Rushdi Shams

Java Data Science Cookbook

Explore the power of MLlib, DL4j, Weka, and more

Packt

Java Data Science Cookbook

If you are looking to build data science models that are good for production, Java has come to the rescue. With the aid of strong libraries such as MLlib, Weka, DL4j, and more, you can efficiently perform all the data science tasks you need to.

This unique book provides modern recipes to solve your common and not-so-common data science-related problems. We start with recipes to help you obtain, clean, index, and search data. Then you will learn a variety of techniques to analyze, learn from, and retrieve information from data. You will also understand how to handle big data, learn deeply from data, and visualize data.

Finally, you will work through unique recipes that solve your problems while taking data science to production, writing distributed data science applications, and much more—things that will come in handy at work.

Things you will learn:

- Find out how to clean and make datasets ready so you can acquire actual insights by removing noise and outliers
- Develop the skills needed to use modern machine learning techniques to retrieve information and transform data to knowledge
- Familiarize yourself with cutting-edge techniques to store and search large volumes of data and retrieve information from large amounts of data in text format
- Develop the basic skills needed to apply big data and deep learning technologies to large volumes of data
- Evolve your data visualization skills and gain valuable insights from your data
- Learn about a step-by-step formula to help you develop an industry-standard, large-scale, real-life data product

Packt›
www.packtpub.com

$ **49.99** US
£ **41.99** UK

Prices do not include local sales
Tax or VAT where applicable

ISBN 978-1-78712-253-6
54999
9 781787 122536

Java Data Science Cookbook

Rushdi Shams

Java Data Science Cookbook

Explore the power of MLlib, DL4j, Weka, and more

Rushdi Shams

BIRMINGHAM - MUMBAI

Java Data Science Cookbook

First published: March 2017

Production reference: 1240317

Published by Packt Publishing Ltd.
Livery Place
35 Livery Street
Birmingham
B3 2PB, UK.
ISBN 978-1-78712-253-6

www.packtpub.com

Credits

Author

Rushdi Shams

Reviewer

Prashant Verma

Commissioning Editor

Veena Pagare

Acquisition Editor

Ajith Menon

Content Development Editor

Cheryl Dsa

Technical Editor

Dharmendra Yadav

Copy Editors

Vikrant Phadke

Manisha Sinha

Project Coordinator

Nidhi Joshi

Proofreader

Safis Editing

Indexer

Aishwarya Gangawane

Graphics

Tania Dutta

Production Coordinator

Arvindkumar Gupta

About the Author

Rushdi Shams has a PhD on application of machine learning in Natural Language Processing (NLP) problem areas from Western University, Canada. Before starting work as a machine learning and NLP specialist in industry, he was engaged in teaching undergrad and grad courses. He has been successfully maintaining his YouTube channel named "Learn with Rushdi" for learning computer technologies.

I would like to acknowledge the Almighty Allah for giving me the strength, support, and knowledge to finish the book.

I extend my thanks to my family members, friends, and colleagues for continuous support, encouragement, and constructive criticism.

I would also like to thank Ajith and Cheryl from Packt for their continuous and spontaneous collaboration with me.

About the Reviewer

Prashant Verma started his IT career in 2011 as a Java developer at Ericsson, working in the telecom domain. After a couple of years of Java EE experience, he moved into the big data domain, and has worked on almost all the popular big data technologies such as Hadoop, Spark, Kafka, Flume, Mongo, Cassandra, and so on. He has also worked in Scala and Python. Currently, he works with QA Infotech as Lead Data Engineer, working on solving e-learning domain problems using data analytics and machine learning.

Prashant has also worked on *Apache Spark for Java Developers*, Packt as a Technical Reviewer.

I want to thank Packt Publishing for giving me the chance to review the book, as well as my employer and my family for their patience while I was busy working on this book.

www.PacktPub.com

For support files and downloads related to your book, please visit `www.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.

`https://www.packtpub.com/mapt`

Get the most in-demand software skills with Mapt. Mapt gives you full access to all Packt books and video courses, as well as industry-leading tools to help you plan your personal development and advance your career.

Why subscribe?

- Fully searchable across every book published by Packt
- Copy and paste, print, and bookmark content
- On demand and accessible via a web browser

Customer Feedback

Thanks for purchasing this Packt book. At Packt, quality is at the heart of our editorial process. To help us improve, please leave us an honest review on this book's Amazon page at `https://www.amazon.com/dp/1787122530`.

If you'd like to join our team of regular reviewers, you can e-mail us at `customerreviews@packtpub.com`. We award our regular reviewers with free eBooks and videos in exchange for their valuable feedback. Help us be relentless in improving our products!

To my lovely wife, Mah-Zereen, and adorable daughter, Ruayda.

Table of Contents