Big Data Analytics with Spark and Hadoop

Venkat Ankam

Big Data Analytics with Spark and Hadoop

Venkat Ankam

Big Data Analytics with Spark and Hadoop Venkat Ankam

A handy reference guide for data analysts and data scientists to fetch "Value" out of big data analytics using Spark on Hadoop Clusters

About This Book

- Practical tutorial with real-world examples that explores Spark on Hadoop clusters
- This book is based on the latest version of Apache Spark and Hadoop integrated with the most commonly used tools
- Learn about all the Spark stack components including the latest topics such as DataFrames, DataSets, and SparkR

Who This Book Is For

Though this book is primarily aimed at data analysts and data scientists, it will also help architects, programmers, and practitioners. Knowledge of either Spark or Hadoop would be beneficial. It is assumed that you have basic programming background in Scala, Python, SQL, or R programming with basic Linux experience. Working experience within big data environments is not mandatory.

What You Will Learn

- Find out about and implement the tools and techniques of big data analytics using Spark on Hadoop clusters
- Understand all the Hadoop and Spark ecosystem components and how Spark replaced MapReduce
- Get to know all the Spark components: Spark Core, Spark SQL, DataFrames, DataSets, Streaming, MLLib, and Graphx
- See batch and real-time data analytics using Spark Core, Spark SQL, and Spark Streaming
- Get to grips with data science and machine learning using MLLib, H2O, Hivemall, Graphx, and SparkR
- Get an introduction to all the new tools (based on Notebooks, Data Flow, and Spark as a Service) and their integrations with Spark and Hadoop

In Detail

This book explains the fundamentals of Apache Spark and Hadoop, and how they are easily integrated together with the most commonly used tools and techniques. All the Spark components—Spark Core, Spark SQL, DataFrames, Data sets, Streaming, MLlib, Graphx, and Hadoop core components—HDFS, MapReduce, and Yarn are explored in greater depth with implementation examples on Spark and Hadoop clusters.

The big data analytics industry is moving away from MapReduce to Spark. In this book, the advantages of Spark over MapReduce are explained at great depth so you can reap the benefits of in-memory speeds. The DataFrames API, Data Sources API, and new Data sets API are explained so you can build big data

analytical applications.

We'll explore real-time data analytics using Spark Streaming with Apache Kafka and HBase to help you build streaming applications. You'll get to know the machine learning techniques using MLLib and SparkR, and Graph Analytics with the GraphX component of Spark.You will also get the opportunity to start working with web-based notebooks such as Jupyter, Apache Zeppelin, and the data flow tool Apache NiFi to analyze and visualize data.

Download Big Data Analytics with Spark and Hadoop ...pdf

Read Online Big Data Analytics with Spark and Hadoop ...pdf

Download and Read Free Online Big Data Analytics with Spark and Hadoop Venkat Ankam

---

From reader reviews:

Barbara Richardson:

Book will be written, printed, or outlined for everything. You can learn everything you want by a publication. Book has a different type. We all know that that book is important issue to bring us around the world. Beside that you can your reading expertise was fluently. A reserve Big Data Analytics with Spark and Hadoop will make you to be smarter. You can feel much more confidence if you can know about everything. But some of you think in which open or reading any book make you bored. It isn't make you fun. Why they could be thought like that? Have you seeking best book or acceptable book with you?

David Binkley:

Information is provisions for anyone to get better life, information presently can get by anyone with everywhere. The information can be a expertise or any news even a problem. What people must be consider whenever those information which is from the former life are challenging be find than now's taking seriously which one works to believe or which one the resource are convinced. If you obtain the unstable resource then you understand it as your main information there will be huge disadvantage for you. All those possibilities will not happen throughout you if you take Big Data Analytics with Spark and Hadoop as the daily resource information.

Karen Jude:

Reading a book tends to be new life style in this era globalization. With examining you can get a lot of information that may give you benefit in your life. With book everyone in this world could share their idea. Textbooks can also inspire a lot of people. A lot of author can inspire their reader with their story or maybe their experience. Not only the storyplot that share in the guides. But also they write about the ability about something that you need instance. How to get the good score toefl, or how to teach your children, there are many kinds of book that exist now. The authors in this world always try to improve their skill in writing, they also doing some research before they write to their book. One of them is this Big Data Analytics with Spark and Hadoop.

Gerald Sosa:

Is it you actually who having spare time then spend it whole day through watching television programs or just lying down on the bed? Do you need something new? This Big Data Analytics with Spark and Hadoop can be the reply, oh how comes? The new book you know. You are thus out of date, spending your time by reading in this completely new era is common not a geek activity. So what these publications have than the others?

Download and Read Online Big Data Analytics with Spark and Hadoop Venkat Ankam #CSA9UZVIK4H

Read Big Data Analytics with Spark and Hadoop by Venkat Ankam for online ebook

Big Data Analytics with Spark and Hadoop by Venkat Ankam 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 Big Data Analytics with Spark and Hadoop by Venkat Ankam books to read online.

Online Big Data Analytics with Spark and Hadoop by Venkat Ankam ebook PDF download

Big Data Analytics with Spark and Hadoop by Venkat Ankam Doc

Big Data Analytics with Spark and Hadoop by Venkat Ankam Mobipocket

Big Data Analytics with Spark and Hadoop by Venkat Ankam EPub