1. Record Nr. UNINA9910809979003321

Autore	Kumar Ashish (Analyst)
Titolo	Learning predictive analytics with Python : gain practical insights into predictive modelling by implementing predictive analytics algorithms on public datasets with Python / / Ashish Kumar ; [foreword by Pradeep Gulipalli, co-founder and head of India operations - Tiger Analytics]
Pubbl/distr/stampa	Birmingham : , : Packt Publishing, , 2016
ISBN	1-78398-327-2
Edizione	[First edition]
Descrizione fisica	1 online resource (354 pages)
Collana	Community experience distilled
Soggetti	Python (Computer program language) R (Computer program language) Decision making - Statistical methods Forecasting - Mathematical models
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Note generali	Includes index.
Nota di contenuto	Cover ; Copyright; Credits; Foreword; About the Author; Acknowledgments; About the Reviewer; www.PacktPub.com; Table of Contents; Preface; Chapter 1: Getting Started with Predictive Modelling ; Introducing predictive modelling; Scope of predictive modelling; Ensemble of statistical algorithms; Statistical tools; Historical data; Mathematical function; Business context; Knowledge matrix for predictive modelling; Task matrix for predictive modelling; Applications and examples of predictive modelling; LinkedIn's ""People also viewed"" feature; What it does?; How is it done? Correct targeting of online adsHow is it done?; Santa Cruz predictive policing; How is it done?; Determining the activity of a smartphone user using accelerometer data; How is it done?; Sport and fantasy leagues; How was it done?; Python and its packages - download and installation; Anaconda; Standalone Python; Installing a Python package; Installing pip; Installing Python packages with pip; Python and its packages for predictive modelling; IDEs for Python; Summary; Chapter 2: Data Cleaning ; Reading the data - variations and examples; Data frames; Delimiters

Various methods of importing data in PythonCase 1 - reading a dataset using the read_csv method; The read_csv method; Use cases of the read_csv method; Case 2 - reading a dataset using the open method of Python; Reading a dataset line by line; Changing the delimiter of a dataset; Case 3 - reading data from a URL; Case 4 - miscellaneous cases; Reading from an .xls or .xlsx file; Writing to a CSV or Excel file; Basics - summary, dimensions, structure; Handling missing values; Checking for missing values; What constitutes missing data?; How missing values are generated and propagated
Treating missing valuesDeletion; Imputation; Creating dummy variables; Visualizing a dataset by basic plotting; Scatter plots; Histograms; Boxplots; Summary; Chapter 3: Data Wrangling ; Subsetting a dataset; Selecting columns; Selecting rows; Selecting a combination of rows and columns; Creating new columns; Generating random numbers and their usage; Various methods for generating random numbers; Seeding a random number; Generating random numbers following probability distributions; Probability density function; Cumulative density function; Uniform distribution; Normal distribution
Using the Monte-Carlo simulation to find the value of piGeometry and mathematics behind the calculation of pi; Generating a dummy data frame; Grouping the data - aggregation, filtering, and transformation; Aggregation; Filtering; Transformation; Miscellaneous operations; Random sampling - splitting a dataset in training and testing datasets; Method 1 - using the Customer Churn Model; Method 2 - using sklearn; Method 3 - using the shuffle function; Concatenating and appending data; Merging/joining datasets; Inner Join; Left Join; Right Join; An example of the Inner Join
An example of the Left Join

Sommario/riassunto

Gain practical insights into predictive modelling by implementing Predictive Analytics algorithms on public datasets with Python About This Book A step-by-step guide to predictive modeling including lots of tips, tricks, and best practices Get to grips with the basics of Predictive Analytics with Python Learn how to use the popular predictive modeling algorithms such as Linear Regression, Decision Trees, Logistic Regression, and Clustering Who This Book Is For If you wish to learn how to implement Predictive Analytics algorithms using Python libraries, then this is the book for you. If you are familiar with coding in Python (or some other programming/statistical/scripting language) but have never used or read about Predictive Analytics algorithms, this book will also help you. The book will be beneficial to and can be read by any Data Science enthusiasts. Some familiarity with Python will be useful to get the most out of this book, but it is certainly not a prerequisite. What You Will Learn Understand the statistical and mathematical concepts behind Predictive Analytics algorithms and implement Predictive Analytics algorithms using Python libraries Analyze the result parameters arising from the implementation of Predictive Analytics algorithms Write Python modules/functions from scratch to execute segments or the whole of these algorithms Recognize and mitigate various contingencies and issues related to the implementation of Predictive Analytics algorithms Get to know various methods of importing, cleaning, sub-setting, merging, joining, concatenating, exploring, grouping, and plotting data with pandas and numpy Create dummy datasets and simple mathematical simulations using the Python numpy and pandas libraries Understand the best practices while handling datasets in Python and creating predictive models out of them In Detail Social Media and the Internet of Things have resulted in an avalanche of data. Data is powerful but not in its

raw form - It needs to be processed and modeled, and Python is one of the most robust tools out there to do so. It has an array of packages for predictive modeling and a suite of IDEs to choose from. Learning to predict who would win, lose, buy, lie, or die with Python is an indispensable skill set to have in this data age. This book is your guide to getting started with Predictive Analytics using Python. You will see how to process data and make predictive models from it. We balance both statistical and mathem...