Preparing Python Ecosystem for Machine Learning

Adel Alaeddini
(adel.alaeddini@utsa.edu),
Syed Hasib Akhter Faruqui
(syed-hasib-akhter.faruqui@my.utsa.edu)

Python is one of the popular and powerful interpreter languages that is being used for research and development of different machine learning algorithms. The primary reason for adopting Python for this course is because of it being a general purpose high level programming language that can be both used for research and production. The purpose of this document is to guide the students to setup python ecosystem for their systems.

Option 1: Python Installation

There are multiple ways to install python ecosystem for machine learning on your computer. The first step is to install Python. I prefer to use Python 3.7 (For those who prefer Python 2.7, it's community support will be discontinued soon, so I would suggest to transfer to a newer version). The installation file can be downloaded directly from https://www.python.org/downloads/ here directly (You can check the detailed installation process here).

Python 3 on Windows

To install python on windows machine,

1. Go to following link and Download the latest python 3 installer for Windows 10 (.exe setup file).
2. Double click on the downloaded file to start the installation process.
3. Make sure to check the check-box on python setup window "Add Python 3 to PATH" before clicking the Install button
4. Once the installation is finished, to verify the installation; open windows command prompt and type-

   ```
   python --version
   ```

   the output should look like the following

   ```
   python 3.7.2
   ```

Python 3 on MAC

1. Open the terminal on your OS terminal emulator and run the following to install Homebrew (This acts as a missing package manager for macOS (or Linux)).

   ```
   $ ruby -e "$(curl -fsSL
   https://raw.githubusercontent.com/Homebrew/install/master/install)"
   ```

2. Input the Homebrew directory at the top of your `PATH` environment variable. This can be done by adding the following line at the bottom of your `~/.profile` file.

   ```
   export PATH="/usr/local/opt/python/libexec/bin:$PATH"
   ```

3. Now to install Python 3 type the following-

   ```
   $ brew install python
   ```

4. Once the installation is finished, to verify the installation; open windows command prompt and type-

   ```
   python --version
   ```

   the output should look like the following

   ```
   python 3.7.2
   ```

Python 3 on Linux

1. Open the command prompt of the Linux you are using.
2. Type in the following command lines (This code was tested on Ubuntu)-

   ```
   $ sudo apt-get update
   $ sudo apt-get install python3.7
   ```

3. Once the installation is finished, to verify the installation; open windows command prompt and type-

   ```
   python --version
   ```

   the output should look like the following

   ```
   python 3.7.2
   ```

A detailed guide to installation for all the systems can be found here.

Installing Necessary Libraries and Tools

We will mostly be using the following libraries to perform different mathematical operations and visualizations.

1. Numpy
2. scipy
3. matplotlib

4. pandas

These libraries can be installed directly by typing the following on command prompt (if on windows please, run it in administrator mode) or terminal.

```
python3 -m pip install numpy
python3 -m pip install scipy
python3 -m pip install matplotlib
python3 -m pip install pandas
```

We will also be needing Jupyter Notebook. To install this, type the following on command prompt / terminal.

```
python3 -m pip install --upgrade pip
python3 -m pip install jupyter
```

Option 2: Setting up the Ecosystem through Anaconda

An easier way to setup a machine learning ecosystem (this is more popular now a days!) is setting up the distribution called Anaconda. Its easier to setup. It comes with all the necessary libraries and tools (e.g. Numpy, scipy, Jupyter Notebook etc.) Just download the latest distribution file from here for your respective OS (Windows/MAC/Linux). For Windows/MAC you can install it directly from the .exe/.pkg file (Make sure to check the check-box on python setup window "Add Python 3 to PATH"). As for Linux, first download the appropriate setup file (e.g. Anaconda3-2019.03-Linux-x86_64.sh) from the link mentioned above. Enter the following on terminal to install Anaconda for Python 3.7 (assuming you downloaded the file on downloads folder).

```
bash ~/Downloads/Anaconda3-2019.03-Linux-x86_64.sh
```

Follow the prompts on the terminal and complete the installation.

A detailed installation instruction can be found here.

Option 3: Utilizing the Free Online Ecosystems

Recently Google™has opened a project called Google Colab. This allows you to run your Jupyter Notebooks stored in Google Drive. If you don't want to setup any Ecosystem on your personal computer but rather prefar to execute your codes on cloud, I suggest you to utilize this system. Google Colab can be accessed through this link: https://colab.research.google.com/notebooks/welcome.ipynb#scrollTo=xitplqMNk_Hc. Make sure you are logged in to your Google account to use the capabilities of Google Colab. Microsoft also provides similar cloud ecosystem. Please, check Microsoft Azure for more information.