**In class Programming Assignment - 5**

**GitHub Link:** <https://github.com/Viditha2277/CS-5720/tree/main/Assignment%205>

**Problem 1:** To implement Naïve Bayes method using scikit-learn library on glass.csv file

**Input**: .csv file

**Solution:**

1. Reading the data.csv file using read\_csv().

2. Entire data is spilt into training and testing data using train\_test\_splot() method imported from sklearn module.

3. Naïve Bayes model is imported from sklearn and loaded into “gnb” object. Then the training data and test data is used to tune the weights of the model.

4. Using predict() method, values are predicted for the test data.

5. Accuracy is calculated between the predicted and actual values using score().

6. Classification report is prepared that gives information on precision, recall, f1score, etc.

**Code:**

A computer screen shot of a program code

Description automatically generated

**Output:**

A screenshot of a computer

Description automatically generated

**Problem 2:** To demonstrate Support vector machine method on the glass.csv file

**Input:** .csv file

**Solution:**

1. Reading the data.csv file using read\_csv() and showing basic statistical description of the data using describe().

2. Entire data is spilt into training and testing data using train\_test\_splot() method imported from sklearn module.

3. SVM model is imported from sklearn and loaded into “svm” object. Then the training data and test data is used to tune the weights of the model.

4. Using predict() method, values are predicted for the test data.

5. Accuracy is calculated between the predicted and actual values using score().

6. Classification report is prepared that gives information on precision, recall, f1score, etc.

**Code:**

A computer screen with green and white text

Description automatically generated

**Output:**

A screenshot of a computer

Description automatically generated

The reason why Naïve Bayes algorithm performs better than Support Vector Machine is that Naïve Bayes algorithm is more resilient to missing values when compared to Support Vector Machine. The glass.csv file contains 0 in most of the places (all the missing values are replaced with 0’s) which reduces the performance of the Support Vector Machine.