School of Computer Science and Engineering, VIT Chennai.

BCSE209P Machine Learning

Lab-4 Naïve Bayes Classifier, KNN

Faculty : Dr. R. Jothi

Due Date : 06/02/2024

Submit your python code (Jupyter notebook): with output for all the questions.

Q1. Suppose you want to build Naïve Bayes classifier for predicting whether a cricket match will be played in the given weather conditions or not. Here the weather conditions are described by features outlook, temperature, humidity and wind. The target is play with two class labels *Yes* and *No*. (Dataset: play.csv)

a. Implement Naïve Bayes classifier which must
   - Print the class prior probabilities form the training set
   - Classify the test sample <Rain, Cool, High, Strong>. Need to print likelihood for <Rain, Cool, High, Strong> and also class conditional probabilities for "YES" and "NO" classes.

b. Use sklearn CategoricalNB to validate the results obtained above.

Q2. Assume you need to build a classifier to predict the graduate admissions into global universities for higher studies using parameters such as UG GPA, GRE score, research expertise, etc. Using the graduate.csv dataset, build a naïve bayes classifier to predict the graduate admissions chances. Report accuracy.

Q3. Apply K-NN for the graduate admissions problem. Compare accuracy of K-NN for different values of K.

Bonus marks

Q4. Plot the bar chart for target variable in graduate.csv. You may observe the class imbalance (ie. Class=1 more than Class=0). You may address this problem either by data augmentation (to create dummy samples for Class=0) or SMOTE algorithm for imbalance.

Those who show improved accuracy with the above said approached tested with both naïve bayes and K-NN will get bonus marks.