

stat4500notes

Yi Wang

2023-09-19

Table of contents

Preface	3
1 Setting up Python Computing Environment	4
1.1 on Your own computer	4
1.2 Use Google Colab	5
1.2.1 How to run a project file from your Google Drive?	5
2 Summary	7
References	8

Preface

This is a Lecture note book written for the course STAT 4500: Machine Learning offered at Auburn University at Montgomery.

This is a book wrtteen by Quarto book.

To learn more about Quarto books visit <https://quarto.org/docs/books>.

1 Setting up Python Computing Environment

1.1 on Your own computer

1. you can either `git clone` or download a zipped file containing the codes from the site: https://github.com/intro-stat-learning/ISLP_labs/tree/stable. If downloaded a zipped file of the codes, unzipped the file to a folder, for example, named `islp`. If `git clone` (preferred, you need to have Git installed on your computer, check this link for how to install Git <https://ywanglab.github.io/stat1010/git.html>), do `git clone https://github.com/intro-stat-learning/ISLP_labs.git`
2. Download and install the following software:
 - **Anaconda**: Download anaconda and install using default installation options
 - **Visual Studio Code (VSC)**: Download VSC and install
 - start VSC and install VSC extensions in VSC: Python, Jupyter, intellicode
 - (optional) **Quarto** for authoring: Download Quarto and install
2. Create a virtual environment named `islp` for Python. Start an anaconda terminal.

```
conda create -n islp python==3.10
conda activate islp
conda install pip ipykernel
pip install -r https://raw.githubusercontent.com/intro-stat-learning/ISLP_labs/v2.1.2/requirements.txt
```

3. You are ready to run the codes using VSC or `jupyter lab`.
 - Activate the venv: `conda activate islp`
 - Start a Anaconda terminal, navigate to the folder using the command `cd path/to/islp`, where `path/to/islp` means the file path to the folder `islp`, such as `\Users\ywang2\islp`. Start VSC by typing `code .` in the anaconda terminal.
 - open/create a `.ipynb` or `.py` file.
 - Select the kernel `islp`
 - Run a code cell by pressing **Shift+Enter** or click the triangular play button.

- Continue to run other cells.
- After finishing using VSC, close the VSC, and deactivate the virtual environment in a conda terminal: `conda deactivate`

1.2 Use Google Colab

All you need is a Google account. Sign in your Google account in a browser, and navigate to Google Colab. Google Colab supports both Python and R. Python is the default engine. Change the engine to R in **Connect->change runtime type**. Then you are all set. Your file will be saved to your Google Drive or you can choose to send it to your GitHub account (recommended).

1.2.1 How to run a project file from your Google Drive?

Many times, when you run a python file in Colab, it needs to access other files, such as data files in a subdirectory. In this case, it would be convenient to have the same file structure in the Google Colab user home directory. To do this, you can use Google Drive to store your project folder, and then mount the Google Drive in Colab.

Let's assume the project folder name, `islp/`. Here are the steps:

1. `git clone` the project folder (example: `git clone https://github.com/intro-stat-learning/ISLP_1`) to your local folder. This step is only needed when you want to clone some remote repo from GitHub.
2. **Upload** the folder (ex: `islp`) to Google Drive.
3. **Open the file using Colab**. In Google Drive, double click on the `ipynb` file, example, `ch06.ipynb` (or click on the three dots on the right end, and choose **open with**, then **Google Colaboratory**), the file will be opened by Google Colab.
4. **Mount the Google Drive**. In Google Colab, with the specific file (example, `ch06.ipynb`) being opened, move your cursor to the first code cell, and then click on the folder icon (this should be the fourth icon) on the upper left border in the Colab browser. This will open the file explorer pane. Typically you would see a folder named `sample_data` shown. On the top of the pane, click on the Google Drive icon to mount the Google Drive. Google Colab will insert the following code below the cursor in your opened `ipynb` file:

```
from google.colab import drive
drive.mount('/content/drive')
```

Run this code cell by pressing **SHIFT+ENTER**, and follow the prompts to complete the authentication. Wait for ~10 seconds, your Google Drive will be mounted in Colab, and it will be displayed as a folder named **drive** in the file explorer pane. You might need to click on the **Refresh** folder icon to see the folder **drive**.

5. Open a new code cell below the above code cell, and type the code

```
%cd /content/drive/MyDrive/islp/
```

This is to change the directory to the project directory on the Google Drive. Run this code cell, and you are ready to run the file **ch06.ipynb** from the folder **islp** on your personal Google Drive, just like it's on your local computer.

2 Summary

In summary, this book has no content whatsoever.

References