setwd("E:\Acadgild\R\project1")

data<-read.csv("Patient Adherence - Data.csv")

head(data)

table(data$Medication)

sum(is.na(data))

which(is.na(data))

#not nessary imputation for this data set

library(Hmisc)

summary(data)

tail(data)

data$AmountPaid<- with(data, impute(AmountPaid, 205))

str(data)

library(dplyr)

data$Date=as.Date(data$Date)

data1<-(data %>%

group\_by(Medication,PatientID) %>%

mutate(datedif = Date - lag(Date, default = Date[1])))

View(data1)

data1<-as.data.frame(data1)

data1$datedif

names(data1)

str(data1)

data1$datedif<- as.numeric(as.character(data1$datedif))

head(data1,10)

View(data1)

data1$adher<- NULL

data1[data1$datedif>data1$For\_How\_Many\_Days,"adher"]<-"no"

data1[data1$datedif<(data1$For\_How\_Many\_Days+2),"adher"]<-"yes"

data1[data1$datedif==0,"adher"]<-"1st"

head(data1)

data1$adher

table(data1$adher)

data1<- data1[ ! data1$adher %in% "1st", ]

data1$adher<-as.factor(data1$adher)

data1$Medication=as.numeric(data1$Medication)

str(data1)

set.seed(123)

View(data1)

#data1$PatientID<-NULL

#data1$Medication<-NULL

#data1$Date<-NULL

#data1$Pharmacy<-NULL

library(Hmisc)

data1$PatientID

table(data1$PatientID)

ind <- sample(2, nrow(data1), replace = TRUE, prob = c(0.7, 0.3))

train <- data1[ind==1,]

test <- data1[ind==2,]

head(train)

train<-subset(data1, PatientID <= 1010)

table(train$PatientID)

test<-subset(data1, PatientID > 1010)

table(test$PatientID)

library(randomForest)

rf<-randomForest(adher~.,data =train)

rf

varImpPlot(rf)

library(caret)

p1<-predict(rf,train)

confusionMatrix(p1,train$adher)

p2<-predict(rf,test)

confusionMatrix(p2,test$adher)

#####cross validation

control <- trainControl(method = 'repeatedcv',

number = 5,

repeats = 3)

set.seed(7)

mtry <- sqrt(ncol(train))

rf\_random <- train(adher~.,

data = train,

method = 'rf',

metric = 'Accuracy',

tuneLength = 7,

trControl = control

)

print(rf\_random)

plot(rf\_random)

predictions<- predict(rf\_random,test)

# append predictions

pred<- cbind(test,predictions)

confusionMatrix<- confusionMatrix(pred$predictions,pred$adher)

confusionMatrix

varImp(rf\_random)

####end############

train$adher

data1<- data1[ ! data1$adher %in% "1st", ]

control <- trainControl(method = 'repeatedcv',

number = 10,

repeats = 3)

library(caret)

seed <-7

metric <- 'Accuracy'

set.seed(seed)

lg\_reg <- train(adher~.,

data = train,

method = 'bayesglm',

metric = metric,

trControl = control,

maxit=100

)

predictions<- predict(lg\_reg,test)

pred<- cbind(test,predictions)

confusionMatrix<- confusionMatrix(pred$predictions,pred$adher)