1001720114

**Sweta Nitinbhai Gadhiya**

**Data Mining**

**Prerequisite To Run The Code:**

● You have python3 installed in your machine and I have used jupyter notebook editor.

● Install dependencies first before running the code (like pandas, numpy and similar which

can be found in the first few lines in the application code)

● Training and Test files should be there in the folder (I have put them in the folder.)

**Function Description for Naïve-Bayes:**

● loadDataset(filename):

○ This function is used to load dataset from space seperated file

● splitDataIntoTrainAndTest(trainingSet, randomValue):

○ This function is used to split train data into train and test data to build our model.

● findUnique(trainDataSet, featureViseUnique):

○ This function is used to find number of unique strings among the data

● calculateFeaturesprobability(trainDataSet,featuresProb, featuresColumn, featureViseUnique, countClassLableFreq):

○ This function is used to calculate all features probability.

● predictTestData(data, conditionalProb, classLableProb, countClassLableFreq, featureViseUnique):

○ Prediction for test data.

**NOTES:**

● It may take some time to execute the code depending on hardware configurations.

● More clear and elaborated comments were written down in the code so for more

information you can see the code comments.