10.2 Exercise Part 2

Justin Wisniewski

2022-05-21

Assignment Instructions:

Fit a logistic regression model to the binary-classifier-data.csv dataset What is the accuracy of the logistic regression classifier?

The accuracy came out to be 58.34%

```
setwd("C:/Users/jwiz3/Desktop/Data Statistics/dsc520")
binary_df <- read.csv("data/binary-classifier-data.csv")
binaryClassifier_df <- read.csv("data/binary-classifier-data.csv")
#Logistic Regression Model
binaryClassifier_glm <- glm(label ~ x + y, data=binaryClassifier_df, family = binomial)
summary(binaryClassifier_glm)
```

```
##
## Call:
## glm(formula = label ~ x + y, family = binomial, data = binaryClassifier_df)
##
## 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
```

```
res_val <- predict(binaryClassifier_glm, type="response")
bcPredictionData <- table(Actual_Value = binaryClassifier_df$label, Predicted_Value = res_val > 0.5)
bcPredictionData
```

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

```
dataModelAccuracy <- (bcPredictionData[[1,1]] + bcPredictionData[[2,2]]) / sum(bcPredictionData)
dataModelAccuracy
```

```
## [1] 0.5834446
```