10.2.2 Exercise

Felipe Rodriguez

2023-02-19

```
setwd('/Users/feliperodriguez/Library/CloudStorage/OneDrive-BellevueUniversity/Github/dsc520//data/')
binary_data <- read.csv('binary-classifier-data.csv')
```

Fit a logistic regression model to the binary-classifier-data.csv dataset

```
label_glm <- glm(label~., data=binary_data, family = binomial(link = "logit"))
summary(label_glm)
```

```
##
## Call:
## glm(formula = label ~ ., family = binomial(link = "logit"), data = binary_data)
##
## 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
```

```
response_label_glm <- predict(label_glm, type='response')
label_prediction <- table(Actual_Value = binary_data$label, Predicted_Value = response_label_glm > .5)
label_prediction
```

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

What is the accuracy of the logistic regression classifier?

```
label_accuracy <- ((label_prediction[[1,1]] + label_prediction[[2,2]]) / sum(label_prediction))
label_accuracy
```

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