

Software Environment Setup

Workshop on Delta Flow-Salinity Modeling Using Machine Learning
Pre-Workshop Setup Meeting, January 25, 2023

Brad Tom, Raymond Hoang, and Peyman Namadi
DWR Delta Modeling Section



CALIFORNIA DEPARTMENT OF
WATER RESOURCES



UCDAVIS
UNIVERSITY OF CALIFORNIA

Part A: Local Computer



CALIFORNIA DEPARTMENT OF
WATER RESOURCES



UC DAVIS
UNIVERSITY OF CALIFORNIA

Prerequisites

- Laptop or Desktop with Windows*
- Administrative privileges

* The ML code will also run on Mac OS, Linux, and Google Colab



CALIFORNIA DEPARTMENT OF
WATER RESOURCES



UCDAVIS
UNIVERSITY OF CALIFORNIA

Overview

- Install Miniconda¹
- Create a conda environment²
- Download ANN code and input data files
- Run the Jupyter Notebook³ ANN code

1. Anaconda is a free, open-source platform that allows you to write and execute code in the Python programming language. Miniconda is a lightweight version of Anaconda. We recommend Miniconda because installation and setup take less time.

2. A conda environment is a folder containing a version of Python, and a specific set of python packages

3. Jupyter Notebook is a web application for creating and sharing computational documents



Install Miniconda/Anaconda

- Miniconda:
<https://docs.conda.io/en/latest/miniconda.html>
- Anaconda:
<https://www.anaconda.com/products/distribution>



Downloading the Miniconda Installer

conda
latest

Search docs

Conda

Conda-build

Miniconda

System requirements

Latest Miniconda Installer Links

Windows installers

macOS installers

Linux installers

Installing

Other resources

Help and support

Contributing

Conda license

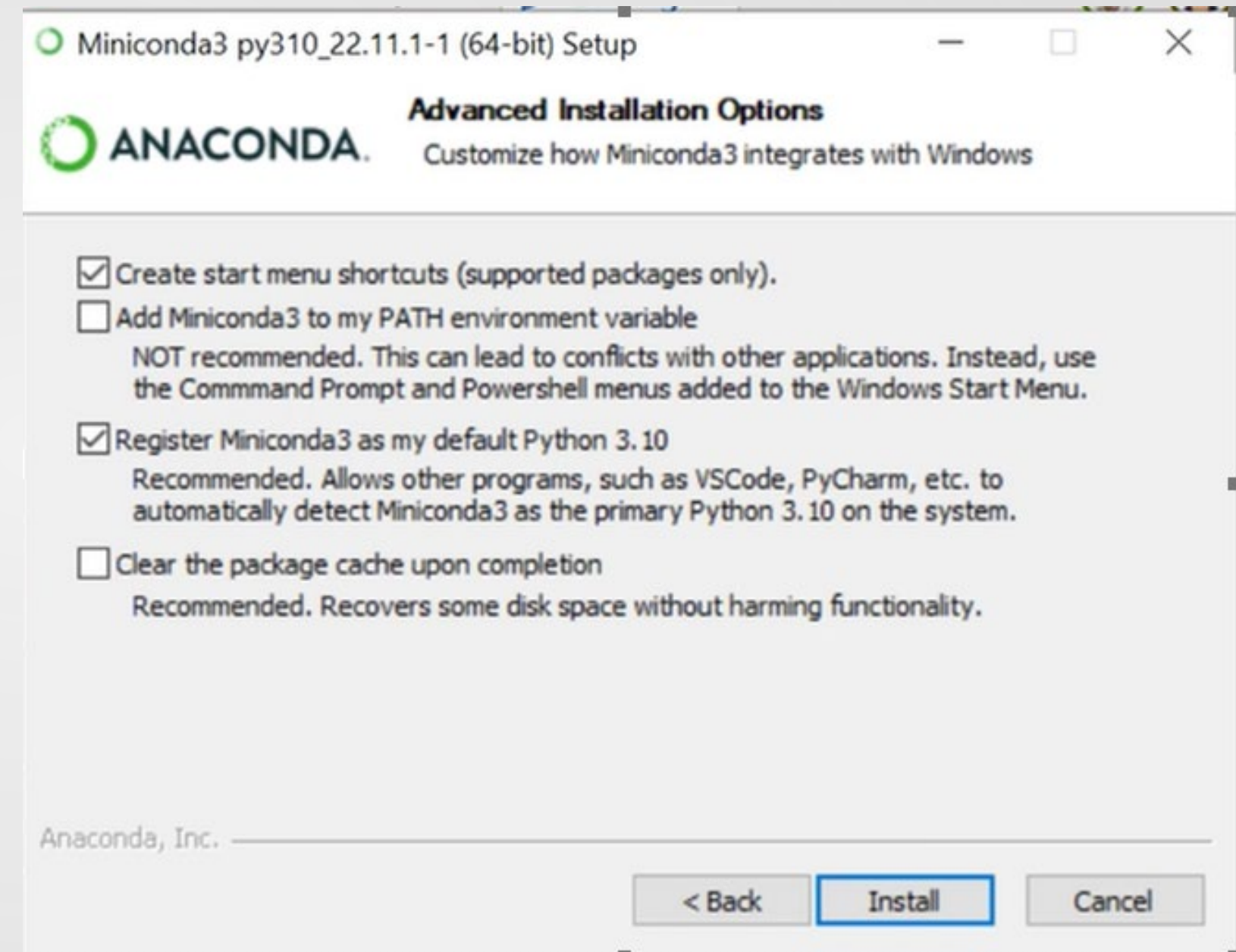
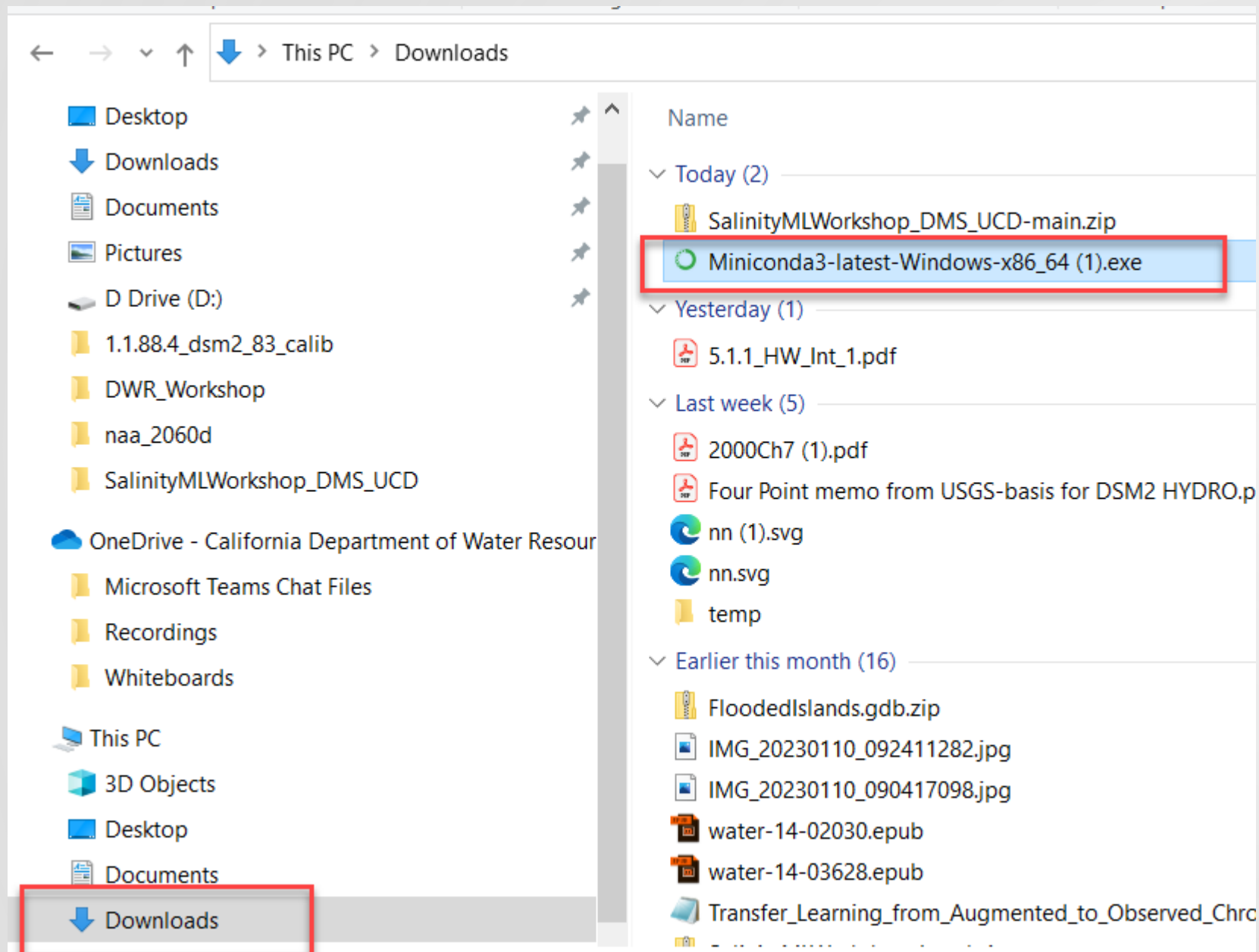
On Windows, macOS, and Linux, it is best to install Miniconda for the local user, which does not require administrator permissions. However, if you need to, you can install Miniconda system wide, which does require administrator permissions.

Latest Miniconda Installer Links

Latest - Conda 22.11.1 Python 3.10.8 released December 22, 2022

Platform	Name	SHA256 hash
Windows	Miniconda3 Windows 64-bit	2e3086630fa3fae7636432a954be530c88d0705fce497120d56e0f5d865b0d51
	Miniconda3 Windows 32-bit	4fb64e6c9c28b88beab16994bfb4829110ea3145baa60bda5344174ab65d462
macOS	Miniconda3 macOS Intel x86 64-bit bash	7406579393427eaf9bc0e094dcd3c66d1e1b93ee9db4e7686d0a72ea5d7c0ce5
	Miniconda3 macOS Intel x86 64-bit pkg	9195ffba1a6984c81c69649ce976a38455ace5b474c24a4363e5ca65fc72e832
	Miniconda3 macOS Apple M1 64-bit bash	22eec9b7d3add25ac3f9b60621d8f3d8df3e63d4aa0ae5eb846b558d7ba68333
	Miniconda3 macOS Apple M1 64-bit pkg	fb33c5770b10a0d5a0deef746e7499bfaf8ff840d0d517175036dd8449357f6
Linux	Miniconda3 Linux 64-bit	00938c3534750a0e4069499baf8f4e6dc1c2e471c86a59caa0dd03f4a9269db6
	Miniconda3 Linux-aarch64 64-bit	48a96df9ff56f7421b6dd7f9f71d548023847ba918c3826059918c08326c2017
	Miniconda3 Linux-ppc64le 64-bit	4c86c3383bb27b44f7059336c3a46c34922df42824577b93eadecefbf7423836
	Miniconda3 Linux-s390x 64-bit	a150511e7fd19d07b770f278fb5dd2df4bc24a8f55f06d6274774f209a36c766

Install Miniconda



Download ML Code/Data

a) If you don't have Git* installed

- https://github.com/CADWRDeltaModeling/SalinityMLWorkshop_DMS_UCD

The screenshot shows the GitHub repository page for `CADWRDeltaModeling / SalinityMLWorkshop_DMS_UCD`. The repository is public. The file list includes:

File Name	Commit Message
Colab_Train_ANN_on_Augmented_Da...	make folder name consisted w
Local_Train_ANN_on_Augmented_Dat...	Rename Train_ANN_on_Augme
README.md	Update README.md
Salinity_DWR.yml	first commit
annutils.py	First commit
dsm2_ann_inputs_base.xlsx	First commit
dsm2_ann_inputs_dcc0.xlsx	First commit
dsm2_ann_inputs_dcc1.xlsx	First commit
dsm2_ann_inputs_rsacminus15day.xlsx	First commit

The 'Code' dropdown menu is open, showing options to clone the repository using HTTPS, SSH, or GitHub CLI, or to download the ZIP file. The 'Download ZIP' option is highlighted by a mouse cursor.

Git* is a free and open source distributed version control system



CALIFORNIA DEPARTMENT OF
WATER RESOURCES



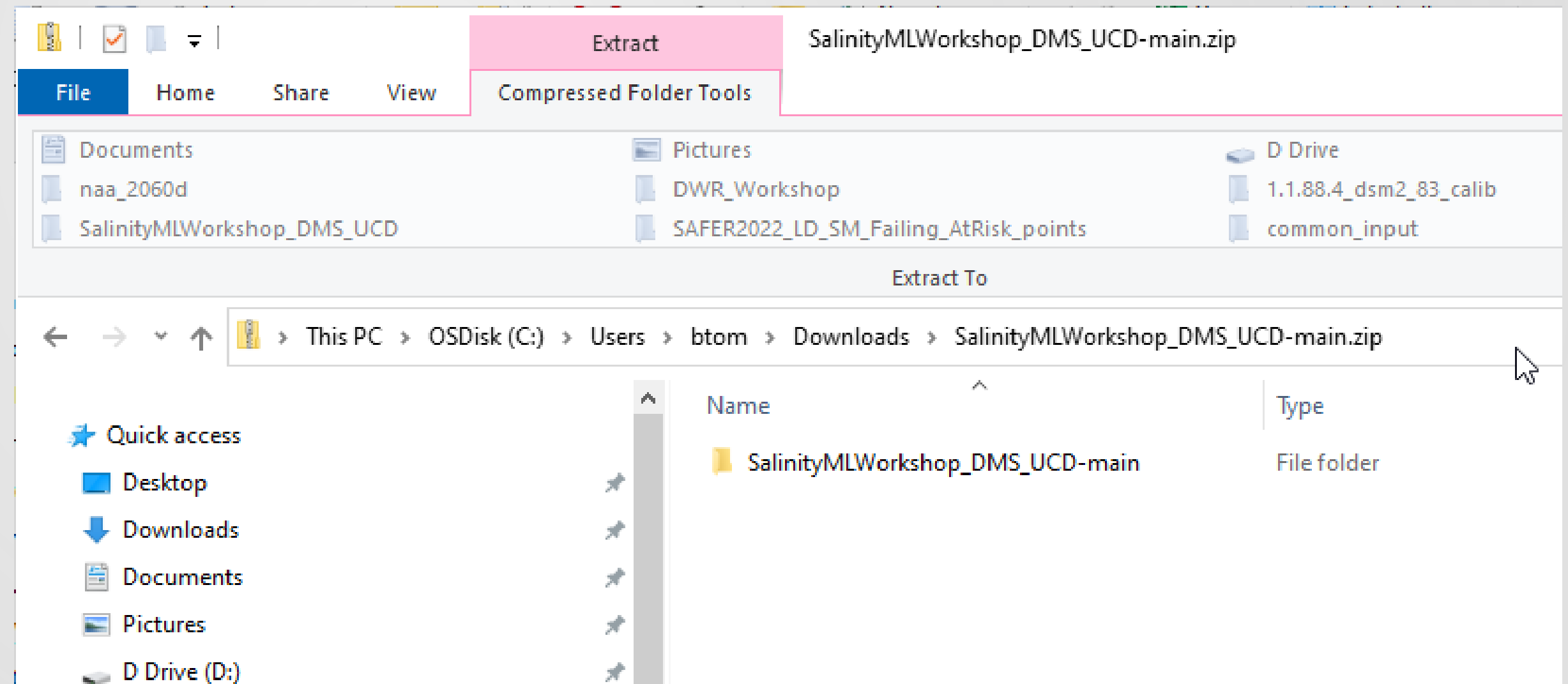
UC DAVIS
UNIVERSITY OF CALIFORNIA

Download ML Code/Data

a) If you don't have Git installed (cont)

Double click the zip file

Copy and paste the folder inside to another location

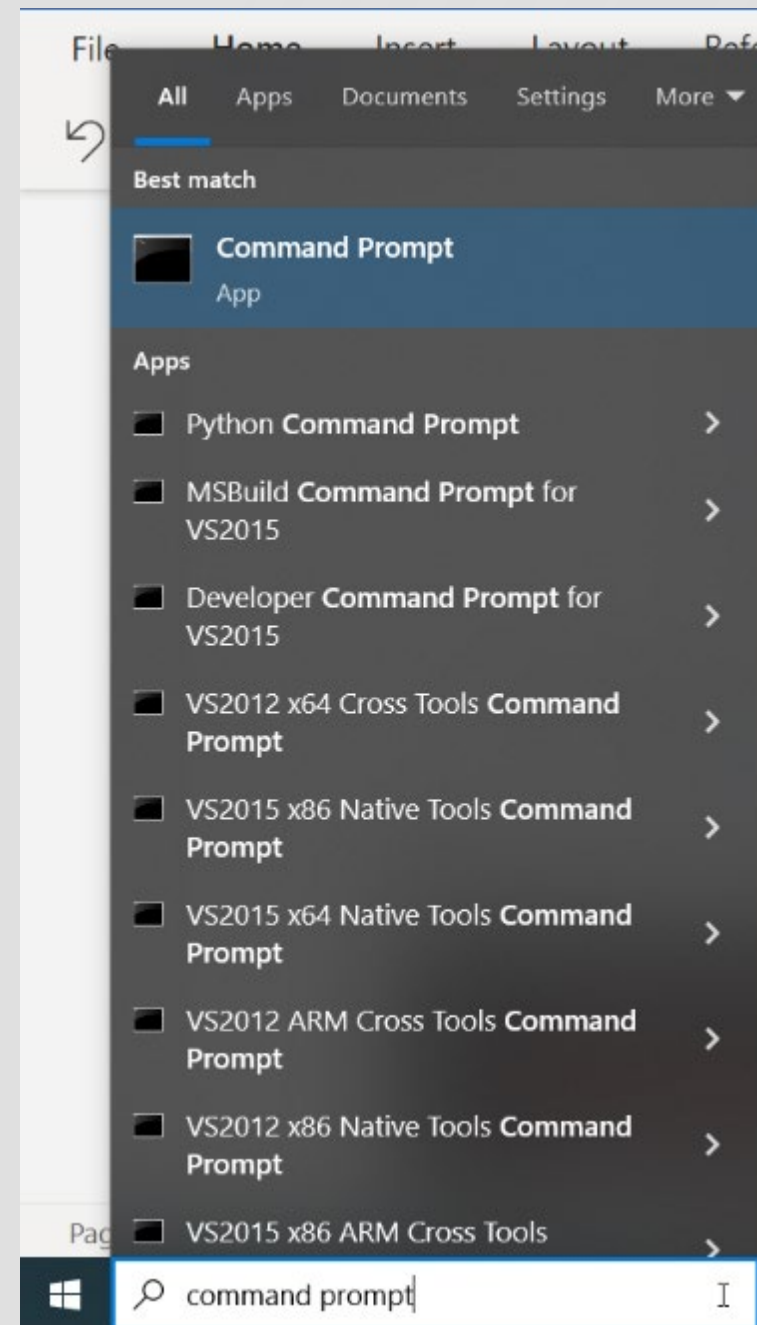


Download ML Code/Data

b) If you have Git installed

Open a command prompt window

Navigate to the location where you want to save files

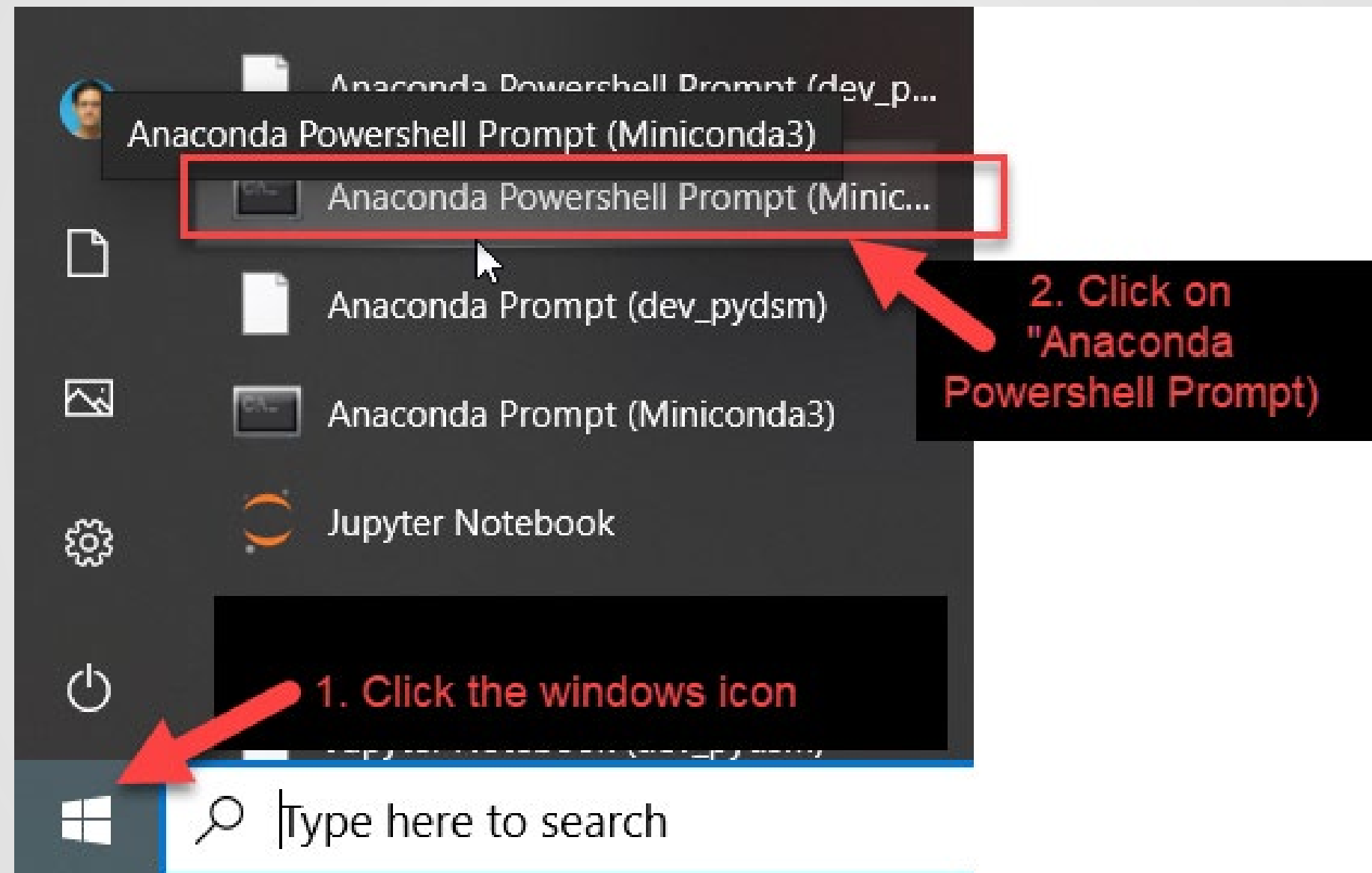


```
Anaconda Powershell Prompt (Miniconda3)

(base) PS C:\Users\btom> e:
(base) PS E:\> cd temp
(base) PS E:\temp> git clone https://github.com/CADWRDeltaModeling/SalinityMLWorkshop_DMS_UCD
Cloning into 'SalinityMLWorkshop_DMS_UCD'...
remote: Enumerating objects: 46, done.
remote: Counting objects: 100% (46/46), done.
remote: Compressing objects: 100% (44/44), done.
remote: Total 46 (delta 14), reused 5 (delta 2), pack-reused 0
Unpacking objects: 100% (46/46), done.
(base) PS E:\temp> _
```

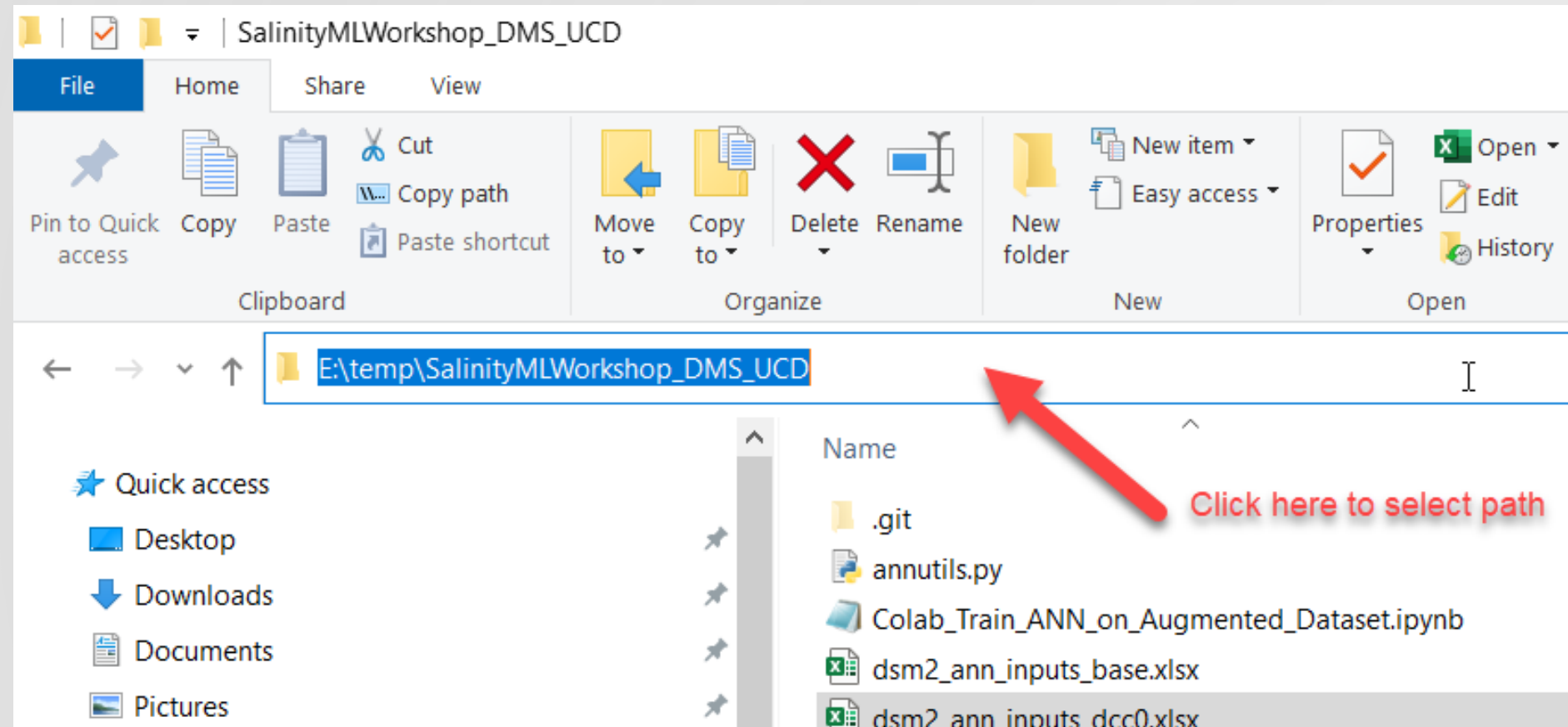
`git clone https://github.com/CADWRDeltaModeling/SalinityMLWorkshop_DMS_UCD`

Open an Anaconda Powershell prompt



Create a conda environment

- Use the “cd” command to navigate to the folder containing your code/data



```
Anaconda Powershell Prompt (Miniconda3)

(base) PS C:\Users\btom> cd E:\temp\SalinityMLWorkshop_DMS_UCD
(base) PS E:\temp\SalinityMLWorkshop_DMS_UCD> 
```


Create a conda environment

- `conda env create -f Salinity_DWR.yml`

```
Anaconda Powershell Prompt (Miniconda3)

(base) PS E:\temp> cd .\SalinityMLWorkshop_DMS_UCD\
(base) PS E:\temp\SalinityMLWorkshop_DMS_UCD> conda env create -f .\Salinity_DWR.yml
Collecting package metadata (repodata.json): done
Solving environment: done

==> WARNING: A newer version of conda exists. <==
  current version: 22.9.0
  latest version: 22.11.1

Please update conda by running

  $ conda update -n base -c defaults conda

Downloading and Extracting Packages
ca-certificates-2023 | 121 KB | #####
Preparing transaction: done
```

```
Anaconda Powershell Prompt (Miniconda3)

1 oauthlib-3.2.2 openpyxl-3.0.7 opt-einsum-3.3.0 protobuf-3.19.6 pyasn1-0.4.8 pyasn1-modules-0.
2.8 requests-oauthlib-1.3.1 rsa-4.9 scikit-learn-1.2.0 scipy-1.10.0 tensorboard-2.11.2 tensorbo
ard-data-server-0.6.1 tensorboard-plugin-wit-1.8.1 tensorflow-2.11.0 tensorflow-estimator-2.11.
0 tensorflow-intel-2.11.0 tensorflow-io-gcs-filesystem-0.30.0 termcolor-2.2.0 threadpoolctl-3.1
.0 werkzeug-2.2.2 wrapt-1.14.1

done

#
# To activate this environment, use
#
#     $ conda activate Salinity_DWR3
#
# To deactivate an active environment, use
#
#     $ conda deactivate

Retrieving notices: ...working... done
(base) PS E:\temp\SalinityMLWorkshop_DMS_UCD>
(base) PS E:\temp\SalinityMLWorkshop_DMS_UCD>
```



Install nb_conda_kernels

- To have access to the new conda environment in a Jupyter Notebook, you need to install “nb_conda_kernels” into the base environment:
- `conda install -y -n base nb_conda_kernels`

```
Anaconda Powershell Prompt (Miniconda3)
(base) PS E:\temp\SalinityMLWorkshop_DMS_UCD> conda install -y -n base nb_conda_kernels
Collecting package metadata (current_repodata.json): done
Solving environment: done

## Package Plan ##

  environment location: C:\Users\btom\Miniconda3

  added / updated specs:
    - nb_conda_kernels

The following packages will be downloaded:
```

package	build	
certifi-2022.12.7	py39haa95532_0	149 KB
conda-22.11.1	py39haa95532_4	892 KB
pluggy-1.0.0	py39haa95532_1	29 KB
ruamel.yaml-0.17.21	py39h2bbff1b_0	174 KB
ruamel.yaml.clib-0.2.6	py39h2bbff1b_1	101 KB
Total:		1.3 MB

```
Anaconda Powershell Prompt (Miniconda3)
ruamel.yaml.clib  pkgs/main/win-64::ruamel.yaml.clib-0.2.6-py39h2bbff1b_1 None

The following packages will be UPDATED:

ca-certificates      2022.07.19-haa95532_0 --> 2023.01.10-haa95532_0 None
certifi              2022.9.24-py39haa95532_0 --> 2022.12.7-py39haa95532_0 None
conda                22.9.0-py39haa95532_0 --> 22.11.1-py39haa95532_4 None
openssl              1.1.1q-h2bbff1b_0 --> 1.1.1s-h2bbff1b_0 None

Downloading and Extracting Packages
certifi-2022.12.7      | 149 KB | ##### | 100%
pluggy-1.0.0          | 29 KB  | ##### | 100%
ruamel.yaml-0.17.21    | 174 KB | ##### | 100%
conda-22.11.1          | 892 KB | ##### | 100%
ruamel.yaml.clib-0.2   | 101 KB | ##### | 100%
Preparing transaction: done
Verifying transaction: done
Executing transaction: done
Retrieving notices: ...working... done
(base) PS E:\temp\SalinityMLWorkshop_DMS_UCD>
```



Open Jupyter Notebook

jupyter notebook

If the above command fails, try these two commands:

call conda activate jupyter

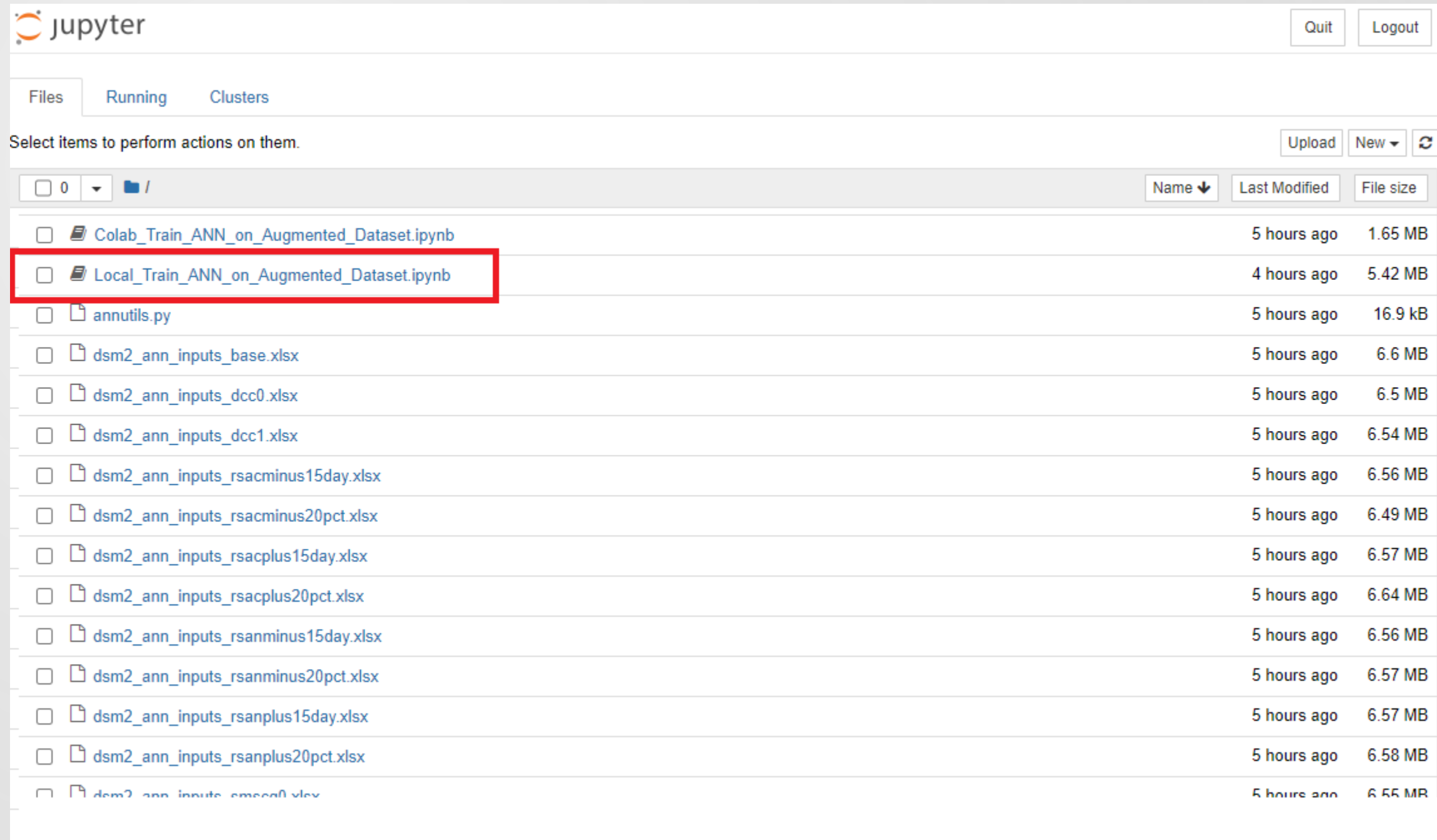
start jupyter notebook



Open the ANN Jupyter Notebook

click on

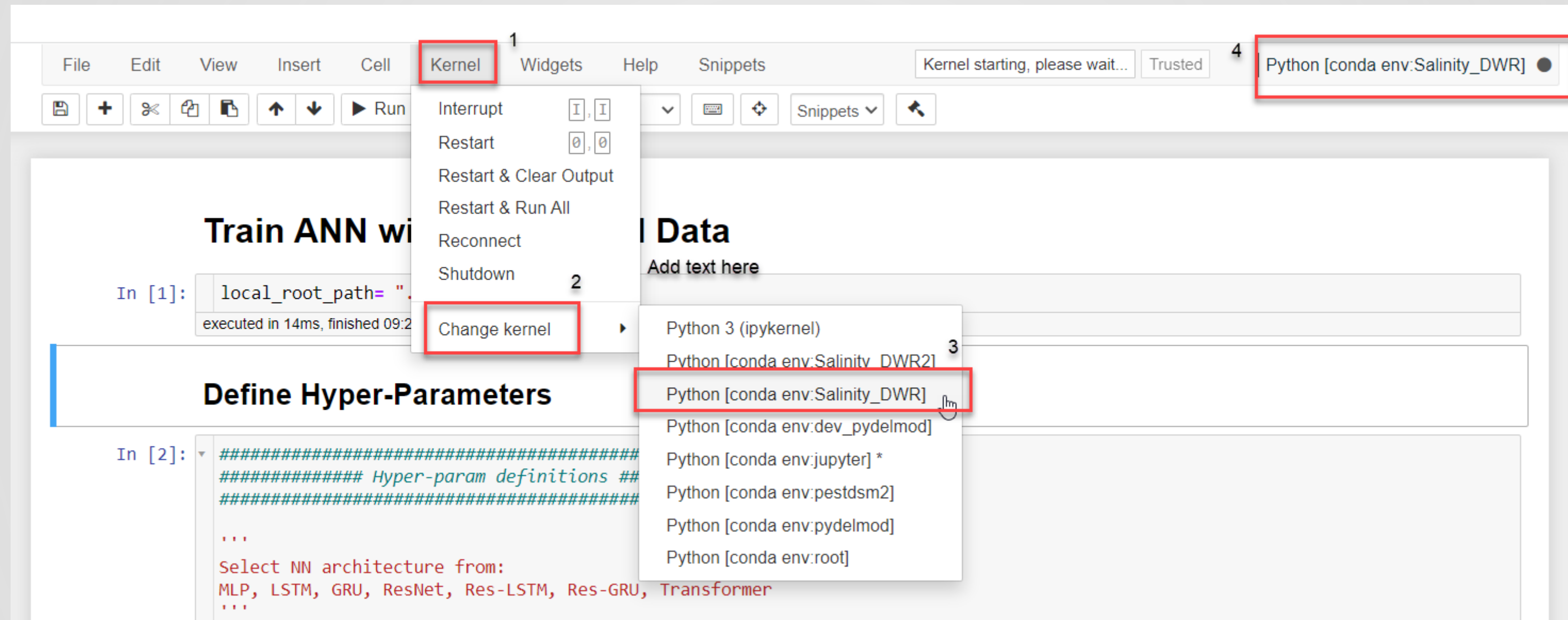
Local_Train_ANN_on_Augmented_Dataset.ipynb



The screenshot shows the JupyterLab interface with the 'Files' tab selected. The file browser displays a list of files in the current directory. The file 'Local_Train_ANN_on_Augmented_Dataset.ipynb' is highlighted with a red box. The table below lists the files and their details:

Name	Last Modified	File size
Colab_Train_ANN_on_Augmented_Dataset.ipynb	5 hours ago	1.65 MB
Local_Train_ANN_on_Augmented_Dataset.ipynb	4 hours ago	5.42 MB
annutils.py	5 hours ago	16.9 kB
dsm2_ann_inputs_base.xlsx	5 hours ago	6.6 MB
dsm2_ann_inputs_dcc0.xlsx	5 hours ago	6.5 MB
dsm2_ann_inputs_dcc1.xlsx	5 hours ago	6.54 MB
dsm2_ann_inputs_rsacminus15day.xlsx	5 hours ago	6.56 MB
dsm2_ann_inputs_rsacminus20pct.xlsx	5 hours ago	6.49 MB
dsm2_ann_inputs_rsacplus15day.xlsx	5 hours ago	6.57 MB
dsm2_ann_inputs_rsacplus20pct.xlsx	5 hours ago	6.64 MB
dsm2_ann_inputs_rsanminus15day.xlsx	5 hours ago	6.56 MB
dsm2_ann_inputs_rsanminus20pct.xlsx	5 hours ago	6.57 MB
dsm2_ann_inputs_rsanplus15day.xlsx	5 hours ago	6.57 MB
dsm2_ann_inputs_rsanplus20pct.xlsx	5 hours ago	6.58 MB
dsm2_ann_inputs_rsanplus20pct.xlsx	5 hours ago	6.55 MB

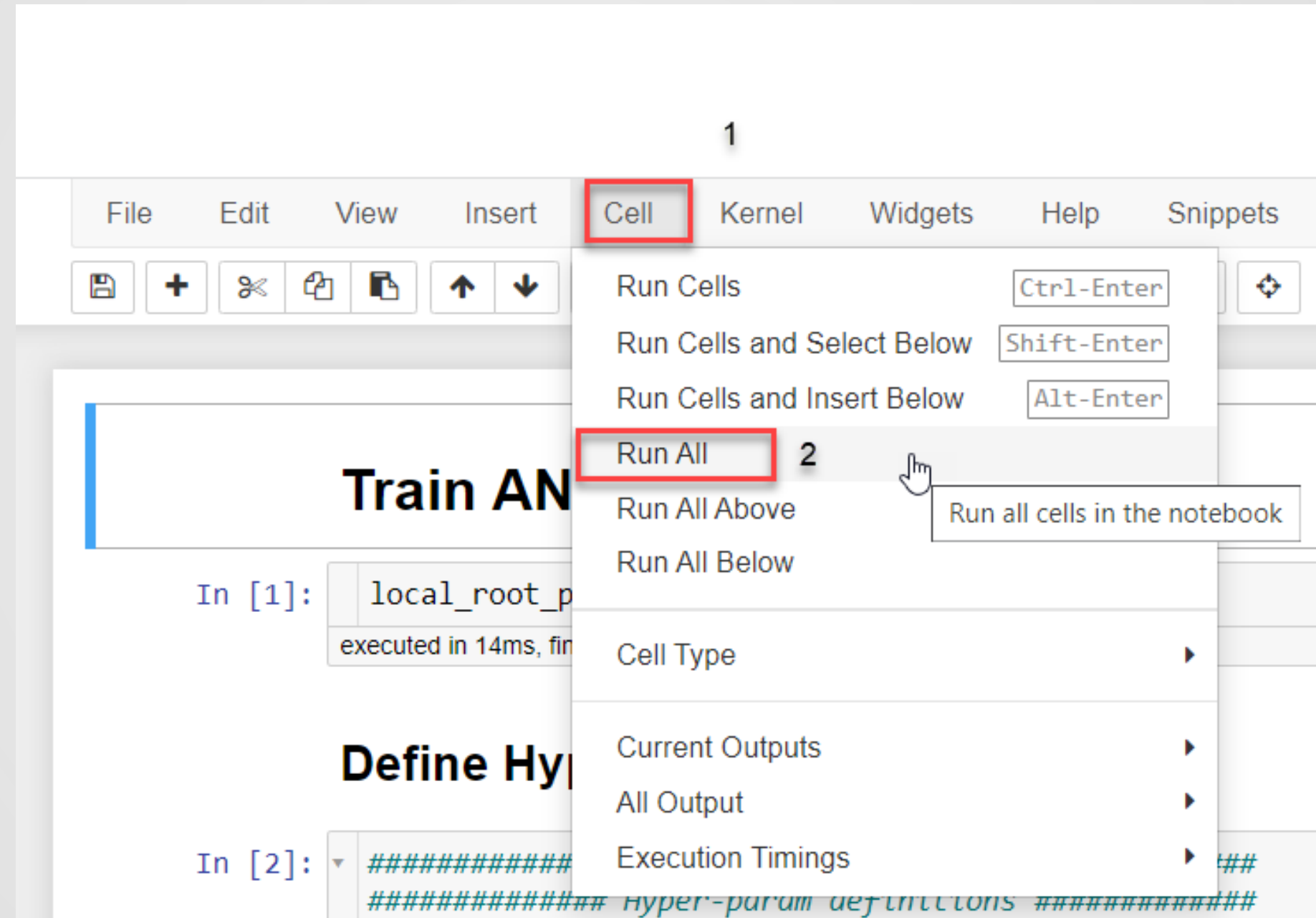
Select notebook kernel



A notebook kernel is a “computational engine” that executes the code contained in a Notebook document.



Run the notebook code



Completed cell vs running cell

Install Packages

```
In [3]: import os
import sys
from sklearn.metrics import r2_score
import matplotlib.pyplot as plt
import pickle
import time

import numpy as np
import pandas as pd
import tensorflow as tf
from tensorflow import keras
from tensorflow.keras.layers.experimental.preprocessing import Normalization
from tensorflow.keras import layers
#import keras
from sklearn.preprocessing import MinMaxScaler
```

executed in 7.84s, finished 18:03:48 2023-01-23

Read Data

```
In [*]: sys.path.append(local_root_path)
import annutils

observed_stations_ordered_by_median = ['RSMKL008', 'RSAN032', 'RSAN037', 'RSAC092', 'SLTRM004', 'ROLD024',
                                         'CHVCT000', 'RSAN018', 'CHSWP003', 'CHDMC006', 'SLDUT007', 'RSAN072',
                                         'OLD_MID', 'RSAN058', 'ROLD059', 'RSAN007', 'RSAC081', 'SLMZU025',
                                         'RSAC075', 'SLMZU011', 'SLSUS012', 'SLCBN002', 'RSAC064']

num_sheets = 9
```



Questions?

Kevin.He@water.ca.gov



CALIFORNIA DEPARTMENT OF
WATER RESOURCES



UC DAVIS
UNIVERSITY OF CALIFORNIA