Scala for Data Science

By Pascal Bugnion

Scala for Data Science By Pascal Bugnion

Leverage the power of Scala with different tools to build scalable, robust data science applications

About This Book

- A complete guide for scalable data science solutions, from data ingestion to data visualization
- Deploy horizontally scalable data processing pipelines and take advantage of web frameworks to build engaging visualizations
- Build functional, type-safe routines to interact with relational and NoSQL databases with the help of tutorials and examples provided

Who This Book Is For

If you are a Scala developer or data scientist, or if you want to enter the field of data science, then this book will give you all the tools you need to implement data science solutions.

What You Will Learn

- Transform and filter tabular data to extract features for machine learning
- Implement your own algorithms or take advantage of MLLib's extensive suite of models to build distributed machine learning pipelines
- Read, transform, and write data to both SQL and NoSQL databases in a functional manner
- Write robust routines to query web APIs
- Read data from web APIs such as the GitHub or Twitter API
- Use Scala to interact with MongoDB, which offers high performance and helps to store large data sets with uncertain query requirements
- Create Scala web applications that couple with JavaScript libraries such as D3 to create compelling interactive visualizations
- Deploy scalable parallel applications using Apache Spark, loading data from HDFS or Hive

In Detail

Scala is a multi-paradigm programming language (it supports both object-oriented and functional programming) and scripting language used to build applications for the JVM. Languages such as R, Python, Java, and so on are mostly used for data science. It is particularly good at analyzing large sets of data without any significant impact on performance and thus Scala is being adopted by many developers and data scientists. Data scientists might be aware that building applications that are truly scalable is hard. Scala, with its powerful functional libraries for interacting with databases and building scalable frameworks will give you the tools to construct robust data pipelines.

This book will introduce you to the libraries for ingesting, storing, manipulating, processing, and visualizing data in Scala.

Packed with real-world examples and interesting data sets, this book will teach you to ingest data from flat files and web APIs and store it in a SQL or NoSQL database. It will show you how to design scalable architectures to process and modelling your data, starting from simple concurrency constructs such as parallel collections and futures, through to actor systems and Apache Spark. As well as Scala's emphasis on functional structures and immutability, you will learn how to use the right parallel construct for the job at hand, minimizing development time without compromising scalability. Finally, you will learn how to build beautiful interactive visualizations using web frameworks.

This book gives tutorials on some of the most common Scala libraries for data science, allowing you to quickly get up to speed with building data science and data engineering solutions.

Style and approach

A tutorial with complete examples, this book will give you the tools to start building useful data engineering and data science solutions straightaway

 Download Scala for Data Science ...pdf

 Read Online Scala for Data Science ...pdf

Scala for Data Science

By Pascal Bugnion

Scala for Data Science By Pascal Bugnion

Leverage the power of Scala with different tools to build scalable, robust data science applications

About This Book

- A complete guide for scalable data science solutions, from data ingestion to data visualization
- Deploy horizontally scalable data processing pipelines and take advantage of web frameworks to build engaging visualizations
- Build functional, type-safe routines to interact with relational and NoSQL databases with the help of tutorials and examples provided

Who This Book Is For

If you are a Scala developer or data scientist, or if you want to enter the field of data science, then this book will give you all the tools you need to implement data science solutions.

What You Will Learn

- Transform and filter tabular data to extract features for machine learning
- Implement your own algorithms or take advantage of MLLib's extensive suite of models to build distributed machine learning pipelines
- Read, transform, and write data to both SQL and NoSQL databases in a functional manner
- Write robust routines to query web APIs
- Read data from web APIs such as the GitHub or Twitter API
- Use Scala to interact with MongoDB, which offers high performance and helps to store large data sets with uncertain query requirements
- Create Scala web applications that couple with JavaScript libraries such as D3 to create compelling interactive visualizations
- Deploy scalable parallel applications using Apache Spark, loading data from HDFS or Hive

In Detail

Scala is a multi-paradigm programming language (it supports both object-oriented and functional programming) and scripting language used to build applications for the JVM. Languages such as R, Python, Java, and so on are mostly used for data science. It is particularly good at analyzing large sets of data without any significant impact on performance and thus Scala is being adopted by many developers and data scientists. Data scientists might be aware that building applications that are truly scalable is hard. Scala, with its powerful functional libraries for interacting with databases and building scalable frameworks will give you the tools to construct robust data pipelines.

This book will introduce you to the libraries for ingesting, storing, manipulating, processing, and visualizing data in Scala.

Packed with real-world examples and interesting data sets, this book will teach you to ingest data from flat files and web APIs and store it in a SQL or NoSQL database. It will show you how to design scalable architectures to process and modelling your data, starting from simple concurrency constructs such as parallel collections and futures, through to actor systems and Apache Spark. As well as Scala's emphasis on functional structures and immutability, you will learn how to use the right parallel construct for the job at hand, minimizing development time without compromising scalability. Finally, you will learn how to build beautiful interactive visualizations using web frameworks.

This book gives tutorials on some of the most common Scala libraries for data science, allowing you to quickly get up to speed with building data science and data engineering solutions.

Style and approach

A tutorial with complete examples, this book will give you the tools to start building useful data engineering and data science solutions straightaway

Scala for Data Science By Pascal Bugnion Bibliography

- Sales Rank: #139125 in eBooks
- Published on: 2016-01-28
- Released on: 2016-01-28
- Format: Kindle eBook

⬇ **Download** Scala for Data Science ...pdf

▤ **Read Online** Scala for Data Science ...pdf

Editorial Review

About the Author

Pascal Bugnion

Pascal Bugnion is a data engineer at the ASI, a consultancy offering bespoke data science services. Previously, he was the head of data engineering at SCL Elections. He holds a PhD in computational physics from Cambridge University. Besides Scala, Pascal is a keen Python developer. He has contributed to NumPy, matplotlib and IPython. He also maintains scikit-monaco, an open source library for Monte Carlo integration. He currently lives in London, UK.

Users Review

From reader reviews:

Margaret Wright:

Have you spare time for just a day? What do you do when you have considerably more or little spare time? Yep, you can choose the suitable activity for spend your time. Any person spent their own spare time to take a wander, shopping, or went to typically the Mall. How about open or read a book allowed Scala for Data Science? Maybe it is to become best activity for you. You already know beside you can spend your time with your favorite's book, you can more intelligent than before. Do you agree with it is opinion or you have other opinion?

Rocio Linville:

Here thing why this particular Scala for Data Science are different and dependable to be yours. First of all reading through a book is good but it really depends in the content of it which is the content is as tasty as food or not. Scala for Data Science giving you information deeper since different ways, you can find any book out there but there is no e-book that similar with Scala for Data Science. It gives you thrill reading through journey, its open up your eyes about the thing which happened in the world which is possibly can be happened around you. It is easy to bring everywhere like in recreation area, café, or even in your approach home by train. When you are having difficulties in bringing the paper book maybe the form of Scala for Data Science in e-book can be your option.

Daniel McDonald:

This Scala for Data Science are usually reliable for you who want to certainly be a successful person, why. The explanation of this Scala for Data Science can be one of the great books you must have is usually giving you more than just simple looking at food but feed anyone with information that perhaps will shock your before knowledge. This book is handy, you can bring it all over the place and whenever your conditions both in e-book and printed versions. Beside that this Scala for Data Science giving you an enormous of experience for example rich vocabulary, giving you trial run of critical thinking that we all know it useful in your day

task. So , let's have it appreciate reading.

Jose Roberts:

That reserve can make you to feel relax. This particular book Scala for Data Science was colorful and of course has pictures on there. As we know that book Scala for Data Science has many kinds or variety. Start from kids until adolescents. For example Naruto or Detective Conan you can read and feel that you are the character on there. So , not at all of book are generally make you bored, any it can make you feel happy, fun and relax. Try to choose the best book in your case and try to like reading this.

Download and Read Online Scala for Data Science By Pascal Bugnion #FEBYSR10H6K

Read Scala for Data Science By Pascal Bugnion for online ebook

Scala for Data Science By Pascal Bugnion 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 Scala for Data Science By Pascal Bugnion books to read online.

Online Scala for Data Science By Pascal Bugnion ebook PDF download

Scala for Data Science By Pascal Bugnion Doc

Scala for Data Science By Pascal Bugnion Mobipocket

Scala for Data Science By Pascal Bugnion EPub

FEBYSR10H6K: Scala for Data Science By Pascal Bugnion