**BIG DATA**

**Group**

**-Carlos Hernandez ID: 986636**

**-Sebastian Valencia ID: 986775**

**Naive Bayes Algorithm Tutorial**

**What you Will see(?)**

* **Handle Data**: Load the data from CSV file and split it into training and test datasets (67%,33% respectively).
* **Summarize Data**: summarize the properties in the training dataset so that we can calculate probabilities and make predictions.
* **Make a Prediction**: Use the summaries of the dataset to generate a single prediction.
* **Make Predictions**: Generate predictions given a test dataset and a summarized training dataset.
* **Evaluate Accuracy**: Evaluate the accuracy of predictions made for a test dataset as the percentage correct out of all predictions made.

**Prerequisites**

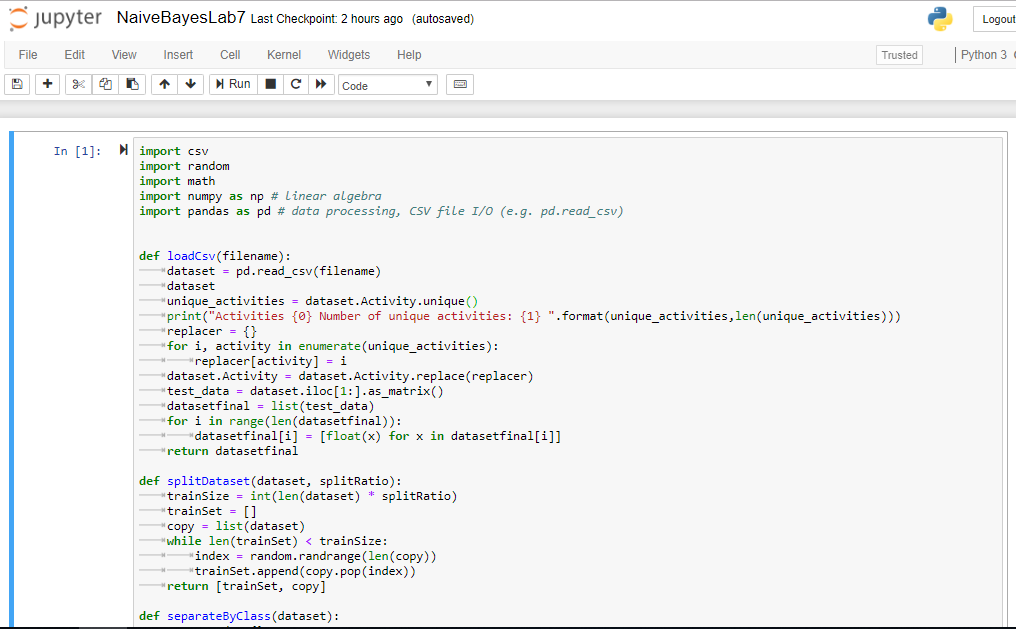
**Be sure that you already has installed the next programs in your machine**

-Anaconda3 (python 3.5.1)

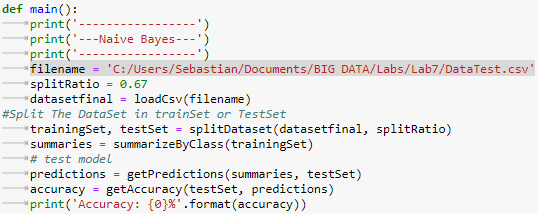
-Jupyter Notebook

-These librarys (csv, random, math, numpy, pandas)

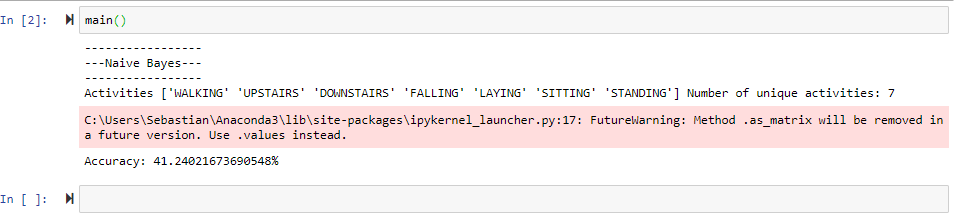
**Open the file 'NaiveBayesLab7' in Jupyter Notebook. (You should see this)**



You should change the path of the file



# Run code (should look something like this)



Enjoy !!! :D