Nataraj Dasgupta

Practical Big Data Analytics

Hands-on techniques to implement enterprise analytics and machine learning using Hadoop, Spark, NoSQL and R

Packt

Practical Big Data Analytics

Big Data analytics relates to the strategies used by organizations to collect, organize, and analyze large amounts of data to uncover valuable business insights that cannot be analyzed through traditional systems. Crafting an enterprise-scale cost-efficient Big Data and machine learning solution to uncover insights and value from your organization's data is a challenge. Today, with hundreds of new Big Data systems, machine learning packages, and BI tools, selecting the right combination of technologies is an even greater challenge. This book will help you do that.

With the help of this guide, you will be able to bridge the gap between the theoretical world of technology and the practical reality of building corporate Big Data and data science platforms. You will get hands-on exposure to Hadoop and Spark, build machine learning dashboards using R and R Shiny, create web-based apps using NoSQL databases such as MongoDB, and even learn how to write R code for neural networks.

By the end of the book, you will have a very clear and concrete understanding of what Big Data analytics means, how it drives revenues for organizations, and how you can develop your own Big Data analytics solution using the different tools and methods articulated in this book.

Things you will learn:

- Get a 360-degree view of the world of Big Data, data science, and machine learning
- Go through a broad range of technical and business Big Data analytics topics that caters to the interests of technical experts as well as corporate IT executives
- Get hands-on experience with industry-standard Big Data and machine learning tools such as Hadoop, Spark, MongoDB, kdb+, and R
- Create production-grade machine learning BI dashboards using R and R Shiny with step-by-step instructions
- Learn how to combine open-source Big Data, machine learning, and BI tools to create low-cost business analytics applications
- Understand corporate strategies for successful Big Data and data science projects
- Go beyond general-purpose analytics to develop cutting-edge Big Data applications using emerging technologies

ISBN 978-1-78355-439-3
90000
9 781783 554393

Packt>
www.packtpub.com

Practical Big Data Analytics

Nataraj Dasgupta

Practical Big Data Analytics

Hands-on techniques to implement enterprise analytics and machine learning using Hadoop, Spark, NoSQL and R

Nataraj Dasgupta

Packt>

BIRMINGHAM - MUMBAI

Practical Big Data Analytics

Commissioning Editor: Veena Pagare
Acquisition Editor: Vinay Argekar
Content Development Editor: Tejas Limkar
Technical Editor: Dinesh Chaudhary
Copy Editor: Safis Editing
Project Coordinator: Manthan Patel
Proofreader: Safis Editing
Indexer: Pratik Shirodkar
Graphics: Tania Dutta
Production Coordinator: Aparna Bhagat

First published: January 2018

Production reference: 1120118

Published by Packt Publishing Ltd.
Livery Place
35 Livery Street
Birmingham
B3 2PB, UK.

ISBN 978-1-78355-439-3

www.packtpub.com

mapt.io

Mapt is an online digital library that gives you full access to over 5,000 books and videos, as well as industry leading tools to help you plan your personal development and advance your career. For more information, please visit our website.

Why subscribe?

- Spend less time learning and more time coding with practical eBooks and Videos from over 4,000 industry professionals
- Improve your learning with Skill Plans built especially for you
- Get a free eBook or video every month
- Mapt is fully searchable
- Copy and paste, print, and bookmark content

PacktPub.com

Did you know that Packt offers eBook versions of every book published, with PDF and ePub files available? You can upgrade to the eBook version at www.PacktPub.com and as a print book customer, you are entitled to a discount on the eBook copy. Get in touch with us at service@packtpub.com for more details.

At www.PacktPub.com, you can also read a collection of free technical articles, sign up for a range of free newsletters, and receive exclusive discounts and offers on Packt books and eBooks.

Contributors

About the author

Nataraj Dasgupta is the vice president of Advanced Analytics at RxDataScience Inc. Nataraj has been in the IT industry for more than 19 years and has worked in the technical and analytics divisions of Philip Morris, IBM, UBS Investment Bank and Purdue Pharma. He led the data science division at Purdue Pharma L.P. where he developed the company's award-winning big data and machine learning platform. Prior to Purdue, at UBS, he held the role of associate director working with high frequency and algorithmic trading technologies in the Foreign Exchange trading division of the bank.

I'd like to thank my wife, Suraiya, for her caring, support, and understanding as I worked during long weekends and evening hours and to my parents, in-laws, sister and grandmother for all the support, guidance, tutelage and encouragement over the years.

I'd also like to thank Packt, especially the editors, Tejas, Dinesh, Vinay, and the team whose persistence and attention to detail has been exemplary.

About the reviewer

Giancarlo Zaccone has more than 10 years experience in managing research projects both in scientific and industrial areas. He worked as a researcher at the C.N.R, the National Research Council, where he was involved in projects on parallel numerical computing and scientific visualization.

He is a senior software engineer at a consulting company, developing and testing software systems for space and defense applications.

He holds a master's degree in physics from the Federico II of Naples and a second level postgraduate master course in scientific computing from La Sapienza of Rome.

Packt is searching for authors like you

If you're interested in becoming an author for Packt, please visit `authors.packtpub.com` and apply today. We have worked with thousands of developers and tech professionals, just like you, to help them share their insight with the global tech community. You can make a general application, apply for a specific hot topic that we are recruiting an author for, or submit your own idea.

Table of Contents