Machine Learning: Programming Assignment 3: Naïve Bayes Classifier

Problem Statement:

Write a program to learn a naïve Bayes classifier and use it to predict class labels of test data. *Laplacian smoothing should be used*. The learned classifier should be tested on test instances and the accuracy of prediction for the test instances should be printed as output. A single program should train the classifier on the training set as well as test it on the test set.

Data Set Description:

The task is to predict whether a citizen is happy to live in a city based on certain parameters of the city as rated by the citizens in a scale of 1-5 during a survey.

Attribute Information:
D = decision/class attribute (D) with values 0 (unhappy) and 1 (happy) (Column 1 of file)
X1 = the availability of information about the city services (Column 2 of file)
X2 = the cost of housing
X3 = the overall quality of public schools
X4 = your trust in the local police
X5 = the maintenance of streets and sidewalks
X6 = the availability of social community events

Attributes X1 to X6 have values 1 to 5.

Training Data Filename: *data2_19.csv,* Test Data Filename: *test2_19.csv*

Submission Guidelines:

You may use one of the following languages: c/C++/Java/Python. You should name your file as <rollnumber_2.extension> (e.g., 15CS10001_2.c). Your program should be standalone and should not use any *special purpose* library. Numpy/pandas may be used. You should submit the program file only and not the output/input file. The submitted single program file *should* have the following header comments:

Roll # Name # Assignment number # Specific compilation/execution flags (if required)

Please submit the program in moodle by ***September 4, 2019 midnight*** *(hard deadline). Copying from friends/web will lead to strict penalties.*