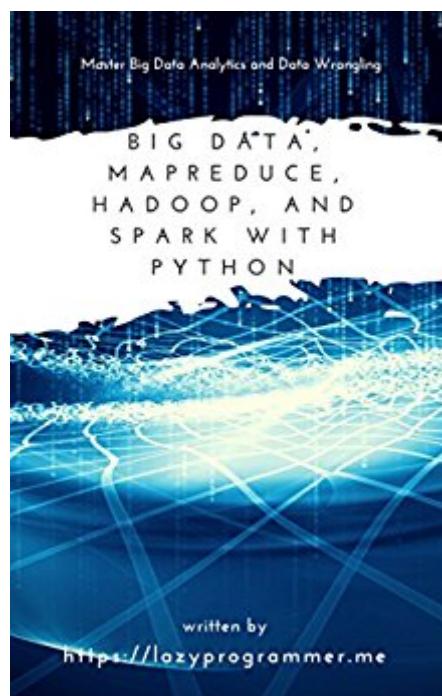


The book was found

Big Data, MapReduce, Hadoop, And Spark With Python: Master Big Data Analytics And Data Wrangling With MapReduce Fundamentals Using Hadoop, Spark, And Python



Synopsis

What's the big deal with big data? It was recently reported in the Wall Street Journal that the government is collecting so much data on its citizens that they can't even use it effectively. A few unicorns have popped up in the past decade or so, promising to help solve the big data problems that billion dollar corporations and the people running your country can't. It goes without saying that programming with frameworks that can do big data processing is a highly-coveted skill. Machine learning and artificial intelligence algorithms, which have garnered increased attention (and fear-mongering) in recent years, mainly due to the rise of deep learning, are completely dependent on data to learn. The more data the algorithm learns from, the smarter it can become. The problem is, the amount of data we collect has outpaced gains in CPU performance. Therefore, scalable methods for processing data are needed. In the early 2000s, Google invented MapReduce, a framework to systematically and methodically process big data in a scalable way by distributing the work across multiple machines. Later, the technology was adopted into an open-source framework called Hadoop, and then Spark emerged as a new big data framework which addressed some problems with MapReduce. In this book we will cover all 3 - the fundamental MapReduce paradigm, how to program with Hadoop, and how to program with Spark. Advance your Career! If Spark is a better version of MapReduce, why are we even talking about it? Good question! Corporations, being slow-moving entities, are often still using Hadoop due to historical reasons. Just search for "big data" and "Hadoop" on LinkedIn and you will see that there are a large number of high-salary openings for developers who know how to use Hadoop. In addition to giving you deeper insight into how big data processing works, learning about the fundamentals of MapReduce and Hadoop first will help you really appreciate how much easier Spark is to work with. Any startup or technical engineering team will appreciate a solid background with all of these technologies. Many will require you to know all of them, so that you can help maintain and patch their existing systems, and build newer and more efficient systems that improve the performance and robustness of the old systems. Amazingly, all the technologies we discuss in this book can be downloaded and installed for FREE. That means all you need to invest after purchasing this book is your effort and your time. The only prerequisites are that you are comfortable with Python coding and the command line shell. For the machine learning chapter you'll want to be familiar with using machine learning libraries. BONUS: At the end of this book, I'll show you a super simple way to train a deep neural network on Spark with the classic MNIST dataset. It will demonstrate how easy it is to apply deep learning to big data.

Book Information

File Size: 276 KB

Print Length: 59 pages

Simultaneous Device Usage: Unlimited

Publication Date: August 15, 2016

Sold by: Digital Services LLC

Language: English

ASIN: B01KH9YWSY

Text-to-Speech: Enabled

X-Ray: Not Enabled

Word Wise: Not Enabled

Lending: Not Enabled

Enhanced Typesetting: Enabled

Best Sellers Rank: #225,686 Paid in Kindle Store (See Top 100 Paid in Kindle Store) #179

in Kindle Store > Kindle Short Reads > 90 minutes (44-64 pages) > Computers & Technology

#179 in Books > Computers & Technology > Databases & Big Data > Data Processing #3022

in Kindle Store > Kindle eBooks > Computers & Technology

[Download to continue reading...](#)

Big Data, MapReduce, Hadoop, and Spark with Python: Master Big Data Analytics and Data Wrangling with MapReduce Fundamentals using Hadoop, Spark, and Python Data Analytics: Practical Data Analysis and Statistical Guide to Transform and Evolve Any Business Leveraging the Power of Data Analytics, Data Science, ... (Hacking Freedom and Data Driven Book 2) Sleisenger and Fordtran's Gastrointestinal and Liver Disease Review and Assessment (Sleisenger and Fordtrans Gastrointestinal and Liver) Eat Fat and Get Thin, Fit, and Healthier Than Ever Before!: Easy Diet and Delicious Cookbook: Recipes for Dramatic and Sustainable Weight Loss (Includes 21 Day Meal Plan) Best Plants for New Mexico Gardens and Landscapes: Keyed to Cities and Regions in New Mexico and Adjacent Areas, Revised and Expanded Edition Amino Acids: The Way to Health and Wellness: Find Health and Healing from Depression, Addictions, Obesity, Anxiety, Sexual Issues, and Fill Nutritional Needs of Vegetarian and Vegan Diets Die-cutting and Tooling: A guide to the manufacture and use of cutting, embossing and foiling dies, anvils and cylinders The Complete Book of Essential Oils and Aromatherapy: Over 800 Natural, Nontoxic, and Fragrant Recipes to Create Health, Beauty, and Safe Home and Work Environments Low-Dimensional and Nanostructured Materials and Devices: Properties, Synthesis, Characterization, Modelling and Applications (NanoScience and Technology) PYTHON: Python in 8 Hours, For Beginners, Learn

Python Fast! A Smart Way to Learn Python, Plain & Simple, Learn Python Programming Language in Easy Steps, A Beginner's Guide, Start Coding Today! Ecommerce Analytics: Analyze and Improve the Impact of Your Digital Strategy (FT Press Analytics) Wheat and Gluten-Free Bread, Rolls, and Biscuits: 50 Exciting Recipes using Healthful and Inspiring Ingredients WHEAT FREE The Mindfulness and Acceptance Workbook for Anxiety: A Guide to Breaking Free from Anxiety, Phobias, and Worry Using Acceptance and Commitment Therapy The Human Side of Science: Edison and Tesla, Watson and Crick, and Other Personal Stories behind Science's Big Ideas Programming and Automating Cisco Networks: A guide to network programmability and automation in the data center, campus, and WAN (Networking Technology) Riddles: Best Riddles For Kids: Short Brain Teasers, Riddle Books Free, Riddle and trick questions, Riddles, Riddles and Puzzles (Jokes and Riddles Book 2) Pure Pulp: THE SHADOW And BLACK MASK - 2 COMPLETE CLASSIC PULP MAGAZINES FROM THE 1930s AND 1940s: OVER 230 PAGES OF MYSTERY, MURDER AND INTRIGUE GoWISE Air Fryer Cookbook: 101 Easy Recipes and How To Instructions for Healthy Low Oil Air Frying and Baking (Air Fryer Recipes and How To Instructions) Healthy Super Shakes and Smoothies: 25 Sugar Free Dairy Free Shakes and Smoothies Recipes (Here's the DEAL - Healthy Weight Loss and Fat Burning Book 3) The Secrets of Baking Soda: Over 40 Recipes and Secret Tips for Cleaning, Cooking and Health on a Budget (DIY Household Hacks and Tips)

[Dmca](#)