Tutorial 1. Getting Started

This tutorial is designed to guide you through setting up and using Anaconda and Google Colab, two powerful tools for Python programming and machine learning.

1. Getting Started with Anaconda

Anaconda is a free and open-source distribution of the Python programming language for scientific computing that aims to simplify package management and deployment.

With the installation of Anaconda, you get the following advantages:

- **Conda:** An application that manages the installation of Python packages and dependencies.
- **Python:** Hundreds of packages and libraries. More importantly, it comes with most of the libraries you will need for data analysis activities (NumPy, Pandas, Matplotlib, etc). If we discover that some packages are missing, you can install the packages and any dependencies using Conda.
- **Jupyter Notebook** : Various IDEs and an interactive Notebook environment, such as the Jupyter Notebook (also commonly referred to as Python Notebook).

Step 1. Download Anaconda

- Visit the official Anaconda website at https://www.anaconda.com/products/individual
- Scroll down to the **Anaconda Installers** section.
- Choose the installer that matches your operating system (Windows, macOS, or Linux).
- Click on the download button.

Step 2. Install Anaconda

Step 3: Find and Run Jupyter Notebook

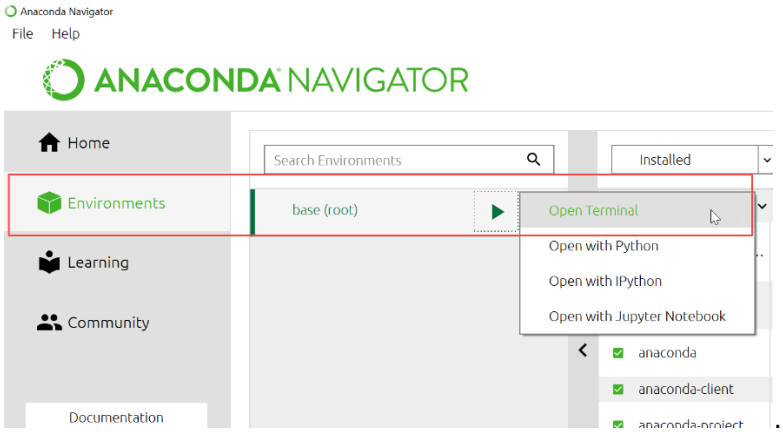
After installing Anaconda, you can open Jupyter Notebook in multiple ways:

- Open Anaconda Navigator, then click Jupyter Notebook to launch it.
- Search for "Jupyter Notebook" in your applications menu and launch it.
- Run the command jupyter notebook in the terminal (Linux/macOS) or command prompt (Windows).

This will open Jupyter Notebook in your default web browser.

Step 4. Change Jupyter Notebook Start Directory

Open the Anaconda Navigator and click on Environments -> base(root) -> Open Terminal



This will open a command prompt window. Execute the following command to generate a Jupyter Notebook configuration file:

jupyter notebook --generate-config

If the jupyter command is not recognised, it may be necessary to specify the full path to the jupyter executable or add it to your system PATH.

Note: Use a double hyphen (--) before ‘generate’ and a single hyphen (-) before ‘config’.

This will create a file named jupyter_notebook_config.py in the following location:

C:\Users\YOUR_USERNAME\.jupyter

Browse to the file location and open it in an editor. Search for the following line in the file:

c.ServerApp.notebook_dir = ''

Replace by

c.ServerApp.notebook_dir = ‘c:/your/preferred/folder/’

Remove the # at the beginning of the line to allow the line to execute.

Step 5. Run Some Python Code

Create a csv file such as:

Name	Age	City	Occupation
John Doe	28	New York	Software Engineer
Jane Smith	34	Los Angeles	Data Scientist
Emily Jones	24	Chicago	Graphic Designer
Chris Lee	40	Houston	Project Manager

Create a new Python Jupyter Notebook. In the notebook, enter the following Python code to read and display the CSV file:

```
import pandas as pd

# Load the CSV data into a DataFrame
df = pd.read_csv('path_to_your_csv_file.csv')

# Display the first few rows of the DataFrame
print(df.head())
```

2. Getting Started with Google Colab

Google Colab, short for Colaboratory, is a free cloud service hosted by Google to encourage machine learning and artificial intelligence research. It provides an environment that enables users to write and execute Python code through their web browsers without any configuration required. Google Colab offers many advantages, including:

- **No Setup**: Users can start coding in Python notebooks without worrying about setting up their computing environment.
- **Free Access to GPUs**: It provides free access to Graphics Processing Units (GPUs), making it easier to perform compute-intensive tasks such as deep learning.
- **Collaboration**: Colab supports real-time collaboration, allowing multiple users to edit notebooks and share their work with others easily.
- **Integration with Google Drive**: It integrates seamlessly with Google Drive, making it simple to save, share, and access notebooks from anywhere.
- **Jupyter Notebook Compatibility**: Colab notebooks are based on Jupyter notebooks, making them familiar to those who have used Jupyter before.

Step 1: Sign in to Google

- Ensure you have a Google account. If not, create one at https://accounts.google.com/SignUp
- Sign in to your Google account.

Step 2: Access Google Colab

- Go directly to the Google Colab website at https://colab.research.google.com.

Step 3: Create a New Notebook

- Once you're in Google Colab, you can start a new notebook by clicking on "New Notebook" in the bottom right corner.
- A new tab will open with a new notebook.

Step 4: Familiarize Yourself with the Interface

- The interface is similar to Jupyter Notebooks, if you're familiar with them.
- You can write Python code in cells and execute them by pressing Shift + Enter.
- Use the menu and toolbar to add new cells, change cell types (code or text), and access other notebook features.

Step 5: Accessing GPU and TPU Resources

- Google Colab provides free access to GPUs and TPUs for accelerated computing.
- To change the runtime type and access GPUs or TPUs:
    - Go to the "Runtime" menu.
    - Select "Change runtime type".
    - Choose "GPU" or "TPU" from the "Hardware accelerator" dropdown list.
    - Click "Save".

Step 6: Mount Google Drive

- If you want to access files from your Google Drive:
    - In a new cell, type and execute the following code:

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

    - Follow the authentication instructions to allow Colab to access your Google Drive.

Step 7. Run Some Python Code

Upload Your CSV File to Google Drive. Then, go to Google Colab and start a new notebook by clicking on 'File' > 'New notebook'. In the notebook, enter the following code snippet to read and display the CSV file.

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

import pandas as pd
```

```
# Ensure you replace the path with your actual CSV file's path
file_path = '/content/drive/My Drive/path/to/your/file.csv'
df = pd.read_csv(file_path)

# Display the first few rows of the dataframe
print(df.head())
```

3. Self Exploration

Embark on a journey of self-discovery within the realm of machine learning by seeking out and experimenting with examples from prominent platforms such as GitHub, Kaggle, and Microsoft's project repositories. Execute the sample code and observe the results.