Installing Python and Jupyter Notebook

All module material will be written in Python 3.

1 I already have Python without Anaconda

You are recommended to use the **Anaconda** Python distribution. If you do not have Anaconda installed, please see Section 2 for how to install it. If you already have Anaconda installed, please go to Section 3 to create a virtual environment with all the required packages for this module.

2 I am new to Python

2.1 Download Anaconda

Anaconda is a Python distribution for data science and machine learning. It can be freely downloaded here.

2.2 Install Anaconda

Follow these instructions to install Anaconda distribution on your computer. The instructions support Windows, Mac and Linux.

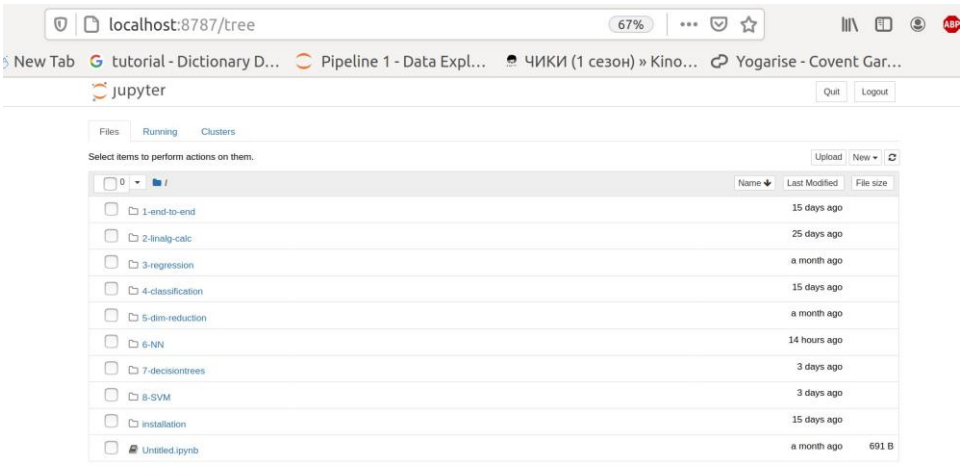

Figure 1: Jupyter notebook in browser

3 Create virtual environment and open notebook

3.1 Open terminal

Anaconda should now be successfully installed on your local machine. Please open a new terminal window with one of the commands below:

Windows: go to the start menu, search for anaconda prompt, and open it.
Mac: press [command]+[space] to open Spotlight. Then type terminal and press [enter].
Linux: in Unity/Gnome press [Ctrl]+[Alt]+[T].

3.2 Create environment

A virtual environment is a named, isolated, working copy of Python that maintains its own files, directories, and paths so that you can work with specific versions of libraries or Python itself without affecting other Python projects.

To create a new virtual environment that contains all packages needed for this module and a stable Python version please run the following two commands (replace *name* with your preferred name):

- conda create --name *name* python=3.7.3 numpy pandas scipy scikit-learn tensorflow keras matplotlib seaborn notebook ipykernel

To activate your newly created virtual environment, run the following command (activation needs to be done every time you want to work in Jupyter):

- conda activate *name*

3.3 Navigate to the folder containing your notebook

Copy the directory containing your notebook:

- For MAC, use the following way: https://support.apple.com/en-gb/guide/mac-help/mchlp1774/mac#:~:text=On%20your%20Mac%2C%20click%20the,bottom%20of%20the%20Finder%20window.
- For Windows, right-click on the folder containing the dataset and click on "copy as path"
- For Linux, use the following way: https://ubuntuhandbook.org/index.php/2020/05/textual-path-ubuntu-20-04-files/.

Navigate to the directory: copy and paste the "/path/to/folder" from the previous step and:

- cd "/path/to/folder"

3.4 Start Jupyter Notebook

Open Jupyter Notebook using the following terminal command:

- Jupyter Notebook

This should open a browser and a webpage that looks similar to the illustration in Figure 1. (It is possible that the webpage shown on your local machine looks slightly different).

Click on the notebook you want to run and start working.

4 Check package versions

You can check the versions of the installed packages by opening up a notebook, importing all the packages and checking the __version__ attribute.

- import numpy as np
- np.__version__

Below you can see the versions for each package that guarantee the notebooks run smoothly (your packages may be different versions):

- numpy 1.19.1
- pandas 1.1.1
- scipy 1.5.2
- sklearn 0.23.2
- tensorflow 2.2.0
- keras 2.4.3
- matplotlib 3.3.1
- seaborn 0.11.0