

Please view the PDF at max 125% zoom. Some of the underscores don't up at higher zoom ratios.

Our goal is to setup the following for the training

1. python 3.6.x with following libraries:
  - a. numpy
  - b. pandas
  - c. matplotlib
  - d. sklearn
  - e. xgboost
  - f. lightgbm
  - g. catboost
  - h. utm
2. Any Python IDE, preferably jupyter notebook

Jupyter Notebook install: <https://jupyter.org/install.html>

Python packages install example: `python -m pip install matplotlib`

One way of managing the above without disturbing any existing python setups in your machine is to install Anaconda python which is very helpful in managing python virtual environments.

You can download and install the open source version from <https://www.anaconda.com/products/individual>

Make sure to choose the right operating system.

After Anaconda is installed, do the following steps:

1. Launch Anaconda Navigator
2. Launch Anaconda Cmd.exe prompt
3. create a virtual environment with right version of python  
call `conda create --name dv_ml_training_spe python=3.6`  
When proceed y/n prompt appears, type y
4. Activate the virtual environment  
call `conda activate dv_ml_training_spe`
5. Check the python version. It should show up as 3.6.xx where xx is a number e.g. 12.  
call `python --version`
6. Navigate to the folder where you have downloaded the requirements.text file and install the required python libraries. E.g.  
`cd C:\Users\sun_g\sgarg\dataVediK\Events\SPE_DSC2022\ML_Workshop\GammaRayPrediction`  
`pip install -r requirements.txt`
7. Select the environment `dv_ml_training_spe` in Anaconda Navigator home and install jupyter Notebook
8. Launch jupyter Notebook

During the class:

1. Navigate to the `GammaRayPrediction` prediction folder
2. Open `Sythentic_Log_Generator.ipynb`