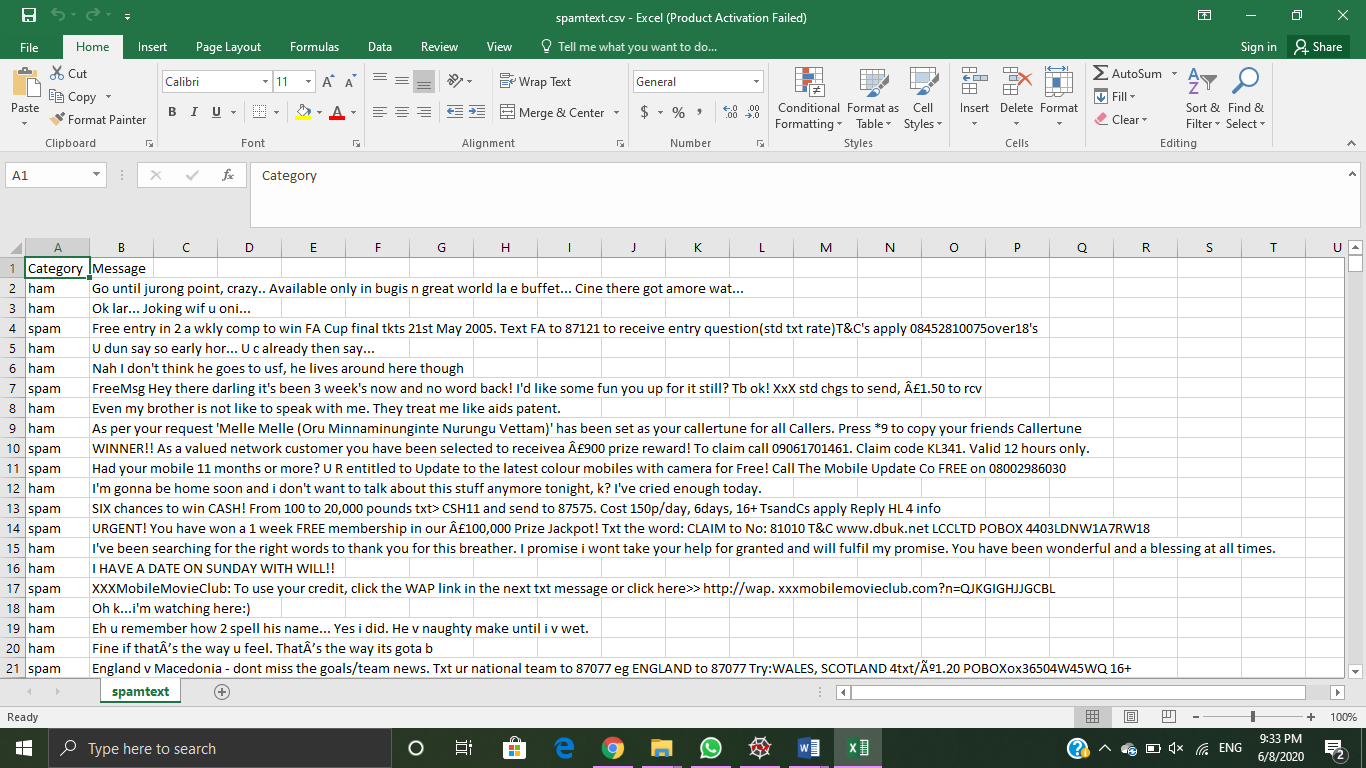
**Implementation steps:**

1. **Download Anaconda latest version from its official site. and configure it.**
2. **Open Spyder IDE**
3. **Create a project folder and copy the dataset into the folder (spamtext.csv)**
4. **Importing dataset in python using pandas**

import pandas as pd

dataset = pd.read\_csv('spamtext.csv')



**1)imported** **data .**

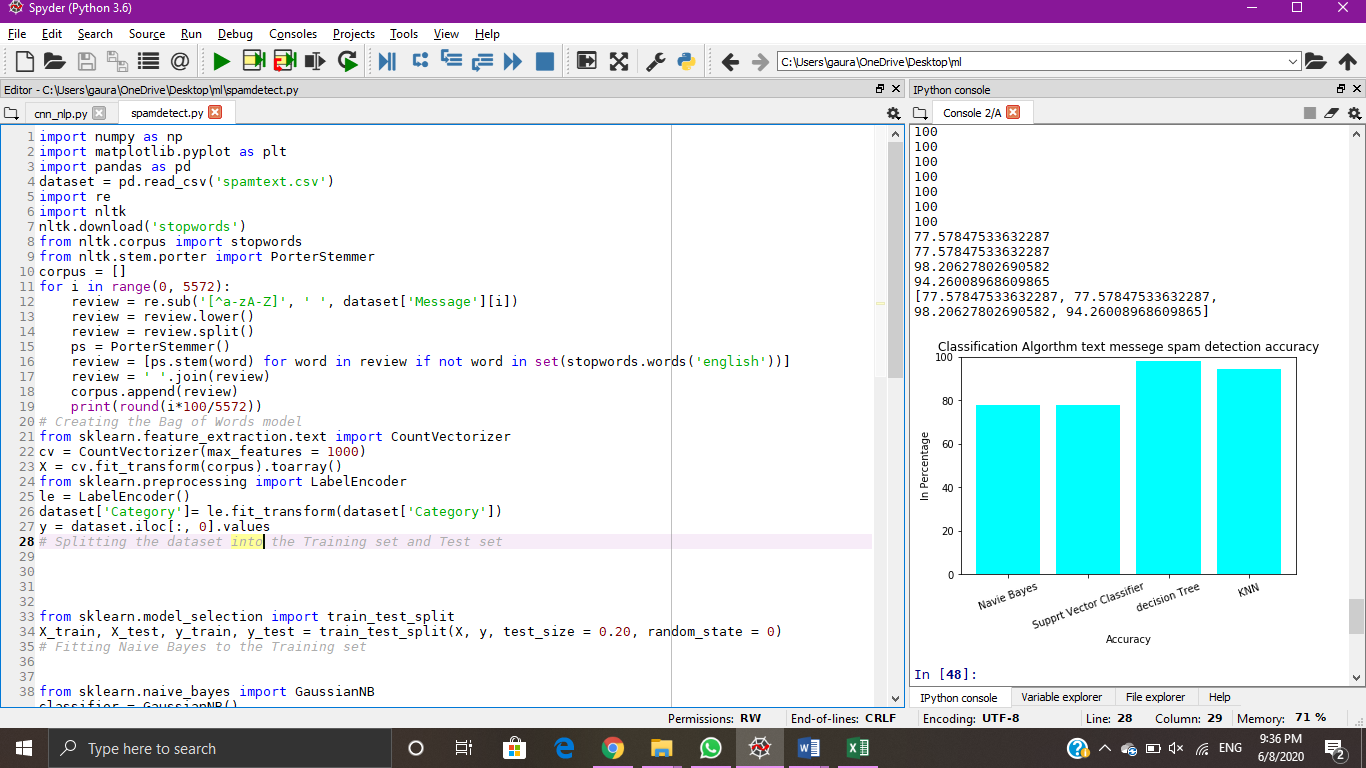
Category: - classes (spam, ham) - 2

Text:- comments to be classified into classes.

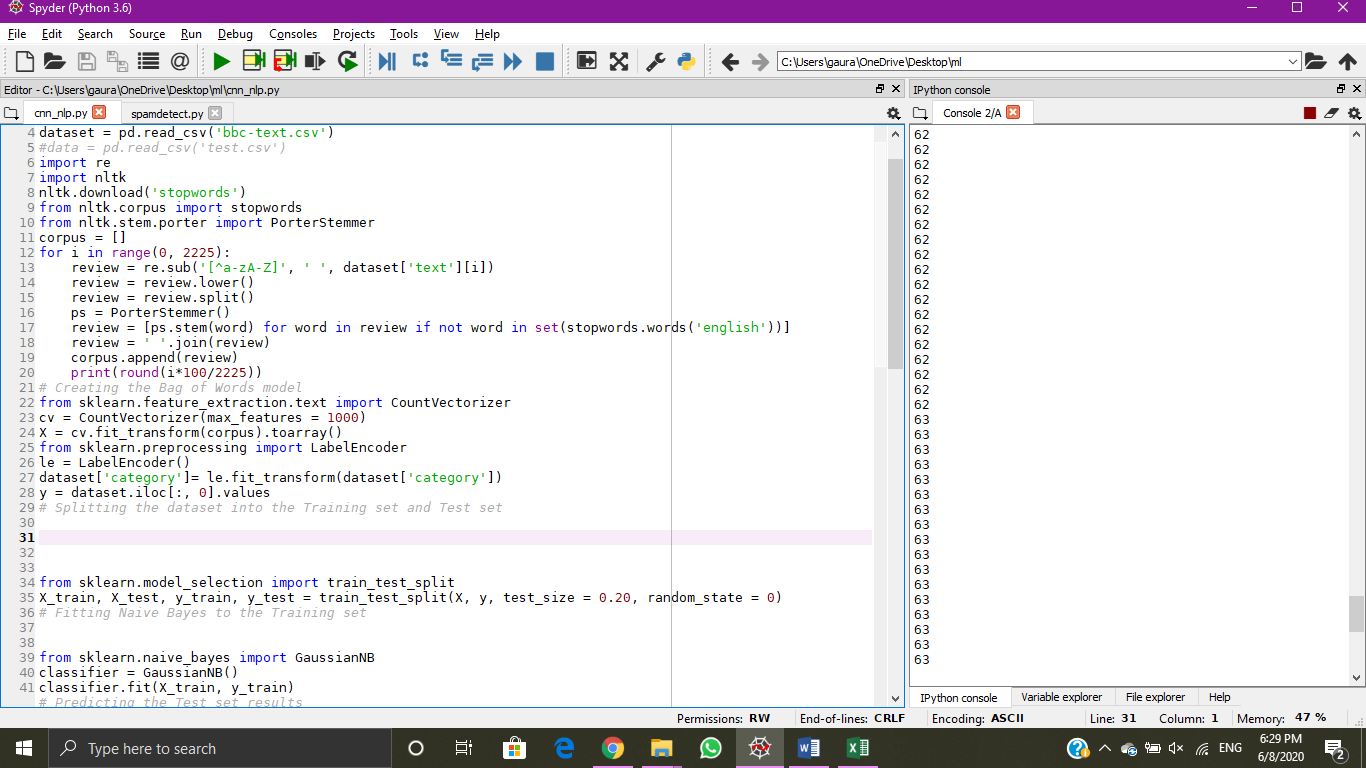
**5.Data preprocessing**

* **Removing symbols -**
* **Considering English words.**

By using natural language processing toolkit.



**To convert categorical Data into numeric data using Feature Extraction.**



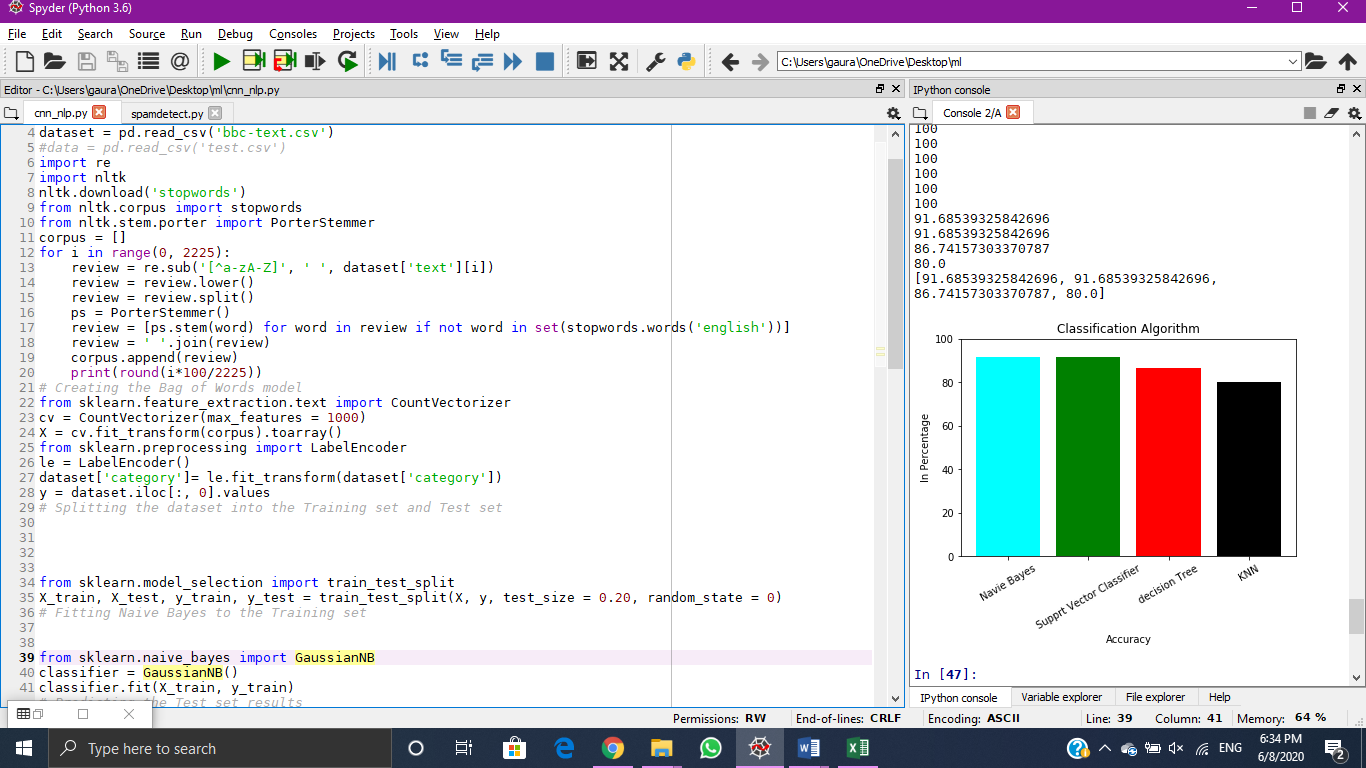
**To convert categorical Data into numeric data using Label Encoding.**

1

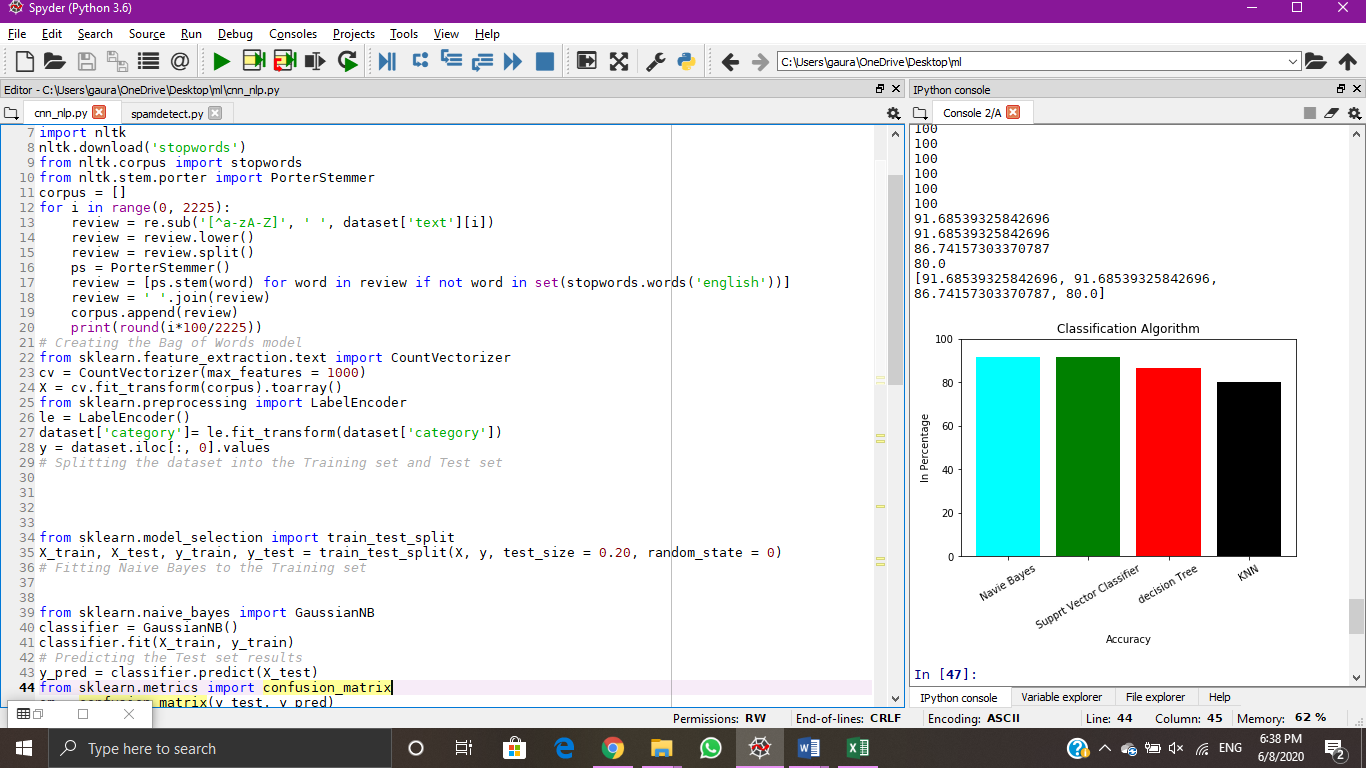
Spam

0

No Spam



**6. Splitting the data into training and testing sets**

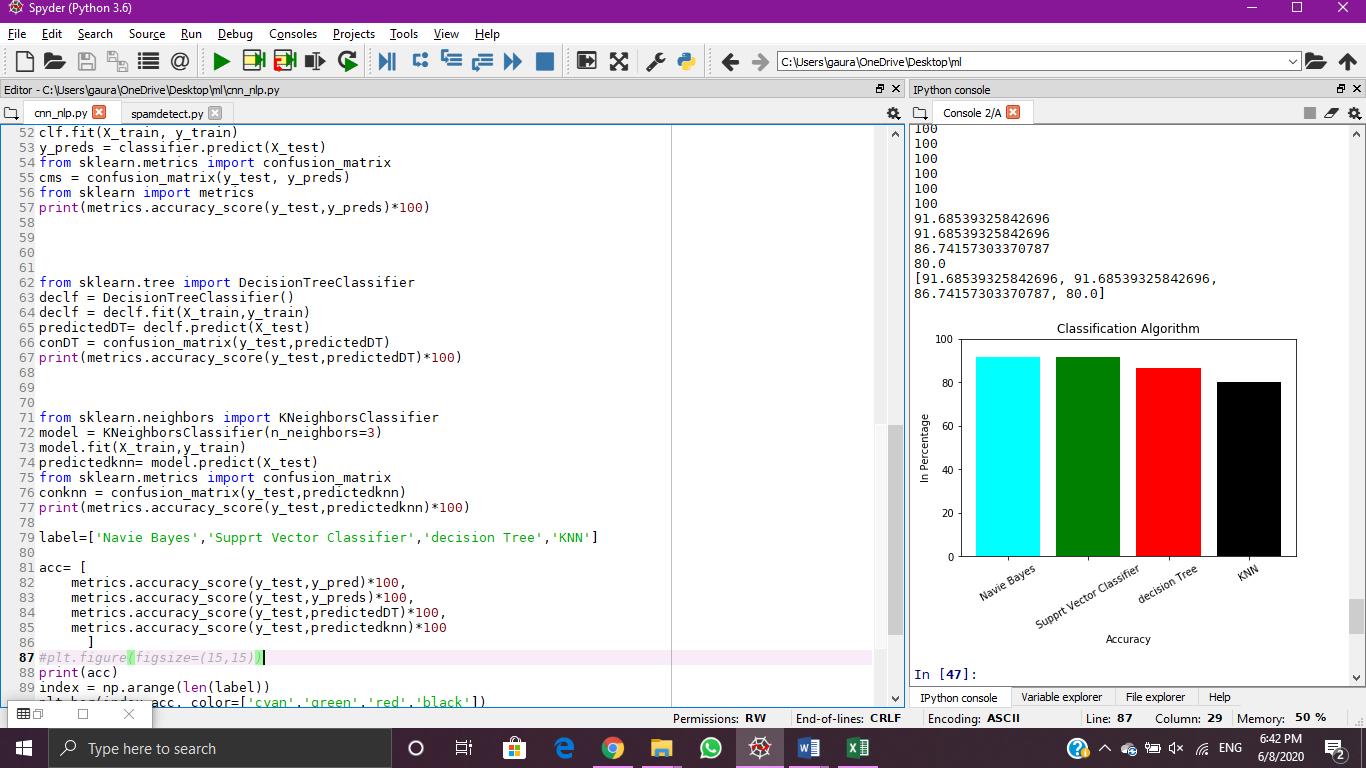


**The dataset is split in ratio of 0.2 means 80% training set and 20% testing set.**

**7. Creating classification models**

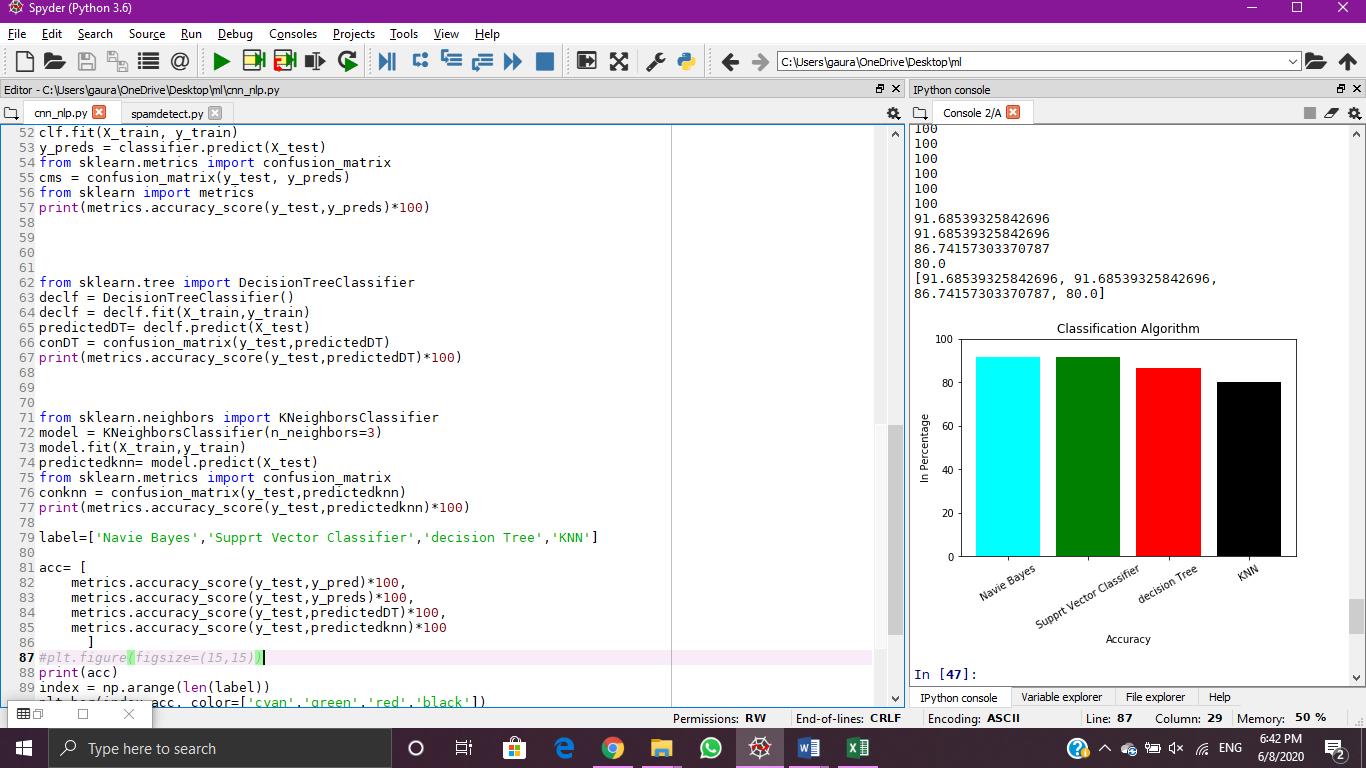
* **KNN classifier**

**Creating classifier**

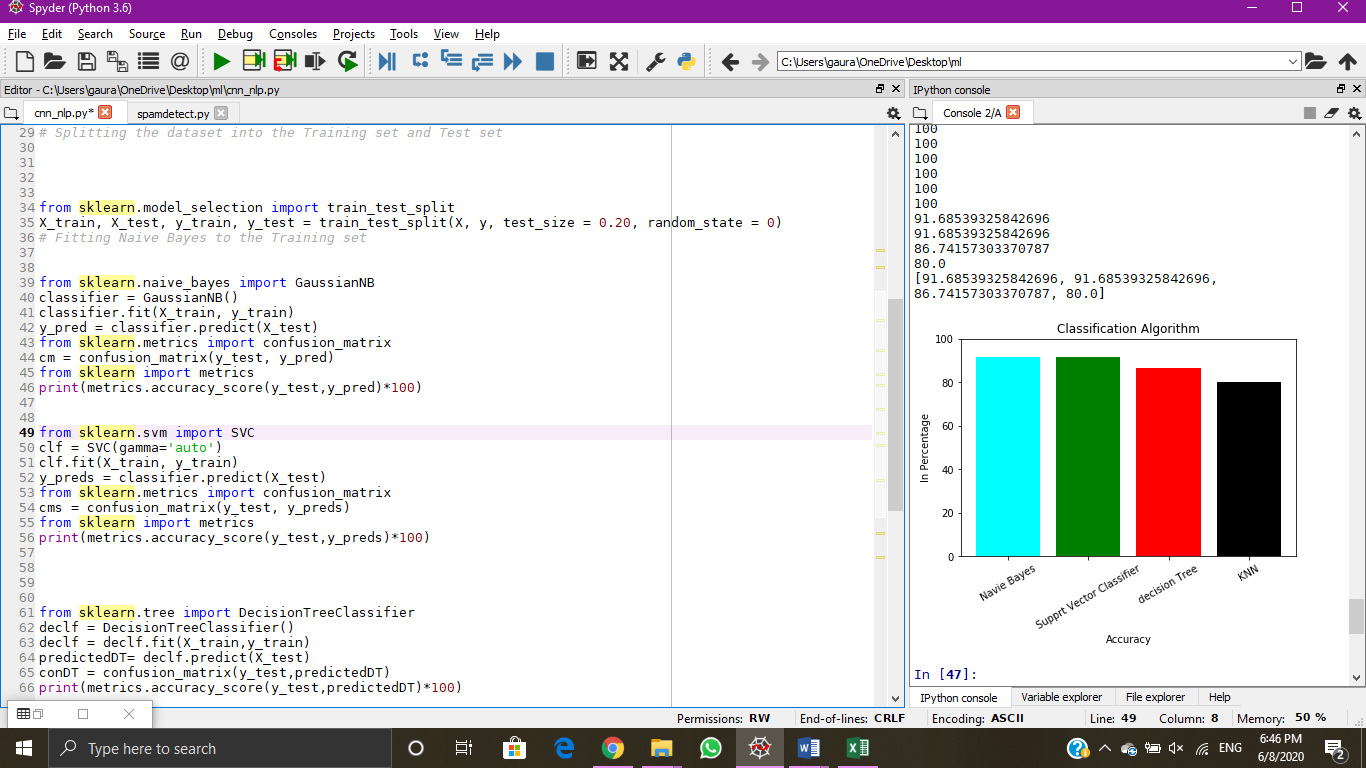


* **Decision Tree:**

**Creating classifier:**

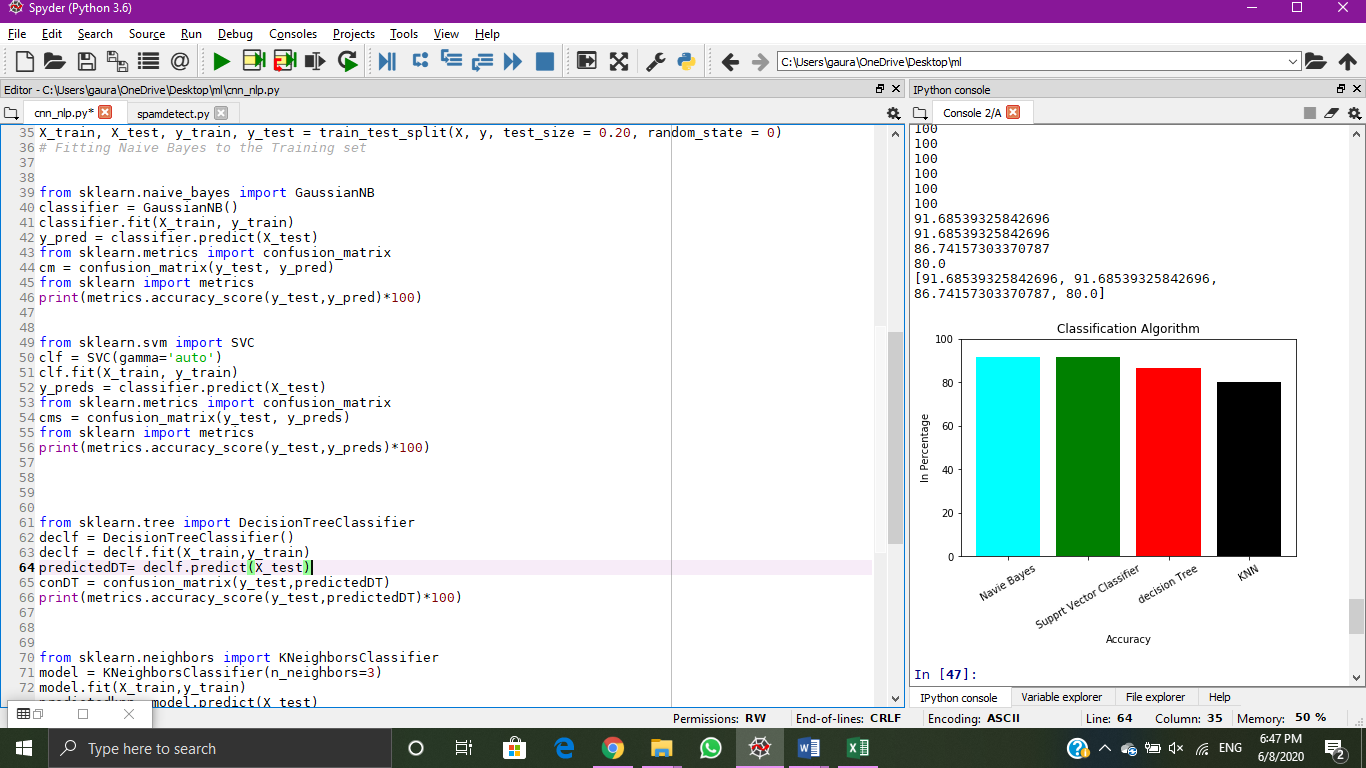


* **Naïve Bayes :  
   Creating classifier :**

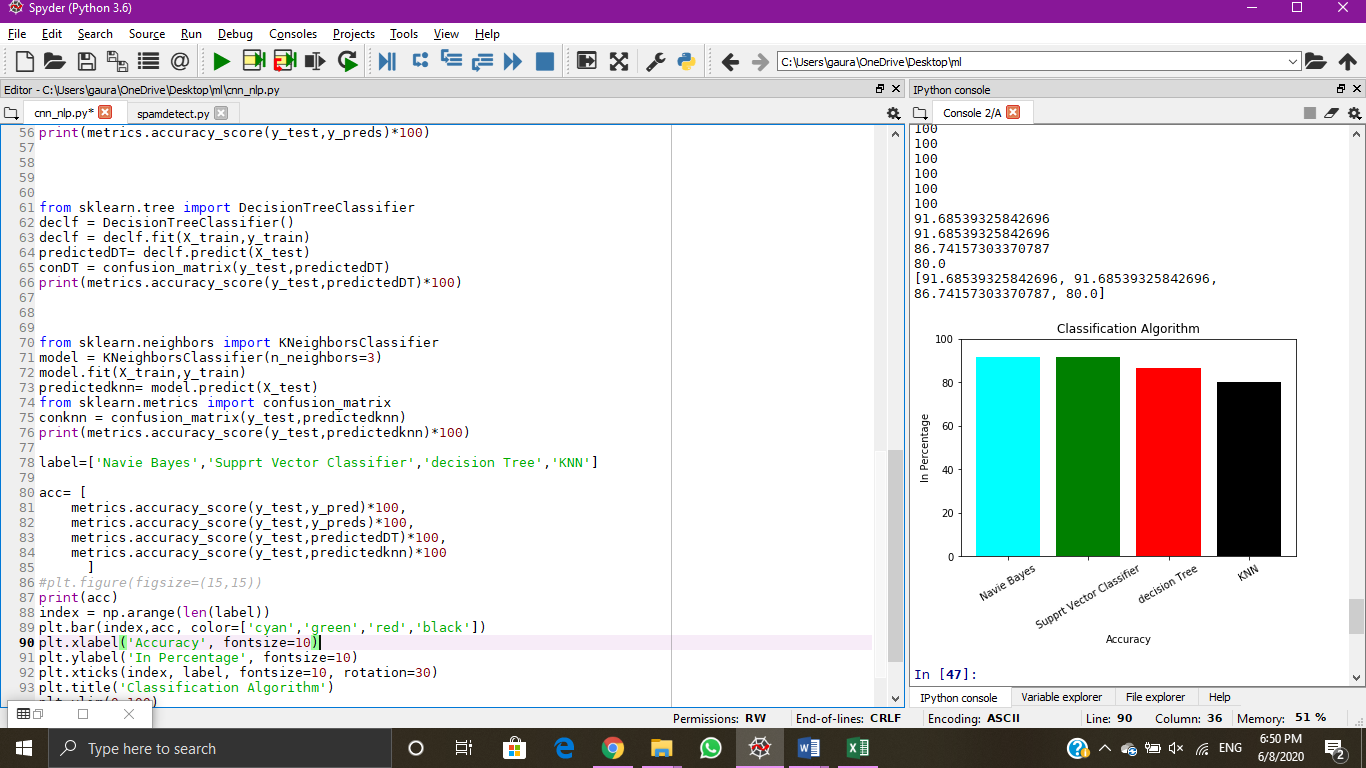


* **Support Vector Machine :**

**Creating Classifier model**



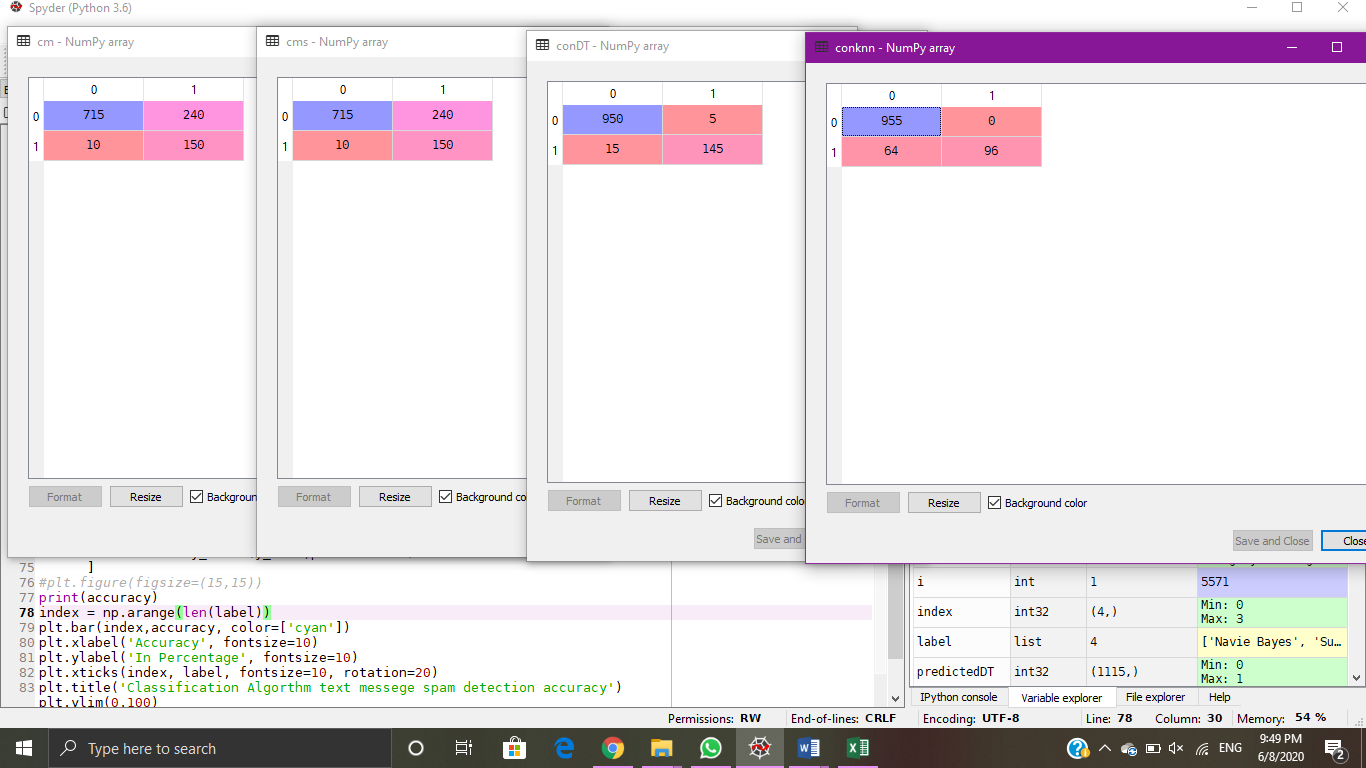
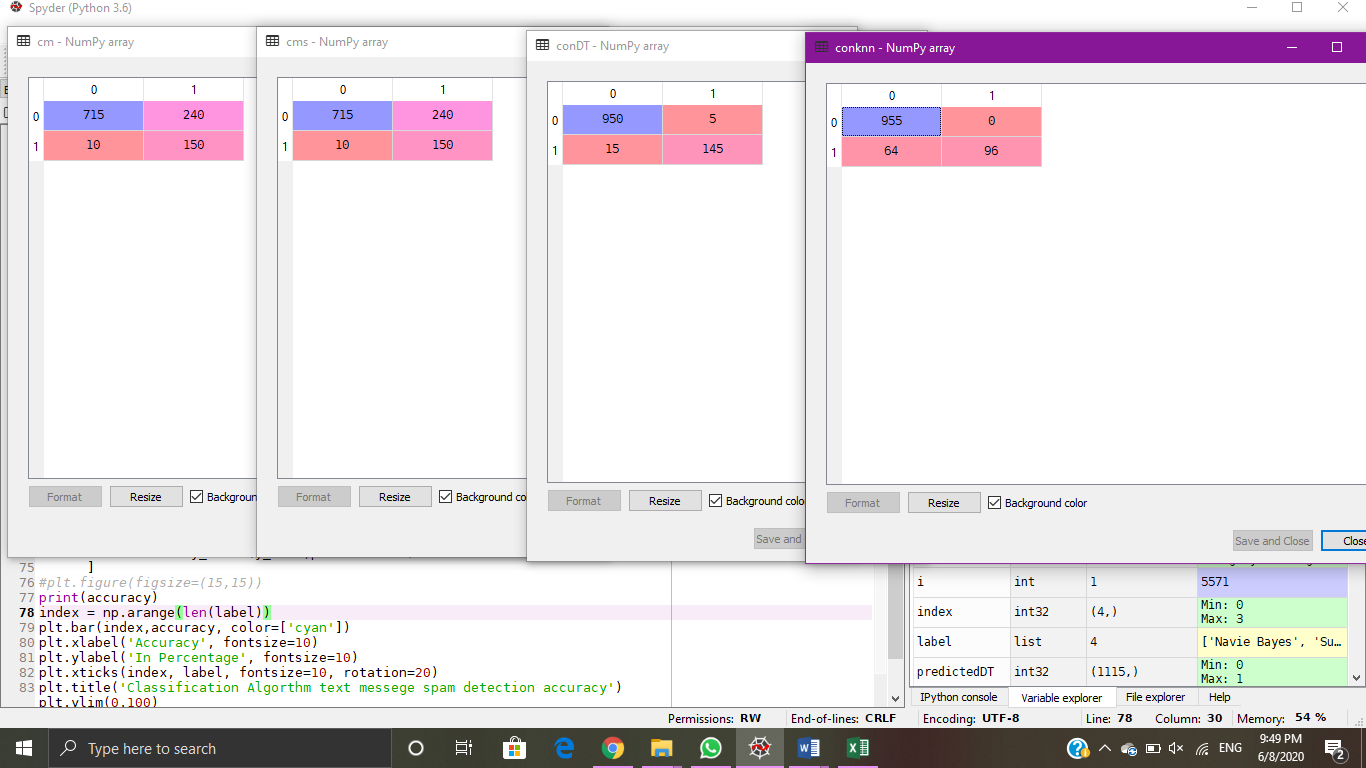
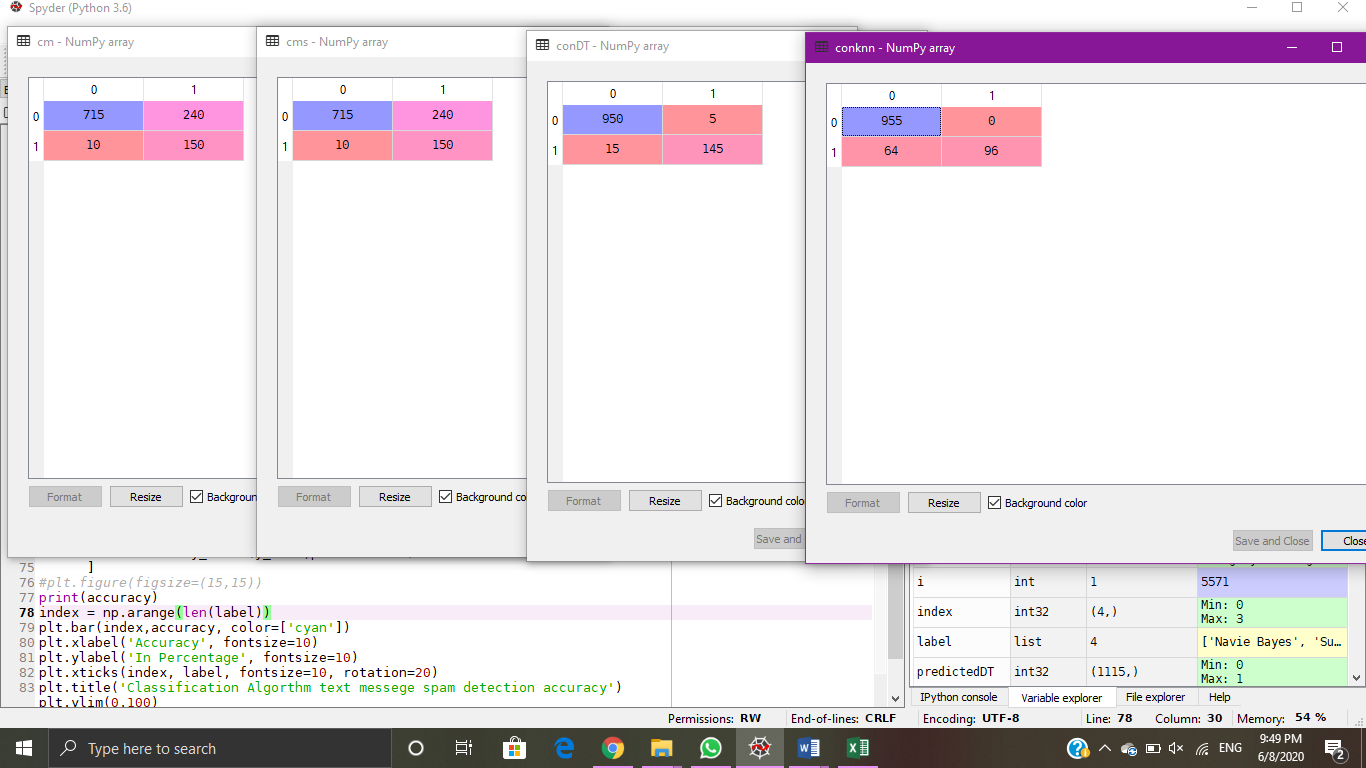
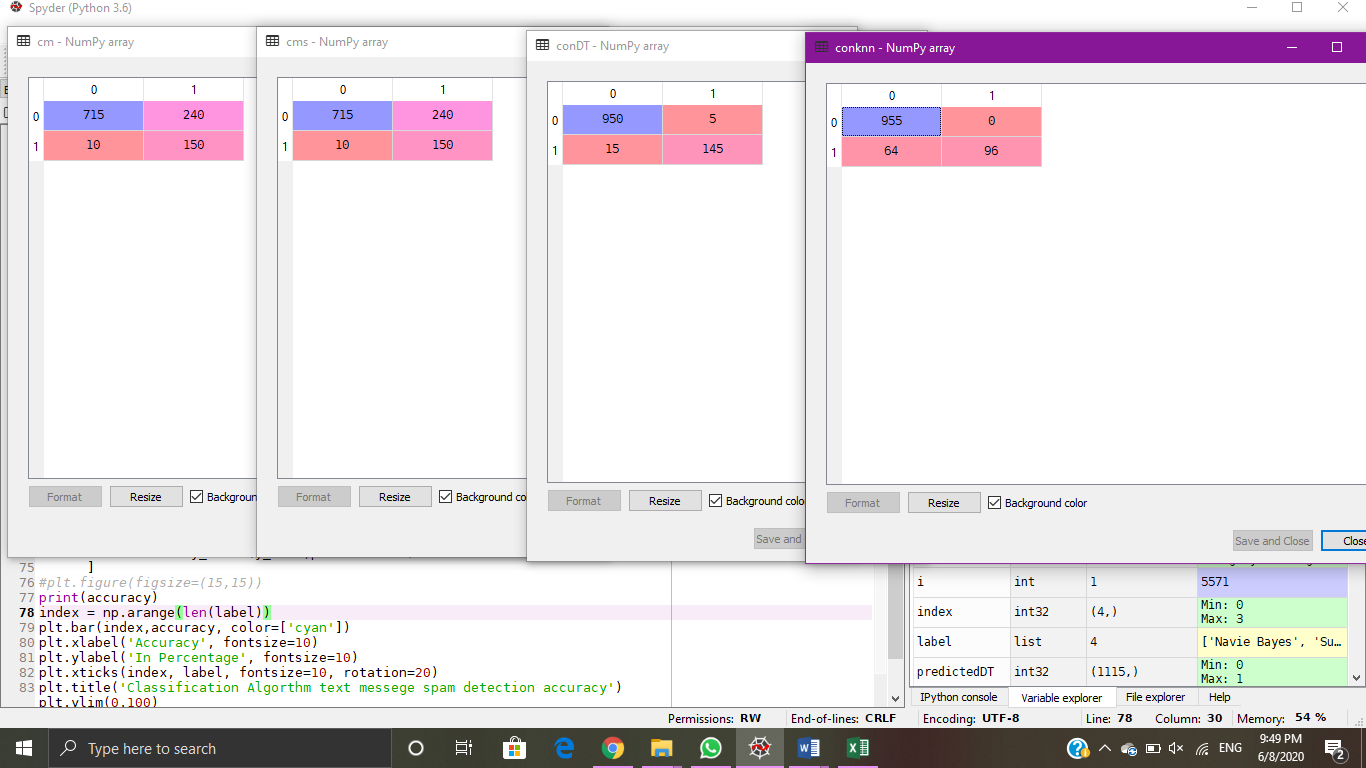
**8.Calculating Accuracy .**



|  |  |
| --- | --- |
| **Method** | **Accuracy** |
| KNN | 94.26% |
| Decision Tree | 98.20% |
| Naïve Bayes | 77.57% |
| SVM | 77.57% |

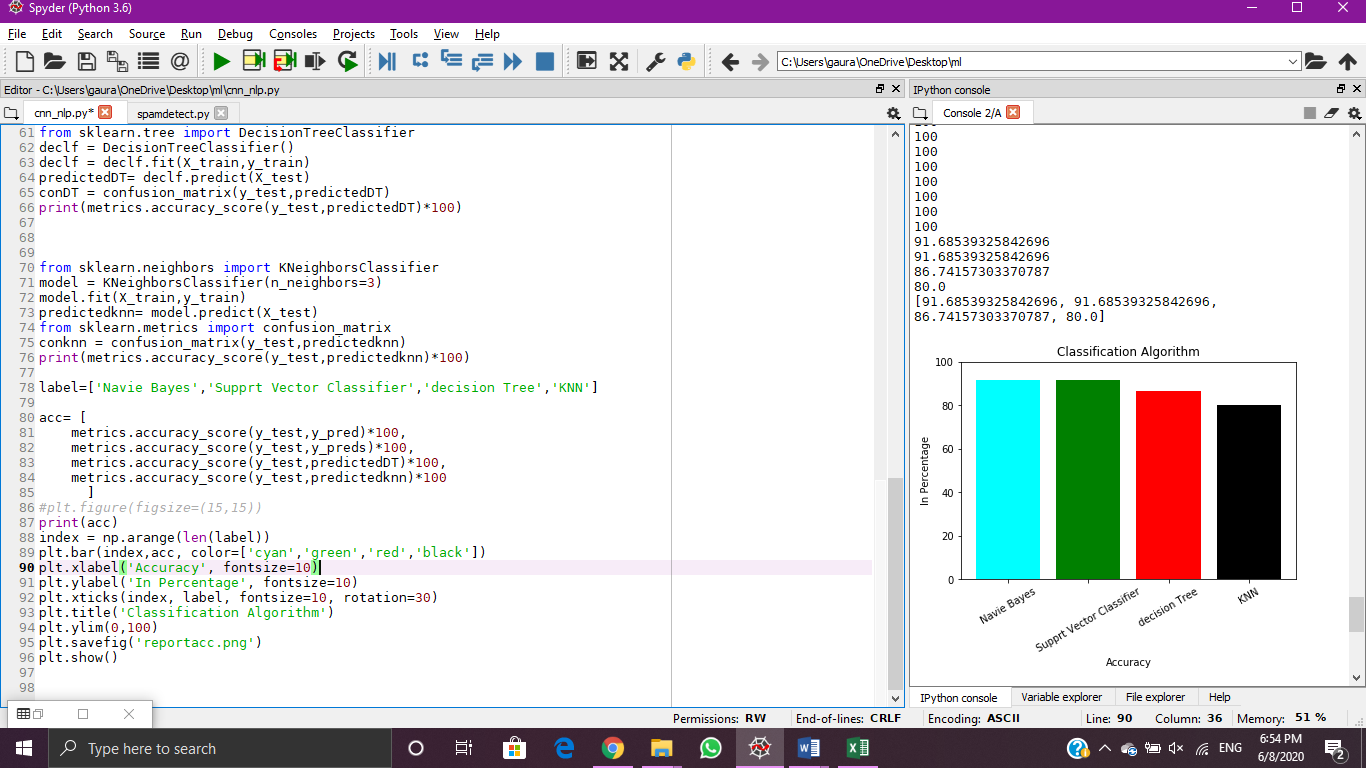
**9 Data Visualization and Results**

* **Data visualization of Confusion matrix**



**Naive Bayes SVM Classifier Decision Tree KNN**

* **Data visualization for comparing classification algorithm**



**Output of the above code :**

