setwd("F:\\files\\r project")

hr.data=read.csv("hr\_train.csv ",stringsAsFactors = FALSE)

hr.test=read.csv("hr\_test.csv ",stringsAsFactors = FALSE)

library(dplyr)

hr.data<- hr.data %>% rename("department" = "sales") %>% tbl\_df()

hr.test<- hr.test %>% rename("department" = "sales") %>% tbl\_df()

glimpse(hr.data)

which(is.na.data.frame(hr.data))

summary(hr.data)

library(ggplot2)

#

library(caret)

set.seed(20)

test.data=hr.data[rbinom(20, 10, 0.5),]

inTrain=createDataPartition(hr.data$left, p=0.6, list=FALSE)

training.data=hr.data[inTrain,]

testing.data=hr.data[-inTrain,]

dim(training.data)

dim(testing.data)

#

fitControl=trainControl(method = "repeatedcv",

number = 3,

repeats = 1)

mod\_BR=train(left ~., training.data, method="gbm", trControl=fitControl, verbose = TRUE)

plot(mod\_BR, main = "model GBM")

predict\_BR=predict(mod\_BR, testing.data)

cm=table(testing.data$left,predict\_BR)

accuracy <- sum(cm[1], cm[4]) / sum(cm[1:4])

precision <- cm[4] / sum(cm[4], cm[2])

sensitivity <- cm[4] / sum(cm[4], cm[3])

fscore <- (2 \* (sensitivity \* precision))/(sensitivity + precision)

specificity <- cm[1] / sum(cm[1], cm[2])

final\_predict=predict(mod\_BR, hr.test)

head(final\_predict, 5)