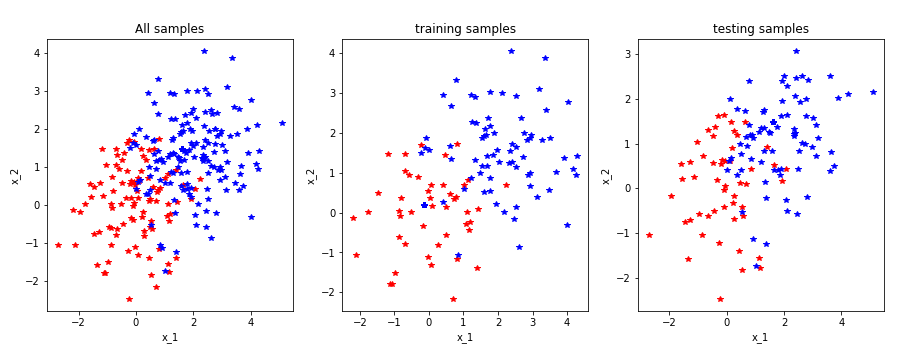
Ross Lewis

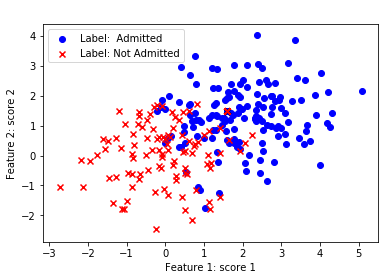
425 - Machine Learning

Problem I:

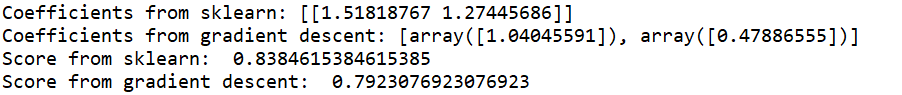
First, I split the data into training and testing samples:



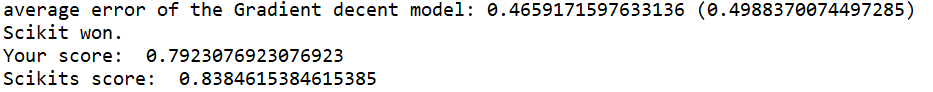
Here I visualize the data and can see there are clear clusters dividing the admitted and non-admitted group based on our features.



We learned the logistic regression coefficients through gradient decent and by using SKlearn. The SKlearn method ran faster and gave us a more accurate model.



After learning the model, I compared the score of each:



Problem II:

Confusion Matrix:

Actual

Cat Dog Monkey

Cat 1 3 2

Predicted Dog 3 3 2

Monkey 1 2 3

Accuracy = Correct predictions / total predictions

= 7 / 20 = .35

Cat Precision: My correct prediction of cat / My total prediction of cat

1/6

Cat Recall: My correct predictions of cat / Total actual cat

1/5

Dog Precision: 3/8

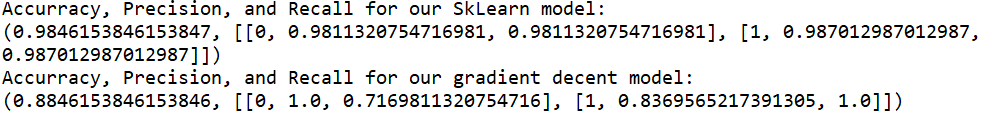
Dog Recall: 3/8

Monkey Precision: 3/6

Monkey Recall: 3/7

Problem III

This function calculates the accuracy of a given set of predictions with their corresponding true values. It also calculates precision and recall for each unique y value.



The first value is the accuracy, and each row of the array contains:

0 – the value

1 – the precision

2 – the recall

In addition to the accuracy, precision, and recall, I print out a small confusion matrix:

