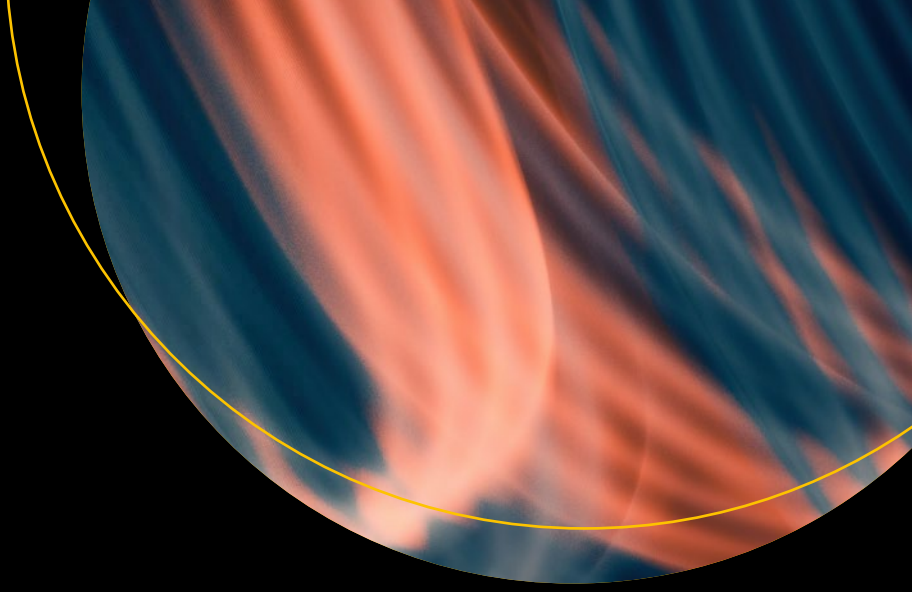

Beginner's Guide to Streamlit with Python

Build Web-Based Data and Machine Learning Applications

Sujay Raghavendra

Apress®

Beginner's Guide to Streamlit with Python

Build Web-Based Data and Machine Learning Applications

Sujay Raghavendra

Apress®

Beginner's Guide to Streamlit with Python: Build Web-Based Data and Machine Learning Applications

Sujay Raghavendra
Dharwad, Karnataka, India

ISBN-13 (pbk): 978-1-4842-8982-2 ISBN-13 (electronic): 978-1-4842-8983-9
https://doi.org/10.1007/978-1-4842-8983-9

Managing Director, Apress Media LLC: Welmoed Spahr
Acquisitions Editor: Celestin Suresh John
Development Editor: Laura Berendson
Coordinating Editor: Mark Powers
Copy Editor: Kim Wimpsett

Cover designed by eStudioCalamar

Cover image by Sunbeam Photography on Unsplash (www.unsplash.com)

Distributed to the book trade worldwide by Apress Media, LLC, 1 New York Plaza, New York, NY 10004, U.S.A. Phone 1-800-SPRINGER, fax (201) 348-4505, e-mail orders-ny@springer-sbm.com, or visit www.springeronline.com. Apress Media, LLC is a California LLC and the sole member (owner) is Springer Science + Business Media Finance Inc (SSBM Finance Inc). SSBM Finance Inc is a **Delaware** corporation.

For information on translations, please e-mail booktranslations@springernature.com; for reprint, paperback, or audio rights, please e-mail bookpermissions@springernature.com.

Apress titles may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Print and eBook Bulk Sales web page at www.apress.com/bulk-sales.

Any source code or other supplementary material referenced by the author in this book is available to readers on GitHub (https://github.com/Apress). For more detailed information, please visit www.apress.com/source-code.

Printed on acid-free paper

To
Prema Amma

Table of Contents

About the Author

Sujay Raghavendra is an IT professional with a master's degree in information technology. His research interests include machine learning, computer vision, NLP, and deep learning. He has been a consultant for multiple research centers at various universities. He has published many research articles in international journals and is the author of the book *Python Testing with Selenium*, published by Apress.

About the Technical Reviewer

Yanxian Lin is a data scientist with a PhD in biochemistry and molecular biology. He has seven years' experience with Python programming. He uses Streamlit for prototyping and demonstrating software developed in Python. Such software includes packages and machine learning models. He also uses Streamlit to develop graphical user interfaces for command-line tools. He is now an applied scientist in fraud detection.

Acknowledgments

I would like to thank my mom, Indumati Raghavendra, and brother, Sumedh Raghavendra, for their constant love, support, and encouragement.

Introduction

Streamlit is an application framework for web development based on Python. Streamlit was designed to reduce the time needed for developing web-based application prototypes for data and machine learning (ML) models. This framework helps to develop data enhanced analytics, build dynamic user experience, and showcase data for data science/ML models. The book provides insight that will help you to create high-quality applications using Python. The book uses a hands-on approach for creating such prototypes quickly.

The book starts by covering the basics of Streamlit by showing how to build a basic application and steps up incrementally, covering visualization techniques and their corresponding features. The book also covers various web application features that can be incorporated into Streamlit applications. You will also learn to handle the flow control of applications and status elements. The book also deals with performance optimization techniques necessary for data modules in an application. Last, you will learn how to deploy Streamlit applications on cloud-based platforms. You will also learn about a few prototype apps in the natural language processing (NLP) and computer vision fields that can be implemented from scratch using Streamlit.

By the end of this book, you will have a complete understanding of all the concepts, functionalities, and performance of Streamlit applications. This will enable you to be confident enough to develop your own dynamic Streamlit web applications.

What Is in the Book

The following is a brief summary of each chapter to help you better understand the structure of this book.

Chapter 1 will introduce you to the Streamlit library. It will walk you through the process of installing and running your first application.

Chapter 2 discusses text and dataframes that are represented in various forms in an application.

Chapter 3 covers visualization. Visualization is one of the important aspects of data science and machine learning. The visualizing techniques help you to understand the data more appropriately. In this chapter, we will implement the different visualizing techniques that are available in Streamlit as well as in other Python libraries for data science and machine learning developers.

Chapter 4 will discuss different media elements such as images, video, audio, etc., that can be implemented in an application.

Chapter 5 will introduce the button feature in Streamlit and how these buttons are used to select the required data to process or visualize data in applications being developed.

Chapter 6 mainly focuses on data provided by the user and how to process that data in an application. We will discuss user data in terms of forms.

Chapter 7 discusses column layouts, containers, and navigation in applications. It will also focus on how to switch between multiple pages in applications using navigation.

Chapter 8 introduces the custom handling of applications using control flow. You will also learn about status elements provided by Streamlit. Additionally, the chapter discusses how to handle huge data and optimize the performance of a Streamlit application.

Chapter 9 will introduce the development and deployment of an NLP application on the cloud-based platform Heroku.

Chapter 10 covers the development of a complete Streamlit application on a computer vision model from scratch. The features covered in the earlier chapters will be used in this application.

Who This Book Is For

This book is for professionals working in the data science and machine learning domains who want to showcase and deploy their work in a web application with no prior knowledge of web development. This book will help you learn basic Streamlit components in order to deploy your research work or prototypes. The book also guides you in developing dynamic data applications.

Source Code

The source code is available on GitHub at `https://github.com/apress/beginners-guide-streamlit-python`.

CHAPTER 1

Introduction to Streamlit

In this first chapter, we will take a look at Streamlit and its core features. We will also discuss why Streamlit is needed and compare it to alternatives. Next, we will discuss the installation steps to run Streamlit applications on various operating systems. Finally, we will develop and run our first Streamlit application.

What Is Streamlit?

Streamlit is an open-source Python framework released in October 2019. It acts as a medium between data and user interaction within an application. The goal of developing the Streamlit library was to enable data scientists and machine learning engineers to interpret and deploy data in a user interface (UI) with no prior knowledge of web development tools such as Flask, Django, Node, etc., using Python. Streamlit is the name of a software company founded by Adrien Treuille, Amanda Kelly, and Thiago Teixeira in 2018 in San Francisco, California.

The Streamlit framework uses the Python language and APIs for deploying built-in objects in the application. Thousands of companies, startups, professionals, and freelancers use Streamlit to deploy models or present data insights to their clients around the globe. We will now look into the reasons for using Streamlit.

S. Raghavendra, *Beginner's Guide to Streamlit with Python*,
https://doi.org/10.1007/978-1-4842-8983-9_1

Why Streamlit?

The Streamlit library was developed to focus on prototype applications for data-driven and machine learning models. The prototype application is easy to use with no prior knowledge of web development. You just need basic knowledge of Python and Streamlit to develop such apps. Traditional development of applications is time-consuming when compared to Streamlit-based development, and Streamlit apps are interactive in nature. An application developed using Streamlit will act as a tailwind for many companies by reducing process time and cost.

Why Streamlit for Data Science and ML Engineers?

We know that data science and machine learning use a technical medium to extract data and present this information with its associated insights to clients/stakeholders. Most of the time Python scripts or programs are developed as custom tools by data scientists and ML engineers. When they want to showcase these insights with their clients/stakeholders, it becomes extremely difficult to share the developed scripts/programs. It is also difficult to represent the data to nontechnical customers/users.

Let's assume that you are a data science or ML employee working in a company. With the client's requirements, you have been asked to develop a customized interactive tool so users can explore data along with associated visualization. When the client wants to check the application being developed, you can use the Streamlit library to develop a customized interactive application and host it on a server to share with your client/stakeholder.

Now you know the perspective of data scientists and ML engineers for developing a Streamlit application. We will now look at the major features of Streamlit in the next section.

Features of Streamlit

We will discuss a few important features that make Streamlit unique compared to other tools and applications.

Open Source

Streamlit is open source, and hence it is free to use. A large community of data scientists and ML engineers use it for developing their applications.

Platforms

Streamlit can run on any platform, which means once applications are developed, they can be modified or altered on any other system.

Ease of Development

You can develop dashboards and data- or ML-driven applications with user interaction in no time as less code is needed.

Interactive Applications

The user interacts with the application by modifying or uploading data to visualize it or apply the implemented model data. You will see the various user interactions that are possible in applications using Streamlit in upcoming chapters.

Reduced Time of Development

The development time for Streamlit applications is less than with various competitors in the domain. The autoreloading feature previews the changes made to the application, reducing the time needed.

No Core Web Development Knowledge

The Streamlit library binds the front end and back end of an application. The developer does not need any prior web development knowledge to develop data-driven applications as Streamlit takes care of it.

Easy to Learn

You need some basic knowledge of the Python programming language along with Streamlit's syntaxes that will be discussed in this book to build data-driven applications.

Model Implementation

You can implement pretrained models in an application and can analyze the model results. This type of application may be client-specific or testing-specific. The Face-GAN explorer is one such example of a model implementation.

Compatibility

Streamlit supports various Python-driven libraries for visualization, computer vision, machine learning, and data science. Table 1-1 lists some of them.