an overview of big data analytics and its importance for organizations and data professionals
- Delve into Spark to see how it is different from existing processing platforms
- Understand the intricacies of various file formats, and how to process them with Apache Spark.
- Realize how to deploy Spark with YARN, MESOS or a Stand-alone cluster manager.
- Learn the concepts of Spark SQL, SchemaRDD, Caching and working with Hive and Parquet file formats
- Understand the architecture of Spark MLlib while discussing some of the off-the-shelf algorithms that come with Spark.
- Introduce yourself to the deployment and usage of SparkR.
- Walk through the importance of Graph computation and the graph processing systems available in the market
- Check the real world example of Spark by building a recommendation engine with Spark using ALS.
- Use a Telco data set, to predict customer churn using Random Forests.



如果想及时了
解Spark、Hadoop或者Hbase相关的文章，欢迎关注微信公共帐号：iteblog_hadoop

本书的章节

1. Architecture and Installation
2. Transformations and Actions with Spark RDDs
3. ETL with Spark
4. Spark SQL
5. Spark Streaming
6. Machine Learning with Spark
7. GraphX
8. Operating in Clustered Mode
9. Building a Recommendation System
10. Customer Churn Prediction
11. There's More with Spark

下载地址

提供了PDF、azw3 以及 epub 二种格式的下載。

本博客文章除特别声明，全部都是原创！
转载本文请加上：转载自过往记忆 (<https://www.iteblog.com/>)
本文链接: 【】 ()