rm(list = ls())

setwd("C:\\Users\\ti1969\\Downloads")

library("readxl")

ds<--read\_excel("Final Data\_Report 3.xlsx")

head(ds)

ds= ds\*-1

ds$`Financial Risk` <- ifelse(ds$`Financial Risk`== 2, 1, 0)

attach(ds)

log.results=glm(`Financial Risk`~.,data=ds,family=binomial)

summary(log.results)

probs<-predict(log.results,type="response")

pred<-rep(0,nrow(ds)) # create a vector of predictions with default value of low risk

pred[probs>0.5]=1

table(`Financial Risk`,pred)

mean(`Financial Risk`==pred)

detach(ds)

set.seed(123)

trainingRowIndex <- sample(1:nrow(ds), 0.8\*nrow(ds))

train<-ds[trainingRowIndex,]

test<-ds[-trainingRowIndex,]

train.glm = glm(`Financial Risk`~.,data=ds,family=binomial)

summary(train.glm)

probs<-predict(train.glm,test,type="response")

m.pred<-rep(0,nrow(test))

m.pred[probs>0.5]<-1

table(test$`Financial Risk`,m.pred)

mean(m.pred==test$`Financial Risk`)

# With reduced # of observations

train.glm = glm(`Financial Risk`~.-RISK\_MORTALITY,data=ds,family=binomial)

summary(train.glm)

probs<-predict(train.glm,test,type="response")

m.pred<-rep(0,nrow(test))

m.pred[probs>0.5]<-1

table(test$`Financial Risk`,m.pred)

mean(m.pred==test$`Financial Risk`)

thresholds<-seq(from=0.1, to=0.9, by=0.05)

acc<-rep(0,19)

for (i in 1:19) {

newpred<-rep(0,nrow(test))

newpred[probs>thresholds[i]]<-1

table(test$`Financial Risk`,newpred)

acc[i]<-mean(newpred==test$`Financial Risk`)

acc[i]

}

acc

max(acc)

which.max(acc)