DS 620 Machine Learning and Deep Learning
HOS01Aa: Setting up the Development Environment

12/21/2020 Developed by Shanshan Yu
03/19/2021 Modified by Minh Nguyen
03/31/2022 Reviewed by Amanda Vaughan
10/03/2022 Reviewed by Alekhya Malla
School of Technology and Computing (STC) @ City University of Seattle (CityU)

Before You Start

- The directory path shown in screenshots may be different from yours.
- The document uses Google Colaboratory as the default compiler. If you want to run the code on a local machine, you need to configure the environment on your own.
- Some steps are not explained in the tutorial**.** If you are not sure what to do:
    1. Consult the resources listed below.
    2. If you cannot solve the problem after a few tries, ask a TA for help.

Learning Outcomes

Students will be able to:

- Setup sklearn on your local machine
- Learn how to use Colab

Resources

- *Installing the development version of scikit-learn*. (n.d.). Scikit-Learn. https://scikit-learn.org/stable/developers/advanced_installation.html#install-bleeding-edge
- *scikit-learn: Machine learning in python*. (n.d.). Scikit-Learn. https://scikit-learn.org/stable/index.html

Pre-requirements for Installation

- Python 3.5 – 3.8
- pip version >19
- macOS 10.12.5 (Sierra) or later
- Windows 7 or later (with C++ redistributables)

Installation

Open Visual Studio Code → Terminal → New Terminal

In the terminal window, type the command:

```
pip install sklearn
```

```
(base) truongminh@Minhs-MacBook-Pro ~ % pip install sklearn
Collecting sklearn
  Downloading sklearn-0.0.tar.gz (1.1 kB)
```

Note: If you meet an error that the system can't detect pip as a command.

```
C:\Users\bcghost pc>pip install tensorflow
'pip' is not recognized as an internal or external command,
operable program or batch file.
```

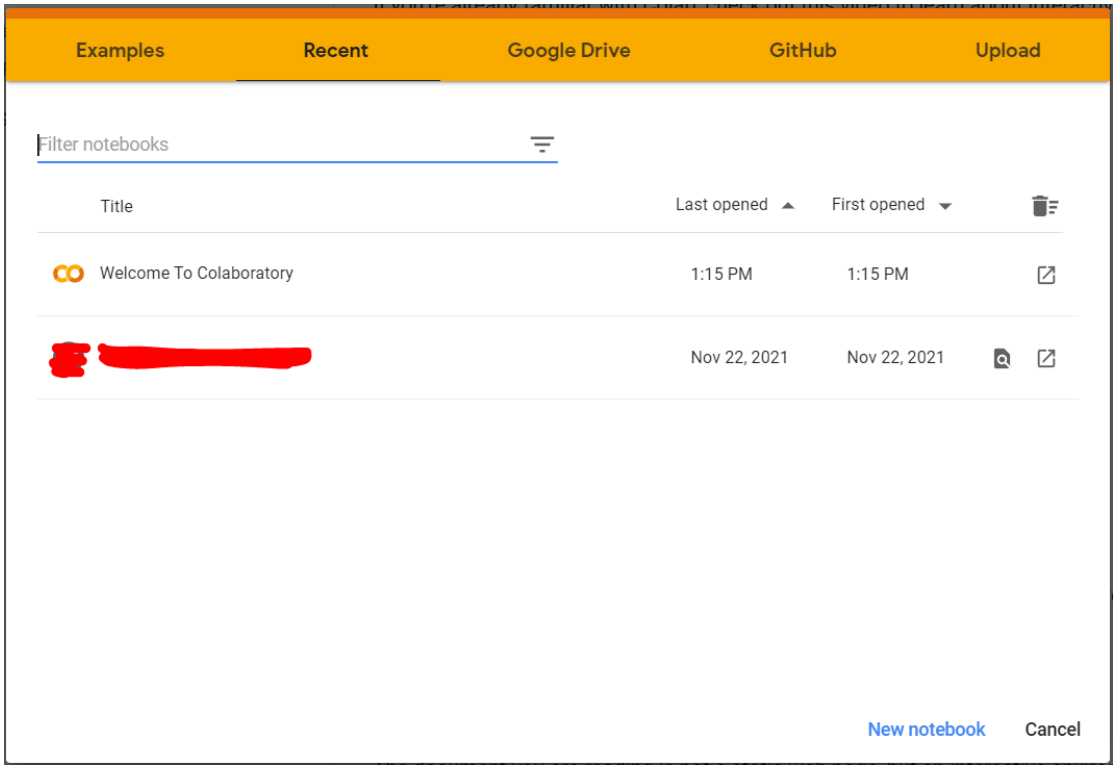
It might be because you have different versions of Python on your computer, or your pip isn't installed in your python. You also can type this command to install scikit-learn:
`python -m pip install sklearn`

If you still cannot install TensorFlow on your local computer, you can go to the next part: "Get started with Colaboratory" and use the tool to finish your assignment. Colab is a google cloud platform notebook and it doesn't require any installation or configuration.

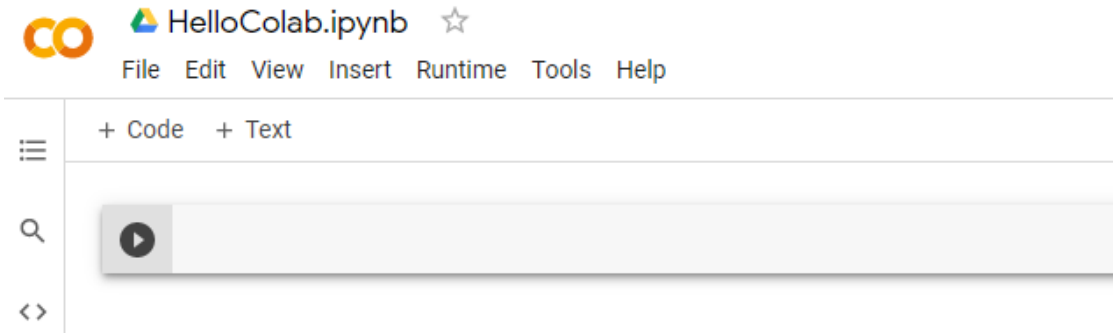
Get started with Colaboratory

Colaboratory (Colab) is a Jupyter notebook-like notebook on Google Cloud. It allows you to write and execute Python in your browser without setting up an environment. Colab provides free access to GPUs and TPUs, which can hugely accelerate the training speed. Training machine learning models on a GPU or a TPU is 5-10 times faster than training them on the CPU. If you are using an old computer with limited memory and computing power, Colab can accelerate your work.

1. Go to https://colab.research.google.com. You need a google account to create a new notebook.

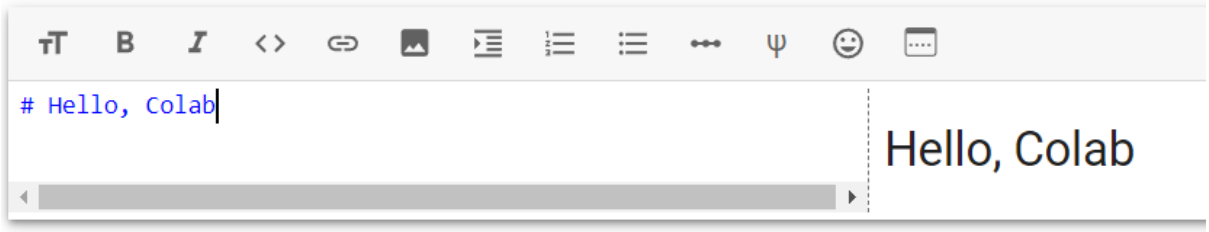


2. Click New Notebook in the bottom right corner. It will generate a blank ipynb file in your Google drive. Double click on the title (it should be something like Untitled0.ipynb) and change the notebook title to **HelloColab.ipynb**:

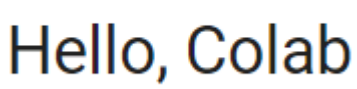


3. Type the following in the first cell then click the play button  or Ctrl + Enter to run the cell.

4. Add a text cell by clicking the + text button or use the keyboard command “b” (but make sure your “esc” the cell first). These buttons will generate a text cell in the notebook, and a text cell is as same as a text editor. Type the following in the text editor and then click outside of it into the original code cell.

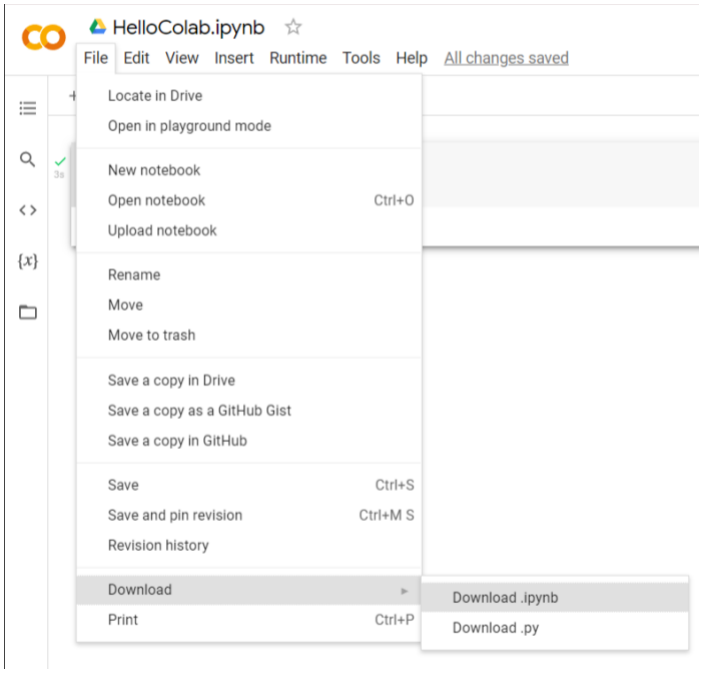


After you click into a different code cell, the text should look as follows:



5. Let’s download the ipynb file to your local computer. Make sure to save your file and then click File → Download → Download .ipynb.

The file will be downloaded to your Download folder (or another folder according to your browser setting).



6. Move the downloaded HelloColab.ipynb file to your Module01 folder in the HOS01A_YourUserName Folder.

If you want to know more details about Colab, watch this video.

Python Review

Use the following resources to practice your Python skills (optional):

- Beazley, D. (2016, August). *Python programming language* [Video]. O'Reilly Online Learning. https://www.oreilly.com/videos/python-programming-language/9780134217314/ (6hr 27m)
- Morris, C. (n.d.). *Python*. Kaggle. https://www.kaggle.com/learn/python
- Taulien, K. (2020, July). *Python for everybody: The ultimate python 3 bootcamp* [Video]. O'Reilly Online Learning. https://www.oreilly.com/videos/python-for-everybody/9781800562196/ (9h 20m)