Learn Autonomous Programming with Python

Utilize Python's capabilities in artificial intelligence, machine learning, deep learning and robotic process automation

Varun P Divadkar

www.bpbonline.com

First Edition 2024

ISBN: 978-93-55517-630

LIMITS OF LIABILITY AND DISCLAIMER OF WARRANTY

The information contained in this book is true to correct and the best of author's and publisher's knowledge. The author has made every effort to ensure the accuracy of these publications, but publisher cannot be held responsible for any loss or damage arising from any information in this book.

All trademarks referred to in the book are acknowledged as properties of their respective owners but BPB Publications cannot guarantee the accuracy of this information.

Dedicated to

Every earnest reader of this book

About the Author

In a diverse career spanning 12+ years, across six multinational organizations, **Varun P Divadkar** holds the distinction of having contributed to a variety of global techno-functional engagements with his delivery expertise in Python programming applications towards cutting edge domains like Artificial Intelligence, Machine Learning, Deep Learning and Robotic Process Automation. Being a qualified Mechanical Engineer with a gold medal from MIT has endowed him with uniquely acquired logical reasoning abilities and quantitative aptitude. This has further enabled him to grasp key mathematical and technical concepts with a steep learning curve and thereby mentor students and professionals in their journey towards exploration of these cutting edge domains.

He began his career with the prestigious Tata group within the manufacturing domain in the engineering services sector and was involved in automation of engineering designs. Thereafter, he had a stint of 5+ years with financial consulting giants Ernst & Young and Deloitte, where he had the privilege to work out of the global financial epicenter of Wall Street, Manhattan, New York and the culturally significant hub of Los Angeles, California. Within India, he has worked across major commercial centers like Mumbai, Bengaluru, Chennai, Hyderabad and Pune, thus adding to his diverse expanse of professional network. Of recent, he has been with three banking organizations and currently leads the Machine Learning and Generative AI initiative at a large global commercial bank as the Vice President.

About the Reviewer

Jyant, a passionate Data Enthusiast with a Postgraduate degree in Artificial Intelligence from Great Lakes, embarked on his IT journey with Accenture, where he honed his skills in SAP BODS, SQL, PYTHON, and Excel, catering to US-based retail clients. Later, at Impact Analytics, he delved into a diverse tech stack, including R, Python, SQL, AWS, GCP, GUROBI, and more, while refining his interpersonal skills.

During his tenure at Sapient, Jyant led teams in delivering a spectrum of products, from price optimization to movie recommendations and Credit Scoring Models, employing a versatile toolkit encompassing Python, R, Advance SQL, Superset, Airflow, Git, Azure, Pyspark, and others.

Currently, as a leader at Zscaler, Jyant spearheads the development of two pivotal models: an attrition model and an innovative chatbot powered by AWS LEX. His expertise extends to LLm, Rag, and ReAct Prompt integration, alongside utilizing SQLAlchemy to enhance the chatbot's capabilities.

Jyant excels in tackling diverse challenges, is adept at creating Regression models for promotion optimization, utilizing Seaborn and Matplotlib for extensive EDA, and applying CNN and RESNET50 for image attribute tagging. He is also experienced at employing NLP for data extraction, developing Movie Recommendation Engines with Lightfm Neural Network models, and crafting intricate Price Optimization and Credit Scoring Models using bank and retail data.

Acknowledgement

I was extremely thrilled when BPB Publications reached out to me with the opportunity to write this wonderful book. I would want to thank all the reviewers, technical experts, and editors of BPB Publications for their valuable support and guidance throughout this journey and making it a worthwhile experience.

I would want to acknowledge the generosity and consistent effort of the worldwide open source community of data science for providing a plethora of libraries and chunks of free reading material online that enables one to levergage their salient learnings and incorporate valuable inputs.

Last but not least, I would wish to thank all the interested readers who have instilled in me, the motivation and inspiration leading to the making of this book.

Preface

The contemporary industry has witnessed an unprecedented growth of technology characterized by the advent of big data, Internet of Things (IoT), Artificial Intelligence, Machine Learning, Deep Learning and Robotic Process Automation. This has led to an inevitable need for professionals to upskill themselves in order to keep themselves relevant with this moving industry trend.

This book has been designed with the sole purpose of providing a launching pad to students and professionals for breaking into these domains by leveraging the salient powers of Python.

As one progresses through this book, one shall embark on an exciting learning curve to get oneself familiarized with preliminary domain specific concepts before picking up the Python editor to write code. The initial chapter would be a refresher that demonstrates the importance of Python in the current scenario and the chapters to follow would concentrate on individual topics. Every chapter would observe a consistent sequence of an initial overview of fundamental concepts followed by corresponding Python exercises.Wherever necessary, detailed guidance has also been provided for setup and installation of required software.

Each chapter would end with an industry relevant use case in Python. These use cases are golden exercises in brushing up ones potential to add value and have been designed keeping in mind the demands of today's industry and the salient skills and abilities required to fulfil them.

This book is intended for both students and professionals who want to take their ability in Python programming to the next level by providing industry relevant value addition. This book should also be useful as an optimum handbook for quick reference of basic concepts of specific domains discussed within the book.

This book shall definitely get one enthused in the arena of Python automation and channelize ones thought process in the appropriate direction for continuing to build value added applications in Python.

Chapter 1: Why Python for Automation? – This chapter provides an overview of why Python is the chosen favorite programming language when it comes to building value added applications in popular domains like Artificial Intelligence, Machine Learning, Deep Learning. It shall discuss the inherent advantages associated with Python as an open source language and then highlight frequently used Python libraries.

Chapter 2: RPA Foundations – This chapter introduces the reader to fundamental concepts of Robotic Process Automation (RPA) and discusses the various components of RPA. Thereafter, an overview of industry leaders in RPA like UiPath, Automation Anywhere and Blueprism has been provided. The chapter shall introduce the user to the Python 'rpa' library and end with an interesting practical use case with it.

Chapter 3: Getting Started with AI/ML in Python – This chapter gets the reader kickstarted with fundamental concepts in Artificial Intelligence, Machine Learning and Deep Learning. Classical machine learning algorithms have been discussed followed by an introduction to deep learning, neural networks, types of neural networks and their applications like Natural Language Processing.

Chapter 4: Automating Web Scraping – This chapter familiarizes the reader with the process of extracting data from the web using Python libraries 'requests' and 'beautiful soup.' Data being a central component of all applications, this chapter presents a useful approach to leverage the web to access data.

Chapter 5: Automating Excel and Spreadsheets – This chapter revisits the arena of Excel automation by highlighting the central role that Python plays by speeding up the process. The libraries 'openpyxl', 'xlwings' and xlsxWriter' shall be discussed individually followed by a practical use case in Python which shall illustrate the beauty of collectively utilizing them in combination.

Chapter 6: Automating Emails and Messaging – This chapter enters the foray of email and social media messaging by demonstrating the powers of Python libraries in sending swift messages through Gmail and Whatsapp.

Chapter 7: Working with PDFs and Images – This chapter expands the capabilities of Python by showcasing the ability to read data from PDF documents. The chapter further discusses the concept of Optical Character Recognition (OCR) and illustrates how 'pytesseract' and 'OpenCV' enable one to achieve it in practice by reading data from images.

Chapter 8: Mechanizing Applications, Folders and Actions – This chapter introduces the 'os' and 'shutil' modules in Python that enable one to automate the process of reading, writing and moving files and folders and is followed by the 'PyAutoGUI' and 'PyWinAuto' libraries that are used to automate mouse operations on a computer.

Chapter 9: Intelligent Automation Part 1: Using Machine Learning – This chapter consolidates into the idea of machine learning by taking a deep dive into various machine learning algorithms and demonstrates their smooth implementation using Python

libraries. This chapter covers both supervised and unsupervised learning methods and ends with a practical use case in Python that makes the user conversant with the process of independently building machine learning Python applications.

Chapter 10: Intelligent Automation Part 2: Using Deep Learning – This chapter revisits the concept of a neural network and implements a basic neural network from scratch in Python. The concept of back propagation has been illustrated using the perceptron. Thereafter the libraries 'TensorFlow' and 'PyTorch' have been discussed followed by a discussion on Natural Language Processing (NLP). An interesting practical use case in Python concludes the chapter.

Chapter 11: Automating Business Process Workflows – This chapter begins with an introduction to the concept of orchestration which is basically workflow automation or can also be thought of as the automation of automations. Thereafter the 'luigi' and 'prefect' modules shall be discussed that enable one to achieve orchestration.

Chapter 12: Hyperautomation – This chapter discusses the novel concept of hyperautomation which is basically the application of Machine Learning, RPA and other AI tools in tandem with regular automation. This chapter shall revisit the topics of NLP and RPA from a different perspective.

Chapter 13: Python and UiPath – This chapter builds on the basic concepts of RPA tool UiPath to illustrate the salient role that Python plays in enhancing its capabilities. This chapter shall be a detailed walkthrough of the setup and environment required to achieve UiPath integration with Python.

Chapter 14: Architecting Automation Projects – This chapter gives an overview of the various components that need to be taken care of while architecting an automation project. The concept of virtual environment and the 'pip' command have been discussed in detail in this chapter.

Chapter 15: The PyScript Framework – This chapter introduces the reader to an ongoing project in Python called the 'PyScript' framework which essentially enhances the potential of an HTML script by allowing the presence of a Python code within it. This chapter shall be exciting for developers who have background in Javascript and shall provide them with a functionality to experiment with a different perspective.

Chapter 16: Test Automation in Python – This chapter concludes the book by discussing a topic that also usually comes at the end of the development cycle, which is testing. The chapter elaborates on 'Selenium', 'PyTest' and the Robot FrameworK that enables smooth automation of the testing process.

Code Bundle and Coloured Images

Please follow the link to download the
Code Bundle and the ***Coloured Images*** of the book:

https://rebrand.ly/nh3wzsj

The code bundle for the book is also hosted on GitHub at **https://github.com/bpbpublications/Learn-Autonomous-Programming-with-Python**. In case there's an update to the code, it will be updated on the existing GitHub repository.

We have code bundles from our rich catalogue of books and videos available at **https://github.com/bpbpublications**. Check them out!

Errata

We take immense pride in our work at BPB Publications and follow best practices to ensure the accuracy of our content to provide with an indulging reading experience to our subscribers. Our readers are our mirrors, and we use their inputs to reflect and improve upon human errors, if any, that may have occurred during the publishing processes involved. To let us maintain the quality and help us reach out to any readers who might be having difficulties due to any unforeseen errors, please write to us at :

errata@bpbonline.com

Your support, suggestions and feedbacks are highly appreciated by the BPB Publications' Family.

Piracy

If you come across any illegal copies of our works in any form on the internet, we would be grateful if you would provide us with the location address or website name. Please contact us at **business@bpbonline.com** with a link to the material.

If you are interested in becoming an author

If there is a topic that you have expertise in, and you are interested in either writing or contributing to a book, please visit **www.bpbonline.com**. We have worked with thousands of developers and tech professionals, just like you, to help them share their insights 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.

Reviews

Please leave a review. Once you have read and used this book, why not leave a review on the site that you purchased it from? Potential readers can then see and use your unbiased opinion to make purchase decisions. We at BPB can understand what you think about our products, and our authors can see your feedback on their book. Thank you!

For more information about BPB, please visit **www.bpbonline.com**.

Join our book's Discord space

Join the book's Discord Workspace for Latest updates, Offers, Tech happenings around the world, New Release and Sessions with the Authors:

https://discord.bpbonline.com

Table of Contents

Chapter 1
Why Python for Automation?

Introduction

This chapter introduces the reader to the significance of Python as an open-source language and further delves deeper to explain why it is the most powerful and formidable tool for autonomous programming. The chapter emphasizes Python's inherent flexibility, adaptability, and user-friendly nature, attributed to its high-level language and straightforward syntax. It also throws insights into Python's flexibility and adaptability considerations and ease of use as a high-level language with a simple syntax. Further into the chapter, Python's rich assortment of libraries like 'Pandas', 'NumPy', 'Matplotlib' would be discussed which are pivotal for machine learning. In subsequent chapters, we will delve deeper into the practical implementations of these libraries.

Structure

The chapter covers the following topics:

- Python as an open-source language
- Python's repository of extensive libraries
- Python as a high-level language

- Portability aspect of Python
- Salient advantages of Python

Objectives

At the end of this chapter, you will clearly understand why Python has been chosen as the language for autonomous programming despite the availability of numerous other alternatives from the myriad pool of programming languages.

Python as an open-source language

Congratulations on choosing the best place to begin your journey of learning autonomous programming with Python! As the book's name implies, the story's crux is built on the Python language. This chapter will tell you precisely why Python has been chosen as an optimum tool to construct the monument of autonomous programming.

Let us keep Python aside for a moment to understand the first advantage it has. Imagine you wish to create a short application that would automatically produce some Excel reports out of a few raw chunks of Excel data. You consult your friend, a professional in writing algorithms in C++, JAVA and Visual Basic. He advises you to use the API of either of these languages and go ahead with building your application. As soon as you agree with your friend, the first roadblock you stumble upon is the realization that this exercise cannot be easily performed free of cost and that you must purchase certain licenses for using the respective platforms. Now in search of hope elsewhere, you research a bit about Python and realize that Python does not have any such cost constraint!

And that is the first gigantic advantage of using Python. Python is an open-source language which means that the source code is freely available, distributable, and modifiable by users making it the most preferred language to develop and maintain code and share it with communities globally. This is exactly the reason why Python is the most preferred tool for building numerous libraries which we shall soon discuss.

Python's repository of extensive libraries

Python provides extensive libraries which are basically reusable pieces of code that data scientists use to further build their applications. The most popular libraries are listed below with some short descriptions:

- **NumPy:** This is one of the most popular libraries of Python that is used to work with large matrices. It is an abbreviation for **Numerical Python**. This library forms the core of scientific computing as well as constitutes a building block for other advanced machine learning libraries like TensorFlow and SciPy.

- **Pandas:** This library is one of the key offerings of Python. The pandas DataFrame is the fundamental unit that enables one to work with data in a tabular format. It allows one to read data from raw sources as a DataFrame, perform transformations on it using multiple inbuilt methods and export the data to various other applications.
- **Matplotlib:** Data visualization has become the need of the hour to have quick insights around raw data as well as transformed data. The library matplotlib does a great job of enabling one to derive valuable insights out of data by providing one with the ability to plot variety of charts like pie charts, histograms, and scatterplots.

Other Python libraries that are primarily used in data science, machine learning and deep learning applications have been summarized below but shall be discussed in detail in relevant chapters to follow. However, this is not an exhaustive list and exercises from actual chapters would have more libraries to discuss.

- **TensorFlow:** This is an advanced deep learning library used to perform complex computations in Mathematics and Physics.
- **Beautiful Soup:** This is a popular Python library for web scraping. Web scraping is the process of automating the data gathering process from the web. Beautiful Soup helps achieve this by enabling data pulling from HTML and XML files.
- **Scrapy:** This is an open-source Python library used for web scraping.
- **Json:** This is a Python package which provides useful tools for working with JSON (JavaScript Object Notation) objects.
- **SciPy:** SciPy is an abbreviation for Scientific Python. It is built on NumPy and is used for scientific computing.
- **Scikit-learn:** This is commonly used for implementing machine learning algorithms like regression, classification, and clustering.
- **PyPDF:** This is used for reading and transforming PDF files.
- **Openpyxl:** Python library used for Excel automation.
- **Pywhatkit:** Python library that is used to send WhatsApp messages.
- **OpenCV:** Python library used for Computer Vision applications.

Python as a high-level language

Returning to JAVA and C++, one needs to be conversant with the syntax and familiar with memory management considerations to develop applications in these languages. Such kinds of languages fall into the category of low level or middle level languages. However, Python tells a different story. The syntax is extremely simple, so much that calling it an extended version of pseudo code would not be an exaggeration!