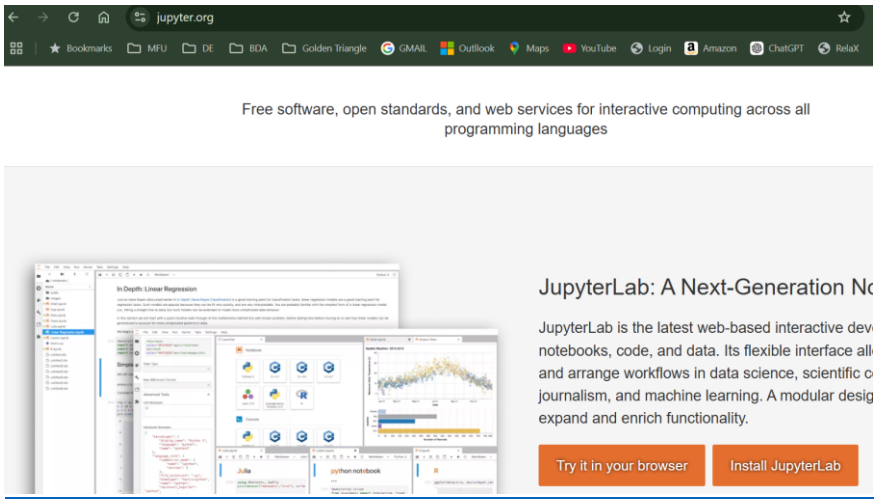
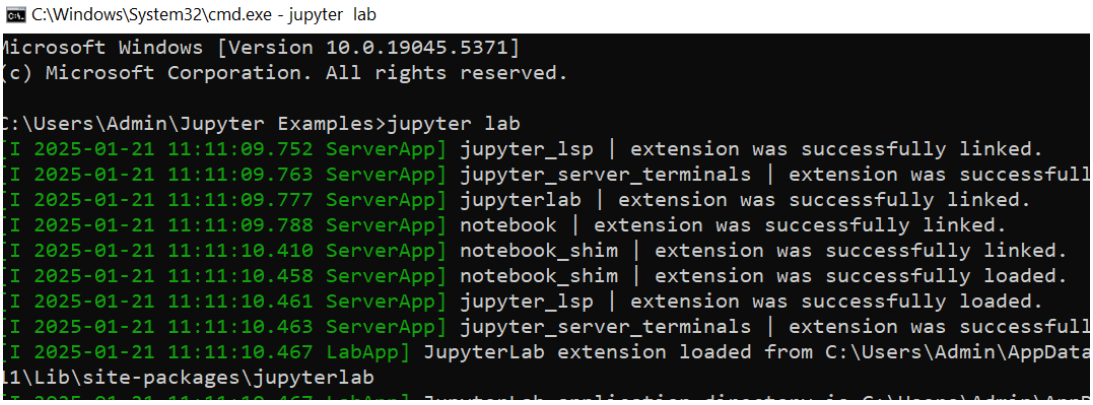
Lab Exercise 4 – Supervised Learning

Naïve Bayseian, Decision Tree, and SVM

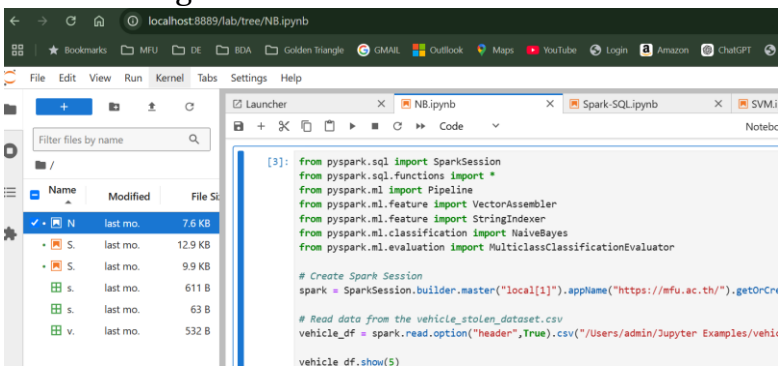
Step 1:
Go to https://jupyter.org/ and install Jupyter Lab.



Step 2:
Create a folder "Jupyter Examples" in C:\ drive and execute **Jupyter Lab**
C:\Users\Admin\Jupyter Examples>jupyter lab

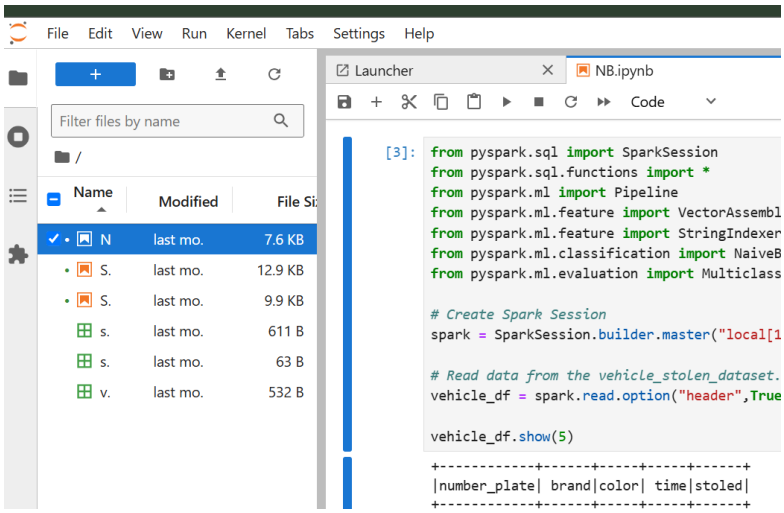


You should see the following site:

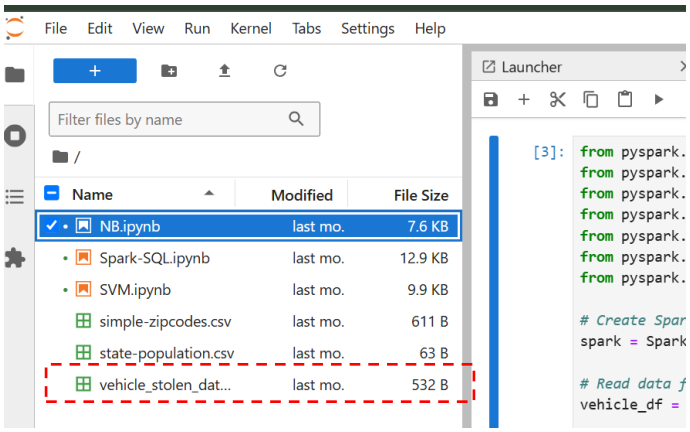


Step 3:

Copy and paste the code from the file: NB.ipynb into a Jupyter notebook.



Step 4:
Make sure to keep the dataset file **vehicle_stolen_dataset.csv** in the working folder.



Step 5:
Load the next two files *spark-SQL.ipynb* and *SVM.ipynb*. The dataset for *spark-SQL.ipynb* is *simple-zipcodes.csv*. For *SVM.ipynb*, there is no need to use any *csv* file, because *breast_cancer()* dataset is downloaded from sklearn.datasets.

Step 6:
Record the accuracy for each program after executing the code.