Robert Layton

Learning Data Mining with Python

Second Edition

Use Python to manipulate data and build predictive models

Packt

Learning Data Mining with Python - Second Edition

The next step in the information age is to gain insights from the deluge of data coming our way. Data mining provides a way of finding these insights, and Python is one of the most popular languages for data mining, providing both power and flexibility in analysis.

This book teaches you how to design and develop data mining applications using a variety of datasets, starting with basic classification and affinity analysis. This book covers a large number of libraries available in Python, including the Jupyter Notebook, pandas, scikit-learn, spaCy and Tensorflow.

You will gain hands-on experience with complex data types including text, images, and graphs. You will also discover object detection using Deep Neural Networks, which is one of the big and difficult areas of machine learning right now.

With restructured examples and code samples updated for the latest edition of Python, each chapter of this book introduces you to new algorithms and techniques. By the end of the book, you will have a great insight into using Python for data mining and understanding of the algorithms as well as implementations.

Things you will learn:

- Apply data mining concepts to real-world problems
- Predict the outcome of sports matches based on past results
- Determine the author of a document based on their writing style
- Use APIs to download datasets from social media and other online services
- Find and extract good features from difficult datasets
- Create models that solve real-world problems
- Design and develop data mining applications using a variety of datasets
- Perform object detection in images using Deep Neural Networks
- Get meaningful insights from your data through intuitive visualizations
- Compute on big data, including real-time data from the Internet

Packt>
www.packtpub.com

$ **44.99** US
£ **37.99** UK

Prices do not include local sales Tax or VAT where applicable

ISBN 978-1-78712-678-7
54499
9 781787 126787

Learning Data Mining with Python - Second Edition

Robert Layton

Learning Data Mining with Python

Second Edition

Use Python to manipulate data and build predictive models

Robert Layton

BIRMINGHAM - MUMBAI

Learning Data Mining with Python

Second Edition

First published: July 2015

Second edition: April 2017

Production reference: 1250417

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

ISBN 978-1-78712-678-7

www.packtpub.com

Credits

Author

Robert Layton

Reviewer

Asad Ahamad

Commissioning Editor

Veena Pagare

Acquisition Editor

Divya Poojari

Content Development Editor

Tejas Limkar

Technical Editor

Danish Shaikh

Copy Editor

Vikrant Phadkay

Project Coordinator

Nidhi Joshi

Proofreader

Safis Editing

Indexer

Mariammal Chettiyar

Graphics

Tania Dutta

Production Coordinator

Aparna Bhagat

About the Author

Robert Layton is a data scientist investigating data-driven applications to businesses across a number of sectors. He received a PhD investigating cybercrime analytics from the Internet Commerce Security Laboratory at Federation University Australia, before moving into industry, starting his own data analytics company dataPipeline (`www.datapipeline.com.au`). Next, he created Eureaktive (`www.eureaktive.com.au`), which works with tech-based startups on developing their proof-of-concepts and early-stage prototypes. Robert also runs `www.learningtensorflow.com`, which is one of the world's premier tutorial websites for Google's TensorFlow library.

Robert is an active member of the Python community, having used Python for more than 8 years. He has presented at PyConAU for the last four years and works with Python Charmers to provide Python-based training for businesses and professionals from a wide range of organisations.

Robert can be best reached via Twitter `@robertlayton`

Thank you to my family for supporting me on this journey, thanks to all the readers of revision 1 for making it a success, and thanks to Matty for his assistance behind-the-scenes with the book.

About the Reviewer

Asad Ahamad is a data enthusiast and loves to work on data to solve challenging problems.

He did his masters in Industrial Mathematics with Computer Application from Jamia Millia Islamia, New Delhi. He admires Mathematics a lot and always tries to use it to gain maximum profit for business.

He has good experience working on data mining, machine learning and data science and worked for various multinationals in India. He mainly uses R and Python to perform data wrangling and modeling. He is fond of using open source tools for data analysis.

He is active social media user. Feel free to connect him on twitter `@asadtaj88`

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/1787126781`.

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!

Table of Contents