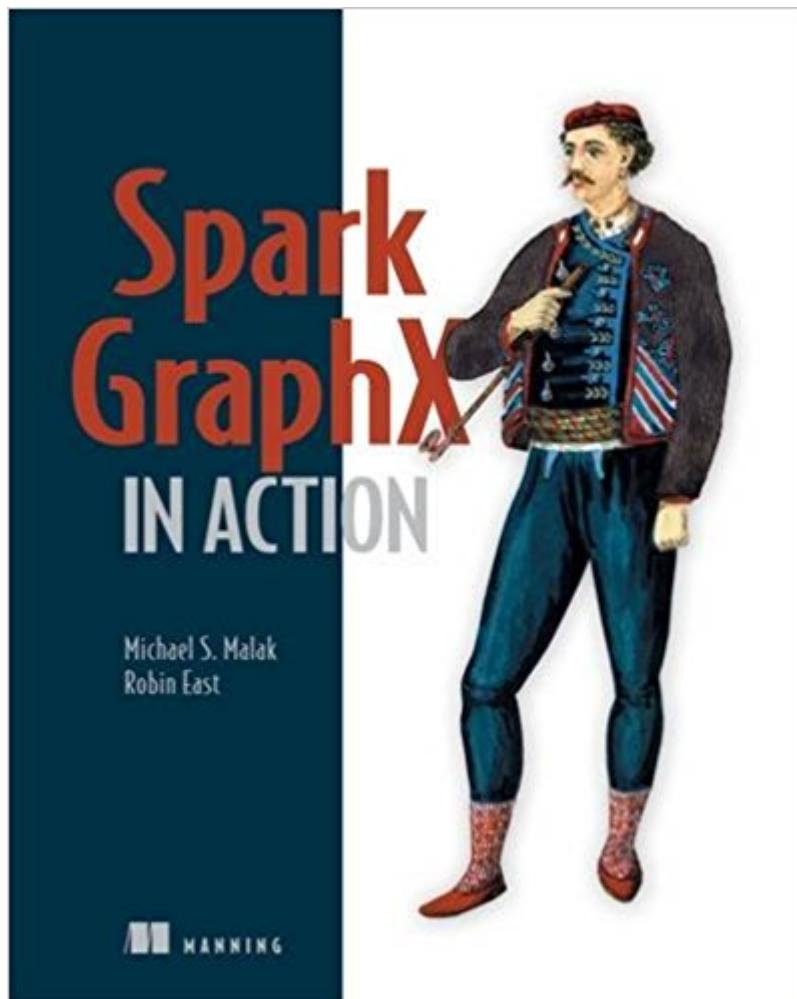The book was found

Spark GraphX In Action

Synopsis

Summary Spark GraphX in Action starts out with an overview of Apache Spark and the GraphX graph processing API. This example-based tutorial then teaches you how to configure GraphX and how to use it interactively. Along the way, you'll collect practical techniques for enhancing applications and applying machine learning algorithms to graph data. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About the Technology GraphX is a powerful graph processing API for the Apache Spark analytics engine that lets you draw insights from large datasets. GraphX gives you unprecedented speed and capacity for running massively parallel and machine learning algorithms. About the Book Spark GraphX in Action begins with the big picture of what graphs can be used for. This example-based tutorial teaches you how to use GraphX interactively. You'll start with a crystal-clear introduction to building big data graphs from regular data, and then explore the problems and possibilities of implementing graph algorithms and architecting graph processing pipelines. Along the way, you'll collect practical techniques for enhancing applications and applying machine learning algorithms to graph data. What's Inside Understanding graph technologyUsing the GraphX APIDeveloping algorithms for big graphsMachine learning with graphsGraph visualization About the Reader Readers should be comfortable writing code. Experience with Apache Spark and Scala is not required. About the Authors Michael Malak has worked on Spark applications for Fortune 500 companies since early 2013. Robin East has worked as a consultant to large organizations for over 15 years and is a data scientist at Worldpay. Table of Contents PART 1 SPARK AND GRAPHSTwo important technologies: Spark and graphsGraphX quick startSome fundamentalsPART 2 CONNECTING VERTICESGraphX BasicsBuilt-in algorithmsOther useful graph algorithmsMachine learningPART 3 OVER THE ARCThe missing algorithmsPerformance and monitoringOther languages and tools

Book Information

Paperback: 280 pages

Publisher: Manning Publications; 1 edition (July 3, 2016)

Language: English

ISBN-10: 1617292524

ISBN-13: 978-1617292521

Product Dimensions: 7.3 x 0.5 x 9.2 inches

Shipping Weight: 15.2 ounces (View shipping rates and policies)

Average Customer Review: 4.8 out of 5 stars   8 customer reviews

Best Sellers Rank: #217,631 in Books (See Top 100 in Books) #37 inÃ Â Books > Computers & Technology > Web Development & Design > Web Services #57 inÃ Â Books > Textbooks > Computer Science > Algorithms #123 inÃ Â Books > Computers & Technology > Databases & Big Data > Data Modeling & Design

Customer Reviews

Michael Malak has worked on Spark applications for Fortune 500 companies since early 2013.Robin East has worked as a consultant to large organizations for over 15 years and is a data scientist at Worldpay.

"Spark GraphX in Action" is the best resource out there for learning this fascinating technology. It is clearly-written, with a lot of hands-on examples, all clearly annotated and explained in the typically superb style of Manning books. The two authors of this book have, in my opinion, succeeded exceptionally well in achieving the two ambitious goals that they had set out for this book to achieve: "to cover everything possible about Spark GraphX, and to assume little to no expertise about any of the technologies represented...[and] the hefty amount of prerequisites to get into GraphXÃƒÂ¢Ã â ¬Ã â •specifically, Spark, Scala, and graphs".The end result is a volume that is pleasant to read. Trust me, the material in the book is *not* watered down. There's plenty in here to keep the reader engaged for a bit. Well done to the authors for demystifying graphs by showing what problems graphs can solve in the context of Spark GraphX!

This book is ideal for anyone wishing to apply graph algorithms to Big Data, esp. those already using Spark or dealing with huge amounts of data. It targets experienced software developers, but reviews the basics of Spark and the Scala language as well as introducing GraphX. The pages are packed with information presented concisely and the book moves right along. The book is very readable and personable (I think the style rather enjoyable to read), but I found it helpful to look up certain terms or acronyms online for more information on unfamiliar topics that the authors sometimes mention in passing.The many charts and graphs were well done and particularly helpful in making the material understandable. There are also plenty of code snippets with great explanations of the pros and cons of various approaches. Topics covered include Spark, Scala, GraphX, graph algorithms, applications (several are discussed in detail), machine learning, and performance tuning.I would say this book is mid-level to advanced; it is not for a novice programmer, but is certainly accessible to someone with much programming experience. Some (or

much) knowledge of Hadoop and/or Spark is certainly helpful. For those needing more to get up to speed, or take what you've learned in the book further, the authors list some excellent online resources in appendix C. The publisher (Manning) has a book just on Spark coming out this fall, "Spark In Action". Some of the other books on Spark are getting dated already, so this is much needed and might make a great companion to this book.Overall this is a well-written, very informative book that will help an experienced software developer become a master at using graph algorithms to solve Big Data problems.

Our team has worked with Spark since 1.0 and I needed to get up to speed. My team recommended this book. The team is made up of experienced people; some have advanced degrees and all have over 20 years experience each. The point is not to brag, but to suggest we have had to evaluate multiple books in the past. But, note this book is accessible to anyone with some programming experience and curiosity.The authors are clearly knowledgeable and presentation and flow are clear and concise. Chapter 3, in particular, which covers fundamentals of Spark and Scala, was very useful. Compared to the large tomes out there on these subjects, this chapter quickly covered what I needed to move forward. The same applied to Chapter 4, which introduces GraphX fundamentals.I have prior experience with machine learning and I appreciated Chapter 7 covering Spark capabilities couched in supervised, unsupervised, and semi-supervised terms. This partitioned GraphX (and some MLlib) capabilities into manageable chunks by application. These concepts are introduced as well, but those familiar with machine learning may feel at home too.Monitoring Spark job progress and performance is not always the easiest. In Chapter 9, the authorsÃƒÂ¢ â ¬Ã â„¢ real-world experience shines with tips and tricks that would only come with actual hands-on trial and error. This is important. When using a new technology, things may not work as expected and learning how to debug at the same time is twice as brutal. So, having real-world operational debugging approaches really helped.I do wish some of the examples were a little more in-depth. The authors appear to have the experience to have done this. But, this book is designed to make people productive quickly, so the authors probably made a conscious choice here.Overall, this book saved me a lot of time and our team has used it. If I calculate the time this book saved and what development time costs, the cost-benefit of this book is clear if you are working with Spark and GraphX.

If you are wanting to dive into GraphX then this is the book to buy. It provides an in-depth look under the hood of SparkÃƒÂ¢Ã â ¬Ã â„¢s graph processing system and will have you running PageRank

and other graph algorithms in only a few lines of code. Graph processing is a powerful tool that is a natural fit for modelling many types of problems. Traditional graph algorithms are covered along with machine learning algorithms. I found this book to be readable by all, but most accessible to Scala developers, since that is what the code snippets are written in. Even so, with no background in Scala, Spark, or GraphX, I found this book to be a worthy read.

<u>Download to continue reading...</u>

Spark GraphX in Action How to Draw Action Figures: Book 2: More than 70 Sketches of Action Figures and Action Poses (Drawing Action Figures, Draw Action Figures Book, How Draw Action Poses, Draw Comic Figures) The Spark Story Bible: Spark a Journey through God's Word High Performance Spark: Best Practices for Scaling and Optimizing Apache Spark Spark in Action FIRST-TIME LESBIAN (5 XXX Hot Action !): FIVE STORY BUNDLE (First-time XXX Lesbian Action) Action Book: Monster Party (Party Action Book) The Action Bible Devotional: 52 Weeks of God-Inspired Adventure (Action Bible Series) The Action Bible New Testament: God's Redemptive Story (Action Bible Series) The Action Bible: God's Redemptive Story (Action Bible Series) Price Action Breakdown: Exclusive Price Action Trading Approach to Financial Markets A Kids' Guide to Climate Change & Global Warming: How to Take Action! (How to Take Action! Series) Technology In Action Introductory (14th Edition) (Evans, Martin & Poatsy, Technology in Action Series) Technology In Action Introductory (13th Edition) (Evans, Martin & Poatsy, Technology in Action Series) BMX Racing (Torque Books: Action Sports) (Torque: Action Sports) BMX Freestyle (Torque Books: Action Sports) (Torque: Action Sports) Technology In Action Complete (14th Edition) (Evans, Martin & Poatsy, Technology in Action Series) Technology In Action Complete (Evans, Martin & Poatsy, Technology in Action Series) Black Sheep - An Action Thriller Novel (A Noah Wolf Novel, Thriller, Action, Mystery Book 6) Vocabulary in Action Level H Teacher Guide: Word Meaning, Pronunciation, Prefixes, Suffixes, Synonyms, Antonyms, and Fun! (Vocabulary in Action 2010)