R code for Remission

#Libraries needed

library(glmnet)

library(dplyr)

library(psych)

library(caret)

#load data

data <- read.csv("/Users/krti1/Documents/823//STARDUNIQUE.csv")

View(data)

str(data)

ncol(data)

names(data)

#data preparation

data = subset(data, select = -c(1,2,4,5,6,7,8,9,10,11,13,14,17,19,20,22,23,27,30,31,32,33,36,40,41,43,44,45,46,47) )

data = subset(data, select = -c(1,2,3,5,6,9,12,13))

#Data partition

set.seed(12345)

remission<-sample(2, nrow(data), replace =T, prob=c(0.8,0.2))

train<- data[remission==1,]

test<-data[remission==2,]

str(train)

str(test)

summary(train)

# Custom Control Parameters

custom<-trainControl(method = "repeatedcv",

number=10,repeats =5,verboseIter=T)

#Lasso

set.seed(12345)

lasso<-train(remission~.,

train,method='glmnet',

tuneGrid=expand.grid(alpha=1,lambda=seq(0.001,1,length=100)),

trControl=custom)

#results

plot(lasso)

lasso

plot(lasso$finalModel,xvar='lambda',label=T)

plot(lasso$finalModel,xvar='dev',label=T)

plot(varImp(lasso,scale=T))

plot(varImp(lasso,scale=F))

#save final model

saveRDS(lasso,"final\_model.rds")

fm <- readRDS("final\_model.rds")

print

summary(fm)

p1 <- predict(fm, train)

a<-(sqrt(mean((train$remission-p1)^2)))

b<-(sqrt(mean((train$remission-mean(train$remission))^2)))

(1-(a/b))\*100

library(nnet)

library(ROCR)

lasso$bestTune

best <- lasso$finalModel

coef(best, s = lasso$bestTune$lambda)

mymodel<-multinom(lasso,data=data)

p<-predict(mymodel,data)

tab<-table(p,data$remission)

tab

sum(diag(tab))/sum(tab)

1-sum(diag(tab))/sum(tab)

pred<-predict(mymodel,data,type='prob')

pred<-prediction(pred,data$remission)

#receiver operating curve

roc<-performance(pred,"tpr","fpr")

plot(roc,colorize=T,main="ROC Curve",ylab="Sensitivity",xlab="1-Specificity")

abline(a=0,b=1)

#Area under curve(AUC)

auc<-performance(pred,"auc")

auc<-unlist(slot(auc,"y.values"))

auc<-round(auc,2)

auc

library(nnet)

mymodel<-multinom(lasso,data=data)

p<-predict(mymodel,data)

tab<-table(p,data$Citalopram)

tab

sum(diag(tab))/sum(tab)

1-sum(diag(tab))/sum(tab)

pred<-prediction(pred,data$remission)

#receiver operating curve

roc<-performance(pred,"tpr","fpr")

plot(roc,colorize=T,main="ROC Curve",ylab="Sensitivity",xlab="1-Specificity")

abline(a=0,b=1)

#Area under curve(AUC)

auc<-performance(pred,"auc")

auc<-unlist(slot(auc,"y.values"))

auc<-round(auc,2)

auc

#save final model

saveRDS(lasso,"final\_model.rds")

fm <- readRDS("final\_model.rds")

print

summary(fm)

#prediction

p1 <- predict(fm, train)

sqrt(mean((train$remission-p1)^2))

p2 <- predict(fm, test)

sqrt(mean((test$remission-p2)^2))