



# Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms

*By Nick Pentreath*

Download now

Read Online ➔

## Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms By Nick Pentreath

### Key Features

- Follow real-world examples to learn how to develop your own machine learning systems with Spark
- A practical tutorial with real-world use cases allowing you to develop your own machine learning systems with Spark
- Combine various techniques and models into an intelligent machine learning system
- Explore and use Spark's powerful range of features to load, analyze, clean, and your data

### Book Description

Apache Spark is a framework for distributed computing that is designed from the ground up to be optimized for low latency tasks and in-memory data storage. It is one of the few frameworks for parallel computing that combines speed, scalability, in-memory processing, and fault tolerance with ease of programming and a flexible, expressive, and powerful API design.

This book guides you through the basics of Spark's API used to load and process data and prepare the data to use as input to the various machine learning models. There are detailed examples and real-world use cases for you to explore common machine learning models including recommender systems, classification, regression, clustering, and dimensionality reduction. You will cover advanced topics such as working with large-scale text data, and methods for online machine learning and model evaluation using Spark Streaming.

### What you will learn

- Create your first Spark program in Scala, Java, and Python
- Set up and configure a development environment for Spark on your own computer, as well as on Amazon EC2

- Access public machine learning datasets and use Spark to load, process, clean, and transform data
- Use Spark's machine learning library to implement programs utilizing well-known machine learning models including collaborative filtering, classification, regression, clustering, and dimensionality reduction
- Write Spark functions to evaluate the performance of your machine learning models
- Deal with large-scale text data, including feature extraction and using text data as input to your machine learning models
- Explore online learning methods and use Spark Streaming for online learning and model evaluation

### About the Author

**Nick Pentreath** is a member of the Apache Spark Project Management Committee. He has a background in financial markets, machine learning, and software development, including experience as a research scientist at the online ad targeting start-up Cognitive Match Limited in London and leading the Data Science and Analytics team at Mxit, Africa's largest social network. He is also one of the cofounders of Graphflow, a big data and machine learning company focused on user-centric recommendations and customer intelligence.

### Table of Contents

1. Getting Up and Running with Spark
2. Designing a Machine Learning System
3. Obtaining, Processing and Preparing Data with Spark
4. Building a Recommendation Engine with Spark
5. Building a Classification Model with Spark
6. Building a Regression Model with Spark
7. Building a Clustering Model with Spark
8. Dimensionality Reduction with Spark
9. Advanced Text Processing with Spark
10. Real-Time Machine Learning with Spark Streaming

 [Download Machine Learning with Spark - Tackle Big Data with ...pdf](#)

 [Read Online Machine Learning with Spark - Tackle Big Data wi ...pdf](#)

# Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms

*By Nick Pentreath*

## Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms

By Nick Pentreath

### Key Features

- Follow real-world examples to learn how to develop your own machine learning systems with Spark
- A practical tutorial with real-world use cases allowing you to develop your own machine learning systems with Spark
- Combine various techniques and models into an intelligent machine learning system
- Explore and use Spark's powerful range of features to load, analyze, clean, and your data

### Book Description

Apache Spark is a framework for distributed computing that is designed from the ground up to be optimized for low latency tasks and in-memory data storage. It is one of the few frameworks for parallel computing that combines speed, scalability, in-memory processing, and fault tolerance with ease of programming and a flexible, expressive, and powerful API design.

This book guides you through the basics of Spark's API used to load and process data and prepare the data to use as input to the various machine learning models. There are detailed examples and real-world use cases for you to explore common machine learning models including recommender systems, classification, regression, clustering, and dimensionality reduction. You will cover advanced topics such as working with large-scale text data, and methods for online machine learning and model evaluation using Spark Streaming.

### What you will learn

- Create your first Spark program in Scala, Java, and Python
- Set up and configure a development environment for Spark on your own computer, as well as on Amazon EC2
- Access public machine learning datasets and use Spark to load, process, clean, and transform data
- Use Spark's machine learning library to implement programs utilizing well-known machine learning models including collaborative filtering, classification, regression, clustering, and dimensionality reduction
- Write Spark functions to evaluate the performance of your machine learning models
- Deal with large-scale text data, including feature extraction and using text data as input to your machine learning models
- Explore online learning methods and use Spark Streaming for online learning and model evaluation

### About the Author

**Nick Pentreath** is a member of the Apache Spark Project Management Committee. He has has a background in financial markets, machine learning, and software development, including experience as a research scientist at the online ad targeting start-up Cognitive Match Limited in London and leading the Data Science and Analytics team at Mxit, Africa's largest social network. He is also one of the cofounders of Graphflow, a

big data and machine learning company focused on user-centric recommendations and customer intelligence.

## Table of Contents

1. Getting Up and Running with Spark
2. Designing a Machine Learning System
3. Obtaining, Processing and Preparing Data with Spark
4. Building a Recommendation Engine with Spark
5. Building a Classification Model with Spark
6. Building a Regression Model with Spark
7. Building a Clustering Model with Spark
8. Dimensionality Reduction with Spark
9. Advanced Text Processing with Spark
10. Real-Time Machine Learning with Spark Streaming

## Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms By Nick Pentreath Bibliography

- Sales Rank: #874703 in Books
- Published on: 2014-12-08
- Released on: 2015-02-20
- Original language: English
- Number of items: 1
- Dimensions: 9.25" h x .77" w x 7.50" l, 1.28 pounds
- Binding: Paperback
- 329 pages

 [Download Machine Learning with Spark - Tackle Big Data with ...pdf](#)

 [Read Online Machine Learning with Spark - Tackle Big Data wi ...pdf](#)

## **Download and Read Free Online Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms By Nick Pentreath**

---

### **Editorial Review**

About the Author

#### **Nick Pentreath**

Nick Pentreath has a background in financial markets, machine learning, and software development. He has worked at Goldman Sachs Group, Inc.; as a research scientist at the online ad targeting start-up Cognitive Match Limited, London; and led the Data Science and Analytics team at Mxit, Africa's largest social network. He is a cofounder of Graphflow, a big data and machine learning company focused on user-centric recommendations and customer intelligence. He is passionate about combining commercial focus with machine learning and cutting-edge technology to build intelligent systems that learn from data to add value to the bottom line. Nick is a member of the Apache Spark Project Management Committee.

### **Users Review**

**From reader reviews:**

#### **Mandy Conway:**

Why don't make it to become your habit? Right now, try to ready your time to do the important act, like looking for your favorite book and reading a reserve. Beside you can solve your short lived problem; you can add your knowledge by the reserve entitled Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms. Try to face the book Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms as your good friend. It means that it can being your friend when you sense alone and beside associated with course make you smarter than ever before. Yeah, it is very fortunated in your case. The book makes you considerably more confidence because you can know every little thing by the book. So , let me make new experience and also knowledge with this book.

#### **Miguel Willis:**

Here thing why this specific Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms are different and reputable to be yours. First of all reading a book is good but it depends in the content of computer which is the content is as delightful as food or not. Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms giving you information deeper and different ways, you can find any reserve out there but there is no guide that similar with Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms. It gives you thrill reading through journey, its open up your own eyes about the thing that happened in the world which is probably can be happened around you. You can actually bring everywhere like in area, café, or even in your means home by train. If you are having difficulties in bringing the branded book maybe the form of Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms in e-book can be your option.

**Patty Shield:**

The book Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms has a lot associated with on it. So when you make sure to read this book you can get a lot of help. The book was published by the very famous author. Tom makes some research ahead of write this book. This particular book very easy to read you will get the point easily after scanning this book.

**Terri Brown:**

Why? Because this Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms is an unordinary book that the inside of the reserve waiting for you to snap that but latter it will surprise you with the secret the idea inside. Reading this book beside it was fantastic author who all write the book in such awesome way makes the content interior easier to understand, entertaining approach but still convey the meaning entirely. So , it is good for you for not hesitating having this any more or you going to regret it. This book will give you a lot of benefits than the other book get such as help improving your ability and your critical thinking means. So , still want to hold up having that book? If I had been you I will go to the guide store hurriedly.

**Download and Read Online Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms By Nick Pentreath #MVBPEA5NYJ2**

# **Read Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms By Nick Pentreath for online ebook**

Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms By Nick Pentreath 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 Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms By Nick Pentreath books to read online.

## **Online Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms By Nick Pentreath ebook PDF download**

**Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms By Nick Pentreath Doc**

**Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms By Nick Pentreath Mobipocket**

**Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms By Nick Pentreath EPub**

**MVBPEA5NYJ2: Machine Learning with Spark - Tackle Big Data with Powerful Spark Machine Learning Algorithms By Nick Pentreath**