



# Learning Apache Spark 2

By Muhammad Asif Abbasi



## Learning Apache Spark 2 By Muhammad Asif Abbasi

### Key Features

- Exclusive guide that covers how to get up and running with fast data processing using Apache Spark
- Explore and exploit various possibilities with Apache Spark using real-world use cases in this book
- Want to perform efficient data processing at real time? This book will be your one-stop solution.

### Book Description

Spark juggernaut keeps on rolling and getting more and more momentum each day. The core challenge are they key capabilities in Spark (Spark SQL, Spark Streaming, Spark ML, Spark R, Graph X) etc. Having understood the key capabilities, it is important to understand how Spark can be used, in terms of being installed as a Standalone framework or as a part of existing Hadoop installation and configuring with Yarn and Mesos.

The next part of the journey after installation is using key components, APIs, Clustering, machine learning APIs, data pipelines, parallel programming. It is important to understand why each framework component is key, how widely it is being used, its stability and pertinent use cases.

Once we understand the individual components, we will take a couple of real life advanced analytics examples like:

- Building a Recommendation system
- Predicting customer churn

The objective of these real life examples is to give the reader confidence of using Spark for real-world problems.

### What you will learn

- Overview 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, Spark UDFs 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 SparkR and walk through the details of data munging including selecting, aggregating and grouping data using R studio.
- 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 collaborative filtering
- Use a telco data set, to predict customer churn using Regression

### **About the Author**

**Asif Abbasi** has worked in the industry for over 15 years, in a variety of roles starting from engineering solutions to selling solutions and everything in between. Asif is currently working with SAS a Market Leader in Analytic Solutions as a Principal Business Solutions Manager for the Global Technologies Practice.

Based out of London, Asif has vast experience in consulting for major organizations & industries across the globe, and running proof-of-concepts across various industries including but not limited to Telecommunications, Manufacturing, Retail, Finance, Services, Utilities and Government.

Asif has presented at various conferences and delivered workshops on topics such as Big Data, Hadoop, Teradata, and Analytics using Aster on Teradata and Hadoop. Asif is a Oracle Certified Java EE 5 Enterprise Architect, Teradata Certified Master, PMP, Hortonworks Hadoop Certified developer and Administrator. Asif also holds a Masters degree in Computer Science and Business Administration.

 [Download Learning Apache Spark 2 ...pdf](#)

 [Read Online Learning Apache Spark 2 ...pdf](#)

# Learning Apache Spark 2

*By Muhammad Asif Abbasi*

**Learning Apache Spark 2** By Muhammad Asif Abbasi

## Key Features

- Exclusive guide that covers how to get up and running with fast data processing using Apache Spark
- Explore and exploit various possibilities with Apache Spark using real-world use cases in this book
- Want to perform efficient data processing at real time? This book will be your one-stop solution.

## Book Description

Spark juggernaut keeps on rolling and getting more and more momentum each day. The core challenge are they key capabilities in Spark (Spark SQL, Spark Streaming, Spark ML, Spark R, Graph X) etc. Having understood the key capabilities, it is important to understand how Spark can be used, in terms of being installed as a Standalone framework or as a part of existing Hadoop installation and configuring with Yarn and Mesos.

The next part of the journey after installation is using key components, APIs, Clustering, machine learning APIs, data pipelines, parallel programming. It is important to understand why each framework component is key, how widely it is being used, its stability and pertinent use cases.

Once we understand the individual components, we will take a couple of real life advanced analytics examples like:

- Building a Recommendation system
- Predicting customer churn

The objective of these real life examples is to give the reader confidence of using Spark for real-world problems.

## What you will learn

- Overview 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, Spark UDFs 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 SparkR and walk through the details of data munging including selecting, aggregating and grouping data using R studio.
- 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 collaborative filtering

- Use a telco data set, to predict customer churn using Regression

### **About the Author**

**Asif Abbasi** has worked in the industry for over 15 years, in a variety of roles starting from engineering solutions to selling solutions and everything in between. Asif is currently working with SAS a Market Leader in Analytic Solutions as a Principal Business Solutions Manager for the Global Technologies Practice.

Based out of London, Asif has vast experience in consulting for major organizations & industries across the globe, and running proof-of-concepts across various industries including but not limited to Telecommunications, Manufacturing, Retail, Finance, Services, Utilities and Government.

Asif has presented at various conferences and delivered workshops on topics such as Big Data, Hadoop, Teradata, and Analytics using Aster on Teradata and Hadoop. Asif is a Oracle Certified Java EE 5 Enterprise Architect, Teradata Certified Master, PMP, Hortonworks Hadoop Certified developer and Administrator. Asif also holds a Masters degree in Computer Science and Business Administration.

### **Learning Apache Spark 2 By Muhammad Asif Abbasi Bibliography**

- Rank: #399008 in eBooks
- Published on: 2017-06-06
- Released on: 2017-06-06
- Format: Kindle eBook

 [Download Learning Apache Spark 2 ...pdf](#)

 [Read Online Learning Apache Spark 2 ...pdf](#)

### Editorial Review

About the Author

#### **Muhammad Asif Abbasi**

Muhammad Asif Abbasi has worked in the industry for over 15 years in a variety of roles from engineering solutions to selling solutions and everything in between. Asif is currently working with SAS a market leader in Analytic Solutions as a Principal Business Solutions Manager for the Global Technologies Practice. Based in London, Asif has vast experience in consulting for major organizations and industries across the globe, and running proof-of-concepts across various industries including but not limited to telecommunications, manufacturing, retail, finance, services, utilities and government. Asif is an Oracle Certified Java EE 5 Enterprise architect, Teradata Certified Master, PMP, Hortonworks Hadoop Certified developer, and administrator. Asif also holds a Master's degree in Computer Science and Business Administration.

### Users Review

**From reader reviews:**

#### **Marcia Eberhart:**

The book Learning Apache Spark 2 can give more knowledge and information about everything you want. So why must we leave the best thing like a book Learning Apache Spark 2? Several of you have a different opinion about publication. But one aim this book can give many details for us. It is absolutely right. Right now, try to closer along with your book. Knowledge or details that you take for that, it is possible to give for each other; you could share all of these. Book Learning Apache Spark 2 has simple shape nevertheless, you know: it has great and massive function for you. You can look the enormous world by start and read a book. So it is very wonderful.

#### **Edith Macklin:**

As people who live in the modest era should be up-date about what going on or facts even knowledge to make these individuals keep up with the era that is always change and move ahead. Some of you maybe will certainly update themselves by studying books. It is a good choice for yourself but the problems coming to you is you don't know which you should start with. This Learning Apache Spark 2 is our recommendation so you keep up with the world. Why, since this book serves what you want and wish in this era.

#### **Linda Livingston:**

Do you like reading a book? Confuse to looking for your preferred book? Or your book ended up being rare? Why so many concern for the book? But just about any people feel that they enjoy to get reading. Some people likes looking at, not only science book but in addition novel and Learning Apache Spark 2 or perhaps others sources were given know-how for you. After you know how the great a book, you feel desire to read more and more. Science book was created for teacher as well as students especially. Those textbooks are

helping them to put their knowledge. In other case, beside science book, any other book likes Learning Apache Spark 2 to make your spare time much more colorful. Many types of book like this one.

**Maria Blanco:**

What is your hobby? Have you heard that question when you got students? We believe that that problem was given by teacher to their students. Many kinds of hobby, Every person has different hobby. And you know that little person such as reading or as examining become their hobby. You need to understand that reading is very important along with book as to be the point. Book is important thing to incorporate you knowledge, except your own personal teacher or lecturer. You discover good news or update concerning something by book. A substantial number of sorts of books that can you decide to try be your object. One of them is niagra Learning Apache Spark 2.

**Download and Read Online Learning Apache Spark 2 By  
Muhammad Asif Abbasi #HZKNU4OID8Y**

## **Read Learning Apache Spark 2 By Muhammad Asif Abbasi for online ebook**

Learning Apache Spark 2 By Muhammad Asif Abbasi Free PDF d0wnl0ad, audio books, books to read, good books to read, cheap books, good books, online books, books online, book reviews epub, read books online, books to read online, online library, greatbooks to read, PDF best books to read, top books to read Learning Apache Spark 2 By Muhammad Asif Abbasi books to read online.

### **Online Learning Apache Spark 2 By Muhammad Asif Abbasi ebook PDF download**

**Learning Apache Spark 2 By Muhammad Asif Abbasi Doc**

**Learning Apache Spark 2 By Muhammad Asif Abbasi Mobipocket**

**Learning Apache Spark 2 By Muhammad Asif Abbasi EPub**

**HZKNU4OID8Y: Learning Apache Spark 2 By Muhammad Asif Abbasi**