COMP5318 Week 1 – Jupyter Notebook

Welcome to COMP5318: Machine Learning and Data Mining!

In this course during the tutorials we will use Python with Jupyter Notebook. Alternatively, you can also use the Colaboratory environment: https://colab.research.google.com/notebooks/intro.ipynb# .

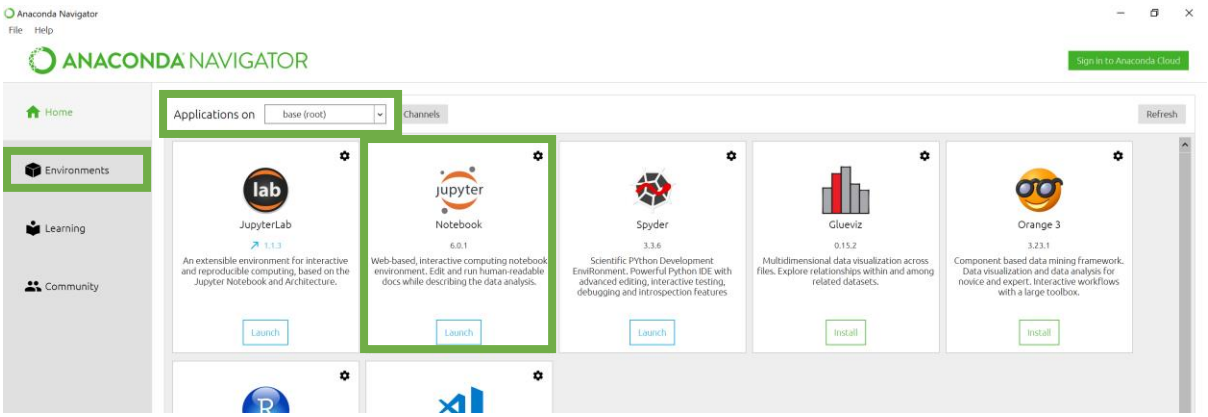
Juptyer Notebook is an application allowing you to display the input and output of a Python script. This document provides instructions how to install it on your computer, together with other packages that we will use in this course.

1. Anaconda

The recommended way to access Juptyer Notebook is installing Anaconda, which is available on Windows, Mac OS and Linux:

https://jupyter.readthedocs.io/en/latest/install.html

https://docs.anaconda.com/anaconda/install/



The default path for Anaconda is:

- **Windows 10:** C:\Users\<your-username>\Anaconda3\
- **macOS:** /Users/<your-username>/anaconda3 for the shell install, ~/opt for the graphical install. See installing on macOS.
- **Linux:** /home/<your-username>/anaconda3

Other useful questions for installing Anaconda can be found here:

https://docs.anaconda.com/anaconda/user-guide/faq/

After installing Anaconda, Jupyter Notebook can be launched from the home tab as above. Notebook will be launched on the specific environment above: base (root) in this example. This allows specific environments to be selected for Python if required for any assignments. To edit or create a new environment, select the environment tab. You can create or clone any environment for use with Jupyter.

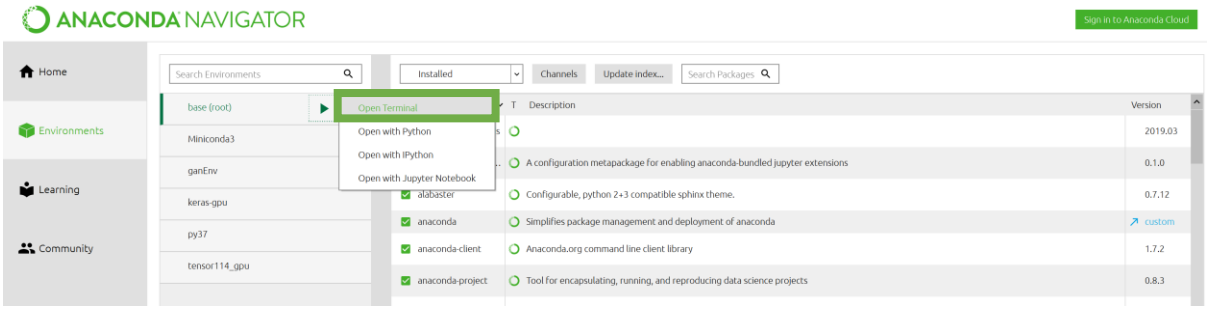
Installing packages in Anaconda

There will be several new packages that will need to be installed during this course. One such package is graphviz (https://www.graphviz.org/), an open source visualisation software, which we will use in week 5. To install new packages (e.g. graphviz):

1. Right click the environment and select “Open terminal”.
2. Search for the required package either with pip or on conda-forge (https://anaconda.org/anaconda/graphviz)
3. Run the command: `conda install -c conda-forge python-graphviz`
4. The package will then download and install itself and any requirements to the current environment

For other examples on managing the Anaconda environment, or more specific details, refer to the Anaconda documentation.

(https://docs.conda.io/projects/conda/en/latest/user-guide/tasks/manage-environments.html)



Notes:

- Be sure to install **python-graphviz** when using the conda command, not **graphviz**, which will not work for the Python environment.
- To check the current path for Anaconda, check the terminal command **echo %PATH%**. **where Python** and **where conda** may be useful in checking where Python and Anaconda is installed on your computer.
- To install pip on Anaconda, the following command should work:

`conda install -c anaconda pip`

2. Jupyter Notebook

Once you launch Jupyter Notebook, create a new Python Notebook, or select a previous Notebook file.

For a quick review of the Notebook interface, go to **Help->User Interface Tour** in the Notebook script or for a more detailed review, refer to the jupyter documentation or examples:

https://jupyter-Notebook.readthedocs.io/en/stable/examples/Notebook/examples_index.html

Testing if graphviz was installed correctly

Here is an example of simple scripts of graphviz, where the Notebook runs live code and can display the results below its cell. To import modules like graphviz, you can either **import graphviz**, or selectively import the packages you need with **from graphviz import Digraph**. Run these examples in your Notebook to test that graphviz was installed correctly.

For other examples of graphviz, refer to https://graphviz.readthedocs.io/en/stable/examples.html .