

Creating a Machine Learning Environment on your local machine

The following instructions will enable you to install a Machine Learning Environment suitable for this course on your local machine.

- We strongly suggest that you don't deviate from the instructions
- The more you stray from the suggestions, the more likely it will be that you encounter problems

Install the basic Machine Learning Environment

We will be installing Anaconda Individual Edition, a complete environment for Machine Learning.

It works on a variety of operating systems.

Click on the link to begin the installation on your local machine.

<https://www.anaconda.com/products/individual#Downloads>
(<https://www.anaconda.com/products/individual#Downloads>)

- Easiest choice: Choose the "64 bit graphical installer"
- If prompted for any choices during installation
 - Accept the default choice, by simply pressing the Enter key
 - Allow the `.bashrc` file to be modified

Potential issue installing on a Mac with the ARM chip (M1/M2)

When installing via the graphical installer: the user should be given the option to create a link on their desktop to start up the Anaconda Navigator.

- Navigator is a convenience tool that presents buttons for starting common tools, e.g., Jupyter
- It is a convenience, not a necessity

At the time of writing, the "M1" graphical installer for the Mac ARM architecture is reported to **not** install Navigator.

The description and solution to the issue is given [here](https://docs.anaconda.com/anaconda/install/mac-os/) (<https://docs.anaconda.com/anaconda/install/mac-os/>).

Install Jupyter extensions (optional)

Although this step is not strictly necessary, installing the following extensions to Jupyter will make you more productive.

Enter the following commands from a terminal:

```
1 # Install contributed extensions
2 conda install -y -c conda-forge jupyter_contrib_nbextensions
3
4 # Install Javascript and CSS files
5 jupyter contrib nbextension install --user
6
7 # Activate extensions
8 jupyter nbextension enable toc2/main
9 jupyter nbextension enable collapsible_headings/main
10 jupyter nbextension enable livemdpreview/livemdpreview
11
12 # Extension for slides
13 conda install -y -c conda-forge rise
14 jupyter nbextension enable rise
15
```

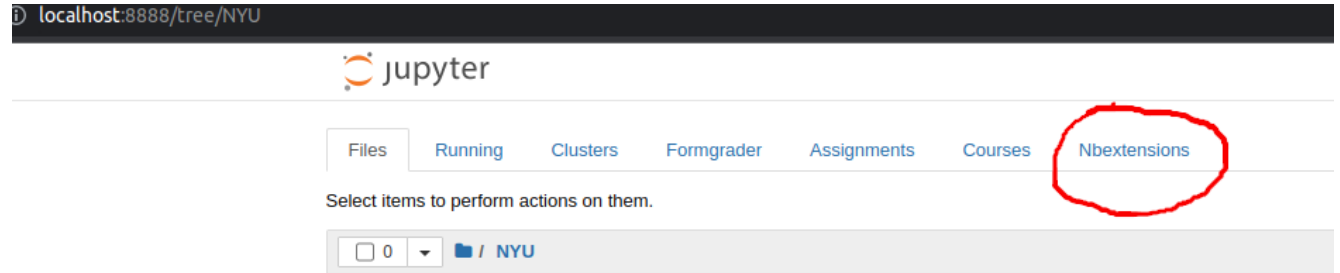
Potential issues in installing Jupyter extensions

Some users have reported difficulties in installing the extensions in new releases of Anaconda.

- Symptom: some of the above extensions are not active

The following steps have been reported as partial solutions.

- When you start Jupyter, you will see the control panel
 - the "Nbextensions" tab should be visible along the top banner



- Clicking on the Nbextensions tab will take you to a screen where you can manually activate/deactivate extensions

Configurable nbextensions

☒ disable configuration for nbextensions without explicit compatibility (they may break your notebook environment, but can be useful to show for nbextension development)

filter: by description, section, or tags			
<ul style="list-style-type: none"> ⌵ (some) LaTeX environments for Jupyter ⌵ AutoSaveTime ⌵ Code prettify ⓘ Collapsible Headings ⌵ Equation Auto Numbering ⌵ Exercise2 ⌵ Help panel ⌵ Highlight selected word ⌵ isort formatter ⌵ Launch QTConsole ⌵ Move selected cells ⌵ nbTranslate <input checked="" type="checkbox"/> RISE ⌵ Ruler in Editor ⌵ Select CodeMirror Keymap ⌵ Snippets Menu ⌵ table_beautifier ⌵ Variable Inspector 	<ul style="list-style-type: none"> ⌵ 2to3 Converter ⌵ Autoscroll ⌵ Codefolding ⌵ Comment/Uncomment Hotkey ⌵ ExecuteTime ⌵ Export Embedded HTML ⌵ Hide Header ⌵ highlighter ⓘ jupyter-js-widgets/extension ⌵ Limit Output ⌵ Navigation-Hotkeys ⌵ Notify ⓘ rise ⌵ Runtools ⌵ SKILL Syntax ⌵ spellchecker ⓘ toc2--sys-prefix ⌵ zenmode 	<ul style="list-style-type: none"> ⌵ AddBefore ⌵ Cell Filter ⌵ Codefolding in Editor ⓘ contrib_nbextensions_help_item ⌵ Execution Dependencies ⌵ Freeze ⌵ Hide input ⌵ Hinterland ⓘ jupyterlab-plotly/extension ⓘ Live Markdown Preview ⓘ Nbextensions dashboard tab ⌵ Printview ⌵ Rubberband ⌵ Scratchpad ⌵ Skip-Traceback ⌵ Split Cells Notebook ⌵ Toggle all line numbers 	<ul style="list-style-type: none"> ⌵ Autopep8 ⌵ Code Font Size ⌵ CodeMirror mode extensions ⌵ datestamper ⌵ Exercise ⌵ Gist-it ⌵ Hide input all ⌵ Initialization cells ⌵ Keyboard shortcut editor ⌵ Load TeX macros ⓘ Nbextensions edit menu item ⌵ Python Markdown ⌵ Ruler ⌵ ScrollDown ⌵ Snippets ⓘ Table of Contents (2) ⌵ Tree Filter

- Make sure to uncheck the checkbox: disable configuration for nbextensions without explicit compatibility ..
- You will then be able to manually activate/deactivate the extensions we tried to install
 - Collapsible headings
 - RISE
 - Table of Contents

Jupyter security (optional)

If you run the Jupyter server on a machine that is connected to a network

- It is possible to allow remote users to connect to your server
- This **will not** happen by default

Should you decide to let remote users connect, it is strongly advised that you protect the notebook server with a password.

See [Public Jupyter Server \(https://jupyter-notebook.readthedocs.io/en/stable/public_server.html\)](https://jupyter-notebook.readthedocs.io/en/stable/public_server.html) for details.

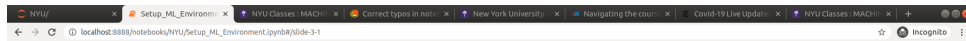
Create a directory for your notebooks

We will refer to this directory as the *notebooks directory*.

- It can be anywhere on your machine
- My convention is to have a directory called "Notebooks" in my home directory.
 - The instructions below will use `Notebooks` as the notebooks directory; modify the instructions if you choose a different location

This is the location to which the course materials will be downloaded.

Enter the following commands from a terminal:



```
1 # Create directory for notebooks
2 cd
3 mkdir Notebooks
4 cd Notebooks
```



Download the course material

We are using a tool called `git` to manage our code.

`git` defines the notion of a *repository* as a collection of material related to a project.

The following will

- Install `git`
- Download the course material from a repository we have created
 - Into the current directory
 - Which is the `Notebooks` directory if you followed the previous step
- Download material from the recommend text books

Enter the following commands from a terminal:

```
1 # Install git
2 conda install -y git
3
4 # Clone git repositories
5 REPO="ML_Spring_2023"
6
7 git clone https://github.com/kenperry-public/$REPO.git
8
9 # Change to the course directory and clone the repos for the external references
10 cd $REPO
11 mkdir external
12 cd external
13
14 git clone https://github.com/jakevdp/PythonDataScienceHandbook.git
15 git clone https://github.com/ageron/handson-ml2.git
16
```

You should verify that your notebooks directory has the following subdirectories

- ML_Spring_2023
 - The course material
 - This is where you will find the lectures
- external subdirectory containing
 - handson-ml
 - Notebooks associated with the recommended textbook by Geron.
 - PythonDataScienceHandbook
 - The *complete text* and associated notebooks of the recommended textbook by VanderPlas

Refreshing your local copy of the repo

In advance of each week's lecture, new material will be published to the repo.

Additionally: corrections might be made for previously published material.

You can always synchronize your local copy with the main repo via the command:

```
git pull
```

If you have made changes to your local repo that conflict with the main repo

- `git pull` will result in an error

Should this occur, you should use

```
git stash  
git pull
```

The `git stash` will "hide" the changes you made and restore your local copy to be consistent with the main repo.

Fear not ! Your changes are hidden, not lost, and you will be able to restore them.

- Please see the `git stash` documentation.

Deep Learning Course: Install additional packages for Deep Learning

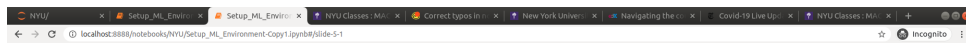
You have installed the "basic" Machine Learning Environment, which suffices for the Classical Machine Learning part of the course.

When you are ready for the Deep Learning part of the course (or want to play with Neural Networks/Deep Learning) you will need to install additional packages.

The installation method varies depending on whether or not you are using an Apple machine with the new ARM based architecture (e.g., M1 or M2 chip).

Instructions if you are ARE NOT using an Apple ARM based machine

Enter the following commands from a terminal:



```
1 # Install Tensorflow
2 conda install -y tensorflow
3
```



Instructions if you ARE using an Apple ARM based machine

See the instructions [here \(Installing TensorFlow on MAC M1 or M2.ipynb\)](#).

Tools for visualization of graphs (optional)

Some of our notebooks uses a particular package (GraphViz) to visualize graphs

- Decision trees
- Deep Learning models (using the Keras `plot_model` command)

If you want to use this, you must install additional packages as follows

```
conda install pydot  
conda install python-graphviz
```

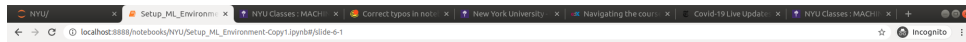
As always: if you have multiple Conda environments, be sure you have activated the desired environment before issuing any `conda install` command.

Start up a Jupyter Notebook server

As part of the installation of Anaconda, a tool called the Anaconda Navigator was installed.

Find and start it

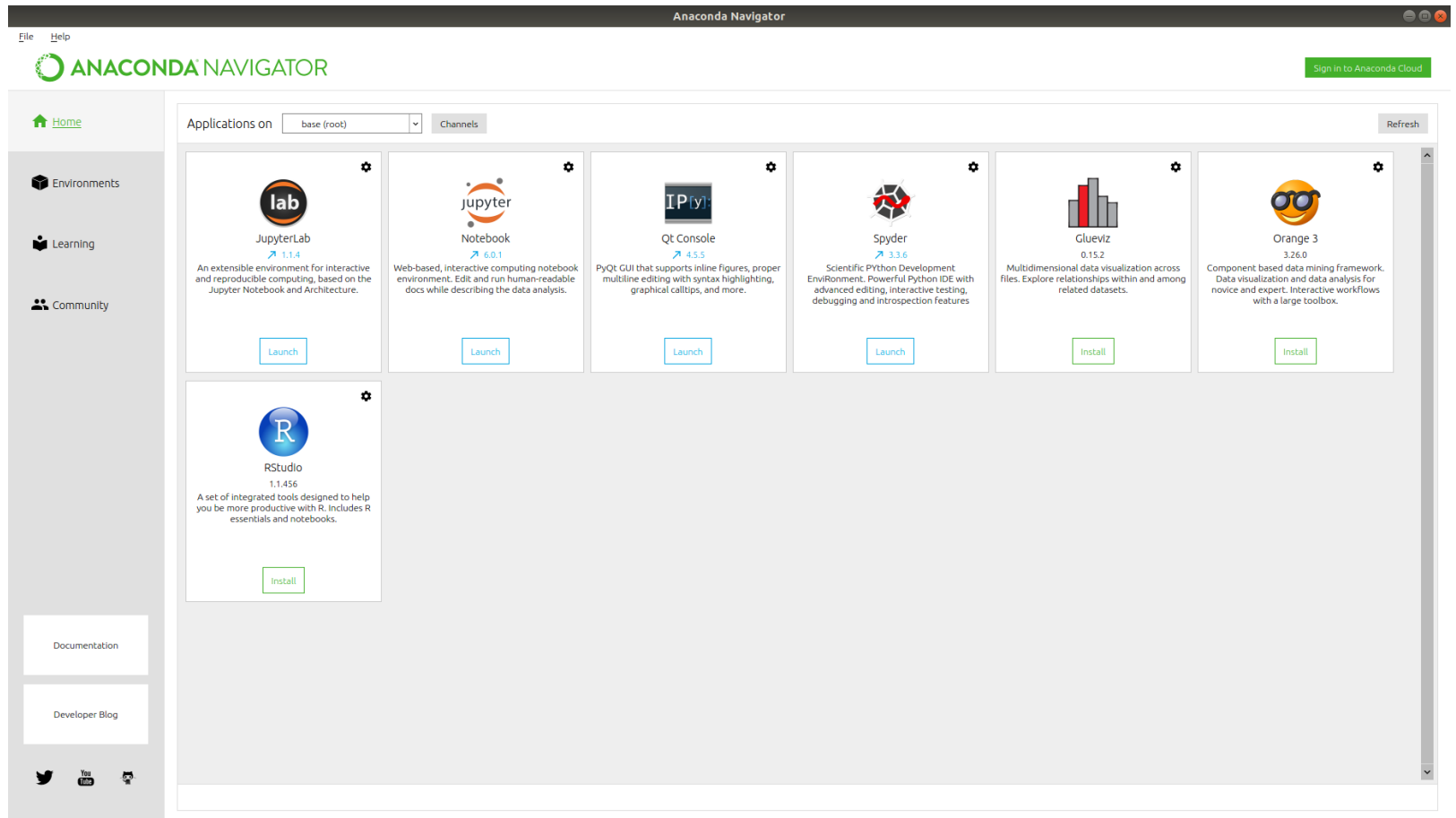
- An icon may have been created on your desktop
- Otherwise, enter the following commands from a terminal:



```
1 cd
2 ~/anaconda3/bin/anaconda-navigator &
```



You will hopefully see something like this



- Click on the icon for Jupyter Notebook to start the Jupyter server.
- Navigate to the ML_Fall_2022 directory to find the course material directory
- Find the Index.ipynb notebook and open it
 - This will take you to the "start page" for the lectures

