TCD Project note III

Suprasanna Pradhan

15 October 2019

```r
#Loading required packages
library(tidyverse)
```

```
## -- Attaching packages --------------------------------- tidyverse 1.2.1 --
```

```
## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   0.8.3     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
```

```
## -- Conflicts ------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(ggplot2)
library(caret)
```

```
## Loading required package: lattice
```

```
##
## Attaching package: 'caret'
```

```
## The following object is masked from 'package:purrr':
##
##     lift
```

```r
library(caretEnsemble)
```

```
##
## Attaching package: 'caretEnsemble'
```

```
## The following object is masked from 'package:ggplot2':
##
##     autoplot
```

```r
library(psych)
```

```
##
## Attaching package: 'psych'
```

```
## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```r
library(Amelia)
```

```
## Loading required package: Rcpp
```

```
## ##
## ## Amelia II: Multiple Imputation
## ## (Version 1.7.5, built: 2018-05-07)
## ## Copyright (C) 2005-2019 James Honaker, Gary King and Matthew Blackwell
## ## Refer to http://gking.harvard.edu/amelia/ for more information
## ##
```

```r
library(mice)
```

```
##
## Attaching package: 'mice'
```

```
## The following object is masked from 'package:tidyr':
##
##     complete
```

```
## The following objects are masked from 'package:base':
##
##     cbind, rbind
```

```r
library(GGally)
```

```
## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```
##
## Attaching package: 'GGally'
```

```
## The following object is masked from 'package:dplyr':
##
##     nasa
```

```
library(gutenbergr)
library(tidytext)
library(dplyr)
library(janeaustenr)
library(stringi)
library(tidyr)
library(rpart)
library(randomForest)
```

```
## randomForest 4.6-14
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
##
## Attaching package: 'randomForest'
```

```
## The following object is masked from 'package:psych':
##
##     outlier
```

```
## The following object is masked from 'package:dplyr':
##
##     combine
```

```
## The following object is masked from 'package:ggplot2':
##
##     margin
```

Importing Data set

```
library(readxl)
setwd("C:/Users/SuprasannaPradhan/Documents/My Files/Great Lakes Projects/Capstone Pro
ject TCD")
train_data1=read.csv("train_new.csv")
test_data1=read.csv("test_data.csv")
```

```
train_bank <- subset (train_data1, select= -c(1))
test_bank <- subset (test_data1, select= -c(1))
str(train_bank)
```

```
## 'data.frame':    9291 obs. of  14 variables:
##  $ LIMIT_BAL      : num  20000 450000 100000 30000 20000 20000 270000 50000 120000
200000 ...
##  $ SEX            : int  1 1 1 2 2 1 2 1 1 1 ...
##  $ EDUCATION      : int  2 1 2 2 2 2 2 3 2 2 ...
##  $ MARRIAGE       : int  2 1 1 2 1 1 2 1 2 1 ...
##  $ AGE            : int  33 45 30 22 24 31 26 53 28 32 ...
##  $ DEFAULT        : int  0 1 1 1 1 1 1 0 0 1 ...
##  $ BILLED._AMT    : num  -0.391 -0.492 0.455 -0.208 -0.647 ...
##  $ REPAY_STATUS.  : num  0.1827 0.3363 0.0201 -0.2731 2.5181 ...
##  $ PAID_AMT       : num  -0.7188 1.278 -0.0873 -0.4301 -0.5607 ...
##  $ TIMELY_PAID_AMT: num  0.543 0.523 0.431 -1.422 -1.695 ...
##  $ RATIO_PADI_AMT1: num  0.419 -0.546 -0.411 -0.44 -0.027 ...
##  $ RATIO_PADI_AMT2: num  0.37743 -2.66432 0.00851 -0.35402 -0.03372 ...
##  $ RATIO_PADI_AMT3: num  0.3778 1.6226 -0.1 -0.0637 0.1059 ...
##  $ RATIO_PADI_AMT4: num  0.62 0.416 0.325 0.922 0.141 ...
```

```
str(test_bank)
```

```
## 'data.frame':    9000 obs. of  14 variables:
##  $ LIMIT_BAL      : num  20000 120000 90000 50000 50000 20000 430000 180000 50000 5
00000 ...
##  $ SEX            : int  2 2 2 2 1 1 2 2 1 1 ...
##  $ EDUCATION      : int  2 2 2 2 2 3 2 1 2 1 ...
##  $ MARRIAGE       : int  1 2 2 1 1 2 2 2 2 1 ...
##  $ AGE            : int  24 26 34 37 57 35 41 29 33 58 ...
##  $ DEFAULT        : int  1 1 0 0 0 0 0 0 1 0 ...
##  $ BILLED._AMT    : num  -0.6551 -0.8545 -0.5227 -0.0487 -0.8533 ...
##  $ REPAY_STATUS.  : num  -1.56607 1.05357 0.09609 -0.00254 0.26657 ...
##  $ PAID_AMT       : num  -0.585 -0.416 -0.3 -0.435 1.929 ...
##  $ TIMELY_PAID_AMT: num  -2.74 0.164 0.487 0.527 -0.191 ...
##  $ RATIO_PADI_AMT1: num  -0.2887 -0.0338 -0.4383 -0.3948 -0.1997 ...
##  $ RATIO_PADI_AMT2: num  -0.334 0.465 0.44 0.403 -1.972 ...
##  $ RATIO_PADI_AMT3: num  0.439 -0.132 0.264 0.548 0.866 ...
##  $ RATIO_PADI_AMT4: num  0.152 1.461 0.74 0.567 1.09 ...
```

```
table(train_bank$DEFAULT)
```

```
## 
##    0    1 
## 4643 4648
```

Scale the data

```
#scale(train_bank)
```

Simple logit model on all variables

```
set.seed(1080)
bank_lg<- glm(DEFAULT ~ ., train_bank, family = "binomial"(link="logit"))
summary(bank_lg)
```

```
##
## Call:
## glm(formula = DEFAULT ~ ., family = binomial(link = "logit"),
##     data = train_bank)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.5842  -0.9796   0.2938   0.9860   2.4416
##
## Coefficients:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)      1.891e-01  1.828e-01   1.034 0.301045
## LIMIT_BAL       -8.208e-07  2.381e-07  -3.448 0.000566 ***
## SEX             -1.092e-01  4.743e-02  -2.303 0.021266 *
## EDUCATION       -3.273e-02  3.225e-02  -1.015 0.310172
## MARRIAGE        -1.460e-01  4.868e-02  -2.999 0.002706 **
## AGE              5.595e-03  2.780e-03   2.013 0.044138 *
## BILLED._AMT     -3.608e-02  2.500e-02  -1.443 0.148979
## REPAY_STATUS.    4.129e-01  2.412e-02  17.122  < 2e-16 ***
## PAID_AMT        -4.515e-01  3.196e-02 -14.126  < 2e-16 ***
## TIMELY_PAID_AMT -6.899e-01  2.354e-02 -29.306  < 2e-16 ***
## RATIO_PADI_AMT1 -9.910e-02  2.796e-02  -3.544 0.000394 ***
## RATIO_PADI_AMT2 -8.048e-02  2.608e-02  -3.086 0.002026 **
## RATIO_PADI_AMT3  8.423e-03  2.774e-02   0.304 0.761400
## RATIO_PADI_AMT4 -1.915e-01  2.616e-02  -7.322 2.44e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 12880  on 9290  degrees of freedom
## Residual deviance: 10973  on 9277  degrees of freedom
## AIC: 11001
##
## Number of Fisher Scoring iterations: 4
```

Revised Logistic Regression Model

```
#View(test_bank)
logit_f1 <- subset(test_bank,select= -c(3,4,7,12,13))
#logit_f1
bank_f1_lg<- glm(DEFAULT ~ ., logit_f1, family = "binomial"(link="logit"))
summary(bank_f1_lg)
```

```
##
## Call:
## glm(formula = DEFAULT ~ ., family = binomial(link = "logit"),
##     data = logit_f1)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.1535  -0.6108  -0.5145  -0.3320   2.8482
##
## Coefficients:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -1.346e+00  1.516e-01  -8.881  < 2e-16 ***
## LIMIT_BAL       -1.199e-06  2.770e-07  -4.329 1.50e-05 ***
## SEX             -1.252e-01  5.754e-02  -2.176  0.02955 *
## AGE              7.655e-03  3.023e-03   2.532  0.01134 *
## REPAY_STATUS.    4.632e-01  2.707e-02  17.112  < 2e-16 ***
## PAID_AMT        -3.757e-01  4.169e-02  -9.011  < 2e-16 ***
## TIMELY_PAID_AMT -7.051e-01  2.598e-02 -27.136  < 2e-16 ***
## RATIO_PADI_AMT1 -9.419e-02  3.626e-02  -2.598  0.00939 **
## RATIO_PADI_AMT4 -1.508e-01  3.003e-02  -5.021 5.13e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 9504.6  on 8999  degrees of freedom
## Residual deviance: 7997.2  on 8991  degrees of freedom
## AIC: 8015.2
##
## Number of Fisher Scoring iterations: 5
```

```
library(car)
```

```
## Loading required package: carData
```

```
## Registered S3 methods overwritten by 'car':
##   method                          from
##   influence.merMod                lme4
##   cooks.distance.influence.merMod lme4
##   dfbeta.influence.merMod         lme4
##   dfbetas.influence.merMod        lme4
```

```
##
## Attaching package: 'car'
```

```
## The following object is masked from 'package:psych':
##
##     logit
```

```
## The following object is masked from 'package:dplyr':
##
##     recode
```

```
## The following object is masked from 'package:purrr':
##
##     some
```

```
vif(bank_f1_lg)
```

```
##       LIMIT_BAL             SEX             AGE   REPAY_STATUS.
##        1.256936        1.010270        1.031040        1.111527
##        PAID_AMT TIMELY_PAID_AMT RATIO_PADI_AMT1 RATIO_PADI_AMT4
##        1.096350        1.018324        1.038484        1.024806
```

Predcting

```
logit_pred_f1 = predict.glm(bank_f1_lg,newdata =test_bank,type = "response")
test_bank1<- cbind(test_bank,logit_pred_f1)
```

Checking Accuracy

```
#View(test_bank1)
table(test_bank$DEFAULT,logit_pred_f1>0.5)
```

```
##
##     FALSE TRUE
##   0  6672  340
##   1  1342  646
```

```
Accuracy = (6672+646)/(6672+340+1342)
Accuracy
```

```
## [1] 0.8759876
```

Validation on test data

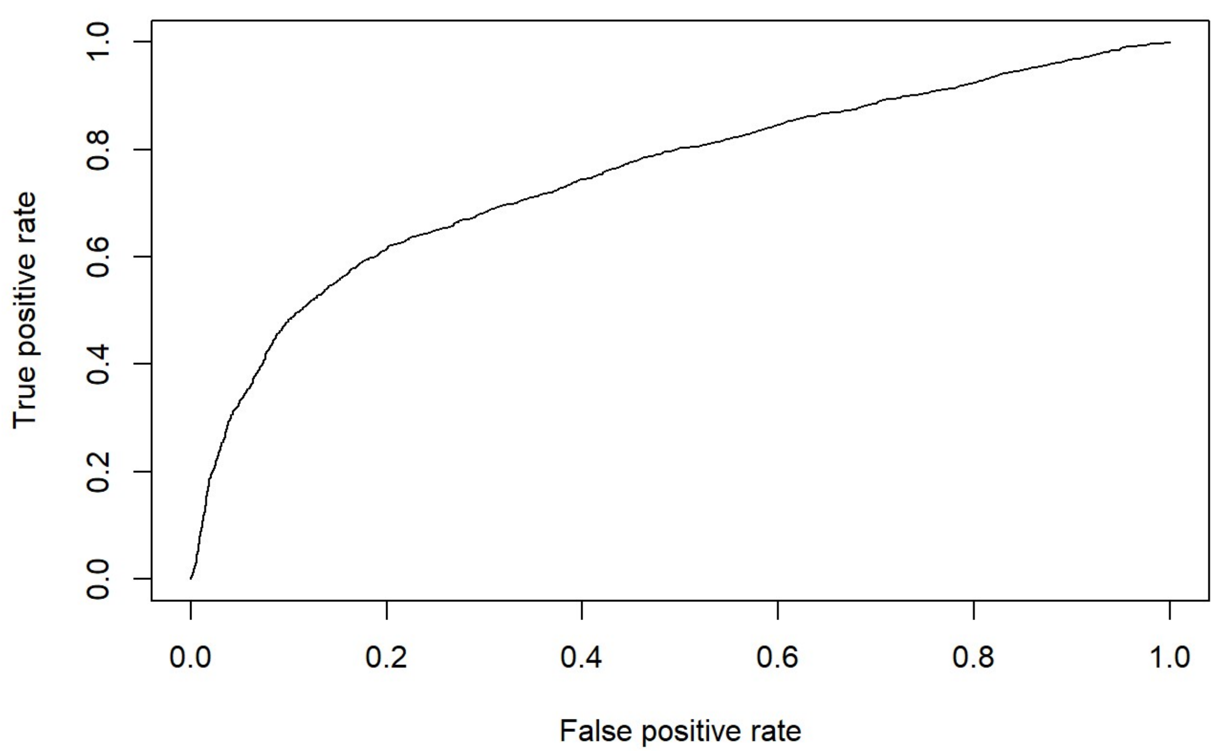
```
library(ROCR)
```

```
## Loading required package: gplots
```

```
##
## Attaching package: 'gplots'
```

```
## The following object is masked from 'package:stats':
##
##     lowess
```

```
DTpredROC1 = ROCR::prediction(logit_pred_f1, test_bank$DEFAULT)
perf1 = performance(DTpredROC1, "tpr", "fpr")
plot(perf1)
```



```
auc_lg <- as.numeric(performance(DTpredROC1, "auc")@y.values)
auc_lg
```

```
## [1] 0.7520219
```

```
KS <- max(attr(perf1, 'y.values')[[1]]-attr(perf1, 'x.values')[[1]])
KS
```

```
## [1] 0.4184996
```

Gini

```
## Gini Coefficient
library(ineq)
gini = ineq(test_bank$DEFAULT, type="Gini")
gini
```

```
## [1] 0.7791111
```

```
###############################################
#CART Model
library(rpart)
library(rpart.plot)
r.ctrl = rpart.control(minsplit = 100, minbucket = 10, cp = 0, xval = 10)
CT_model = rpart(DEFAULT ~ ., data = train_bank, method = "class", control = r.ctrl)
CT_model
```

```
## n= 9291
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##      1) root 9291 4643 1 (0.49973092 0.50026908)
##        2) TIMELY_PAID_AMT>=-0.6144733 6272 2319 0 (0.63026148 0.36973852)
##          4) REPAY_STATUS.< 1.000148 5686 1940 0 (0.65881112 0.34118888)
##            8) PAID_AMT>=0.7048767 1012  216 0 (0.78656126 0.21343874)
##             16) PAID_AMT>=2.234394 228   17 0 (0.92543860 0.07456140) *
##             17) PAID_AMT< 2.234394 784  199 0 (0.74617347 0.25382653)
##               34) RATIO_PADI_AMT1>=-0.8005472 671  155 0 (0.76900149 0.23099851)
##                 68) RATIO_PADI_AMT1< -0.3804398 186   27 0 (0.85483871 0.14516129)
*
##                 69) RATIO_PADI_AMT1>=-0.3804398 485  128 0 (0.73608247 0.26391753)
##                  138) RATIO_PADI_AMT1>=-0.3492855 471  118 0 (0.74946921 0.2505307
9)
##                    276) TIMELY_PAID_AMT>=0.2063298 298   60 0 (0.79865772 0.2013422
8) *
##                    277) TIMELY_PAID_AMT< 0.2063298 173   58 0 (0.66473988 0.3352601
2)
##                      554) RATIO_PADI_AMT2>=-0.1305164 122   33 0 (0.72950820 0.2704
9180)
##                       1108) RATIO_PADI_AMT4>=-0.4187486 112   27 0 (0.75892857 0.24
107143) *
##                       1109) RATIO_PADI_AMT4< -0.4187486 10    4 1 (0.40000000 0.600
00000) *
##                      555) RATIO_PADI_AMT2< -0.1305164 51   25 0 (0.50980392 0.49019
608) *
##                  139) RATIO_PADI_AMT1< -0.3492855 14    4 1 (0.28571429 0.71428571)
*
##               35) RATIO_PADI_AMT1< -0.8005472 113   44 0 (0.61061947 0.38938053)
##                 70) AGE< 38.5 74   22 0 (0.70270270 0.29729730) *
##                 71) AGE>=38.5 39   17 1 (0.43589744 0.56410256) *
##            9) PAID_AMT< 0.7048767 4674 1724 0 (0.63115105 0.36884895)
##             18) BILLED._AMT>=-0.4255529 2411  747 0 (0.69017005 0.30982995)
##               36) LIMIT_BAL>=75000 1658  452 0 (0.72738239 0.27261761)
##                 72) RATIO_PADI_AMT1< -0.1935291 998  226 0 (0.77354709 0.22645291)
##                  144) RATIO_PADI_AMT3< 0.6952995 888  182 0 (0.79504505 0.20495495)
##                    288) RATIO_PADI_AMT4>=0.210228 582  100 0 (0.82817869 0.1718213
1) *
##                    289) RATIO_PADI_AMT4< 0.210228 306   82 0 (0.73202614 0.2679738
6)
##                      578) LIMIT_BAL>=145000 202   42 0 (0.79207921 0.20792079)
##                       1156) REPAY_STATUS.>=-0.3865024 177   28 0 (0.84180791 0.1581
9209) *
##                       1157) REPAY_STATUS.< -0.3865024 25   11 1 (0.44000000 0.56000
000) *
```

```
##                        579) LIMIT_BAL< 145000 104   40 0 (0.61538462 0.38461538)
##                         1158) RATIO_PADI_AMT4< 0.1351998 81   25 0 (0.69135802 0.3086
4198) *
##                         1159) RATIO_PADI_AMT4>=0.1351998 23    8 1 (0.34782609 0.6521
7391) *
##                    145) RATIO_PADI_AMT3>=0.6952995 110   44 0 (0.60000000 0.40000000)
##                      290) RATIO_PADI_AMT1>=-0.3811395 57   11 0 (0.80701754 0.1929824
6) *
##                      291) RATIO_PADI_AMT1< -0.3811395 53   20 1 (0.37735849 0.6226415
1) *
##                   73) RATIO_PADI_AMT1>=-0.1935291 660  226 0 (0.65757576 0.34242424)
##                    146) REPAY_STATUS.< -0.4930504 148   33 0 (0.77702703 0.22297297)
##                      292) RATIO_PADI_AMT2>=-0.9765527 133   25 0 (0.81203008 0.187969
92) *
##                      293) RATIO_PADI_AMT2< -0.9765527 15    7 1 (0.46666667 0.5333333
3) *
##                    147) REPAY_STATUS.>=-0.4930504 512  193 0 (0.62304688 0.37695312)
##                      294) RATIO_PADI_AMT4>=-0.7237819 469  168 0 (0.64179104 0.358208
96)
##                        588) PAID_AMT>=-1.279013 455  158 0 (0.65274725 0.34725275)
##                         1176) RATIO_PADI_AMT2< -0.3540127 51    9 0 (0.82352941 0.176
47059) *
##                         1177) RATIO_PADI_AMT2>=-0.3540127 404  149 0 (0.63118812 0.36
881188)
##                           2354) RATIO_PADI_AMT2>=-0.2564896 384  135 0 (0.64843750 0.
35156250)
##                             4708) REPAY_STATUS.>=-0.4200141 371  126 0 (0.66037736 0.
33962264)
##                               9416) TIMELY_PAID_AMT< 0.7034911 331  105 0 (0.68277946
0.31722054)
##                                18832) PAID_AMT< -0.6879101 20    1 0 (0.95000000 0.05
000000) *
##                                18833) PAID_AMT>=-0.6879101 311  104 0 (0.66559486 0.3
3440514)
##                                  37666) RATIO_PADI_AMT1< 2.907036 287   89 0 (0.68989
547 0.31010453)
##                                    75332) REPAY_STATUS.>=-0.0978989 52    6 0 (0.8846
1538 0.11538462) *
##                                    75333) REPAY_STATUS.< -0.0978989 235   83 0 (0.646
80851 0.35319149)
##                                     150666) PAID_AMT>=-0.3456294 189   59 0 (0.687830
69 0.31216931)
##                                       301332) TIMELY_PAID_AMT>=0.5027466 15    0 0
(1.00000000 0.00000000) *
##                                       301333) TIMELY_PAID_AMT< 0.5027466 174   59 0
(0.66091954 0.33908046)
##                                         602666) TIMELY_PAID_AMT< 0.4392857 148   42 0
(0.71621622 0.28378378) *
##                                         602667) TIMELY_PAID_AMT>=0.4392857 26    9 1
```

```
(0.34615385 0.65384615) *
##                                    150667) PAID_AMT< -0.3456294 46   22 1 (0.4782608
7 0.52173913) *
##                                 37667) RATIO_PADI_AMT1>=2.907036 24    9 1 (0.375000
00 0.62500000) *
##                              9417) TIMELY_PAID_AMT>=0.7034911 40   19 1 (0.47500000
0.52500000) *
##                            4709) REPAY_STATUS.< -0.4200141 13    4 1 (0.30769231 0.6
9230769) *
##                          2355) RATIO_PADI_AMT2< -0.2564896 20    6 1 (0.30000000 0.7
0000000) *
##                       589) PAID_AMT< -1.279013 14    4 1 (0.28571429 0.71428571) *
##                     295) RATIO_PADI_AMT4< -0.7237819 43   18 1 (0.41860465 0.5813953
5) *
##                 37) LIMIT_BAL< 75000 753  295 0 (0.60823373 0.39176627)
##                   74) REPAY_STATUS.>=-0.02558607 577  210 0 (0.63604853 0.36395147)
##                    148) RATIO_PADI_AMT4>=0.5302308 310   97 0 (0.68709677 0.31290323)
##                      296) PAID_AMT< -0.4174805 149   34 0 (0.77181208 0.22818792) *
##                      297) PAID_AMT>=-0.4174805 161   63 0 (0.60869565 0.39130435)
##                        594) BILLED._AMT>=-0.1486893 42    9 0 (0.78571429 0.21428571)
*
##                        595) BILLED._AMT< -0.1486893 119   54 0 (0.54621849 0.4537815
1)
##                         1190) BILLED._AMT< -0.3171202 54   17 0 (0.68518519 0.3148148
1) *
##                         1191) BILLED._AMT>=-0.3171202 65   28 1 (0.43076923 0.5692307
7) *
##                    149) RATIO_PADI_AMT4< 0.5302308 267  113 0 (0.57677903 0.42322097)
##                      298) LIMIT_BAL>=35000 191   71 0 (0.62827225 0.37172775)
##                        596) BILLED._AMT< -0.244037 30    4 0 (0.86666667 0.13333333)
*
##                        597) BILLED._AMT>=-0.244037 161   67 0 (0.58385093 0.41614907)
##                         1194) REPAY_STATUS.< 0.0508605 90   28 0 (0.68888889 0.311111
11) *
##                         1195) REPAY_STATUS.>=0.0508605 71   32 1 (0.45070423 0.549295
77) *
##                      299) LIMIT_BAL< 35000 76   34 1 (0.44736842 0.55263158) *
##                   75) REPAY_STATUS.< -0.02558607 176   85 0 (0.51704545 0.48295455)
##                    150) AGE>=53.5 10    1 0 (0.90000000 0.10000000) *
##                    151) AGE< 53.5 166   82 1 (0.49397590 0.50602410)
##                      302) REPAY_STATUS.< -0.03234571 141   66 0 (0.53191489 0.4680851
1)
##                        604) BILLED._AMT< 0.07650831 114   48 0 (0.57894737 0.4210526
3)
##                         1208) RATIO_PADI_AMT3< -0.02379037 14    2 0 (0.85714286 0.14
285714) *
##                         1209) RATIO_PADI_AMT3>=-0.02379037 100   46 0 (0.54000000 0.4
6000000)
##                           2418) BILLED._AMT>=-0.3023209 67   26 0 (0.61194030 0.38805
```

```
970) *
##                           2419) BILLED._AMT< -0.3023209 33   13 1 (0.39393939 0.60606
061) *
##                        605) BILLED._AMT>=0.07650831 27    9 1 (0.33333333 0.66666667)
*
##                      303) REPAY_STATUS.>=-0.03234571 25    7 1 (0.28000000 0.7200000
0) *
##               19) BILLED._AMT< -0.4255529 2263  977 0 (0.56827221 0.43172779)
##                 38) PAID_AMT>=-1.078965 2087  867 0 (0.58457115 0.41542885)
##                   76) LIMIT_BAL>=145000 978  356 0 (0.63599182 0.36400818)
##                    152) PAID_AMT>=-0.4779268 581  180 0 (0.69018933 0.30981067)
##                      304) BILLED._AMT>=-0.7307623 398   99 0 (0.75125628 0.24874372)
##                        608) RATIO_PADI_AMT2>=0.05438328 183   33 0 (0.81967213 0.1803
2787) *
##                        609) RATIO_PADI_AMT2< 0.05438328 215   66 0 (0.69302326 0.3069
7674)
##                         1218) LIMIT_BAL< 470000 203   58 0 (0.71428571 0.28571429)
##                           2436) BILLED._AMT< -0.4726066 191   51 0 (0.73298429 0.2670
1571) *
##                           2437) BILLED._AMT>=-0.4726066 12    5 1 (0.41666667 0.58333
333) *
##                         1219) LIMIT_BAL>=470000 12    4 1 (0.33333333 0.66666667) *
##                      305) BILLED._AMT< -0.7307623 183   81 0 (0.55737705 0.44262295)
##                        610) AGE< 44.5 159   63 0 (0.60377358 0.39622642)
##                         1220) REPAY_STATUS.< 0.07471067 100   30 0 (0.70000000 0.3000
0000) *
##                         1221) REPAY_STATUS.>=0.07471067 59   26 1 (0.44067797 0.55932
203) *
##                        611) AGE>=44.5 24    6 1 (0.25000000 0.75000000) *
##                    153) PAID_AMT< -0.4779268 397  176 0 (0.55667506 0.44332494)
##                      306) REPAY_STATUS.< -1.065462 124   35 0 (0.71774194 0.28225806)
*
##                      307) REPAY_STATUS.>=-1.065462 273  132 1 (0.48351648 0.51648352)
##                        614) RATIO_PADI_AMT3< -0.5827417 35   10 0 (0.71428571 0.28571
429) *
##                        615) RATIO_PADI_AMT3>=-0.5827417 238  107 1 (0.44957983 0.5504
2017)
##                         1230) BILLED._AMT>=-0.6571899 209  100 1 (0.47846890 0.521531
10)
##                           2460) BILLED._AMT< -0.6496236 12    2 0 (0.83333333 0.16666
667) *
##                           2461) BILLED._AMT>=-0.6496236 197   90 1 (0.45685279 0.5431
4721)
##                             4922) MARRIAGE>=1.5 84   38 0 (0.54761905 0.45238095) *
##                             4923) MARRIAGE< 1.5 113   44 1 (0.38938053 0.61061947)
##                               9846) RATIO_PADI_AMT1< -0.3895532 13    4 0 (0.69230769
0.30769231) *
##                               9847) RATIO_PADI_AMT1>=-0.3895532 100   35 1 (0.3500000
0 0.65000000)
```

```
##                                19694) RATIO_PADI_AMT2< -0.7615761 19    7 0 (0.631578
95 0.36842105) *
##                                19695) RATIO_PADI_AMT2>=-0.7615761 81   23 1 (0.283950
62 0.71604938) *
##                         1231) BILLED._AMT< -0.6571899 29    7 1 (0.24137931 0.7586206
9) *
##                   77) LIMIT_BAL< 145000 1109  511 0 (0.53922453 0.46077547)
##                    154) RATIO_PADI_AMT4>=-1.089607 1009  444 0 (0.55996036 0.4400396
4)
##                      308) RATIO_PADI_AMT4>=0.6008163 223   76 0 (0.65919283 0.3408071
7)
##                        616) RATIO_PADI_AMT1>=-0.5215779 179   53 0 (0.70391061 0.2960
8939)
##                         1232) LIMIT_BAL>=95000 27    2 0 (0.92592593 0.07407407) *
##                         1233) LIMIT_BAL< 95000 152   51 0 (0.66447368 0.33552632)
##                           2466) RATIO_PADI_AMT1< -0.4813234 34    5 0 (0.85294118 0.1
4705882) *
##                           2467) RATIO_PADI_AMT1>=-0.4813234 118   46 0 (0.61016949 0.
38983051)
##                             4934) BILLED._AMT>=-0.6187031 77   23 0 (0.70129870 0.298
70130) *
##                             4935) BILLED._AMT< -0.6187031 41   18 1 (0.43902439 0.560
97561) *
##                        617) RATIO_PADI_AMT1< -0.5215779 44   21 1 (0.47727273 0.52272
727) *
##                      309) RATIO_PADI_AMT4< 0.6008163 786  368 0 (0.53180662 0.4681933
8)
##                        618) AGE< 24.5 139   47 0 (0.66187050 0.33812950) *
##                        619) AGE>=24.5 647  321 0 (0.50386399 0.49613601)
##                         1238) TIMELY_PAID_AMT< 0.7801033 600  286 0 (0.52333333 0.476
66667)
##                           2476) PAID_AMT>=-0.7027773 473  209 0 (0.55813953 0.4418604
7)
##                             4952) RATIO_PADI_AMT4< 0.5665761 400  164 0 (0.59000000 
0.41000000)
##                               9904) BILLED._AMT>=-0.7393585 336  126 0 (0.62500000 0.
37500000)
##                                19808) TIMELY_PAID_AMT>=-0.04511526 312  111 0 (0.6442
3077 0.35576923)
##                                  39616) RATIO_PADI_AMT2>=0.5945616 53    9 0 (0.83018
868 0.16981132) *
##                                  39617) RATIO_PADI_AMT2< 0.5945616 259  102 0 (0.6061
7761 0.39382239)
##                                    79234) TIMELY_PAID_AMT>=0.4893128 179   61 0 (0.65
921788 0.34078212) *
##                                    79235) TIMELY_PAID_AMT< 0.4893128 80   39 1 (0.487
50000 0.51250000) *
##                                19809) TIMELY_PAID_AMT< -0.04511526 24    9 1 (0.37500
000 0.62500000) *
```

```
##                              9905) BILLED._AMT< -0.7393585 64   26 1 (0.40625000 0.5
9375000) *
##                            4953) RATIO_PADI_AMT4>=0.5665761 73   28 1 (0.38356164 0.
61643836) *
##                          2477) PAID_AMT< -0.7027773 127   50 1 (0.39370079 0.6062992
1)
##                            4954) TIMELY_PAID_AMT< -0.05351894 15    4 0 (0.73333333
0.26666667) *
##                            4955) TIMELY_PAID_AMT>=-0.05351894 112   39 1 (0.34821429
0.65178571) *
##                        1239) TIMELY_PAID_AMT>=0.7801033 47   12 1 (0.25531915 0.7446
8085) *
##                   155) RATIO_PADI_AMT4< -1.089607 100   33 1 (0.33000000 0.67000000)
##                     310) RATIO_PADI_AMT3< -0.5977489 11    2 0 (0.81818182 0.1818181
8) *
##                     311) RATIO_PADI_AMT3>=-0.5977489 89   24 1 (0.26966292 0.7303370
8) *
##                39) PAID_AMT< -1.078965 176   66 1 (0.37500000 0.62500000)
##                  78) LIMIT_BAL< 245000 111   54 0 (0.51351351 0.48648649)
##                   156) AGE< 33.5 45   17 0 (0.62222222 0.37777778) *
##                   157) AGE>=33.5 66   29 1 (0.43939394 0.56060606) *
##                  79) LIMIT_BAL>=245000 65    9 1 (0.13846154 0.86153846) *
##           5) REPAY_STATUS.>=1.000148 586  207 1 (0.35324232 0.64675768)
##            10) REPAY_STATUS.< 3.373648 520  200 1 (0.38461538 0.61538462)
##              20) REPAY_STATUS.< 1.778853 212   98 1 (0.46226415 0.53773585)
##                40) REPAY_STATUS.>=1.033171 176   87 0 (0.50568182 0.49431818)
##                  80) RATIO_PADI_AMT4>=-1.170842 132   57 0 (0.56818182 0.43181818)
##                   160) RATIO_PADI_AMT4< -0.6053733 17    2 0 (0.88235294 0.11764706)
*
##                   161) RATIO_PADI_AMT4>=-0.6053733 115   55 0 (0.52173913 0.4782608
7)
##                     322) RATIO_PADI_AMT2< 1.108128 90   36 0 (0.60000000 0.40000000)
*
##                     323) RATIO_PADI_AMT2>=1.108128 25    6 1 (0.24000000 0.76000000)
*
##                  81) RATIO_PADI_AMT4< -1.170842 44   14 1 (0.31818182 0.68181818) *
##                41) REPAY_STATUS.< 1.033171 36    9 1 (0.25000000 0.75000000) *
##              21) REPAY_STATUS.>=1.778853 308  102 1 (0.33116883 0.66883117)
##                42) RATIO_PADI_AMT1>=-0.6084661 275   99 1 (0.36000000 0.64000000)
##                  84) TIMELY_PAID_AMT>=0.9348192 199   83 1 (0.41708543 0.58291457)
##                   168) RATIO_PADI_AMT4>=0.5695125 71   31 0 (0.56338028 0.43661972)
*
##                   169) RATIO_PADI_AMT4< 0.5695125 128   43 1 (0.33593750 0.66406250)
##                     338) RATIO_PADI_AMT1>=0.2565464 10    3 0 (0.70000000 0.3000000
0) *
##                     339) RATIO_PADI_AMT1< 0.2565464 118   36 1 (0.30508475 0.6949152
5) *
##                  85) TIMELY_PAID_AMT< 0.9348192 76   16 1 (0.21052632 0.78947368) *
##                43) RATIO_PADI_AMT1< -0.6084661 33    3 1 (0.09090909 0.90909091) *
```

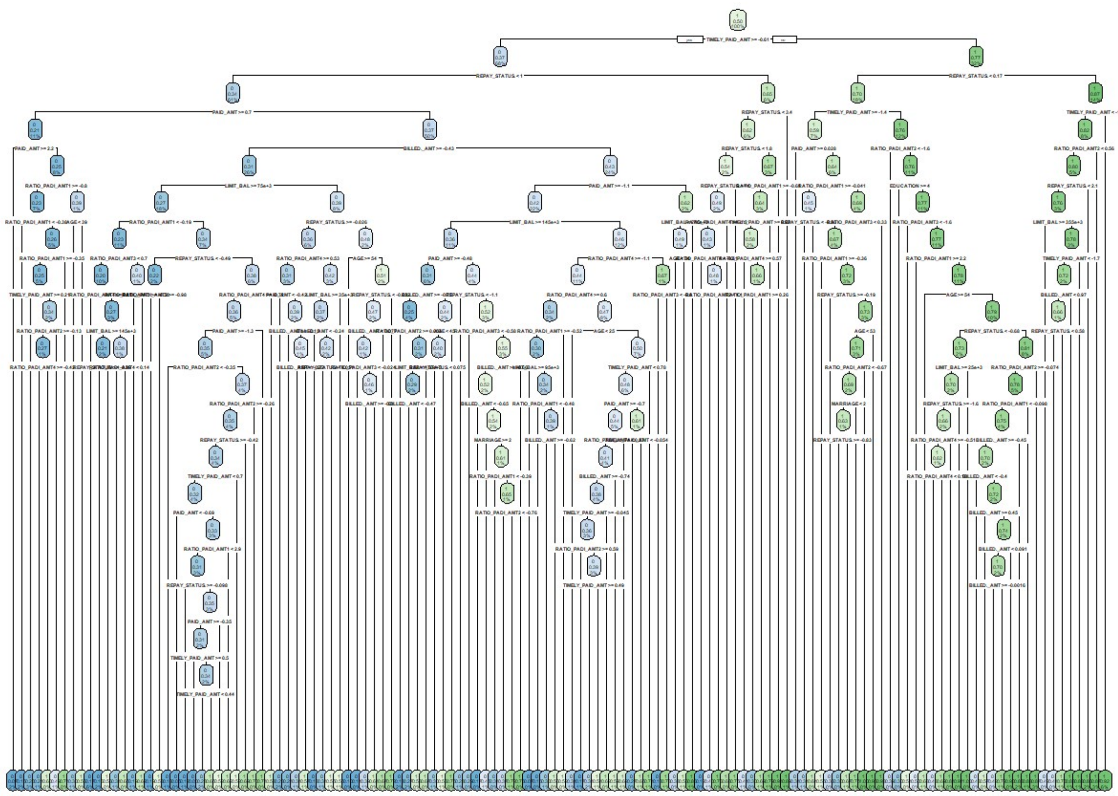
```
##           11) REPAY_STATUS.>=3.373648 66    7 1 (0.10606061 0.89393939) *
##        3) TIMELY_PAID_AMT< -0.6144733 3019  690 1 (0.22855250 0.77144750)
##          6) REPAY_STATUS.< 0.172774 1685  510 1 (0.30267062 0.69732938)
##           12) TIMELY_PAID_AMT>=-1.367929 611  250 1 (0.40916530 0.59083470)
##             24) PAID_AMT>=0.02797305 95   32 0 (0.66315789 0.33684211) *
##             25) PAID_AMT< 0.02797305 516  187 1 (0.36240310 0.63759690)
##               50) RATIO_PADI_AMT1>=-0.04064672 118   53 0 (0.55084746 0.44915254)
##                100) REPAY_STATUS.< -0.561787 74   25 0 (0.66216216 0.33783784) *
##                101) REPAY_STATUS.>=-0.561787 44   16 1 (0.36363636 0.63636364) *
##               51) RATIO_PADI_AMT1< -0.04064672 398  122 1 (0.30653266 0.69346734)
##                102) RATIO_PADI_AMT3< 0.3343695 369  120 1 (0.32520325 0.67479675)
##                  204) RATIO_PADI_AMT1>=-0.3647836 96   43 1 (0.44791667 0.55208333)
*
##                  205) RATIO_PADI_AMT1< -0.3647836 273   77 1 (0.28205128 0.7179487
2)
##                    410) REPAY_STATUS.>=-0.1900732 11    4 0 (0.63636364 0.36363636)
*
##                    411) REPAY_STATUS.< -0.1900732 262   70 1 (0.26717557 0.7328244
3)
##                      822) AGE< 52.5 239   69 1 (0.28870293 0.71129707)
##                       1644) RATIO_PADI_AMT2< -0.6723654 224   69 1 (0.30803571 0.69
196429)
##                         3288) MARRIAGE< 1.5 112   41 1 (0.36607143 0.63392857)
##                           6576) REPAY_STATUS.>=-0.8312857 10    3 0 (0.70000000 0.3
0000000) *
##                           6577) REPAY_STATUS.< -0.8312857 102   34 1 (0.33333333 0.
66666667) *
##                         3289) MARRIAGE>=1.5 112   28 1 (0.25000000 0.75000000) *
##                       1645) RATIO_PADI_AMT2>=-0.6723654 15    0 1 (0.00000000 1.000
00000) *
##                      823) AGE>=52.5 23    1 1 (0.04347826 0.95652174) *
##                103) RATIO_PADI_AMT3>=0.3343695 29    2 1 (0.06896552 0.93103448) *
##           13) TIMELY_PAID_AMT< -1.367929 1074  260 1 (0.24208566 0.75791434)
##             26) RATIO_PADI_AMT2< -1.644013 18    7 0 (0.61111111 0.38888889) *
##             27) RATIO_PADI_AMT2>=-1.644013 1056  249 1 (0.23579545 0.76420455)
##               54) EDUCATION>=3.5 10    3 0 (0.70000000 0.30000000) *
##               55) EDUCATION< 3.5 1046  242 1 (0.23135755 0.76864245)
##                110) RATIO_PADI_AMT3< -1.627155 26   12 1 (0.46153846 0.53846154) *
##                111) RATIO_PADI_AMT3>=-1.627155 1020  230 1 (0.22549020 0.77450980)
##                  222) RATIO_PADI_AMT1>=2.167438 18    9 0 (0.50000000 0.50000000) *
##                  223) RATIO_PADI_AMT1< 2.167438 1002  221 1 (0.22055888 0.77944112)
##                    446) AGE>=53.5 59   22 1 (0.37288136 0.62711864) *
##                    447) AGE< 53.5 943  199 1 (0.21102863 0.78897137)
##                      894) REPAY_STATUS.< -0.6823706 223   61 1 (0.27354260 0.726457
40)
##                       1788) LIMIT_BAL>=25000 201   61 1 (0.30348259 0.69651741)
##                         3576) REPAY_STATUS.>=-1.571277 169   58 1 (0.34319527 0.656
80473)
##                           7152) RATIO_PADI_AMT4>=-0.5079225 137   52 1 (0.37956204
```

```
0.62043796)
##                            14304) RATIO_PADI_AMT4< 0.1817267 37   16 0 (0.56756757
0.43243243) *
##                            14305) RATIO_PADI_AMT4>=0.1817267 100   31 1 (0.31000000
0.69000000) *
##                           7153) RATIO_PADI_AMT4< -0.5079225 32    6 1 (0.18750000
0.81250000) *
##                         3577) REPAY_STATUS.< -1.571277 32    3 1 (0.09375000 0.9062
5000) *
##                       1789) LIMIT_BAL< 25000 22    0 1 (0.00000000 1.00000000) *
##                      895) REPAY_STATUS.>=-0.6823706 720  138 1 (0.19166667 0.808333
33)
##                       1790) RATIO_PADI_AMT2>=-0.07447176 471  105 1 (0.22292994 0.7
7707006)
##                         3580) RATIO_PADI_AMT1< -0.09830139 375   93 1 (0.24800000
0.75200000)
##                           7160) BILLED._AMT>=-0.4495703 228   68 1 (0.29824561 0.70
175439)
##                            14320) BILLED._AMT< -0.4035051 14    6 0 (0.57142857 0.4
2857143) *
##                            14321) BILLED._AMT>=-0.4035051 214   60 1 (0.28037383 0.
71962617)
##                              28642) BILLED._AMT>=0.4529917 30   13 1 (0.43333333 0.
56666667) *
##                              28643) BILLED._AMT< 0.4529917 184   47 1 (0.25543478
0.74456522)
##                                57286) BILLED._AMT< 0.09121162 142   43 1 (0.3028169
0 0.69718310)
##                                 114572) BILLED._AMT>=-0.001614114 12    5 0 (0.5833
3333 0.41666667) *
##                                 114573) BILLED._AMT< -0.001614114 130   36 1 (0.276
92308 0.72307692) *
##                                57287) BILLED._AMT>=0.09121162 42    4 1 (0.09523810
0.90476190) *
##                           7161) BILLED._AMT< -0.4495703 147   25 1 (0.17006803 0.82
993197) *
##                         3581) RATIO_PADI_AMT1>=-0.09830139 96   12 1 (0.12500000 0.
87500000) *
##                       1791) RATIO_PADI_AMT2< -0.07447176 249   33 1 (0.13253012 0.8
6746988) *
##          7) REPAY_STATUS.>=0.172774 1334  180 1 (0.13493253 0.86506747)
##           14) TIMELY_PAID_AMT< -1.351992 760  133 1 (0.17500000 0.82500000)
##             28) RATIO_PADI_AMT2< 0.5565909 506  101 1 (0.19960474 0.80039526)
##               56) REPAY_STATUS.< 2.111825 304   72 1 (0.23684211 0.76315789)
##                112) LIMIT_BAL>=355000 12    5 0 (0.58333333 0.41666667) *
##                113) LIMIT_BAL< 355000 292   65 1 (0.22260274 0.77739726)
##                  226) TIMELY_PAID_AMT< -1.674265 187   52 1 (0.27807487 0.72192513)
##                    452) BILLED._AMT< 0.9709496 136   46 1 (0.33823529 0.66176471)
##                      904) REPAY_STATUS.< 0.5836448 17    7 0 (0.58823529 0.4117647
```

```
1) *
##                        905) REPAY_STATUS.>=0.5836448 119   36 1 (0.30252101 0.6974789
9) *
##                      453) BILLED._AMT>=0.9709496 51    6 1 (0.11764706 0.88235294) *
##                    227) TIMELY_PAID_AMT>=-1.674265 105   13 1 (0.12380952 0.87619048)
*
##                 57) REPAY_STATUS.>=2.111825 202   29 1 (0.14356436 0.85643564) *
##               29) RATIO_PADI_AMT2>=0.5565909 254   32 1 (0.12598425 0.87401575) *
##             15) TIMELY_PAID_AMT>=-1.351992 574   47 1 (0.08188153 0.91811847) *
```

```
rpart.plot(CT_model)
```

```
## Warning: labs do not fit even at cex 0.15, there may be some overplotting
```



```
attributes(CT_model)
```

```
## $names
##  [1] "frame"               "where"               "call"
##  [4] "terms"               "cptable"             "method"
##  [7] "parms"               "control"             "functions"
## [10] "numresp"             "splits"              "variable.importance"
## [13] "y"                   "ordered"
##
## $xlevels
## named list()
##
## $ylevels
## [1] "0" "1"
##
## $class
## [1] "rpart"
```

```
CT_model$cptable
```

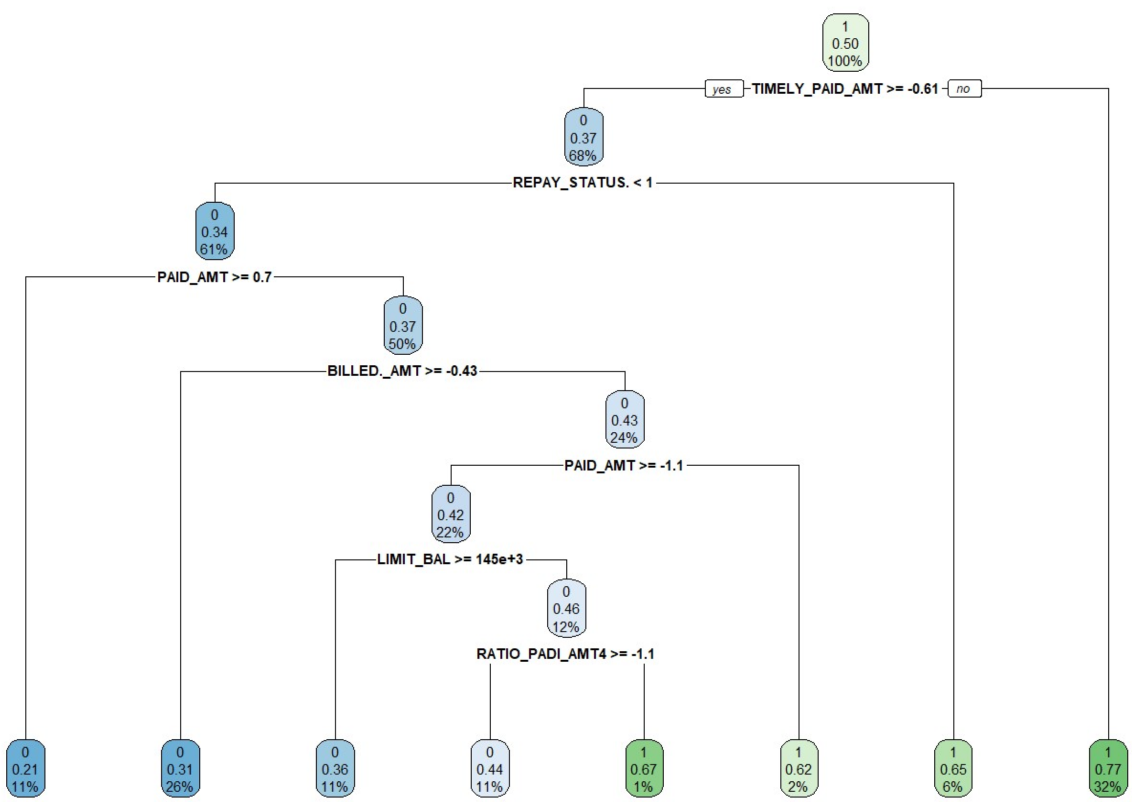
```
##              CP nsplit rel error    xerror        xstd
## 1  0.3519276330      0 1.0000000 1.0230454 0.010377498
## 2  0.0370450140      1 0.6480724 0.6500108 0.009722263
## 3  0.0031588772      2 0.6110274 0.6164118 0.009584658
## 4  0.0026922249      7 0.5942279 0.6172733 0.009588369
## 5  0.0025845359     12 0.5797975 0.6129658 0.009569715
## 6  0.0022255725     13 0.5772130 0.6142580 0.009575337
## 7  0.0017230239     18 0.5653672 0.6149042 0.009578140
## 8  0.0015076459     21 0.5601981 0.6142580 0.009575337
## 9  0.0012922679     23 0.5571829 0.6222270 0.009609520
## 10 0.0011845789     27 0.5517984 0.6183502 0.009592995
## 11 0.0011486826     29 0.5494292 0.6187810 0.009594840
## 12 0.0010768899     32 0.5459832 0.6194271 0.009597605
## 13 0.0009692009     34 0.5438294 0.6213655 0.009605865
## 14 0.0008884342     40 0.5371527 0.6248115 0.009620428
## 15 0.0008615120     48 0.5300452 0.6250269 0.009621333
## 16 0.0007538230     52 0.5265992 0.6284730 0.009635731
## 17 0.0006461340     62 0.5179841 0.6332113 0.009655279
## 18 0.0005743413     69 0.5132457 0.6347189 0.009661437
## 19 0.0005384450     72 0.5115227 0.6407495 0.009685781
## 20 0.0005025486     76 0.5093689 0.6407495 0.009685781
## 21 0.0004738316     80 0.5072152 0.6409649 0.009686642
## 22 0.0004307560     85 0.5048460 0.6450571 0.009702884
## 23 0.0003589633     89 0.5031230 0.6444109 0.009700334
## 24 0.0002153780     92 0.5020461 0.6472109 0.009711348
## 25 0.0001435853    102 0.4998923 0.6513030 0.009727267
## 26 0.0001346112    105 0.4994616 0.6541030 0.009738037
## 27 0.0001076890    113 0.4983847 0.6541030 0.009738037
## 28 0.0000000000    127 0.4964463 0.6571182 0.009749526
```

Pruning the tree

```
ptree = prune(CT_model, .0031, "CP")
ptree
```

```
## n= 9291
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##   1) root 9291 4643 1 (0.4997309 0.5002691)
##     2) TIMELY_PAID_AMT>=-0.6144733 6272 2319 0 (0.6302615 0.3697385)
##       4) REPAY_STATUS.< 1.000148 5686 1940 0 (0.6588111 0.3411889)
##         8) PAID_AMT>=0.7048767 1012  216 0 (0.7865613 0.2134387) *
##         9) PAID_AMT< 0.7048767 4674 1724 0 (0.6311510 0.3688490)
##          18) BILLED._AMT>=-0.4255529 2411  747 0 (0.6901701 0.3098299) *
##          19) BILLED._AMT< -0.4255529 2263  977 0 (0.5682722 0.4317278)
##            38) PAID_AMT>=-1.078965 2087  867 0 (0.5845712 0.4154288)
##              76) LIMIT_BAL>=145000 978  356 0 (0.6359918 0.3640082) *
##              77) LIMIT_BAL< 145000 1109  511 0 (0.5392245 0.4607755)
##               154) RATIO_PADI_AMT4>=-1.089607 1009  444 0 (0.5599604 0.4400396) *
##               155) RATIO_PADI_AMT4< -1.089607 100   33 1 (0.3300000 0.6700000) *
##            39) PAID_AMT< -1.078965 176   66 1 (0.3750000 0.6250000) *
##       5) REPAY_STATUS.>=1.000148 586  207 1 (0.3532423 0.6467577) *
##     3) TIMELY_PAID_AMT< -0.6144733 3019  690 1 (0.2285525 0.7714475) *
```

```
rpart.plot(ptree)
```



```
CT_model$variable.importance
```

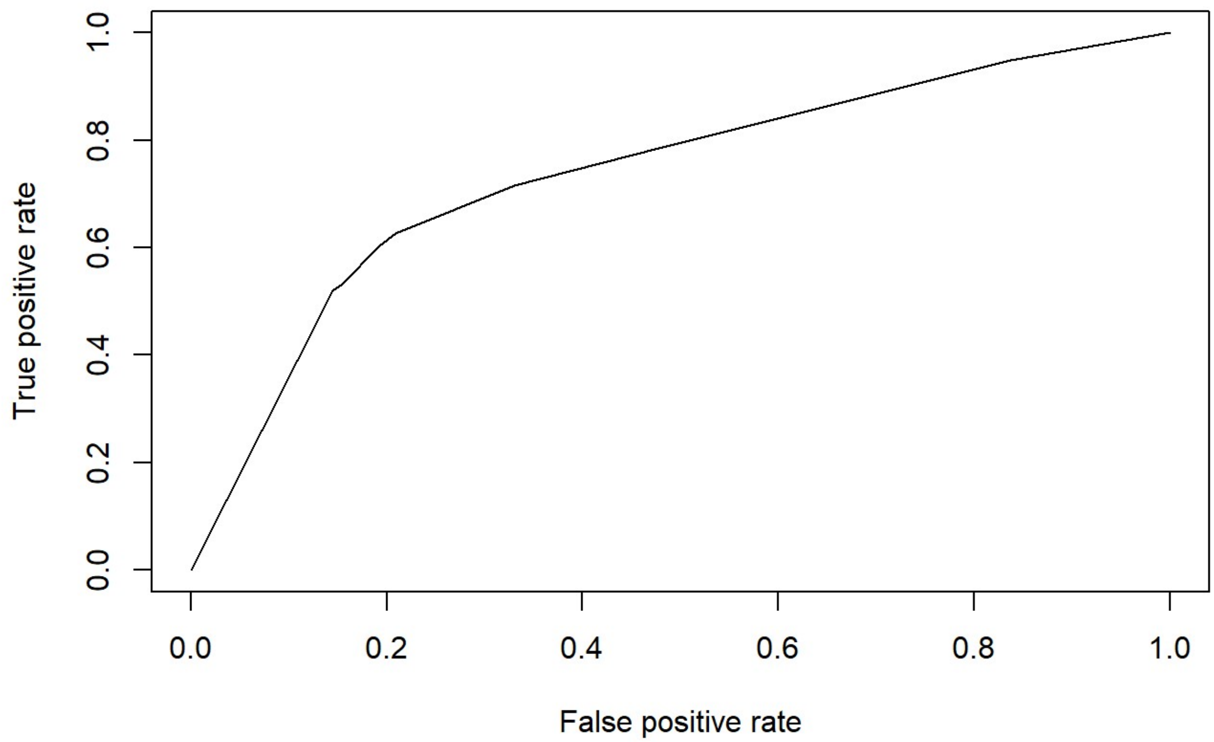
```
## TIMELY_PAID_AMT   REPAY_STATUS.        PAID_AMT     BILLED._AMT
##      806.215602      382.899461      194.203815      164.608151
## RATIO_PADI_AMT4 RATIO_PADI_AMT2 RATIO_PADI_AMT1 RATIO_PADI_AMT3
##      152.255511      125.071627      109.634039       74.565264
##       LIMIT_BAL             AGE       EDUCATION        MARRIAGE
##       68.850349       30.739832        7.163986        5.436529
```

CART validation on test data

```
predTrain = predict(ptree, newdata = train_bank)
pred_class = predict(ptree, newdata = train_bank[,-6], type = "class")
predDT = predict(ptree, newdata = test_bank)
```

Validation on test data

```
library(ROCR)
DTpredROC_CT = prediction(predDT[,2],test_bank$DEFAULT)
perf2 =performance(DTpredROC_CT, "tpr", "fpr")
plot(perf2)
```




```
Auc <- as.numeric(performance(DTpredROC_CT, "auc")@y.values)
Auc
```

```
## [1] 0.7369729
```

KS

```
KS1 <- max(attr(perf2, 'y.values')[[1]]-attr(perf2, 'x.values')[[1]])
KS1
```

```
## [1] 0.4181934
```

Gini

```
## Gini Coefficient
library(ineq)
gini1 = ineq(test_bank$DEFAULT, type="Gini")
gini1
```

```
## [1] 0.7791111
```

```
#Check classification error using confusion matrix
#table(test_bank$DEFAULT,pred_class )

accuarcy <- (5715+1184)/(5715+1184+1297+804)
accuarcy
```

```
## [1] 0.7665556
```

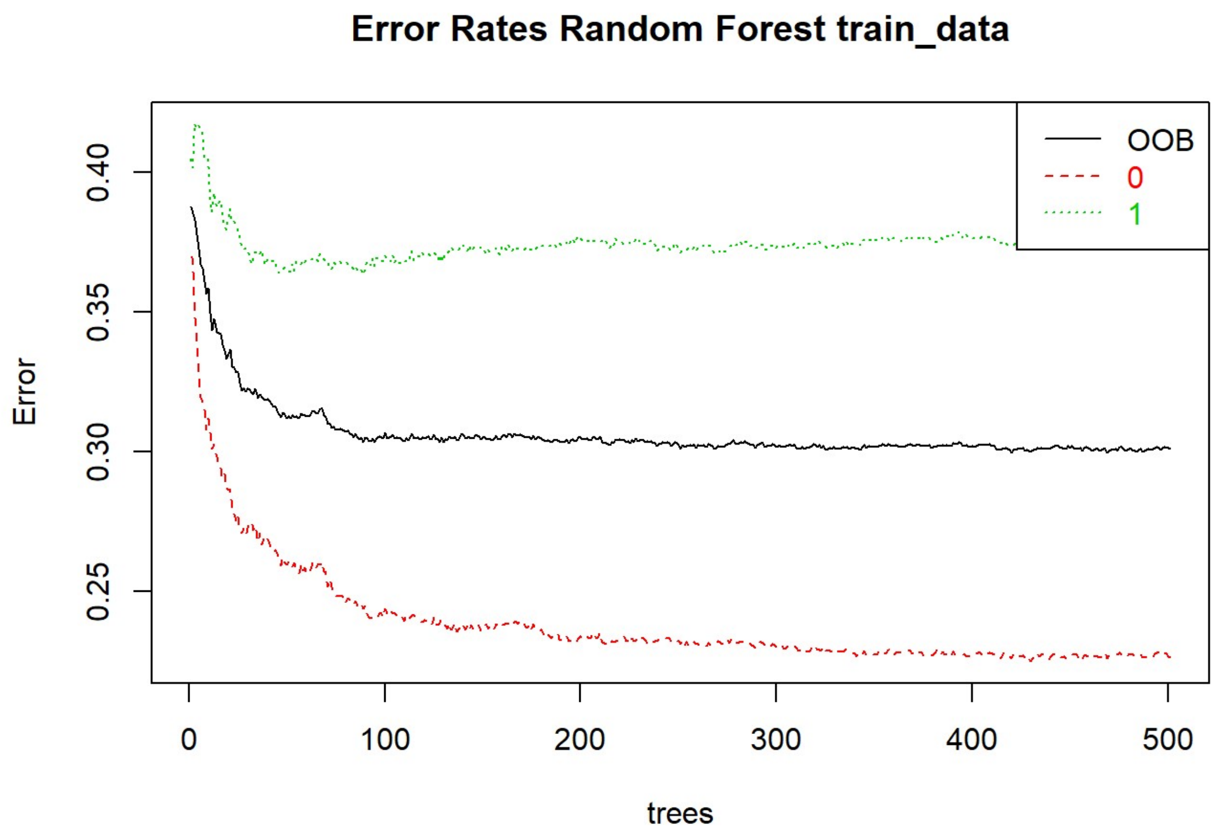
Random Forest

```
train_bank$DEFAULT=as.factor(train_bank$DEFAULT)
library(randomForest)
seed=101
set.seed(seed)
RF_model = randomForest(DEFAULT ~ ., data = train_bank, mtry = 3, nodesize =10, ntree
=501, importance = TRUE)
print(RF_model)
```

```
##
## Call:
##  randomForest(formula = DEFAULT ~ ., data = train_bank, mtry = 3,      nodesize = 1
0, ntree = 501, importance = TRUE)
##                Type of random forest: classification
##                      Number of trees: 501
## No. of variables tried at each split: 3
##
##         OOB estimate of  error rate: 30.09%
## Confusion matrix:
##      0    1 class.error
## 0 3591 1052   0.2265776
## 1 1744 2904   0.3752151
```

ploting RF moden

```
plot(RF_model, main="")
legend("topright", c("OOB", "0", "1"), text.col=1:6, lty=1:3, col=1:3)
title(main="Error Rates Random Forest train_data")
```



Checking OOB

```
rf_err_rate <-  as.data.frame(RF_model$err.rate)
rf_err_rate$ID <- seq.int(nrow(rf_err_rate))
rf_err_rate[which(rf_err_rate$OOB==min(rf_err_rate$OOB)),]
```

```
##           OOB         0        1  ID
## 420 0.2998601 0.2261469 0.373494 420
## 469 0.2998601 0.2261469 0.373494 469
```

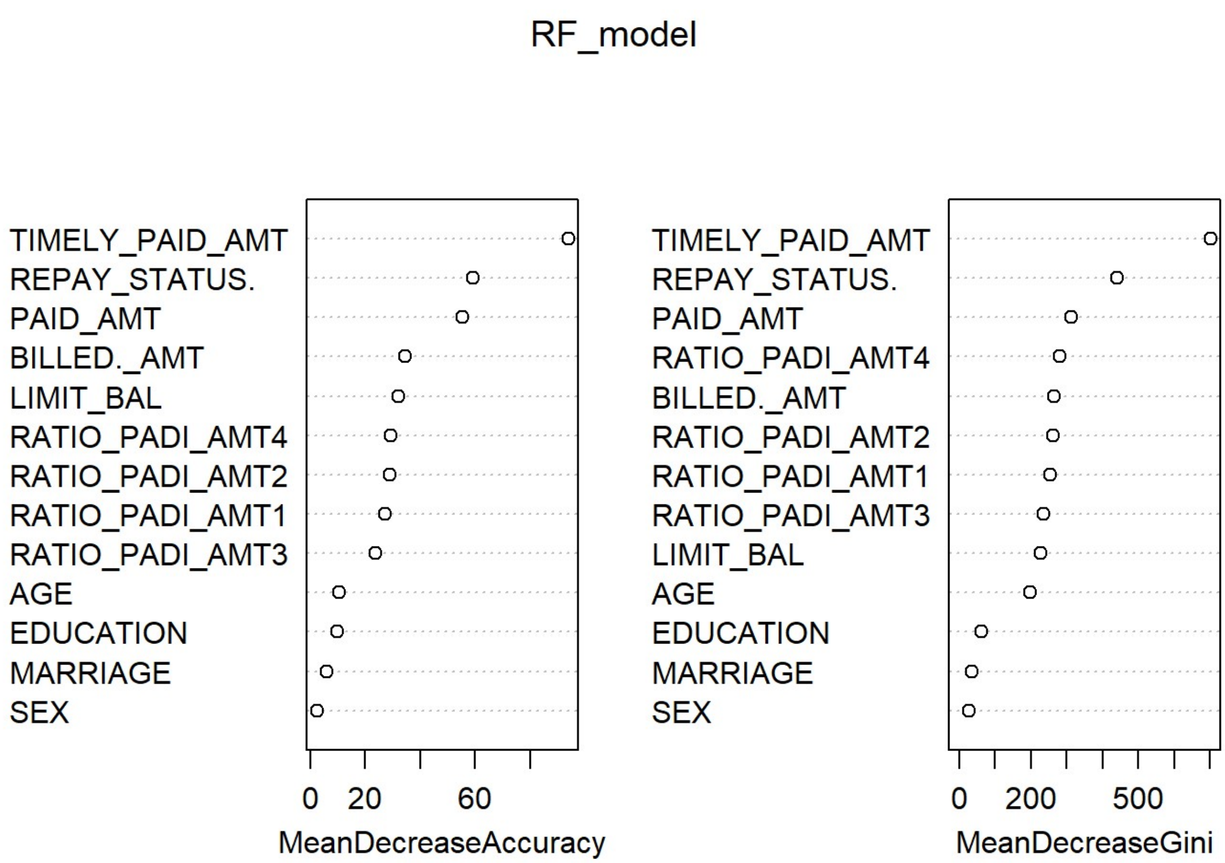
```
min_tree<-min(rf_err_rate[which(rf_err_rate$OOB==min(rf_err_rate$OOB)),]$ID)
```

```
## List the importance of the variables.
impVar <- round(randomForest::importance(RF_model), 2)
impVar[order(impVar[,3],decreasing = TRUE),]
```

```
##                    0     1 MeanDecreaseAccuracy MeanDecreaseGini
## TIMELY_PAID_AMT 76.59 41.05                93.85           700.48
## REPAY_STATUS.   41.56 23.38                59.15           439.35
## PAID_AMT        18.16 42.96                55.36           312.82
## BILLED._AMT     16.83 20.24                34.56           265.57
## LIMIT_BAL       15.94 23.68                32.12           226.66
## RATIO_PADI_AMT4 16.31 15.72                29.22           280.35
## RATIO_PADI_AMT2 10.11 20.30                28.91           263.15
## RATIO_PADI_AMT1  2.46 29.72                27.23           254.00
## RATIO_PADI_AMT3  7.73 17.35                23.93           235.53
## AGE             13.06  0.56                10.64           198.71
## EDUCATION        6.01  7.93                10.06            62.33
## MARRIAGE         9.33 -2.24                 5.91            35.17
## SEX              1.98  1.38                 2.47            27.94
```

Variable Importance: Graphical representation

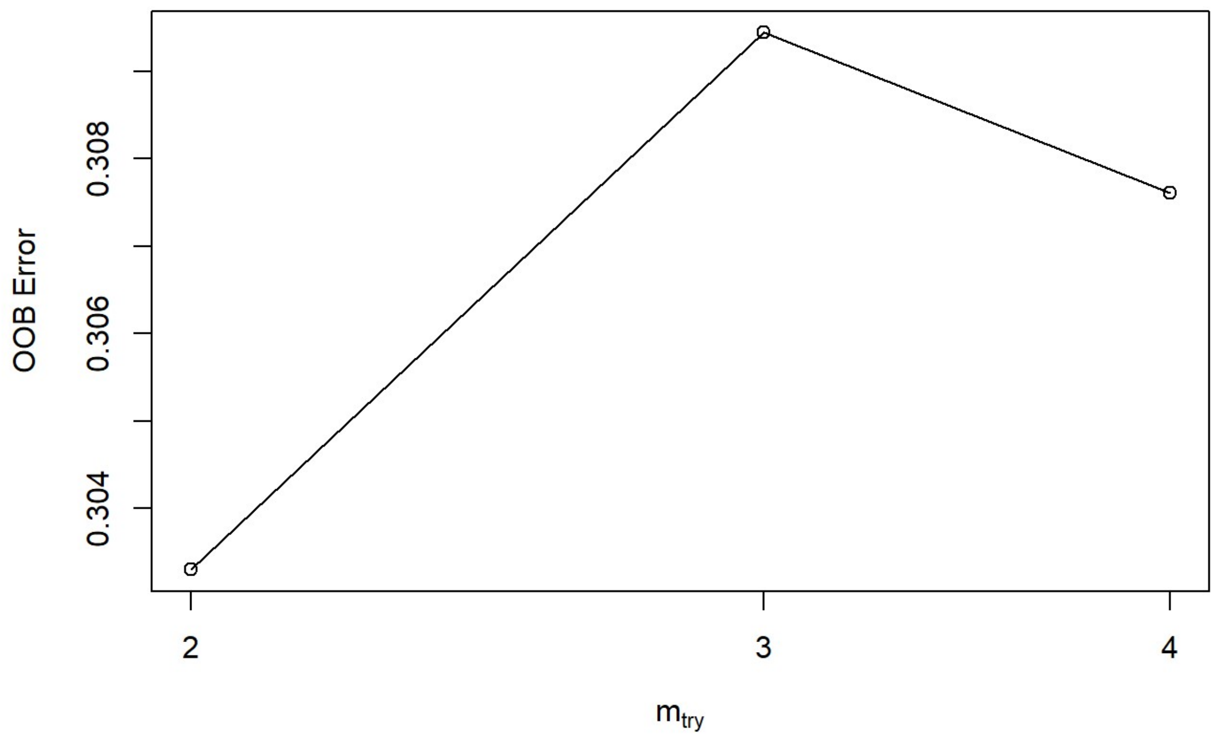
```
varImpPlot(RF_model)
```



Optimal mtry value

```
#View(train_bank)
tune_rf_model <- tuneRF(x =train_bank[,-c(2,6)],
              y=as.factor(train_bank$DEFAULT),
              mtryStart = 3,
              ntreeTry=100,
              stepFactor = 1.5,
              improve =0.0001,
              trace=TRUE,
              plot = TRUE,
              doBest = TRUE,
              nodesize = 10,
              importance=TRUE
)
```

```
## mtry = 3  OOB error = 30.94%
## Searching left ...
## mtry = 2     OOB error = 30.33%
## 0.01982609 1e-04
## Searching right ...
## mtry = 4     OOB error = 30.76%
## -0.01419446 1e-04
```



#Validate RF model on test data

```
pred_RF =predict(tune_rf_model, newdata = test_bank[,-6], type="prob")
pred_RF_CL=predict(tune_rf_model, newdata = test_bank[,-6], type="class")
head(pred_RF)
```

```
##       0     1
## 1 0.294 0.706
## 2 0.408 0.592
## 3 0.884 0.116
## 4 0.826 0.174
## 5 0.704 0.296
## 6 0.672 0.328
```

```
class(pred_RF)
```

```
## [1] "matrix" "votes"
```

Deciling andrank order table

```
# deciling
decile <- function(x){
  deciles <- vector(length=10)
  for (i in seq(0.1,1,.1)){
    deciles[i*10] <- quantile(x, i, na.rm=T)
  }
  return (
    ifelse(x<deciles[1], 1,
           ifelse(x<deciles[2], 2,
                  ifelse(x<deciles[3], 3,
                         ifelse(x<deciles[4], 4,
                                ifelse(x<deciles[5], 5,
                                       ifelse(x<deciles[6], 6,
                                              ifelse(x<deciles[7], 7,
                                                     ifelse(x<deciles[8], 8,
                                                            ifelse(x<deciles[9], 9, 10
                                                            ))))))))))
}


test_bank$deciles <- decile(pred_RF[,2])
```

Rank order

```
library(tidyverse)
library(magrittr)
```

```
##
## Attaching package: 'magrittr'
```

```
## The following object is masked from 'package:purrr':
##
##     set_names
```

```
## The following object is masked from 'package:tidyr':
##
##     extract
```

```
library(data.table)
```

```
##
## Attaching package: 'data.table'
```

```
## The following objects are masked from 'package:dplyr':
##
##     between, first, last
```

```
## The following object is masked from 'package:purrr':
##
##     transpose
```

```
library(scales)
```

```
##
## Attaching package: 'scales'
```

```
## The following objects are masked from 'package:psych':
##
##     alpha, rescale
```

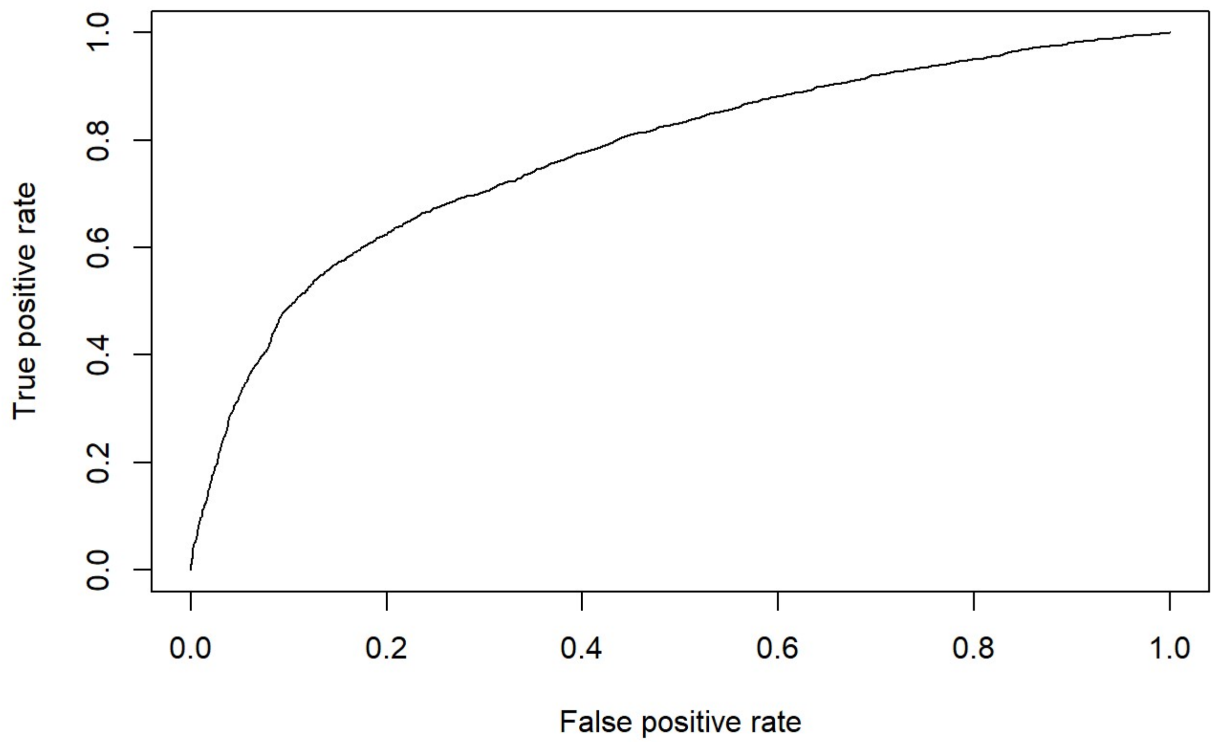
```
## The following object is masked from 'package:purrr':
##
##     discard
```

```
## The following object is masked from 'package:readr':
##
##     col_factor
```

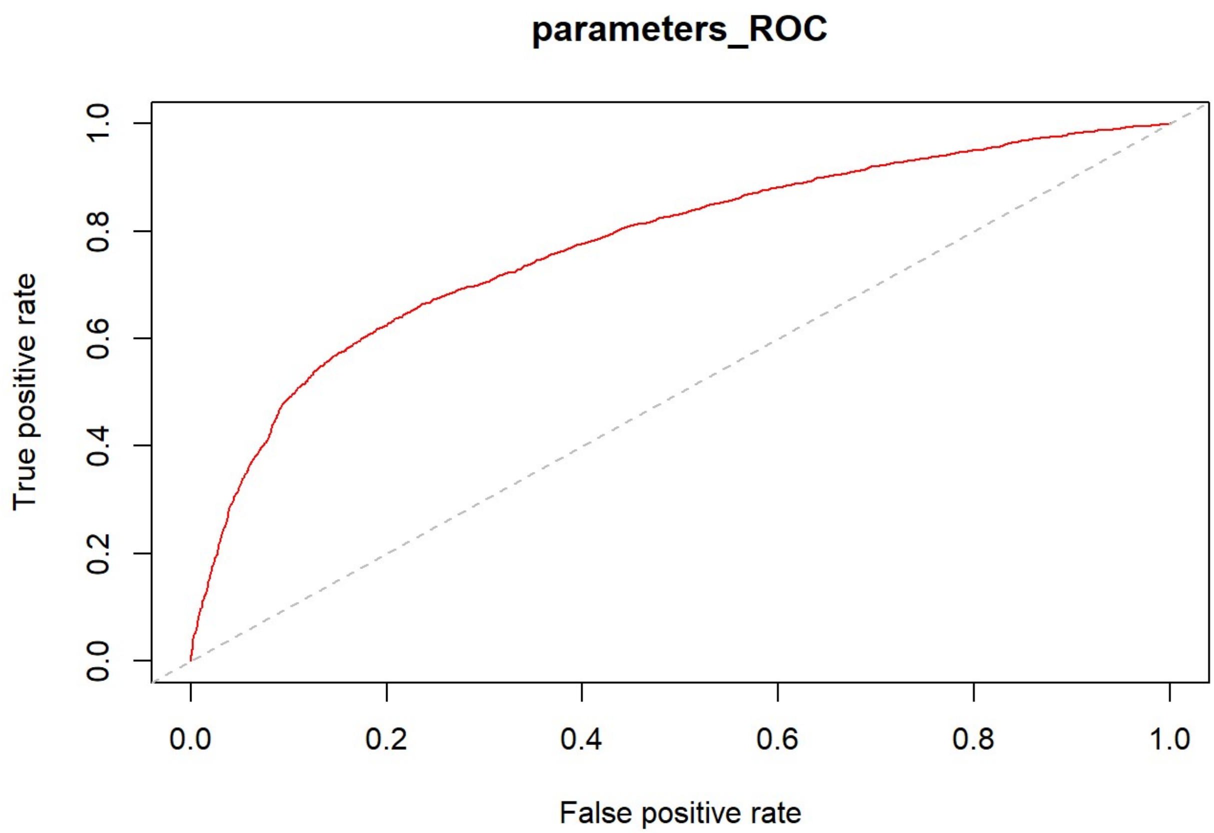
```r
tmp_DT = data.table(test_bank)
rank <- tmp_DT[, list(
  cnt = length(DEFAULT),
  cnt_resp = sum(DEFAULT),
  cnt_non_resp = sum(DEFAULT ==0)) ,
  by=deciles][order(-deciles)]
rank$rrate <- round (rank$cnt_resp / rank$cnt,4);
rank$cum_resp <- cumsum(rank$cnt_resp)
rank$cum_non_resp <- cumsum(rank$cnt_non_resp)
rank$cum_rel_resp <- round(rank$cum_resp / sum(rank$cnt_resp),4);
rank$cum_rel_non_resp <- round(rank$cum_non_resp /sum(rank$cnt_non_resp),4);
rank$ks <- abs(rank$cum_rel_resp - rank$cum_rel_non_resp);

library(scales)
rank$rrate <- percent(rank$rrate)
rank$cum_rel_resp <- percent(rank$cum_rel_resp)
rank$cum_rel_non_resp <- percent(rank$cum_rel_non_resp)

View(rank)
```

Validation with test data

```r
library(ROCR)
RF_pred = ROCR::prediction(pred_RF[,2], test_bank$DEFAULT)
perfx = performance(RF_pred, "tpr", "fpr")
plot(perfx)
```




```
plot(perfx,col="red", main="parameters_ROC")
abline(0,1, lty = 8, col = "grey")
```



Performance Measures

Area Under Curve

```
Auc <- as.numeric(performance(RF_pred, "auc")@y.values)
Auc
```

```
## [1] 0.7739189
```

```
KS <- max(attr(perfx, 'y.values')[[1]]-attr(perfx, 'x.values')[[1]])
KS
```

```
## [1] 0.428749
```

```
## Gini Coefficient
library(ineq)
gini = ineq(pred_RF[,2], type="Gini")
gini
```

```
## [1] 0.3066809
```

```
## Classification Error
with(test_bank,table(DEFAULT,pred_RF_CL))
```

```
##        pred_RF_CL
## DEFAULT    0    1
##       0 5500 1512
##       1  711 1277
```

```
accuracy <-(5509+1273)/(5509+1273+1503+715)
accuracy
```

```
## [1] 0.7535556
```

Machine learning approach with Ensemble Methods

```
#knn compare
library(class)
#View(train_bank)
#View(test_bank)
knn_fit<- knn(train = train_bank[,-6], test = test_bank[,-c(6,15)], cl=train_bank$DEFA
ULT,k =5,prob=TRUE)
knn_chk= table(test_bank$DEFAULT,knn_fit)
knn_chk
```

```
##    knn_fit
##        0    1
##   0 4786 2226
##   1  805 1183
```

```
accuracy.knn = sum(diag(knn_chk))/sum(knn_chk)
accuracy.knn
```

```
## [1] 0.6632222
```

Naive Bayes

```
#naive bayes
library(e1071)
library(caret)
train_bank$DEFAULT = as.factor(train_bank$DEFAULT)
test_bank$DEFAULT = as.factor(test_bank$DEFAULT)
NB = naiveBayes(x =train_bank[-6], y =train_bank$DEFAULT)
pred.NB = predict(NB, newdata =test_bank[-6])
pred.NB
```

```
##    [1] 1 1 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 1 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 1
##   [35] 1 0 0 0 1 1 0 1 1 1 0 1 0 1 1 0 1 1 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0
##   [69] 0 0 1 0 0 0 0 1 0 1 1 1 1 1 0 0 1 0 0 0 0 0 0 0 1 1 0 0 0 1 0 0 1 1
##  [103] 0 0 1 0 1 0 0 0 0 0 1 1 1 1 0 0 1 0 0 1 1 0 0 1 1 0 1 0 0 0 1 0 0 0
##  [137] 0 0 0 0 0 0 1 0 0 0 0 1 0 1 1 0 0 1 0 0 1 0 0 1 1 0 0 0 0 0 0 0 0 0
##  [171] 0 0 1 1 0 0 0 0 1 1 1 0 1 1 0 0 0 0 0 1 0 1 0 1 1 0 0 0 0 1 1 0 1 1
##  [205] 0 0 0 0 0 0 1 1 0 0 1 1 0 0 1 1 1 0 1 1 1 0 0 1 1 0 0 0 1 0 0 0 0 0
##  [239] 1 0 1 0 0 0 0 1 1 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 1 1 0 0 0 0 0 1 0 0
##  [273] 0 1 0 0 0 1 0 1 0 0 1 1 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0
##  [307] 0 0 1 0 1 0 1 1 1 0 0 0 0 0 0 0 0 1 0 1 0 1 1 1 0 1 0 1 1 1 0 0 0 0
##  [341] 1 1 0 1 0 1 1 1 1 0 0 0 1 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 1 1
##  [375] 0 0 0 0 0 0 1 0 1 0 0 0 1 0 1 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 1 1 0 0
##  [409] 1 1 1 0 0 0 0 0 0 0 0 1 0 1 0 1 1 0 0 0 0 0 1 0 1 0 0 0 0 1 0 0 1 0
##  [443] 0 0 1 1 1 0 0 0 1 1 0 1 1 0 0 0 0 1 0 1 0 1 0 0 0 0 1 1 0 0 0 0 0 1
##  [477] 1 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 1 0 0 0 0 0 0 1 1 0 0
##  [511] 0 0 0 0 0 0 1 0 0 1 0 0 0 1 1 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0
##  [545] 1 1 0 0 0 1 1 0 0 0 0 1 1 0 1 0 0 1 1 1 0 1 0 1 0 1 1 1 1 0 0 1 0 0
##  [579] 1 0 1 0 0 0 0 1 1 1 1 0 1 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 0 0 0 0 1 1
##  [613] 0 0 0 1 0 0 0 1 1 1 0 0 1 1 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
##  [647] 1 0 0 0 0 0 1 0 1 0 1 0 0 0 1 1 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1
##  [681] 0 0 0 0 1 1 1 0 0 1 0 1 1 1 1 0 1 0 0 1 1 0 1 0 0 1 1 1 0 1 0 0 0 1
##  [715] 1 1 0 0 0 1 0 1 1 1 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 1 0 0 0 0 1
##  [749] 0 0 1 0 0 0 0 1 0 1 1 1 0 0 0 0 0 0 0 0 0 1 0 1 1 1 0 0 1 0 1 1 1 0
##  [783] 0 0 1 1 0 1 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 1 1 0 0 0 1 0 1 1
##  [817] 1 0 0 0 0 1 1 0 0 0 1 0 1 0 0 0 0 0 1 1 0 0 0 0 0 0 1 0 1 0 0 0 1 0
##  [851] 1 0 1 0 0 1 0 0 0 1 0 1 1 0 0 0 0 0 1 0 0 0 1 1 0 0 1 0 0 1 0 0 1 0
##  [885] 1 1 0 1 0 0 0 0 1 1 1 0 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0
##  [919] 1 0 0 1 0 1 1 0 1 1 0 0 1 0 1 1 0 0 1 0 1 1 0 0 0 0 0 1 0 0 1 1 1 0
##  [953] 0 1 0 1 1 1 1 0 0 1 0 1 0 0 0 0 0 1 1 1 1 1 0 0 0 0 0 1 1 1 0 0 0 1
##  [987] 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 1
## [1021] 0 1 1 1 0 0 0 1 0 0 1 1 0 1 0 0 1 1 1 0 0 0 1 0 1 0 0 0 1 0 1 0 0 0
## [1055] 0 0 0 0 1 1 0 0 0 0 0 1 1 1 0 0 0 0 1 0 1 0 1 0 0 1 0 1 0 0 0 1 0 0
## [1089] 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0 0 0 0 1 0 1 1 0 0 0
## [1123] 0 1 0 0 0 1 0 1 0 1 1 0 1 0 0 0 1 0 1 0 0 1 0 0 1 1 1 1 0 0 0 0 1 0
## [1157] 1 0 0 1 0 0 0 0 0 0 1 0 1 1 0 0 0 0 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0 1
## [1191] 1 0 0 1 0 0 0 1 1 0 0 1 0 0 0 0 0 1 0 1 0 1 1 0 1 0 0 0 1 0 0 0 0 0
## [1225] 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 1 0 1 0 0 1 0 0 0 0 0 0 0 0 0 1
## [1259] 0 1 1 0 0 0 0 0 1 0 1 1 0 1 0 0 0 0 0 0 1 1 1 0 0 0 1 1 0 0 1 0 0 0
## [1293] 0 1 0 1 0 1 0 1 0 1 0 1 0 1 1 0 0 1 0 0 0 0 1 0 0 0 0 1 1 0 1 1 0 0
## [1327] 1 0 1 0 1 1 1 0 1 1 0 1 0 0 0 1 0 0 1 0 1 0 1 1 0 0 0 0 0 0 0 1 1 0
## [1361] 0 0 1 0 1 0 0 0 0 0 0 1 0 1 1 0 0 0 0 1 1 1 0 0 0 1 0 0 0 0 0 0 1 1
## [1395] 1 1 0 1 0 1 1 1 1 0 0 1 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 0 1 0 0 0 0
## [1429] 0 1 0 0 1 1 1 1 1 1 0 0 1 0 0 1 0 0 0 0 0 0 0 1 1 0 0 0 0 1 0 0 0 0
## [1463] 0 0 0 1 1 1 0 0 0 0 1 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 1
## [1497] 0 0 1 1 0 0 0 0 0 0 0 1 1 1 0 0 1 0 1 0 1 1 0 0 0 0 0 1 0 1 0 0 0 0
## [1531] 1 1 0 1 0 0 0 0 0 0 1 0 0 1 1 0 0 0 0 0 1 1 1 1 0 0 1 0 0 1 0 0 0 0
## [1565] 0 0 0 0 0 1 0 0 0 0 0 1 0 1 0 0 0 1 1 0 1 0 0 1 0 1 1 1 0 1 0 0 0 1
## [1599] 1 0 0 0 0 0 1 0 1 1 0 0 1 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 0 1 1 0 1 1
```

```
## [1633] 1 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 1 1 0 1 1 0 0 1 0 0 1 0 0 1 1 0 0
## [1667] 1 0 0 0 0 0 1 1 0 0 0 1 0 1 0 0 1 0 0 1 1 1 0 1 0 0 0 0 0 0 0 0 1 0
## [1701] 0 0 0 0 1 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 1 0 1 1
## [1735] 0 1 1 1 0 0 1 0 0 0 0 0 0 0 1 1 1 0 0 0 0 1 0 1 0 0 0 0 1 0 0 0 0 1
## [1769] 0 0 0 1 0 0 1 0 1 0 0 0 1 0 0 1 0 0 1 1 0 1 0 1 0 0 1 0 0 0 0 1 0 0
## [1803] 1 0 1 0 0 1 0 1 1 0 0 0 0 1 0 1 0 0 0 0 1 0 0 1 1 1 1 0 0 0 0 0 0 0
## [1837] 0 0 0 1 0 0 1 0 0 1 0 1 1 0 0 0 0 1 0 1 0 1 0 0 1 1 1 0 0 0 0 0 1 0
## [1871] 1 1 0 1 0 1 0 1 0 0 0 0 1 0 1 1 0 1 1 1 0 0 1 1 0 0 1 0 0 1 0 0 1 0
## [1905] 1 0 0 1 0 1 0 0 1 0 0 0 0 0 1 0 0 0 1 1 0 0 0 1 0 0 0 0 0 0 0 0 1 0
## [1939] 0 0 1 0 0 0 0 1 1 0 0 1 0 0 0 0 1 0 1 1 1 0 0 0 1 0 0 0 0 0 0 1 1 0
## [1973] 0 0 0 1 0 0 0 1 0 0 0 1 1 0 0 0 0 1 0 0 0 0 1 1 0 0 0 1 0 1 0 1 0 0
## [2007] 1 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 1 0 1 0 1 0 0 0 0 1 0 0 0 1
## [2041] 1 0 0 0 0 1 0 0 0 1 1 1 1 0 0 0 1 0 1 1 1 0 0 0 1 1 1 0 0 0 0 0 0 1
## [2075] 0 0 0 0 1 1 1 0 0 0 1 0 0 1 0 0 0 1 0 1 1 0 1 0 1 1 0 0 0 1 0 0 1 1
## [2109] 0 0 0 1 0 0 0 0 0 1 1 0 0 1 0 0 1 0 1 1 0 0 1 0 0 0 0 1 1 1 0 0 1 0
## [2143] 0 1 0 1 1 0 1 1 0 1 0 0 1 1 0 1 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1
## [2177] 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0 1 0
## [2211] 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1 1 1 0 0 1 0 0 0 0 0 1 1 0 0 0 0 1 1 0
## [2245] 0 1 1 0 0 1 1 0 1 0 1 0 0 1 1 0 1 1 0 0 0 0 0 1 1 0 0 0 1 0 1 0 0 0
## [2279] 1 0 1 0 0 0 0 1 0 0 0 1 0 0 1 1 1 1 1 0 1 0 0 1 0 0 0 0 1 1 0 0 0 0
## [2313] 0 1 1 0 0 1 0 1 0 0 0 1 0 0 0 0 0 1 1 0 0 1 0 0 1 0 0 0 0 1 0 1 0 1
## [2347] 0 0 1 1 0 0 0 1 1 0 1 0 1 1 1 0 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0
## [2381] 0 1 1 1 1 0 0 0 1 0 0 0 1 1 0 0 0 0 0 1 0 1 1 0 0 0 1 1 1 1 0 0 1 0
## [2415] 0 0 0 1 0 0 1 1 0 0 1 0 1 0 1 1 1 1 0 1 1 1 0 0 0 0 1 1 1 1 1 0 0 0
## [2449] 0 0 0 1 0 1 1 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1
## [2483] 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0 0 0 1 0 1 0 1 0
## [2517] 0 0 1 0 0 0 1 1 0 0 0 1 1 1 0 0 0 1 0 1 0 0 1 0 1 0 1 0 0 0 0 1 0 0
## [2551] 0 1 1 0 0 0 1 0 0 1 0 0 0 0 0 1 1 0 1 1 0 0 1 0 0 0 0 0 0 0 0 1 1 0
## [2585] 0 1 1 0 0 1 0 1 0 0 0 1 0 1 1 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 1 0
## [2619] 0 0 0 0 0 0 1 0 0 1 1 0 1 0 0 0 0 0 0 0 0 1 1 0 0 0 0 1 0 0 1 0 0 0
## [2653] 0 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 1 1 0 0 0 0
## [2687] 1 0 0 0 1 1 0 1 1 1 0 1 0 0 0 1 0 0 0 1 1 0 1 0 0 1 0 0 0 0 1 0 0 1
## [2721] 0 0 0 1 1 1 0 0 0 0 0 0 0 1 0 1 0 1 1 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0
## [2755] 1 1 0 0 1 1 0 0 0 1 0 0 1 0 0 1 0 0 0 1 0 1 0 0 0 1 1 1 0 0 1 0 0 0
## [2789] 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0 1 0 0 1 1 0
## [2823] 0 0 0 0 0 0 0 0 1 1 0 1 0 0 0 0 1 0 1 0 1 1 0 1 0 0 1 0 0 0 1 0 0 0
## [2857] 0 1 1 0 0 0 0 1 1 0 0 0 1 1 1 1 0 0 0 0 0 0 0 1 1 0 0 1 0 0 0 1 0 0
## [2891] 1 0 1 0 0 0 0 0 0 1 0 1 0 0 1 1 0 0 0 0 0 1 1 1 0 1 0 1 1 1 1 0 0 0
## [2925] 0 1 0 1 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0
## [2959] 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 1 1 1 0 0 0 0 1 0 0 1
## [2993] 0 1 0 0 0 1 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 1 0 1 0 0 0 0 1 1 0 1
## [3027] 0 1 0 0 1 0 0 0 1 1 1 0 0 1 0 0 0 0 1 0 0 0 0 0 1 1 1 0 0 0 1 0 1 0
## [3061] 0 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 1 1 1 0 1 1 1 0 1 0 0 1 0 0 0 0 0 0
## [3095] 1 0 0 0 1 1 1 0 0 0 0 1 0 0 0 1 0 1 1 0 0 0 0 0 0 0 1 0 1 1 0 0 0 0
## [3129] 1 1 0 0 1 1 0 0 0 0 0 1 1 0 1 1 0 1 0 1 1 0 0 0 0 0 1 1 0 0 1 1 1 1
## [3163] 1 0 0 1 0 1 0 0 0 0 0 1 0 0 0 0 1 1 1 0 0 0 0 1 0 0 1 1 0 1 0 1 0 1
## [3197] 0 0 1 1 0 1 1 0 0 0 1 1 0 1 1 0 1 1 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0
## [3231] 0 1 1 0 1 0 1 0 1 1 0 0 1 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0 1 0 1 1 0 1
## [3265] 0 1 0 0 0 0 0 0 0 1 1 1 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 1
```

```
## [3299] 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 1 0 0 0 1 1 0 0 0 0
## [3333] 0 1 1 1 0 0 1 1 1 0 0 0 0 0 1 0 0 1 0 1 1 1 0 0 0 0 1 0 0 1 0 0 1 0
## [3367] 0 0 0 0 0 0 1 1 1 1 0 0 0 1 1 0 1 0 0 1 0 1 0 0 1 1 1 1 1 0 1 0 1 0
## [3401] 0 0 0 1 1 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 1 1 1
## [3435] 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 1 1 0 0
## [3469] 0 1 0 0 0 0 1 0 1 1 0 0 0 1 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0 0 0
## [3503] 1 0 1 1 1 0 0 1 0 0 0 0 0 1 0 0 0 0 1 1 0 0 0 1 0 0 0 1 0 1 0 1 0 1
## [3537] 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 1 1 1 0 0 1 1 0 1 0 0 0 0 0 1 0
## [3571] 1 0 0 0 0 0 0 0 0 0 1 0 1 1 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 1 0 1
## [3605] 0 1 1 0 1 0 0 1 1 1 0 0 1 1 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 0 1 0 1 0
## [3639] 1 0 0 0 0 0 0 1 0 0 1 0 1 0 0 0 1 0 1 1 0 0 0 1 0 1 1 0 1 0 0 1 1 0
## [3673] 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 1 1 1 0 0 1 0 1 1 0 1 0 1 0 0
## [3707] 0 1 0 0 0 0 0 0 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 0 0 0 0 1 0 0 1 0 0 0
## [3741] 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 1 0 1 0 0 1 0 0 1 0 0 1 0 1 0 1 1 0 1
## [3775] 1 0 0 1 0 1 0 0 1 1 0 0 0 0 1 0 1 0 0 0 1 1 0 1 0 0 1 0 0 0 0 0 1 0
## [3809] 1 0 0 1 1 0 0 1 0 1 0 0 1 0 1 0 0 1 1 0 1 0 0 0 0 0 0 1 0 0 1 1 0 0
## [3843] 0 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0
## [3877] 0 1 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 1 1 0 0 0 0 1 0
## [3911] 0 1 0 1 0 0 0 0 1 0 0 1 0 1 0 0 1 0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0
## [3945] 0 1 0 0 0 1 0 1 1 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 1 0 1 0 0 0 0 0 1 0
## [3979] 0 1 1 0 0 0 0 0 0 0 1 0 1 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0
## [4013] 0 1 0 1 0 0 0 0 0 1 1 1 0 0 1 1 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0
## [4047] 0 0 0 1 0 1 0 1 1 1 0 0 0 0 1 0 1 0 0 0 0 0 1 0 0 0 1 0 0 1 1 1 1 0
## [4081] 0 0 1 0 0 0 0 0 0 1 1 0 1 0 0 0 0 1 0 0 0 0 0 1 1 1 1 0 0 1 0 0 0 1
## [4115] 0 0 0 0 0 0 1 1 0 1 0 1 1 1 1 1 0 1 1 0 0 1 1 1 0 0 0 0 0 1 0 1 0 0
## [4149] 1 0 0 0 1 1 0 1 0 0 1 0 0 0 1 0 0 0 0 1 1 1 1 1 0 1 0 1 0 0 1 0 0 0
## [4183] 0 0 0 1 1 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0
## [4217] 0 0 0 0 1 0 1 0 0 0 0 1 1 0 1 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0 1 0
## [4251] 0 0 0 1 1 1 0 1 0 0 0 1 0 1 0 1 0 1 0 0 0 0 0 0 1 1 0 1 1 1 0 1 1 0
## [4285] 0 0 0 0 0 1 0 1 0 1 0 0 1 0 1 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 1 1 0 0
## [4319] 0 0 0 0 1 0 0 0 1 0 1 0 0 0 1 1 0 0 0 0 0 0 1 0 1 1 0 1 0 1 0 1 1 0
## [4353] 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 1 0 1 1 0 1 0 1 0 0 0 0
## [4387] 0 0 0 1 1 0 0 0 1 0 1 1 1 0 0 1 1 0 1 0 0 0 1 0 1 0 0 0 1 0 0 1 1 1
## [4421] 0 0 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 0 1 0 0 0 1 1 1 1 0 0
## [4455] 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0
## [4489] 1 0 1 1 0 1 0 0 0 0 0 0 1 1 0 1 1 1 0 0 0 1 0 0 1 1 1 0 0 1 0 1 0 0
## [4523] 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 1 0
## [4557] 0 0 1 0 0 0 0 1 0 0 0 1 1 0 1 0 0 0 0 0 0 1 0 0 1 1 1 1 1 0 0 0 0 0
## [4591] 1 1 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 1 1 0 0 0
## [4625] 0 0 0 1 1 1 0 0 1 1 0 0 0 0 0 1 0 0 1 0 0 1 0 1 0 1 0 1 0 0 1 0 0 0
## [4659] 1 0 0 0 0 0 1 0 1 0 1 1 1 1 1 1 1 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1
## [4693] 0 0 0 0 0 1 1 0 1 1 1 0 0 1 0 1 1 1 0 1 0 1 0 1 0 0 0 1 0 0 0 0 0 0
## [4727] 1 0 0 0 0 0 0 1 0 1 0 0 0 1 1 1 1 1 0 0 0 1 0 1 0 0 0 1 1 1 1 0 0 1
## [4761] 0 1 1 0 1 1 0 0 1 1 0 0 1 1 1 0 1 1 0 0 0 1 1 1 0 1 1 1 0 0 1 0 0 1
## [4795] 0 0 0 1 0 1 1 0 0 1 1 0 0 0 1 0 1 0 0 0 0 1 1 0 0 1 1 1 1 1 1 1 0 0
## [4829] 0 0 0 0 0 0 1 0 1 0 1 0 1 1 1 0 0 1 1 1 0 1 0 1 0 0 0 1 1 1 0 0 1 0
## [4863] 1 0 1 0 0 0 0 1 0 0 1 0 0 0 1 0 0 1 1 1 0 1 1 1 0 1 1 1 0 0 1 1 0 1
## [4897] 1 0 0 1 0 1 0 0 0 0 1 1 1 0 0 0 1 1 1 1 0 1 0 1 0 0 0 1 0 1 0 1 0 0
## [4931] 0 0 0 0 0 1 0 0 1 1 0 1 0 0 1 0 0 1 0 0 1 1 0 1 0 0 1 1 0 1 0 0 0 0
```

```
## [4965] 1 1 0 1 1 0 1 0 0 0 0 0 0 0 1 1 1 0 1 0 0 0 0 1 1 1 0 1 1 0 0 1 0 0
## [4999] 1 1 0 1 1 1 1 0 0 1 0 1 0 0 1 1 0 1 0 0 0 0 0 0 0 1 0 1 1 1 0 1 0 0
## [5033] 0 1 1 0 0 1 0 0 0 0 1 0 0 0 1 1 0 1 0 0 1 0 1 0 0 1 0 0 1 1 0 0 1 1
## [5067] 1 1 1 1 0 0 0 1 1 0 1 0 1 0 1 0 0 0 1 1 0 0 0 0 1 1 0 1 0 0 1 0 0 0
## [5101] 0 0 1 0 0 0 0 0 1 0 0 0 0 1 1 0 0 1 1 0 0 1 1 0 1 0 0 1 1 0 0 1 0 1
## [5135] 1 0 0 1 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 1 1 0 1 1 0 0 0 1 0 0 1 1 0 0
## [5169] 1 0 0 0 1 0 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 1 0 0
## [5203] 0 1 0 0 0 1 1 0 1 1 1 0 0 1 0 0 0 1 0 0 0 1 1 0 1 0 0 1 0 1 0 0 0 0
## [5237] 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 1 1 0 1 0 1 1 1 1 0 1 0 1
## [5271] 1 1 1 0 0 0 0 0 0 1 1 0 0 1 0 1 0 0 0 1 0 0 1 0 1 1 0 0 1 0 1 0 1 0
## [5305] 1 0 1 0 1 0 1 0 0 0 1 0 1 1 1 0 1 1 0 0 0 1 0 0 0 1 0 0 1 1 0 1 0 0
## [5339] 0 0 1 1 1 1 0 1 0 0 0 1 0 0 0 1 0 0 1 0 1 1 0 0 0 0 0 1 1 1 1 1 1 0
## [5373] 1 0 0 0 0 0 1 1 0 0 1 1 0 0 0 1 1 0 0 1 0 0 0 0 1 1 0 0 1 0 0 0 1 0
## [5407] 0 1 0 0 1 0 1 1 1 0 0 1 0 1 1 1 0 0 0 0 1 0 0 0 1 0 1 0 0 1 1 1 0 0
## [5441] 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 1 1 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 0
## [5475] 1 0 0 1 1 0 1 0 0 0 1 0 0 1 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0
## [5509] 0 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 0 0 1 1 1 0 1 0 1 1 0 0 0 1 0 0 1
## [5543] 1 0 0 0 0 0 0 0 1 1 0 1 1 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0
## [5577] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0
## [5611] 1 1 0 0 1 0 1 0 1 0 0 0 1 0 0 1 0 0 1 0 0 1 0 0 1 1 1 0 0 1 0 0 0 0
## [5645] 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 1 0 0 1 0 0 0 1 0 0 0 0 1
## [5679] 0 1 1 0 1 0 0 1 1 1 0 0 0 0 0 0 0 0 1 0 0 0 1 1 0 0 0 0 0 0 0 0 1 0
## [5713] 0 1 0 0 1 0 1 1 0 1 0 0 0 0 0 0 1 1 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0
## [5747] 1 0 0 0 0 0 0 0 0 1 1 0 1 0 0 1 0 0 1 1 0 0 1 0 0 0 0 1 0 1 0 1 0 0
## [5781] 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 1 0 1 0 0 0 0 0 1 1 0 0 1 0 0
## [5815] 1 0 0 1 0 0 0 0 0 1 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [5849] 0 1 0 0 0 1 0 0 0 1 1 1 0 0 0 0 0 0 1 1 0 0 0 0 1 1 1 0 0 0 0 0 0 0
## [5883] 0 1 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 1 1 1 1 1 0 0 1 0 1 1 0 0 0 0 0 0
## [5917] 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 1 1 0 1 0 0 1 1 1 0 0 1 1 0 0 0 0 0
## [5951] 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0
## [5985] 0 1 1 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 1 1 1 0 0 0 1 0 0 0 1 0 1
## [6019] 0 1 1 0 1 0 1 1 0 0 0 0 1 1 0 0 0 1 1 0 0 0 1 0 1 0 0 0 1 0 0 1 1 0
## [6053] 1 1 0 1 0 0 1 1 0 0 0 0 1 1 0 0 1 1 1 0 1 0 0 0 0 0 0 1 0 0 0 0 1 0
## [6087] 1 0 1 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 1 1 0 1 1 1 0 0 0 0
## [6121] 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 1 1 0 1 0 1 0 1 1 1 0 1 0 1 1 0 1
## [6155] 0 0 0 0 0 0 1 1 1 0 1 1 1 1 1 0 0 0 0 1 0 0 1 1 0 0 0 0 1 0 0 0 0 0
## [6189] 1 1 0 0 0 0 0 0 0 1 0 0 0 0 1 1 1 1 1 1 1 0 0 1 1 1 1 1 0 1 0 1 1 0
## [6223] 1 1 0 0 0 1 0 1 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0
## [6257] 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 1 1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0
## [6291] 0 1 0 0 1 0 1 1 0 0 0 0 1 0 0 0 0 1 0 0 1 0 1 1 0 1 1 1 0 0 0 0 0 0
## [6325] 1 0 1 1 0 0 1 0 0 0 1 0 0 0 1 0 1 1 1 0 1 0 0 0 0 0 0 1 0 1 1 1 0 0
## [6359] 1 0 1 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0
## [6393] 0 0 0 1 1 1 1 0 1 1 1 1 0 1 1 0 1 0 0 1 1 0 1 0 1 0 0 0 0 0 1 0 1 0
## [6427] 0 0 0 0 0 0 1 0 1 1 0 0 0 0 1 0 0 0 0 0 1 0 0 1 0 0 1 0 1 0 1 0 0 0
## [6461] 0 0 0 1 0 0 1 0 0 0 1 0 1 0 0 1 0 0 1 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0
## [6495] 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
## [6529] 0 0 1 1 1 0 0 0 0 1 1 0 0 0 0 0 1 0 1 0 1 0 1 0 1 0 0 0 1 0 0 0 0 0
## [6563] 1 1 0 0 0 0 1 0 0 1 0 0 0 1 1 0 0 0 0 0 0 1 1 0 1 0 1 0 0 0 0 1 0 0
## [6597] 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 1
```

```
## [6631] 0 0 0 0 0 0 1 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
## [6665] 0 1 0 0 0 0 0 0 1 0 0 0 0 1 1 1 0 0 1 1 0 0 0 1 0 0 0 0 1 0 1 0 0 1
## [6699] 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 1 0 1 1
## [6733] 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
## [6767] 1 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 1 0 0 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0
## [6801] 1 1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 1 0 1 0 0 1 0 0 0 1 1 0 0 0 0
## [6835] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0 1 0 0 0
## [6869] 0 1 0 0 0 1 0 1 0 0 0 0 1 0 0 0 0 0 0 0 1 1 0 0 0 0 0 1 1 0 0 0 0 0
## [6903] 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0
## [6937] 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 1
## [6971] 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0 0 1 1 0 0 0 1 0 0
## [7005] 0 0 0 0 1 1 1 0 0 1 0 0 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0
## [7039] 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0
## [7073] 1 0 0 0 0 0 1 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0
## [7107] 0 1 1 1 1 1 1 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 1 0 0 1 1 0 0 0 0 1
## [7141] 1 1 0 0 0 1 1 1 1 1 1 0 0 1 1 0 1 0 0 0 0 0 1 1 0 0 1 1 0 0 0 0 0 1
## [7175] 0 1 1 0 0 1 0 1 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 1 0 0 1 0 1 0 1 0
## [7209] 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0 1 1 0 0 0 1 1 0 0 0 0 1 1 0 1 0 1
## [7243] 1 0 0 0 1 0 0 0 1 0 0 1 1 0 1 0 1 1 1 0 0 1 0 1 0 1 0 0 0 0 0 1 0 0
## [7277] 0 0 0 0 1 0 0 0 1 1 1 0 0 0 0 1 1 1 1 1 0 0 1 0 0 0 0 0 1 0 0 0 1 0
## [7311] 1 1 1 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0 1 1 0 0 0 0 0 1 1 1 0 0
## [7345] 0 1 0 0 0 0 1 1 0 0 1 1 0 1 0 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0 1 1 1 0
## [7379] 0 0 1 1 1 0 1 1 1 0 0 1 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0 0 1 0 0 1 0 0
## [7413] 1 1 0 0 0 0 0 1 0 0 1 0 1 0 0 1 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 1 0 0
## [7447] 0 1 1 0 1 0 0 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0 1 1 0 0 1 0 0 0 1 0 0 0
## [7481] 0 0 0 0 0 1 0 1 1 1 0 0 1 0 0 1 0 0 1 1 1 1 0 1 1 0 0 0 1 0 0 1 0 1
## [7515] 0 0 0 0 0 1 1 1 0 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 1
## [7549] 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 1 0
## [7583] 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0 1 1 0 0 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0
## [7617] 0 0 1 0 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
## [7651] 1 0 0 0 0 0 1 1 0 0 0 0 1 1 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 1 1 0 0 0
## [7685] 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0 0 0 1 0 0 0 1 0 0 0 0 0
## [7719] 0 0 1 0 1 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 1 0 0 1 1 1 0 0 0 1 0 0
## [7753] 1 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 0 0
## [7787] 0 0 1 0 1 1 0 0 1 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0
## [7821] 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0
## [7855] 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0
## [7889] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 1 1 1 0 1 0 1 1 0 1 1 0 1 0 0 0
## [7923] 0 0 1 0 0 0 1 1 1 0 0 1 0 0 1 0 0 0 0 0 0 1 1 1 1 0 0 0 1 0 1 1 0 1
## [7957] 1 1 1 0 1 0 1 0 1 0 1 0 0 0 1 1 0 1 0 1 0 1 1 1 0 0 0 1 0 0 0 1 1 1
## [7991] 0 1 0 1 0 1 1 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0 0 0 1 0
## [8025] 0 1 1 1 0 0 0 0 1 0 1 0 0 1 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0
## [8059] 0 0 1 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 0 0 0 1 0 0 0 0 1 0 1 0 0 1 0 1
## [8093] 0 0 1 1 0 0 0 0 1 1 1 0 0 1 0 0 0 0 0 1 0 0 0 0 0 1 0 1 1 0 0 0 0 1
## [8127] 0 0 0 1 1 0 1 1 0 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 1 1 1 0 0 0 0 0
## [8161] 0 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 1 0 0 1
## [8195] 1 1 0 0 1 0 0 0 1 1 0 0 1 1 0 0 1 0 1 0 1 1 1 1 1 1 1 1 1 0 1 0 1 0
## [8229] 0 1 1 0 1 1 1 0 1 0 1 0 1 0 0 0 0 0 1 1 0 0 1 1 0 1 0 0 0 0 0 0 1 1
## [8263] 0 0 1 0 0 0 1 0 0 1 0 0 0 0 1 1 0 0 1 0 1 0 1 0 0 0 0 0 1 0 1 1 0 0
```

```
## [8297] 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 1 1 0 0 1 0 0 0
## [8331] 0 0 1 0 1 0 0 0 1 0 0 0 1 0 0 0 0 1 1 1 0 1 0 0 0 0 0 0 1 1 0 0 0 0
## [8365] 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1 0 1 0 0 0 1 0 0 1 0 1 1 1 1 0 0 0
## [8399] 0 0 1 0 0 0 0 0 0 1 0 0 1 0 1 1 1 0 0 0 1 0 1 1 0 0 0 0 0 1 0 0 0 0
## [8433] 0 0 0 0 0 0 0 0 0 1 1 0 1 0 0 1 0 1 0 1 1 0 0 0 0 1 0 0 1 1 0 0 0 0
## [8467] 0 1 1 0 0 0 1 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 1 1 0 1 0 0 0 1 0 1 0 0
## [8501] 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 0 0 1 1 1 1 0 0 0 0 0 1 0 1
## [8535] 0 0 1 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 1 1 1 1
## [8569] 1 0 0 1 0 0 0 0 1 0 1 0 0 0 1 0 0 0 1 1 0 1 0 0 0 0 0 0 0 1 0 0 0 0
## [8603] 0 0 0 0 0 0 0 0 1 1 0 0 0 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0
## [8637] 0 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0 1 0 0 1 0 1 0 1 0 0 1 0 0 0 0 0 0 0
## [8671] 0 0 1 0 1 0 1 0 0 0 1 0 0 1 0 0 0 1 1 0 0 1 0 1 0 0 0 1 0 1 0 1 1 1
## [8705] 1 1 0 0 1 0 0 1 1 0 1 0 1 1 0 0 1 0 0 1 1 1 1 0 1 0 0 1 1 0 1 0 1 0
## [8739] 0 0 0 1 0 0 0 1 0 1 0 0 1 0 1 1 1 0 0 0 0 0 0 1 0 1 1 0 0 1 0 1 0 0
## [8773] 0 0 1 0 1 1 0 1 1 1 0 0 1 0 0 0 0 1 0 0 1 1 0 0 1 0 0 0 1 1 0 0 0 0
## [8807] 0 0 1 1 0 0 1 1 0 0 1 0 0 1 0 0 0 1 1 0 0 0 1 0 0 0 0 1 1 0 0 0 0 0
## [8841] 1 0 0 1 1 1 1 0 1 1 1 0 0 0 0 0 0 0 0 1 0 1 0 1 1 1 1 1 0 0 0 0 1 1
## [8875] 1 0 0 1 1 0 1 0 1 0 0 0 1 1 0 0 1 0 0 0 1 0 0 1 0 1 0 0 1 0 1 0 0 0
## [8909] 0 1 1 0 0 1 1 1 1 0 0 0 1 1 0 0 1 0 0 0 0 0 1 0 1 1 0 1 1 0 0 1 1 1
## [8943] 0 0 0 1 1 1 0 0 0 0 1 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 1 1 0 0 0 0 0
## [8977] 0 0 0 0 0 1 0 1 0 0 0 0 1 1 0 0 0 0 0 0 0 1 1 0 0
## Levels: 0 1
```

```
tab.NB =table(test_bank[,6], pred.NB)
tab.NB
```

```
##    pred.NB
##        0    1
##   0 5370 1642
##   1  729 1259
```

```
accuracy.NB = sum(diag(tab.NB))/sum(tab.NB)
accuracy.NB
```

```
## [1] 0.7365556
```

```
confusionMatrix(pred.NB, test_bank$DEFAULT)
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##          0 5370  729
##          1 1642 1259
##
##                Accuracy : 0.7366
##                  95% CI : (0.7273, 0.7456)
##     No Information Rate : 0.7791
##     P-Value [Acc > NIR] : 1
##
##                   Kappa : 0.3427
##
##  Mcnemar's Test P-Value : <2e-16
##
##             Sensitivity : 0.7658
##             Specificity : 0.6333
##          Pos Pred Value : 0.8805
##          Neg Pred Value : 0.4340
##              Prevalence : 0.7791
##          Detection Rate : 0.5967
##    Detection Prevalence : 0.6777
##       Balanced Accuracy : 0.6996
##
##        'Positive' Class : 0
##
```

Naive Bayes classifier

```
#names(train_bank)
#train_bank$DEFAULT = as.integer(train_bank$DEFAULT)
#test_bank$DEFAULT = as.integer(test_bank$DEFAULT)
#install.packages(mlr)
library(mlr)
```

```
## Loading required package: ParamHelpers
```

```
##
## Attaching package: 'mlr'
```

```
## The following object is masked from 'package:e1071':
##
##     impute
```

```
## The following object is masked from 'package:ROCR':
##
##     performance
```

```
## The following object is masked from 'package:caret':
##
##     train
```

```
#Create a classification task for learning on bank Dataet and specify default feature
task = makeClassifTask(data = train_bank,target ="DEFAULT")
#Initialize the Naive Bayes classifier
selected_model = makeLearner("classif.naiveBayes")
#Train the model
NB_mlr= train(selected_model,task)
#Read the model learned
NB_mlr$learner.model
```

```
##
## Naive Bayes Classifier for Discrete Predictors
##
## Call:
## naiveBayes.default(x = X, y = Y, laplace = laplace)
##
## A-priori probabilities:
## Y
##         0         1
## 0.4997309 0.5002691
##
## Conditional probabilities:
##    LIMIT_BAL
## Y       [,1]     [,2]
##   0 174127.9 126738.5
##   1 131340.4 114400.6
##
##    SEX
## Y       [,1]      [,2]
##   0 1.606935 0.4884836
##   1 1.568417 0.4953504
##
##    EDUCATION
## Y       [,1]      [,2]
##   0 1.835236 0.7937926
##   1 1.903614 0.7253323
##
##    MARRIAGE
## Y       [,1]      [,2]
##   0 1.563429 0.5205949
##   1 1.521945 0.5256006
##
##    AGE
## Y       [,1]     [,2]
##   0 35.46866 9.093063
##   1 35.80443 9.723742
##
##    BILLED._AMT
## Y           [,1]      [,2]
##   0  0.008420614 0.9801492
##   1 -0.042941510 0.9762284
##
##    REPAY_STATUS.
## Y          [,1]      [,2]
##   0 -0.08402265 0.7870855
##   1  0.40146671 1.3678757
##
##    PAID_AMT
```

```
## Y          [,1]      [,2]
##   0  0.07222019 0.9913057
##   1 -0.22903501 0.6413568
##
##    TIMELY_PAID_AMT
## Y         [,1]      [,2]
##   0  0.1593017 0.8065326
##   1 -0.5536048 1.1702067
##
##    RATIO_PADI_AMT1
## Y          [,1]      [,2]
##   0  0.02844571 0.9316682
##   1 -0.07652734 0.7439935
##
##    RATIO_PADI_AMT2
## Y           [,1]      [,2]
##   0  0.006558947 0.9229444
##   1 -0.009970515 0.9391812
##
##    RATIO_PADI_AMT3
## Y           [,1]      [,2]
##   0 -0.014470130 0.9407554
##   1  0.005906947 0.7912172
##
##    RATIO_PADI_AMT4
## Y          [,1]      [,2]
##   0  0.04260285 0.8371578
##   1 -0.10099714 1.0038388
```

```
#Predict on the dataset without passing the target feature
predictions_mlr = as.data.frame(predict(NB_mlr, newdata = test_bank[,1:13]))
```

```
## Warning in predict.naiveBayes(.model$learner.model, newdata = .newdata, :
## Type mismatch between training and new data for variable 'RATIO_PADI_AMT4'.
## Did you use factors with numeric labels for training, and numeric values
## for new data?
```

```
table(test_bank$DEFAULT)
```

```
##
##    0    1
## 7012 1988
```

```
##Confusion matrix to check accuracy
table(predictions_mlr[,1],test_bank$DEFAULT)
```

```
##
##         0    1
##    0 7012    0
##    1    0 1988
```

```
predictions_mlr = as.data.frame(predict(NB_mlr, newdata = test_bank[,-6]))

table(predictions_mlr[,1],test_bank$DEFAULT)
```

```
##
##         0    1
##    0 5370  729
##    1 1642 1259
```

Bagging

```
#Bagging#
library(gbm)
```

```
## Loaded gbm 2.1.5
```

```
#install.packages('xgboost')
library(xgboost)
```

```
##
## Attaching package: 'xgboost'
```

```
## The following object is masked from 'package:dplyr':
##
##     slice
```

```
#install.packages('caret')
library(caret)
library(ipred)
library(rpart)

bank_bagging <- bagging(DEFAULT ~.,data=train_bank,
                         control=rpart.control(maxdepth=5, minsplit=4))

pred_class <- predict(bank_bagging, test_bank)

tab.bg <- table(test_bank$DEFAULT,pred_class)
accuracy.bg = sum(diag(tab.bg))/sum(tab.bg)
accuracy.bg
```

```
## [1] 0.7672222
```

```
table(test_bank$DEFAULT,pred_class)
```

```
##    pred_class
##        0    1
##   0 5724 1288
##   1  807 1181
```

XGBoost

```
# XGBoost
#install.packages('xgboost')
library(xgboost)
set.seed(123)
classifier = xgboost(data = as.matrix(train_bank[,-6]), label = train_bank$DEFAULT, nr
ounds = 10)
```

```
## [1]  train-rmse:0.843655
## [2]  train-rmse:0.666488
## [3]  train-rmse:0.557688
## [4]  train-rmse:0.492921
## [5]  train-rmse:0.456142
## [6]  train-rmse:0.436011
## [7]  train-rmse:0.423342
## [8]  train-rmse:0.415939
## [9]  train-rmse:0.411865
## [10] train-rmse:0.408762
```

```
#Predicting the Test set results
y_pred <- predict(classifier, newdata = as.matrix(test_bank[-c(6,15)]))
y_pred = (y_pred >= 0.5)

# Making the Confusion Matrix
cm = table(test_bank$DEFAULT, y_pred)
cm
```

```
##    y_pred
##     TRUE
##   0 7012
##   1 1988
```

```
accuracy.bs = sum(diag(cm))/sum(cm)
accuracy.bs
```

```
## [1] 0.7791111
```

K-Fold Cross Validation

```
library(caret)
folds_bank = createFolds(train_bank$DEFAULT, k =10)
cv = lapply(folds_bank, function(x) {
tr_fold = train_bank[-x, ]
tt_fold = test_bank[x, ]
classifier = xgboost(data = as.matrix(train_bank[-6]), label = train_bank$DEFAULT, nro
unds = 10)
y_pred = predict(classifier, newdata = as.matrix(tt_fold[-c(6,15)]))
y_pred = (y_pred >= 0.5)
cmx= table(tt_fold[,6], y_pred)
accuracy = (cmx[1,1] + cmx[2,1])  / (cmx[1,1] + cmx[2,1] + cmx[1,1] + cmx[2,1])
return(accuracy)
})
```

```
## [1]  train-rmse:0.843655
## [2]  train-rmse:0.666488
## [3]  train-rmse:0.557688
## [4]  train-rmse:0.492921
## [5]  train-rmse:0.456142
## [6]  train-rmse:0.436011
## [7]  train-rmse:0.423342
## [8]  train-rmse:0.415939
## [9]  train-rmse:0.411865
## [10] train-rmse:0.408762
## [1]  train-rmse:0.843655
## [2]  train-rmse:0.666488
## [3]  train-rmse:0.557688
## [4]  train-rmse:0.492921
## [5]  train-rmse:0.456142
## [6]  train-rmse:0.436011
## [7]  train-rmse:0.423342
## [8]  train-rmse:0.415939
## [9]  train-rmse:0.411865
## [10] train-rmse:0.408762
## [1]  train-rmse:0.843655
## [2]  train-rmse:0.666488
## [3]  train-rmse:0.557688
## [4]  train-rmse:0.492921
## [5]  train-rmse:0.456142
## [6]  train-rmse:0.436011
## [7]  train-rmse:0.423342
## [8]  train-rmse:0.415939
## [9]  train-rmse:0.411865
## [10] train-rmse:0.408762
## [1]  train-rmse:0.843655
## [2]  train-rmse:0.666488
## [3]  train-rmse:0.557688
## [4]  train-rmse:0.492921
## [5]  train-rmse:0.456142
## [6]  train-rmse:0.436011
## [7]  train-rmse:0.423342
## [8]  train-rmse:0.415939
## [9]  train-rmse:0.411865
## [10] train-rmse:0.408762
## [1]  train-rmse:0.843655
## [2]  train-rmse:0.666488
## [3]  train-rmse:0.557688
## [4]  train-rmse:0.492921
## [5]  train-rmse:0.456142
## [6]  train-rmse:0.436011
## [7]  train-rmse:0.423342
## [8]  train-rmse:0.415939
```

```
## [9]  train-rmse:0.411865
## [10] train-rmse:0.408762
## [1]  train-rmse:0.843655
## [2]  train-rmse:0.666488
## [3]  train-rmse:0.557688
## [4]  train-rmse:0.492921
## [5]  train-rmse:0.456142
## [6]  train-rmse:0.436011
## [7]  train-rmse:0.423342
## [8]  train-rmse:0.415939
## [9]  train-rmse:0.411865
## [10] train-rmse:0.408762
## [1]  train-rmse:0.843655
## [2]  train-rmse:0.666488
## [3]  train-rmse:0.557688
## [4]  train-rmse:0.492921
## [5]  train-rmse:0.456142
## [6]  train-rmse:0.436011
## [7]  train-rmse:0.423342
## [8]  train-rmse:0.415939
## [9]  train-rmse:0.411865
## [10] train-rmse:0.408762
## [1]  train-rmse:0.843655
## [2]  train-rmse:0.666488
## [3]  train-rmse:0.557688
## [4]  train-rmse:0.492921
## [5]  train-rmse:0.456142
## [6]  train-rmse:0.436011
## [7]  train-rmse:0.423342
## [8]  train-rmse:0.415939
## [9]  train-rmse:0.411865
## [10] train-rmse:0.408762
## [1]  train-rmse:0.843655
## [2]  train-rmse:0.666488
## [3]  train-rmse:0.557688
## [4]  train-rmse:0.492921
## [5]  train-rmse:0.456142
## [6]  train-rmse:0.436011
## [7]  train-rmse:0.423342
## [8]  train-rmse:0.415939
## [9]  train-rmse:0.411865
## [10] train-rmse:0.408762
## [1]  train-rmse:0.843655
## [2]  train-rmse:0.666488
## [3]  train-rmse:0.557688
## [4]  train-rmse:0.492921
## [5]  train-rmse:0.456142
## [6]  train-rmse:0.436011
## [7]  train-rmse:0.423342
```

```
## [8]  train-rmse:0.415939
## [9]  train-rmse:0.411865
## [10] train-rmse:0.408762
```

```
accuracy = mean(as.numeric(cv))
accuracy
```

```
## [1] 0.5
```