ASSIGNMENT 7.2

Blaine Blasdell

2020-07-19

Assignment 7.2

Question 7.2a

a. What is the accuracy of the logistic regression classifier?

Answer - Using glm of the existing data set, I have a logistic regression classifier value of 58.3%

```
##
## Call:
## glm(formula = label ~ x + y, family = binomial(), data = bin_train)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.3728  -1.1697  -0.9575   1.1646   1.3989
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.424809   0.117224   3.624  0.00029 ***
## x           -0.002571   0.001823  -1.411  0.15836
## y           -0.007956   0.001869  -4.257 2.07e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 2075.8  on 1497  degrees of freedom
## Residual deviance: 2052.1  on 1495  degrees of freedom
## AIC: 2058.1
##
## Number of Fisher Scoring iterations: 4


##              Predicted_Value
## Actual_Value FALSE TRUE
##            0   429  338
##            1   286  445


## [1] 58.34446
```

Question 7.1b

    b. How does the accuracy of the logistic regression classifier compare to the nearest neighbors algorithm?

Answer - The accuracy of my KNN model is 96.97 %, which compared to the logistic regression classigfier (glm) was 58.34%. The KNN model appears to be more accurate (although it is probably introducing additional bias).

```
##          bin_df_target_test
## pred_bin   0   1
##        0 259   5
##        1  10 221
```

```
## [1] 96.9697
```

Question 7.1c

    c. To compute the accuracy of your model, use the dataset to predict the outcome variable. The percent of correct predictions is the accuracy of your model. What is the accuracy of your model?

Answer - KNN and logistic regression use two different algorithms for classification. KNN is useful when the data is closer in scale, as well as a smaller number of parameters or variables. When you increase the value of k, in the KNN classification, you are decreasing the variance and increasing bias. With a smaller value for k, you are increasing variance but decreasing bias. While you increase bias, you can negatively affect your accuracy. Becuase of these factors, KNN and logistic regression will result in different accuracy of the same model.

"'

References