

```

install.packages("caret")
install.packages("ROCR")
install.packages("e1071")

library(caret)
library(ROCR)
library(e1071)

credito_dataset <- read.csv("original.csv", header= TRUE, sep = ",")
head(credito_dataset)
summary(credito_dataset)
str(credito_dataset)

credito_dataset <- credito_dataset[,-1]

credito_dataset[['age']] <- as.integer(credito_dataset[['age']])

to.factors <- function(df, variables){
  for (variable in variables){
    df[[variable]] <- as.factor(df[[variable]])
  }
  return(df)
}

scale.features <- function(df, variables){
  for (variable in variables){
    df[[variable]] <- scale(df[[variable]], center = T, scale = T)
  }
  return(df)
}

credito_dataset[credito_dataset < 0] <- NA

credito_dataset[['loan']][credito_dataset[['loan']] < 1000] <- NA

credito_dataset <- na.omit(credito_dataset)

head(credito_dataset)
summary(credito_dataset)
str(credito_dataset)

numerics.vars <- c('income', 'age', 'loan')
credito_scaled <- scale.features(credito_dataset, numerics.vars)

categoric.vars <- c('default')
credito_dataset_final <- to.factors(df= credito_scaled,
                                   variables = categoric.vars)

str(credito_dataset_final)

indexes <- sample(1:nrow(credito_dataset_final),
                 size = 0.6 * nrow(credito_dataset_final))
train.data <- credito_dataset_final[indexes,]
test.data <- credito_dataset_final[-indexes,]

```

```

class(train.data)
class(test.data)

test.feature.vars <- test.data[,-4]
test.class.var <- test.data[,4]
class(test.feature.vars)
class(test.class.var)

formula.init <- 'default ~ .'
formula.init <- as.formula(formula.init)

modelo_v1 <- glm(formula = formula.init, data = train.data, family =
'binomial')

summary(modelo_v1)

previsoes <- predict(modelo_v1, test.data, type = 'response')
previsoes <- round(previsoes)

View(previsoes)

confusionMatrix(table(data = previsoes,
                      reference = test.class.var), positive = '1')

modelo_final <- modelo_v1

previsoes <- predict(modelo_final, test.feature.vars, type = 'response')
avaliacao <- prediction(previsoes, test.class.var)

plot.roc.curve <- function(predictions, title.text){
  perf <- performance(avaliacao, 'tpr', 'fpr')
  plot(perf, col = 'black', lty = 1, lwd = 2, main = title.text,
       cex.main = 0.6, cex.lab = 0.8, xaxs = 'i', yaxas = 'i')
  abline(0,1, col = 'red')
  auc <- performance(avaliacao, 'auc')
  auc <- unlist(slot(auc, 'y.values'))
  auc <- round(auc, 2)
  legend(0.4,0.4, legend = c(paste0('AUC: ', auc)),
        cex = 0.6, bty = 'n', box.col = 'white')
}

par(mfrow = c(1,2))
plot.roc.curve(avaliacao, title.text = 'Curva ROC')

income <- c(23456, 12034, 18765, 39086)
age <- c(31, 23, 44, 29)
loan <- c(3420, 3987, 1987, 4300)

novo_data_set <- data.frame(income, age, loan)

class(novo_data_set)
View(novo_data_set)

new.numeric.vars <- c('income', 'age', 'loan')

novo_dataset_final <- scale.features(novo_data_set, new.numeric.vars)
str(novo_dataset_final)

```

```
previsoes_novo_cliente <- predict(modelo_final, newdata =  
  novo_dataset_final,  
                                   type = 'response')  
round(previsoes_novo_cliente)
```