ams580\_hw1

Mustafa Yigit Isik

2023-02-12

#Read the Data

data <- read.csv("/Users/mustafayigitisik/Desktop/stuff/semesters/spring 2023/ams 580/hw1\_glm/banknote.csv")

#Quick Look and Fix $class

str(data)

## 'data.frame': 1372 obs. of 5 variables:  
## $ variance: num 3.622 4.546 3.866 3.457 0.329 ...  
## $ skewness: num 8.67 8.17 -2.64 9.52 -4.46 ...  
## $ curtosis: num -2.81 -2.46 1.92 -4.01 4.57 ...  
## $ entropy : num -0.447 -1.462 0.106 -3.594 -0.989 ...  
## $ class : int 1 1 1 1 1 1 1 1 1 1 ...

head(data)

## variance skewness curtosis entropy class  
## 1 3.62160 8.6661 -2.8073 -0.44699 1  
## 2 4.54590 8.1674 -2.4586 -1.46210 1  
## 3 3.86600 -2.6383 1.9242 0.10645 1  
## 4 3.45660 9.5228 -4.0112 -3.59440 1  
## 5 0.32924 -4.4552 4.5718 -0.98880 1  
## 6 4.36840 9.6718 -3.9606 -3.16250 1

data$class <- as.factor(data$class)  
str(data)

## 'data.frame': 1372 obs. of 5 variables:  
## $ variance: num 3.622 4.546 3.866 3.457 0.329 ...  
## $ skewness: num 8.67 8.17 -2.64 9.52 -4.46 ...  
## $ curtosis: num -2.81 -2.46 1.92 -4.01 4.57 ...  
## $ entropy : num -0.447 -1.462 0.106 -3.594 -0.989 ...  
## $ class : Factor w/ 2 levels "0","1": 2 2 2 2 2 2 2 2 2 2 ...

is.factor(data$class)

## [1] TRUE

#Splitting the Data

library(lattice)  
library(ggplot2)  
library(caret)  
set.seed(123)  
cursor <- createDataPartition(data$class,p=0.8,list=FALSE)  
training <- data[cursor,]  
testing <- data[-cursor,]

#Regression Model

model <- glm (class ~ ., data = training, family ='binomial')

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

summary(model)

##   
## Call:  
## glm(formula = class ~ ., family = "binomial", data = training)  
##   
## Deviance Residuals:   
## Min 1Q Median 3Q Max   
## -2.49913 -0.00005 0.00000 0.00000 1.44236   
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -8.5895 2.1863 -3.929 8.54e-05 \*\*\*  
## variance 9.3611 2.4703 3.789 0.000151 \*\*\*  
## skewness 4.6968 1.2673 3.706 0.000210 \*\*\*  
## curtosis 6.1372 1.6414 3.739 0.000185 \*\*\*  
## entropy 0.5193 0.4208 1.234 0.217156   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 1508.568 on 1097 degrees of freedom  
## Residual deviance: 33.991 on 1093 degrees of freedom  
## AIC: 43.991  
##   
## Number of Fisher Scoring iterations: 13

All coefs are significant except entropy since it gave out 0.217156. This gives regression model of: P = 1 / (e^(-8.5895+9.3611(intercept)+4.6968(skewness)+6.1372(curtosis))) where P indicates the probability for genuinity.

#Confusion Matrix

predict\_testing<-predict(model, testing, type= 'response' )  
predict\_training<-predict(model, training, type= 'response' )  
predict\_testing<-predict(model, testing, type= 'response' )  
confusion\_matrix\_training <- table(actual = training$class, prediction = predict\_training > 0.8)  
confusion\_matrix\_testing <- table(actual = testing$class, prediction = predict\_testing > 0.8)  
confusion\_matrix\_training

## prediction  
## actual FALSE TRUE  
## 0 487 1  
## 1 10 600

#Accuracy

accuracy\_training<-sum(diag(confusion\_matrix\_training))/sum(confusion\_matrix\_training)  
accuracy\_training

## [1] 0.9899818

accuracy\_testing<-sum(diag(confusion\_matrix\_testing))/sum(confusion\_matrix\_testing)  
accuracy\_testing

## [1] 0.9854015

#Sensitivity

sensitivity\_training <-confusion\_matrix\_training[2,2]/(confusion\_matrix\_training[2,2]+confusion\_matrix\_training[2,1])  
sensitivity\_training

## [1] 0.9836066

sensitivity\_testing <-confusion\_matrix\_testing[2,2]/(confusion\_matrix\_testing[2,2]+confusion\_matrix\_testing[2,1])  
sensitivity\_testing

## [1] 0.9802632

#Specificity

specificity\_training = confusion\_matrix\_training[1,1]/(confusion\_matrix\_training[1,1]+confusion\_matrix\_training[1,2])  
specificity\_training

## [1] 0.9979508

specificity\_testing = confusion\_matrix\_testing[1,1]/(confusion\_matrix\_testing[1,1]+confusion\_matrix\_testing[1,2])  
specificity\_testing

## [1] 0.9918033