William Woodard
801069657

Homework 2

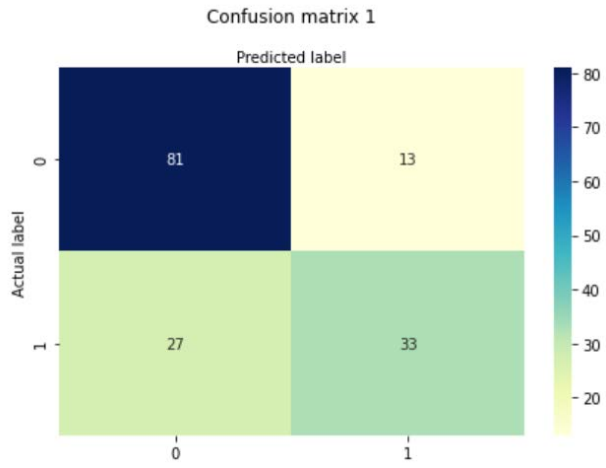
Problem 1

Accuracy: 0.7402597402597403

Precision: 0.717391304347826

Recall: 0.55

Confusion Matrix 1:



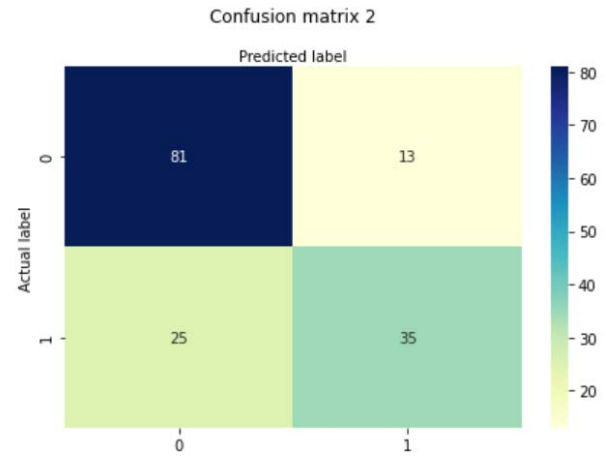
Problem 2

Accuracy: 0.7532467532467533

Precision: 0.7291666666666666

Recall: 0.5833333333333334

Confusion Matrix 2:



The values for accuracy, precision, and recall all seem to be slightly higher using the Naïve Bays binary classifier in problem 2 than they were in problem 1.

Problem 3

K = 5

Accuracy: 0.7703052112488338

Precision: 0.7153861192570871

Recall: 0.5439828606324659

K = 10

Accuracy: 0.7769169751454257

Precision: 0.7255430242272348

Recall: 0.559557984286245

As the value of K increases, the accuracy value increases. The accuracy value using k-fold method is already higher than problem 1 and it continues to increase as the K value is increased. The precision for K = 5 is lower than the precision from problem 1, but the precision for K = 10 is higher than problem 1. The recall is a similar case as the recall for K = 5 is less than problem 1 but the recall for K = 10 is greater than problem 1.

Problem 4

K = 5

Accuracy: 0.7556444089031055

Precision: 0.6546389287422286

Recall: 0.6004574788244792

K = 10

Accuracy: 0.7606292966684294

Precision: 0.6697937218365607

Recall: 0.6003783113022243

Like problem 3, the accuracy of the k-fold method is higher than the accuracy of problem 2. Furthermore, as the value of K-increases the accuracy continues to increase. The precision of the k-fold method seems to be less than the precision of problem 2, but it does seem to increase as the value of K increases. The recall of the k-fold method on the other hand seems to be higher than the recall, however it seems to decrease as the value of K increases.

This makes sense as k-fold method tends to be more precise and less biased. As the value of K increase, that means there are more random samples to analyze, thus making the results more accurate.

Github: https://github.com/willwoodard16/Intro-to-Ml-4105