Bryan Faryadi (801178567) - Homework 3

Github: <https://github.com/bfaryadi/ECGR4105_hw3_bf>

Problem 1

Results displayed in notebook

Problem 2

Adding the weight penalty in part 2 increased the performance of the model in every aspect. Notably, the number of false negatives decreased from 3 to 0. However, the number of false positives stayed the same. As such, the recall is 100%. Accuracy and F1 score also increased, and loss decreased by roughly 50%. Strangely, the precision increased by 0.002, despite the number of true positives and false positives staying the same.

Problem 3

Repeating problem 2 with a naive Bayes classifier resulted in poorer performance on all metrics compared to logistic regression (with a weight penalty), although not by much. Accuracy, precision, recall, and F1 score were all still above 0.9.

Problem 4

PCA with K=1 through K=12 generally results in worse performance than no PCA. Starting at K=13 and onward, however, the results are the same as problem 2.2 (logistic regression with weight penalties).

Problem 5

The Bayes classifier with PCA performed worse than without PCA for all values of K. The best performing number was K=5, which resulted in 2 more false positives than the model in problem 3, and the same number of false negatives. As K increased, all metrics went down.