Homework 2
Zhaoxing Wu
Stat 443

In the homework, I want to predict the probability of the patients using right heart catheterization using a logistical regression model. First, I need to deal with the missing values. For the missing values in cat2, I created a new level called "Missing" for this variable, because there are more than 4000 data missing for this variable and it would be hard to predict the missing value based on the remaining value. For the missing values in meanbp1, hrt1, resp1, and wtkilo1, I used the median value to replace NA in the dataset. This is because there are only around a hundred data missing in these variables, and median will not be affected by outliers. Second, I used the important scores produced by GUIDE as a reference to build the logistical regression model. Some of the variables I chose represent the health condition about the patient, like cat1, cat2, aps1, pafi1, meanbp1, alb1, neuro, card, hema1, seps1, resp, dnr1, bili1, paco21, hrt1, transhx, chrpulhx, resp1, dementhx, ph1, psychhx, renal, gastr, age, wtkilo1 and from my perspective, the weaker the patient, the higher likelihood that the patient will do the operation. I also include variables such as income and ninsclas, since they represent the patient's affordability to rhc operation.

The following is the summary of the logistical regression model in r. All of the splitting variables used in the tree node are all statistically significant in the regression model, except some of the levels in the cat1 and ninsclas variable. Variable crea1 is the third most important variable according to the plot of importance scores, but it is not statistically significant in the regression model. Alb1, seps, bili1, chrpulhx, renal, and income are also highly important variables in the plot of importance scores, but not statistically significant in the regression model. According to the regression model, patients with CHF, MOSF w/sepsis, high APACHE score, Cardiovascular Diagnosis, high weight, high Heart rate, being transferred from another hospital have a higher probability to have done rhc operation. Some of the variables do not affect the probability of getting rhc, such as sex, income, sepsis, bilirubin and so on.

I used the confusion matrix to check the accuracy of the regression model and it is 0.7289. In the forest model, the resubstitution estimate of mean misclassification cost is only .1357. So, the forest model is more accurate. And in the regression model, there is might be potential interaction or dependency that hasn't been dealt with.

```
##
## Call:
## glm(formula = swang1 ~ cat1 + cat2 + age + aps1 + crea1 + pafi1 +
##     meanbp1 + sex + alb1 + neuro + card + hema1 + wtkilo1 + seps +
##     resp + dnr1 + bili1 + paco21 + hrt1 + transhx + chrpulhx +
##     resp1 + ninsclas + dementhx + ph1 + psychhx + renal + gastr +
##     income, family = "binomial", data = rhc)
##
## Deviance Residuals:
##     Min      1Q   Median      3Q      Max
## -2.5479  -0.8498  -0.4272   0.9276   2.7191
##
## Coefficients:
##                              Estimate Std. Error z value Pr(>|z|)
## (Intercept)                 10.4923288  2.8915649   3.629 0.000285 ***
```

```
## cat1CHF                       0.7662971  0.1343659   5.703 1.18e-08 ***
## cat1Cirrhosis                -1.1875691  0.2012319  -5.901 3.60e-09 ***
## cat1Colon Cancer             -0.8098403  1.1029319  -0.734 0.462789
## cat1Coma                     -0.3987114  0.1544601  -2.581 0.009842 **
## cat1COPD                     -0.5934339  0.1869666  -3.174 0.001503 **
## cat1Lung Cancer              -0.7271691  0.5114540  -1.422 0.155094
## cat1MOSF w/Malignancy        -0.4600714  0.1363432  -3.374 0.000740 ***
## cat1MOSF w/Sepsis             0.5747715  0.0993191   5.787 7.16e-09 ***
## cat2Colon Cancer              2.0182096  1.5821048   1.276 0.202080
## cat2Coma                      0.1247385  0.4894883   0.255 0.798851
## cat2Lung Cancer              -0.4448814  0.8813987  -0.505 0.613738
## cat2Missing                   0.6688793  0.4023745   1.662 0.096447 .
## cat2MOSF w/Malignancy         0.0213216  0.4310095   0.049 0.960546
## cat2MOSF w/Sepsis             1.1765263  0.4070411   2.890 0.003847 **
## age                           0.0002531  0.0026040   0.097 0.922581
## aps1                          0.0124199  0.0028313   4.387 1.15e-05 ***
## crea1                        -0.0039256  0.0190055  -0.207 0.836362
## pafi1                        -0.0048171  0.0003344 -14.404  < 2e-16 ***
## meanbp1                      -0.0071503  0.0010092  -7.085 1.39e-12 ***
## sexMale                      -0.0109786  0.0669880  -0.164 0.869819
## alb1                         -0.0819671  0.0467797  -1.752 0.079741 .
## neuroYes                     -0.5222866  0.1315844  -3.969 7.21e-05 ***
## cardYes                       0.5615296  0.0830290   6.763 1.35e-11 ***
## hema1                        -0.0127585  0.0046026  -2.772 0.005571 **
## wtkilo1                       0.0102853  0.0017033   6.038 1.56e-09 ***
## sepsYes                       0.0979139  0.0960007   1.020 0.307762
## respYes                      -0.3176959  0.0813271  -3.906 9.37e-05 ***
## dnr1Yes                      -0.6008507  0.1145042  -5.247 1.54e-07 ***
## bili1                         0.0093454  0.0072513   1.289 0.197470
## paco21                       -0.0244687  0.0035734  -6.847 7.52e-12 ***
## hrt1                          0.0043669  0.0010254   4.259 2.06e-05 ***
## transhx                       0.4359999  0.0982472   4.438 9.09e-06 ***
## chrpulhx                     -0.0022255  0.1000208  -0.022 0.982248
## resp1                        -0.0268091  0.0027701  -9.678  < 2e-16 ***
## ninsclasMedicare              0.2314148  0.1340998   1.726 0.084403 .
## ninsclasMedicare & Medicaid   0.3693986  0.1672153   2.209 0.027166 *
## ninsclasNo insurance          0.5117753  0.1656576   3.089 0.002006 **
## ninsclasPrivate               0.4376032  0.1237659   3.536 0.000407 ***
## ninsclasPrivate & Medicare    0.3625996  0.1372789   2.641 0.008258 **
## dementhx                     -0.4242212  0.1184504  -3.581 0.000342 ***
## ph1                          -1.3676117  0.3735161  -3.661 0.000251 ***
## psychhx                      -0.4477258  0.1379162  -3.246 0.001169 **
## renalYes                      0.2062996  0.1467169   1.406 0.159693
## gastrYes                      0.2610842  0.1022334   2.554 0.010655 *
## income$11-$25k               -0.0745562  0.1333433  -0.559 0.576073
## income$25-$50k               -0.0082732  0.1351751  -0.061 0.951197
## incomeUnder $11k             -0.0784984  0.1249599  -0.628 0.529880
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 7621.4  on 5734  degrees of freedom
## Residual deviance: 6025.7  on 5687  degrees of freedom
## AIC: 6121.7
##
## Number of Fisher Scoring iterations: 5
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##          0 2948  952
##          1  603 1232
##
##               Accuracy : 0.7289
##                 95% CI : (0.7172, 0.7403)
##    No Information Rate : 0.6192
##    P-Value [Acc > NIR] : < 2.2e-16
##
```

GUIDE importance scores

RHC
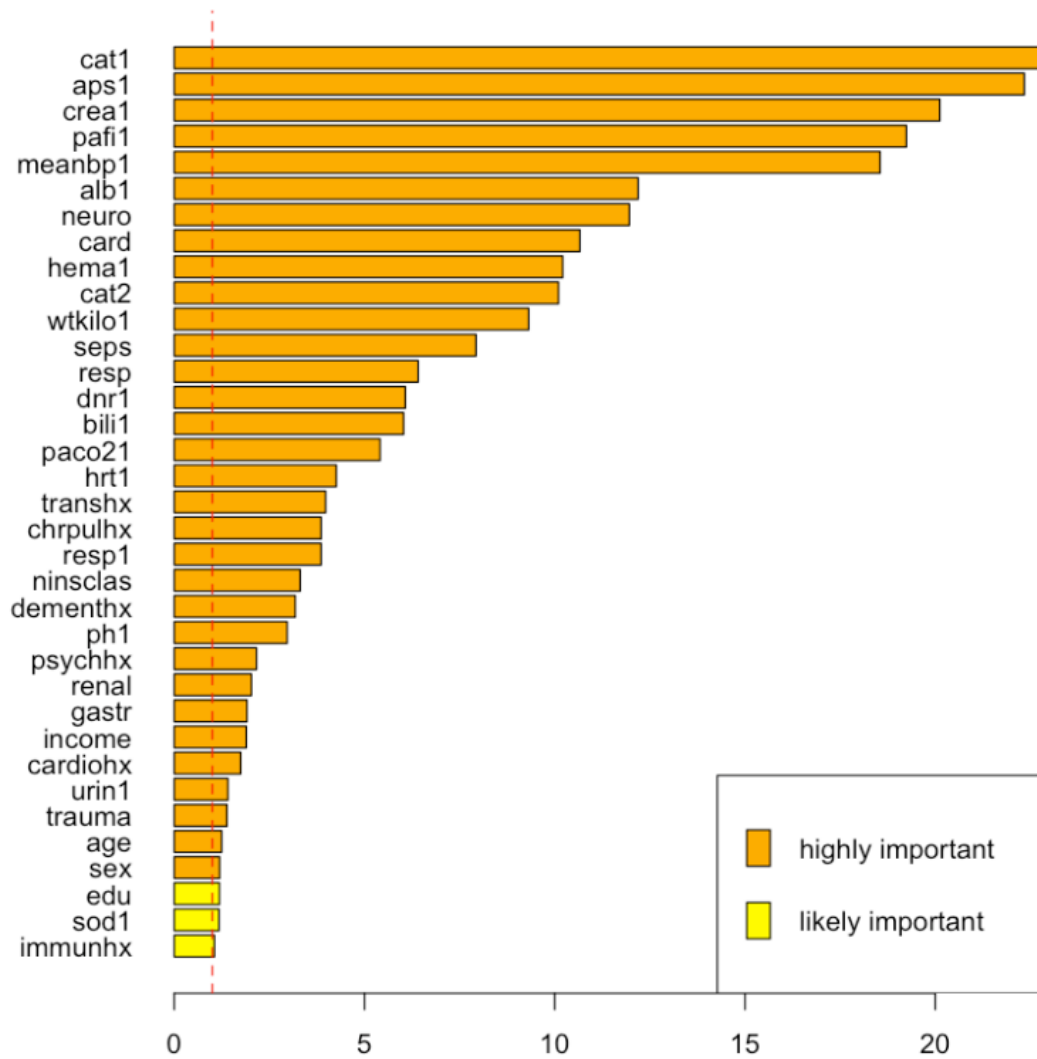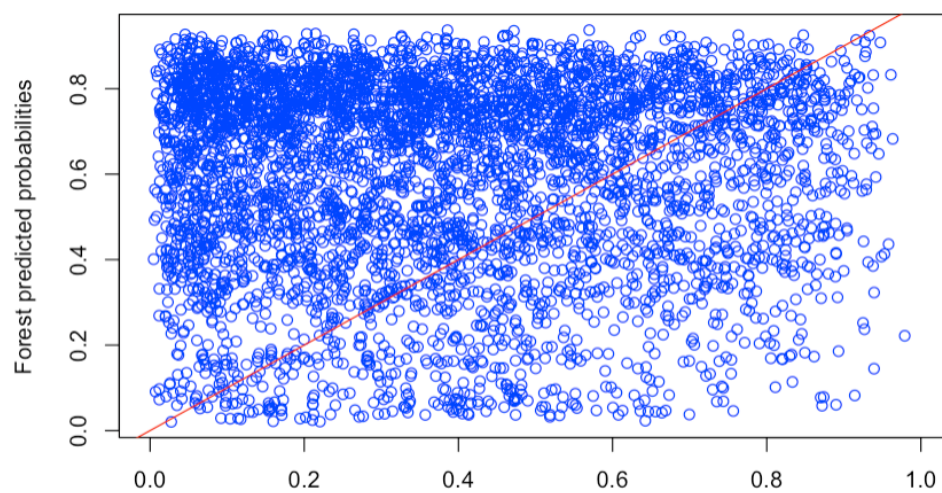
```r
library(ggplot2)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --
```

```
## v tibble  3.0.4     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.0
## v purrr   0.3.4
```

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

```r
library(caret)
```

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

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

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

```r
rhc = read.csv("rhcdata.txt", sep = " ")
#summary(rhc)
#sapply(rhc, function(x) sum(is.na(x)))
#sapply(rhc, function(x) length(unique(x)))
```

```r
par(las=1,mar=c(5,12,4,2),cex=1)
leg.col <- c("orange","yellow")
leg.txt <- c("highly important","likely important")
x <- read.table("impscr.txt",header=TRUE)
score <- x$Score
vars <- x$Variable
type <- x$Type
barcol <- rep("orange",length(vars))
barcol[type == "L"] <- "yellow"
barcol[type == "U"] <- "cyan"
n <- sum(x$Type != "U")
barplot(rev(score[1:n]),names.arg=rev(vars[1:n]),col=rev(barcol[1:n]),horiz=TRUE,xlab="GUIDE importance
abline(v=1,col="red",lty=2)
legend("bottomright",legend=leg.txt,fill=leg.col)
```
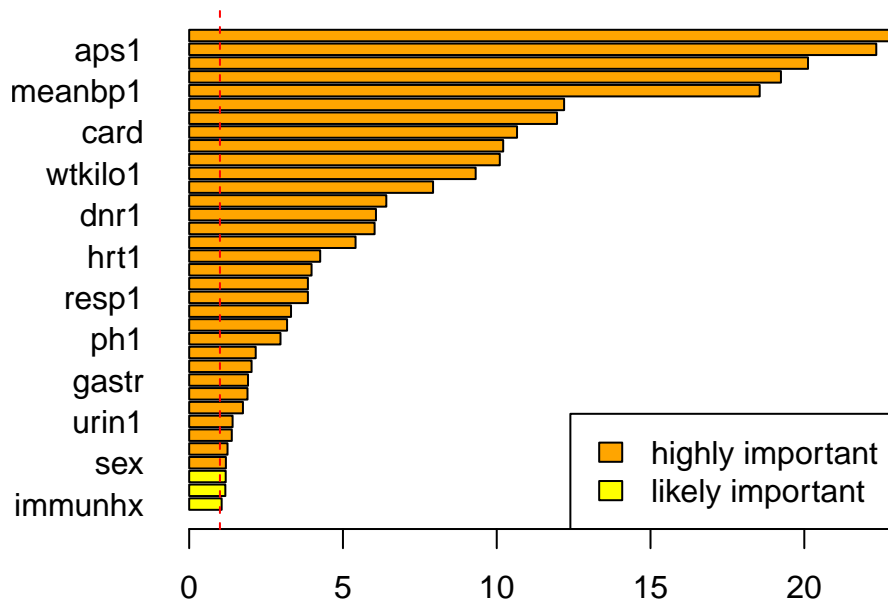
GUIDE importance scores

```
rhc = rhc %>%
  mutate(swang1 = replace(swang1, swang1 == "RHC", "1"),
         swang1 = replace(swang1, swang1 == "NoRHC", 0)) %>%
  replace_na(list(cat2 = "Missing",
                  meanbp1 = median(rhc$meanbp1, na.rm = TRUE),
                  hrt1 = median(rhc$hrt1, na.rm = TRUE),
                  resp1 = median(rhc$resp1, na.rm = TRUE),
                  wtkilo1 = median(rhc$wtkilo1, na.rm = TRUE)))


#cat2, wtkilo1, hema1, meanbp1
rhc$swang1 = as.integer(rhc$swang1)
rhc$cat1 = as.factor(rhc$cat1)
rhc$cat2 = as.factor(rhc$cat2)
rhc$aps1 = as.numeric(rhc$aps1)
rhc$crea1 = as.numeric(rhc$crea1)
rhc$pafi1 = as.numeric(rhc$pafi1)
rhc$meanbp1 = as.numeric(rhc$meanbp1)
rhc$alb1 = as.numeric(rhc$alb1)
rhc$neuro = as.factor(rhc$neuro)
rhc$card = as.factor(rhc$card)
rhc$hema1 = as.numeric(rhc$hema1)
rhc$wtkilo1 = as.numeric(rhc$wtkilo1)
rhc$seps = as.factor(rhc$seps)
rhc$resp = as.factor(rhc$resp)
rhc$dnr1 = as.factor(rhc$dnr1)
rhc$bili1 = as.numeric(rhc$bili1)
rhc$paco21 = as.numeric(rhc$paco21)
rhc$hrt1 = as.integer(rhc$hrt1)
rhc$transhx = as.integer(rhc$transhx)
rhc$chrpulhx = as.integer(rhc$chrpulhx)
```

```
rhc$resp1 = as.integer(rhc$resp1)
rhc$ninsclas = as.factor(rhc$ninsclas)
rhc$dementhx = as.integer(rhc$dementhx)
rhc$ph1 = as.numeric(rhc$ph1)
rhc$psychhx = as.integer(rhc$psychhx)
rhc$renal = as.factor(rhc$renal)
rhc$gastr = as.factor(rhc$gastr)
rhc$income = as.factor(rhc$income)

rhc$age = as.numeric(rhc$age)
rhc$sex = as.factor(rhc$sex)
logistic = glm(swang1~cat1+cat2+age+aps1+crea1+pafi1+meanbp1+sex+alb1+neuro+card+hema1+wtkilo1+seps+resp
summary(logistic)
```

```
##
## Call:
## glm(formula = swang1 ~ cat1 + cat2 + age + aps1 + crea1 + pafi1 +
##     meanbp1 + sex + alb1 + neuro + card + hema1 + wtkilo1 + seps +
##     resp + dnr1 + bili1 + paco21 + hrt1 + transhx + chrpulhx +
##     resp1 + ninsclas + dementhx + ph1 + psychhx + renal + gastr +
##     income, family = "binomial", data = rhc)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.5479  -0.8498  -0.4272   0.9276   2.7191
##
## Coefficients:
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)            10.4923288  2.8915649   3.629 0.000285 ***
## cat1CHF                 0.7662971  0.1343659   5.703 1.18e-08 ***
## cat1Cirrhosis          -1.1875691  0.2012319  -5.901 3.60e-09 ***
## cat1Colon Cancer       -0.8098403  1.1029319  -0.734 0.462789
## cat1Coma               -0.3987114  0.1544601  -2.581 0.009842 **
## cat1COPD               -0.5934339  0.1869666  -3.174 0.001503 **
## cat1Lung Cancer        -0.7271691  0.5114540  -1.422 0.155094
## cat1MOSF w/Malignancy  -0.4600714  0.1363432  -3.374 0.000740 ***
## cat1MOSF w/Sepsis       0.5747715  0.0993191   5.787 7.16e-09 ***
## cat2Colon Cancer        2.0182096  1.5821048   1.276 0.202080
## cat2Coma                0.1247385  0.4894883   0.255 0.798851
## cat2Lung Cancer        -0.4448814  0.8813987  -0.505 0.613738
## cat2Missing             0.6688793  0.4023745   1.662 0.096447 .
## cat2MOSF w/Malignancy   0.0213216  0.4310095   0.049 0.960546
## cat2MOSF w/Sepsis       1.1765263  0.4070411   2.890 0.003847 **
## age                     0.0002531  0.0026040   0.097 0.922581
## aps1                    0.0124199  0.0028313   4.387 1.15e-05 ***
## crea1                  -0.0039256  0.0190055  -0.207 0.836362
## pafi1                  -0.0048171  0.0003344 -14.404  < 2e-16 ***
## meanbp1                -0.0071503  0.0010092  -7.085 1.39e-12 ***
## sexMale                -0.0109786  0.0669880  -0.164 0.869819
## alb1                   -0.0819671  0.0467797  -1.752 0.079741 .
## neuroYes               -0.5222866  0.1315844  -3.969 7.21e-05 ***
## cardYes                 0.5615296  0.0830290   6.763 1.35e-11 ***
## hema1                  -0.0127585  0.0046026  -2.772 0.005571 **
```

```
## wtkilo1                         0.0102853  0.0017033   6.038 1.56e-09 ***
## sepsYes                         0.0979139  0.0960007   1.020 0.307762
## respYes                        -0.3176959  0.0813271  -3.906 9.37e-05 ***
## dnr1Yes                        -0.6008507  0.1145042  -5.247 1.54e-07 ***
## bili1                           0.0093454  0.0072513   1.289 0.197470
## paco21                         -0.0244687  0.0035734  -6.847 7.52e-12 ***
## hrt1                            0.0043669  0.0010254   4.259 2.06e-05 ***
## transhx                         0.4359999  0.0982472   4.438 9.09e-06 ***
## chrpulhx                       -0.0022255  0.1000208  -0.022 0.982248
## resp1                          -0.0268091  0.0027701  -9.678  < 2e-16 ***
## ninsclasMedicare                0.2314148  0.1340998   1.726 0.084403 .
## ninsclasMedicare & Medicaid     0.3693986  0.1672153   2.209 0.027166 *
## ninsclasNo insurance            0.5117753  0.1656576   3.089 0.002006 **
## ninsclasPrivate                 0.4376032  0.1237659   3.536 0.000407 ***
## ninsclasPrivate & Medicare      0.3625996  0.1372789   2.641 0.008258 **
## dementhx                       -0.4242212  0.1184504  -3.581 0.000342 ***
## ph1                            -1.3676117  0.3735161  -3.661 0.000251 ***
## psychhx                        -0.4477258  0.1379162  -3.246 0.001169 **
## renalYes                        0.2062996  0.1467169   1.406 0.159693
## gastrYes                        0.2610842  0.1022334   2.554 0.010655 *
## income$11-$25k                 -0.0745562  0.1333433  -0.559 0.576073
## income$25-$50k                 -0.0082732  0.1351751  -0.061 0.951197
## incomeUnder $11k               -0.0784984  0.1249599  -0.628 0.529880
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 7621.4  on 5734  degrees of freedom
## Residual deviance: 6025.7  on 5687  degrees of freedom
## AIC: 6121.7
##
## Number of Fisher Scoring iterations: 5
```

```r
logModel <- data.frame(predict(logistic, type = "response"))
logModel = logModel %>%
  mutate(pred = ifelse(logModel$predict.logistic..type....response..>0.5, 1, 0))
forest <- read.table("gfpro.txt",header=TRUE)
forest.p <- forest[,3]
plot(forest.p ~ logModel$predict.logistic..type....response..[1:4693], xlab="Logistic regression predict
abline(c(0,1),col="red")
```

Forest predicted probabilities (y-axis) vs Logistic regression predicted probabilities (x-axis)
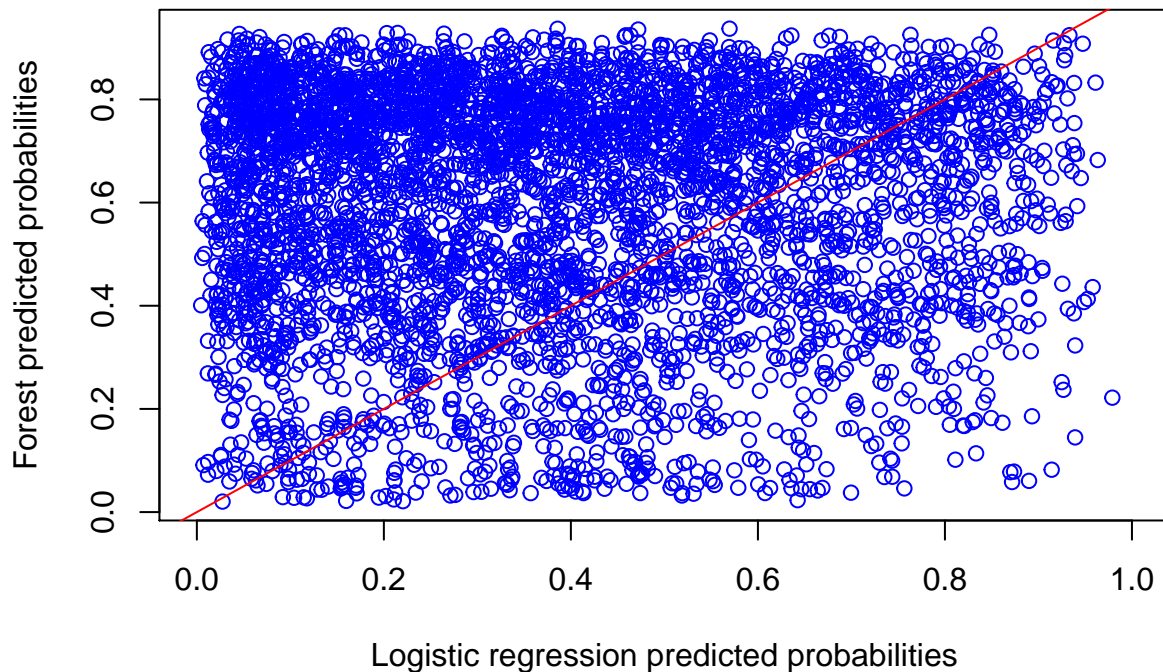
```
pdata = as.factor(as.integer(predict(logistic, newdata = rhc, type = "response")>0.5))
confusionMatrix(data = pdata,
               reference  = as.factor(rhc$swang1))
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    0    1
##          0 2948  952
##          1  603 1232
##
##                Accuracy : 0.7289
##                  95% CI : (0.7172, 0.7403)
##     No Information Rate : 0.6192
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.4068
##
##  Mcnemar's Test P-Value : < 2.2e-16
##
##             Sensitivity : 0.8302
##             Specificity : 0.5641
##          Pos Pred Value : 0.7559
##          Neg Pred Value : 0.6714
##              Prevalence : 0.6192
##          Detection Rate : 0.5140
##    Detection Prevalence : 0.6800
##       Balanced Accuracy : 0.6971
##
##        'Positive' Class : 0
##
```

```
GUIDE         (do not edit this file unless you know what you are doing)
  36.2        (version of GUIDE that generated this file)
 2            (1=model fitting, 2=importance or DIF scoring, 3=data
conversion)
"imp.out"  (name of output file)
 1            (1=classification, 2=regression, 3=propensity score
grouping)
 1            (1=univariate and interaction splits, 2=skip interactions)
"rhcdsc1.txt"  (name of data description file)
 1            (1=estimated priors, 2=equal priors, 3=other priors)
 1            (1=unit misclassification costs, 2=other)
 2            (1=split point from quantiles, 2=use exhaustive search)
 1            (1=default max. number of split levels, 2=specify no. in
next line)
 1            (1=default min. node size, 2=specify min. value in next
line)
 2            (0=no LaTeX code, 1=tree without node numbers, 2=tree with
node numbers)
"imp.tex" (latex file name)
 1            (1=color terminal nodes, 2=no colors)
 2            (0=#errors, 1=sample sizes, 2=sample proportions,
3=posterior probs, 4=nothing)
 1            (1=no storage, 2=store fit and split variables, 3=store
split variables and values)
 1            (1=do not create description file for selected variables,
2=create the file)
 1            (1=create file for importance scores, 2=do not create)
"imp.scr" (file name for importance scores)
 1            (rank of top variable to split root node)
```

```
              GGG   U   U  I  DDDD   EEEE
              G   G U   U  I  D   D  E
              G     U   U  I  D   D  E
              G  GG U   U  I  D   D  EEE
              G   G U   U  I  D   D  E
              G   G U   U  I  D   D  E
               GGG    UUU  I  DDDD   EEEE
```

 GUIDE Classification and Regression Trees and Forests
 Version 36.2 (Build date: January 8, 2021)
 Compiled with NAG Fortran 6.2.0 on macOS Catalina 10.15.7
 Copyright (c) 1997-2020 Wei-Yin Loh. All rights reserved.
 This software is based upon work supported by the U.S. Army Research
Office,
 the National Science Foundation and the National Institutes of
Health.

 This job was started on 02/16/21 at 20:48

 Classification tree
 No pruning
 Data description file: rhcdsc1.txt
 Training sample file: rhcdata.txt
 Missing value code: NA
 Records in data file start on line 2
 20 N variables changed to S
 D variable is swang1
 Number of records in data file: 5735
 Length of longest entry in data file: 19
 Missing values found among categorical variables
 Separate categories will be created for missing categorical variables
 Missing values found among non-categorical variables
 Number of classes: 2
 Training sample class proportions of D variable swang1:
 Class #Cases   Proportion
 NoRHC   3551   0.61918047
 RHC   2184   0.38081953

     Total #cases w/   #missing
    #cases   miss. D ord. vals   #X-var   #N-var   #F-var   #S-var
     5735   0   3443   13   0   0   20
    #P-var   #M-var   #B-var   #C-var   #I-var
         0   0   0   30   0
 Number of cases used for training: 5735
 Number of split variables: 50
 Number of cases excluded due to 0 weight or missing D: 0

 Importance scoring of variables
 Simple node models
 Estimated priors
```

Unit misclassification costs
Univariate split highest priority
Interaction splits 2nd priority; no linear splits
Split values for N and S variables based on exhaustive search
Maximum number of split levels: 4
Minimum node sample size: 57
Starting 300 permutations to standardize means of importance scores
Finished permutations to standardize means of importance scores
95 and 99% thresholds for unadjusted importance scores =   21.276
25.602

 Note: final tree is shorter due to pruning of sibling nodes with same
predicted values.
 Structure of final tree. Each terminal node is marked with a T.

 Node cost is node misclassification cost divided by number of
training cases

| Node label | Total cases | Train cases | Predicted class | Node cost | Split variables | Interacting variable |
|---|---|---|---|---|---|---|
| 1 | 5735 | 5735 | NoRHC | 3.808E−01 | cat1 | |
| 2 | 1683 | 1683 | RHC | 4.599E−01 | meanbp1 | |
| 4 | 1117 | 1117 | RHC | 3.796E−01 | pafi1 | |
| 8T | 655 | 655 | RHC | 3.038E−01 | resp1 | |
| 9 | 462 | 462 | RHC | 4.870E−01 | ninsclas | |
| 18T | 244 | 244 | RHC | 3.730E−01 | bili1 | |
| 19T | 218 | 218 | NoRHC | 3.853E−01 | card | |
| 5 | 566 | 566 | NoRHC | 3.816E−01 | alb1 | |
| 10 | 158 | 158 | RHC | 4.810E−01 | pafi1 | |
| 20T | 72 | 72 | RHC | 3.194E−01 | − | |
| 21T | 86 | 86 | NoRHC | 3.837E−01 | − | |
| 11T | 408 | 408 | NoRHC | 3.284E−01 | bili1 | |
| 3 | 4052 | 4052 | NoRHC | 3.147E−01 | pafi1 | |
| 6 | 1292 | 1292 | NoRHC | 4.837E−01 | resp | |
| 12 | 581 | 581 | RHC | 4.200E−01 | dnr1 | |
| 24T | 515 | 515 | RHC | 3.903E−01 | cat1 | |
| 25T | 66 | 66 | NoRHC | 3.485E−01 | − | |
| 13 | 711 | 711 | NoRHC | 4.051E−01 | seps | |
| 26T | 110 | 110 | RHC | 3.636E−01 | − | |
| 27T | 601 | 601 | NoRHC | 3.627E−01 | aps1 | |
| 7T | 2760 | 2760 | NoRHC | 2.355E−01 | aps1 | |

 Number of terminal nodes of final tree: 11
 Total number of nodes of final tree: 21
 Second best split variable (based on curvature test) at root node is
aps1

 Classification tree:
 For categorical variable splits, values not in training data go to

the right

```
 Node 1: cat1 = "CHF", "MOSF w/Sepsis"
    Node 2: meanbp1 <= 68.500000 or NA
      Node 4: pafi1 <= 266.15625
         Node 8: RHC
      Node 4: pafi1 > 266.15625 or NA
         Node 9: ninsclas = "No insurance", "Private", "Private &
Medicare"
            Node 18: RHC
         Node 9: ninsclas /= "No insurance", "Private", "Private &
Medicare"
            Node 19: NoRHC
    Node 2: meanbp1 > 68.500000
      Node 5: alb1 <= 2.9499511
         Node 10: pafi1 <= 216.25000
            Node 20: RHC
         Node 10: pafi1 > 216.25000 or NA
            Node 21: NoRHC
      Node 5: alb1 > 2.9499511 or NA
         Node 11: NoRHC
 Node 1: cat1 /= "CHF", "MOSF w/Sepsis"
    Node 3: pafi1 <= 142.35938
      Node 6: resp = "No"
         Node 12: dnr1 = "No"
            Node 24: RHC
         Node 12: dnr1 /= "No"
            Node 25: NoRHC
      Node 6: resp /= "No"
         Node 13: seps = "Yes"
            Node 26: RHC
         Node 13: seps /= "Yes"
            Node 27: NoRHC
    Node 3: pafi1 > 142.35938 or NA
      Node 7: NoRHC


    ****************************************************************

  Predictor means below are means of cases with no missing values.

Node 1: Intermediate node
 A case goes into Node 2 if cat1 = "CHF", "MOSF w/Sepsis"
 cat1 mode = "ARF"
 Class       Number   Posterior
 NoRHC         3551   0.6192E+00
 RHC           2184   0.3808E+00
 Number of training cases misclassified = 2184
 Predicted class is NoRHC
 _____
Node 2: Intermediate node
```

```
GUIDE        (do not edit this file unless you know what you are doing)
  36.2       (version of GUIDE that generated this file)
 1           (1=model fitting, 2=importance or DIF scoring, 3=data
conversion)
"gfout.txt"  (name of output file)
 2           (1=one tree, 2=ensemble)
 2           (1=bagging, 2=rforest)
 2           (1=random splits of missing values, 2=nonrandom)
 1           (1=classification, 2=regression)
 2           (1=interaction tests, 2=skip them)
"ceclass.dsc"  (name of data description file)
 1           (1=accept default number of trees, 2=change)
 1           (1=accept default number of variables for splitting,
2=change it)
 1           (1=estimated priors, 2=equal priors, 3=other priors)
 1           (1=unit misclassification costs, 2=other)
 1           (1=split point from quantiles, 2=use exhaustive search)
 1           (1=accept default splitting fraction, 2=change it)
 1           (1=default max. number of split levels, 2=specify no. in
next line)
 1           (1=default min. node size, 2=specify min. value in next
line)
"gfpro.txt" (file name for predicted class and probability estimates)
 1           (rank of top variable to split root node)
```

```
          GGG   U   U  I  DDDD   EEEE
         G   G  U   U  I  D   D  E
         G      U   U  I  D   D  E
         G  GG  U   U  I  D   D  EEE
         G   G  U   U  I  D   D  E
         G   G  U   U  I  D   D  E
          GGG    UUU   I  DDDD   EEEE
```

 GUIDE Classification and Regression Trees and Forests
 Version 36.2 (Build date: January 8, 2021)
 Compiled with NAG Fortran 6.2.0 on macOS Catalina 10.15.7
 Copyright (c) 1997-2020 Wei-Yin Loh. All rights reserved.
 This software is based upon work supported by the U.S. Army Research
Office,
 the National Science Foundation and the National Institutes of
Health.

 This job was started on 02/22/21 at 16:40

 Random forest of classification trees
 No pruning
 Data description file: ceclass.dsc
 Training sample file: cedata.txt
 Missing value code: NA
 Records in data file start on line 2
 Number of M variables associated with C variables: 33
 422 N variables changed to S
 D variable is INTRDVX_
 Number of records in data file: 4693
 Length of longest entry in data file: 11
 Missing values found among categorical variables
 Separate categories will be created for missing categorical variables
 Missing values found among non-categorical variables
 Number of classes: 3
 Warning: S variable MISC2PQ is constant
 Warning: S variable MISC2CQ is constant
 Warning: S variable TCARTRKP is constant
 Warning: S variable TCARTRKC is constant
 Warning: S variable TOTHVHRP is constant
 Warning: S variable TOTHVHRC is constant
 Warning: S variable VMISCHEP is constant
 Warning: S variable VMISCHEC is constant
 Warning: S variable ROTHRFLP is constant
 Warning: S variable ROTHRFLC is constant
 Smallest and largest positive weights are 1.3507E+03 and 7.0269E+04
 Training sample class proportions of D variable INTRDVX_:
 Class #Cases   Proportion
 C   1771   0.37737055
 D   2838   0.60473045
 T   84   0.01789900
```

```
 Summary information for training sample of size 4693
 d=dependent, b=split and fit cat variable using indicator variables,
 c=split-only categorical, i=fit-only categorical (via indicators),
 s=split-only numerical, n=split and fit numerical, f=fit-only
 numerical,
 m=missing-value flag variable, p=periodic variable, w=weight
 Levels of M variables are for missing values in associated variables
                                                    #Codes/
                                                    Levels/
    Column   Name            Minimum    Maximum     Periods    #Missing
         1   DIRACC     c                                2         155
         2   DIRACC_    m                                1
         3   AGE_REF    s    18.00      87.00
         4   AGE_REF_   m                                0
         5   AGE2       s    22.00      87.00                     1903
         6   AGE2_      m                                1
         7   AS_COMP1   s    0.000      4.000
         8   AS_C_MP1   m                                0
         9   AS_COMP2   s    0.000      4.000
        10   AS_C_MP2   m                                0
        11   AS_COMP3   s    0.000      4.000
        12   AS_C_MP3   m                                0
        13   AS_COMP4   s    0.000      4.000
        14   AS_C_MP4   m                                0
        15   AS_COMP5   s    0.000      2.000
        16   AS_C_MP5   m                                0
        17   BATHRMQ    s    1.000      8.000                       21
        18   BATHRMQ_   m                                2
        19   BEDROOMQ   s    0.000      9.000                       25
        20   BEDR_OMQ   m                                2
        21   BLS_URBN   s    1.000      2.000
        22   BUILDING   c                               11
        23   BUIL_ING   m                                0
        24   CUTENURE   c                                6
        25   CUTE_URE   m                                0
        26   EARNCOMP   c                                9
        27   EARN_OMP   m                                0
        28   EDUC_REF   s    10.00      16.00
        29   EDUC0REF   m                                0
        30   EDUCA2     s    10.00      16.00                     1903
        31   EDUCA2_    m                                1
        32   FAM_SIZE   s    1.000      9.000
        33   FAM__IZE   m                                0
        34   FAM_TYPE   c                               10
        35   FAM__YPE   m                                0
        36   FAMTFEDX   s    0.000      0.9928E+05
        37   FAMT_EDX   m                                0
        38   FEDRFNDX   s    4.000      0.1428E+05                2597
        39   FEDR_NDX   m                                2
```

40	FEDTAXX	s	2.000	0.8223E+05		3793
41	FEDTAXX_	m			2	
42	FGOVRETX	s	0.000	0.2469E+05		
43	FGOV_ETX	m			0	
44	FINCATAX	s	−0.3380E+06	0.1410E+07		
45	FINCAT_X	m			0	
46	FINCBTAX	s	−0.3430E+06	0.1410E+07		
47	FINCBT_X	m			0	
48	FINDRETX	s	0.000	0.1272E+06		
49	FIND_ETX	m			0	
50	FINLWT21	w	1351.	0.7027E+05		
51	FJSSDEDX	s	0.000	0.3042E+05		
52	FJSS_EDX	m			0	
53	FPRIPENX	s	0.000	0.5902E+05		
54	FPRI_ENX	m			0	
55	FRRDEDX	s	0.000	9980.		
56	FRRDEDX_	m			0	
57	FRRETIRX	s	0.000	0.5807E+05		
58	FRRE_IRX	m			0	
59	FSALARYX	s	0.000	0.5301E+06		
60	FSAL_RYX	m			0	
61	FSLTAXX	s	0.000	0.3010E+05		
62	FSLTAXX_	m			0	
63	FSSIX	s	0.000	0.3048E+05		
64	FSSIX_	m			0	
65	HLFBATHQ	s	0.000	4.000		23
66	HLFB_THQ	m			2	
67	INC_HRS1	s	1.000	93.00		1744
68	INC__RS1	m			1	
69	INC_HRS2	s	1.000	99.00		2886
70	INC__RS2	m			1	
71	INC_RANK	s	0.1000E−03	1.000		367
72	INC__ANK	m			1	
73	INCNONW1	c			6	2949
74	INCN_NW1	m			1	
75	INCNONW2	c			6	3710
76	INCN_NW2	m			1	
77	INCOMEY1	c			6	1744
78	INCO_EY1	m			1	
79	INCOMEY2	c			6	2886
80	INCO_EY2	m			1	
81	INCWEEK1	s	0.000	52.00		
82	INCW_EK1	m			0	
83	INCWEEK2	s	0.000	52.00		1903
84	INCW_EK2	m			1	
85	MISCTAXX	s	5.000	0.2524E+05		4601
86	MISC_AXX	m			3	
87	LUMPSUMX	s	4.000	0.5492E+06		4454
88	LUMP_UMX	m			2	
89	MARITAL1	c			6	

```
 90  MARI_AL1   m                                         0
 91  NO_EARNR   s    0.000        6.000
 92  NO_E_RNR   m                                         0
 93  NONINCMX   s    0.000        0.5492E+06
 94  NONI_CMX   m                                         0
 95  NUM_AUTO   s    0.000        7.000
 96  NUM__UTO   m                                         0
 97  OCCUCOD1   c                                        15       1744
 98  OCCU_OD1   m                                         1
 99  OCCUCOD2   c                                        15       2886
100  OCCU_OD2   m                                         1
101  OTHRINCX   s    2.000        0.5788E+05                      4566
102  OTHR_NCX   m                                         2
103  PERSLT18   s    0.000        7.000
104  PERS_T18   m                                         0
105  PERSOT64   s    0.000        4.000
106  PERS_T64   m                                         0
107  POPSIZE    s    1.000        5.000                             38
108  PRINEARN   c                                         8
109  PRIN_ARN   m                                         0
110  QINTRVMO   c                                        12
111  QINTRVYR   c                                         2
112  RACE2      c                                         6       1903
113  RACE2_     m                                         1
114  REF_RACE   c                                         7
115  REF__ACE   m                                         0
116  REGION     c                                         4         38
117  RENTEQVX   s    1.000        4694.                            667
118  RENT_QVX   m                                         1
119  RESPSTAT   c                                         3
120  RESP_TAT   m                                         0
121  ROOMSQ     s    1.000        19.00                             30
122  ROOMSQ_    m                                         2
123  SEX_REF    c                                         3
124  SEX_REF_   m                                         0
125  SEX2       c                                         2       1903
126  SEX2_      m                                         1
127  SLOCTAXX   s    1.000        0.2657E+05                      4053
128  SLOC_AXX   m                                         2
129  SLRFUNDX   s    1.000        4169.                           3242
130  SLRF_NDX   m                                         2
131  SMSASTAT   c                                         2
132  ST_HOUS    c                                         3
133  ST_HOUS_   m                                         0
134  TOTTXPDX   s   -0.1845E+05   0.1467E+06
135  TOTT_PDX   m                                         0
136  VEHQ       s    0.000        17.00
137  VEHQ_      m                                         0
138  WELFAREX   s    300.0        4344.                           4680
139  WELF_REX   m                                         2
```

```
140    TOTEXPPQ    s     233.2        0.2782E+06
141    TOTEXPCQ    s    −3759.        0.9669E+05
142    FOODPQ      s     0.000        0.2358E+05
143    FOODCQ      s     0.000         7363.
144    FDHOMEPQ    s     0.000         8450.
145    FDHOMECQ    s     0.000         6067.
146    FDAWAYPQ    s     0.000        0.2098E+05
147    FDAWAYCQ    s     0.000         5660.
148    FDXMAPPQ    s     0.000        0.2098E+05
149    FDXMAPCQ    s     0.000         5660.
150    FDMAPPQ     s     0.000         900.0
151    FDMAPCQ     s     0.000         666.7
152    ALCBEVPQ    s     0.000         3152.
153    ALCBEVCQ    s     0.000         2550.
154    HOUSPQ      s     0.000        0.1811E+06
155    HOUSCQ      s    −2196.        0.3466E+05
156    SHELTPQ     s     0.000        0.4074E+05
157    SHELTCQ     s     0.000        0.3354E+05
158    OWNDWEPQ    s     0.000        0.3070E+05
159    OWNDWECQ    s     0.000        0.3321E+05
160    MRTINTPQ    s     0.000        0.2531E+05
161    MRTINTCQ    s     0.000        0.1112E+05
162    PROPTXPQ    s     0.000         5870.
163    PROPTXCQ    s     0.000         4247.
164    MRPINSPQ    s     0.000        0.2110E+05
165    MRPINSCQ    s     0.000        0.2373E+05
166    RENDWEPQ    s     0.000         8546.
167    RENDWECQ    s     0.000         6742.
168    RNTXRPPQ    s     0.000         8546.
169    RNTXRPCQ    s     0.000         6742.
170    RNTAPYPQ    s     0.000         2922.
171    RNTAPYCQ    s     0.000         3000.
172    OTHLODPQ    s     0.000        0.3342E+05
173    OTHLODCQ    s     0.000        0.1367E+05
174    UTILPQ      s     0.000         7581.
175    UTILCQ      s     0.000         3921.
176    NTLGASPQ    s     0.000         2306.
177    NTLGASCQ    s     0.000         885.0
178    ELCTRCPQ    s     0.000         4473.
179    ELCTRCCQ    s     0.000         3261.
180    ALLFULPQ    s     0.000         2752.
181    ALLFULCQ    s     0.000         3081.
182    FULOILPQ    s     0.000         2752.
183    FULOILCQ    s     0.000         3081.
184    OTHFLSPQ    s     0.000         1981.
185    OTHFLSCQ    s     0.000         2269.
186    TELEPHPQ    s     0.000         1638.
187    TELEPHCQ    s     0.000         1907.
188    WATRPSPQ    s     0.000         1880.
189    WATRPSCQ    s     0.000         1035.
```

190	HOUSOPPQ	s	−37.00	0.2493E+05
191	HOUSOPCQ	s	−4868.	0.1815E+05
192	DOMSRVPQ	s	−37.00	0.2003E+05
193	DOMSRVCQ	s	−4960.	0.1805E+05
194	DMSXCCPQ	s	−37.00	0.2003E+05
195	DMSXCCCQ	s	−4960.	0.1000E+05
196	BBYDAYPQ	s	0.000	0.1500E+05
197	BBYDAYCQ	s	0.000	0.1740E+05
198	OTHHEXPQ	s	0.000	0.2493E+05
199	OTHHEXCQ	s	0.000	5653.
200	HOUSEQPQ	s	0.000	0.1544E+06
201	HOUSEQCQ	s	0.000	0.2268E+05
202	TEXTILPQ	s	0.000	4000.
203	TEXTILCQ	s	0.000	2946.
204	FURNTRPQ	s	0.000	0.7500E+05
205	FURNTRCQ	s	0.000	0.1811E+05
206	FLRCVRPQ	s	0.000	0.1000E+05
207	FLRCVRCQ	s	0.000	5500.
208	MAJAPPPQ	s	0.000	0.1802E+05
209	MAJAPPCQ	s	0.000	0.1200E+05
210	SMLAPPPQ	s	0.000	3000.
211	SMLAPPCQ	s	0.000	944.0
212	MISCEQPQ	s	0.000	0.6510E+05
213	MISCEQCQ	s	0.000	7155.
214	APPARPQ	s	0.000	0.2440E+05
215	APPARCQ	s	0.000	4604.
216	MENBOYPQ	s	0.000	4200.
217	MENBOYCQ	s	0.000	1797.
218	MENSIXPQ	s	0.000	4200.
219	MENSIXCQ	s	0.000	1797.
220	BOYFIFPQ	s	0.000	2150.
221	BOYFIFCQ	s	0.000	448.0
222	WOMGRLPQ	s	0.000	4540.
223	WOMGRLCQ	s	0.000	2958.
224	WOMSIXPQ	s	0.000	4474.
225	WOMSIXCQ	s	0.000	2958.
226	GRLFIFPQ	s	0.000	1799.
227	GRLFIFCQ	s	0.000	1624.
228	CHLDRNPQ	s	0.000	717.0
229	CHLDRNCQ	s	0.000	961.0
230	FOOTWRPQ	s	0.000	2162.
231	FOOTWRCQ	s	0.000	1148.
232	OTHAPLPQ	s	0.000	0.2048E+05
233	OTHAPLCQ	s	0.000	4076.
234	TRANSPQ	s	0.000	0.8778E+05
235	TRANSCQ	s	0.000	0.6490E+05
236	CARTKNPQ	s	0.000	0.8700E+05
237	CARTKNCQ	s	0.000	0.6480E+05
238	CARTKUPQ	s	0.000	0.4200E+05
239	CARTKUCQ	s	0.000	0.4163E+05

240	OTHVEHPQ	s	0.000	0.1417E+05
241	OTHVEHCQ	s	0.000	0.1800E+05
242	GASMOPQ	s	0.000	4832.
243	GASMOCQ	s	0.000	6400.
244	VEHFINPQ	s	0.000	1201.
245	VEHFINCQ	s	0.000	716.0
246	MAINRPPQ	s	0.000	0.1400E+05
247	MAINRPCQ	s	0.000	8060.
248	VEHINSPQ	s	0.000	4236.
249	VEHINSCQ	s	0.000	3800.
250	VRNTLOPQ	s	0.000	0.2200E+05
251	VRNTLOCQ	s	0.000	0.2223E+05
252	PUBTRAPQ	s	0.000	0.2287E+05
253	PUBTRACQ	s	0.000	0.1198E+05
254	TRNTRPPQ	s	0.000	0.2287E+05
255	TRNTRPCQ	s	0.000	0.1198E+05
256	TRNOTHPQ	s	0.000	1448.
257	TRNOTHCQ	s	0.000	1386.
258	HEALTHPQ	s	−2402.	0.1665E+05
259	HEALTHCQ	s	−0.1281E+05	0.2189E+05
260	HLTHINPQ	s	0.000	0.1426E+05
261	HLTHINCQ	s	0.000	8789.
262	MEDSRVPQ	s	−3290.	0.1543E+05
263	MEDSRVCQ	s	−0.1330E+05	0.1368E+05
264	PREDRGPQ	s	−940.0	6844.
265	PREDRGCQ	s	−260.0	2800.
266	MEDSUPPQ	s	−3600.	7000.
267	MEDSUPCQ	s	−449.0	7530.
268	ENTERTPQ	s	0.000	0.6318E+05
269	ENTERTCQ	s	0.000	0.4249E+05
270	FEEADMPQ	s	0.000	0.1958E+05
271	FEEADMCQ	s	0.000	0.1577E+05
272	TVRDIOPQ	s	0.000	7007.
273	TVRDIOCQ	s	0.000	5143.
274	OTHEQPPQ	s	0.000	0.6300E+05
275	OTHEQPCQ	s	0.000	0.4204E+05
276	PETTOYPQ	s	0.000	0.1165E+05
277	PETTOYCQ	s	0.000	5657.
278	OTHENTPQ	s	0.000	0.6300E+05
279	OTHENTCQ	s	0.000	0.4204E+05
280	PERSCAPQ	s	0.000	1550.
281	PERSCACQ	s	0.000	973.3
282	READPQ	s	0.000	3304.
283	READCQ	s	0.000	1100.
284	EDUCAPQ	s	0.000	0.3850E+05
285	EDUCACQ	s	0.000	0.3500E+05
286	TOBACCPQ	s	0.000	2253.
287	TOBACCCQ	s	0.000	2600.
288	MISCPQ	s	0.000	0.2305E+05
289	MISCCQ	s	0.000	0.1703E+05

290	MISC1PQ	s	0.000	0.2305E+05		
291	MISC1CQ	s	0.000	0.1703E+05		
292	MISC2PQ	s	0.000	0.000		
293	MISC2CQ	s	0.000	0.000		
294	CASHCOPQ	s	0.000	0.8109E+05		
295	CASHCOCQ	s	0.000	0.2150E+05		
296	PERINSPQ	s	0.000	0.7000E+05		
297	PERINSCQ	s	0.000	0.3337E+05		
298	LIFINSPQ	s	0.000	0.7000E+05		
299	LIFINSCQ	s	0.000	0.3100E+05		
300	RETPENPQ	s	0.000	0.2584E+05		
301	RETPENCQ	s	0.000	0.2298E+05		
302	HH_CU_Q	s	1.000	5.000		
303	HH_CU_Q_	m			0	
304	HHID	c			46	4614
305	HHID_	m			1	
306	POV_CY	c			2	378
307	POV_CY_	m			1	
308	POV_PY	c			2	378
309	POV_PY_	m			1	
310	SWIMPOOL	c			1	4115
311	SWIM_OOL	m			2	
312	APTMENT	c			1	4616
313	APTMENT_	m			2	
314	OFSTPARK	c			1	1181
315	OFST_ARK	m			2	
316	WINDOWAC	c			1	4049
317	WIND_WAC	m			2	
318	CNTRALAC	c			1	1491
319	CNTR_LAC	m			2	
320	CHILDAGE	s	0.000	7.000		
321	CHIL_AGE	m			0	
322	INCLASS	s	1.000	9.000		
323	STATE	c			39	492
324	ERANKH	s	0.4735E-02	1.000		367
325	ERANKH_	m			1	
326	TOTEX4PQ	s	233.2	0.2782E+06		
327	TOTEX4CQ	s	-3759.	0.9669E+05		
328	MISCX4PQ	s	0.000	0.2305E+05		
329	MISCX4CQ	s	0.000	0.1703E+05		
330	VEHQL	s	0.000	4.000		
331	VEHQL_	m			0	
332	NUM_TVAN	s	0.000	9.000		
333	NUM__VAN	m			0	
334	TTOTALP	s	0.000	0.3821E+05		
335	TTOTALC	s	0.000	0.2303E+05		
336	TFOODTOP	s	0.000	5600.		
337	TFOODTOC	s	0.000	4305.		
338	TFOODAWP	s	0.000	5500.		
339	TFOODAWC	s	0.000	4180.		

```
340   TFOODHOP    s    0.000        3300.
341   TFOODHOC    s    0.000        1050.
342   TALCBEVP    s    0.000        2252.
343   TALCBEVC    s    0.000        1220.
344   TOTHRLOP    s    0.000        0.1013E+05
345   TOTHRLOC    s    0.000        7498.
346   TTRANPRP    s    0.000        0.2296E+05
347   TTRANPRC    s    0.000        0.1198E+05
348   TGASMOTP    s    0.000        1750.
349   TGASMOTC    s    0.000        2200.
350   TVRENTLP    s    0.000        445.0
351   TVRENTLC    s    0.000        514.0
352   TCARTRKP    s    0.000        0.000
353   TCARTRKC    s    0.000        0.000
354   TOTHVHRP    s    0.000        0.000
355   TOTHVHRC    s    0.000        0.000
356   TOTHTREP    s    0.000        445.0
357   TOTHTREC    s    0.000        514.0
358   TTRNTRIP    s    0.000        0.2287E+05
359   TTRNTRIC    s    0.000        0.1198E+05
360   TFAREP      s    0.000        0.2202E+05
361   TFAREC      s    0.000        0.1126E+05
362   TAIRFARP    s    0.000        0.2086E+05
363   TAIRFARC    s    0.000        0.1000E+05
364   TOTHFARP    s    0.000        9800.
365   TOTHFARC    s    0.000        6238.
366   TLOCALTP    s    0.000        853.0
367   TLOCALTC    s    0.000        1000.
368   TENTRMNP    s    0.000        7400.
369   TENTRMNC    s    0.000        4131.
370   TFEESADP    s    0.000        7400.
371   TFEESADC    s    0.000        4131.
372   TOTHENTP    s    0.000        1400.
373   TOTHENTC    s    0.000        2400.
374   OWNVACP     s    0.000        0.2329E+05
375   OWNVACC     s    0.000        0.1367E+05
376   VOTHRLOP    s    0.000        0.2329E+05
377   VOTHRLOC    s    0.000        0.1367E+05
378   VMISCHEP    s    0.000        0.000
379   VMISCHEC    s    0.000        0.000
380   UTILOWNP    s    0.000        2077.
381   UTILOWNC    s    0.000        1523.
382   VFUELOIP    s    0.000        682.0
383   VFUELOIC    s    0.000        625.0
384   VOTHRFLP    s    0.000        547.0
385   VOTHRFLC    s    0.000        907.0
386   VELECTRP    s    0.000        1360.
387   VELECTRC    s    0.000        988.0
388   VNATLGAP    s    0.000        2077.
389   VNATLGAC    s    0.000        201.0
```

390	VWATERPP	s	0.000	475.0		
391	VWATERPC	s	0.000	571.0		
392	MRTPRNOP	s	0.000	0.2643E+05		
393	MRTPRNOC	s	0.000	0.1322E+05		
394	UTILRNTP	s	0.000	1157.		
395	UTILRNTC	s	0.000	628.0		
396	RFUELOIP	s	0.000	565.0		
397	RFUELOIC	s	0.000	553.0		
398	ROTHRFLP	s	0.000	0.000		
399	ROTHRFLC	s	0.000	0.000		
400	RELECTRP	s	0.000	558.0		
401	RELECTRC	s	0.000	209.0		
402	RNATLGAP	s	0.000	254.0		
403	RNATLGAC	s	0.000	89.00		
404	RWATERPP	s	0.000	552.0		
405	RWATERPC	s	0.000	242.0		
406	POVLEVCY	s	0.1145E+05	0.5184E+05		
407	POVL_VCY	m			0	
408	POVLEVPY	s	0.1122E+05	0.5078E+05		
409	POVL_VPY	m			0	
410	PORCH	c			1	1006
411	PORCH_	m			2	
412	ETOTALP	s	233.2	0.2782E+06		
413	ETOTALC	s	−2683.	0.7288E+05		
414	ETOTAPX4	s	233.2	0.2782E+06		
415	ETOTACX4	s	−2683.	0.7288E+05		
416	EHOUSNGP	s	0.000	0.1811E+06		
417	EHOUSNGC	s	−2196.	0.3897E+05		
418	ESHELTRP	s	0.000	0.4456E+05		
419	ESHELTRC	s	0.000	0.3786E+05		
420	EOWNDWLP	s	0.000	0.4456E+05		
421	EOWNDWLC	s	0.000	0.3752E+05		
422	EOTHLODP	s	0.000	0.3342E+05		
423	EOTHLODC	s	0.000	0.1433E+05		
424	EMRTPNOP	s	0.000	0.3516E+05		
425	EMRTPNOC	s	0.000	0.2247E+05		
426	EMRTPNVP	s	0.000	0.2643E+05		
427	EMRTPNVC	s	0.000	0.1322E+05		
428	ETRANPTP	s	0.000	0.8868E+05		
429	ETRANPTC	s	0.000	0.5436E+05		
430	EVEHPURP	s	0.000	0.8790E+05		
431	EVEHPURC	s	0.000	0.5400E+05		
432	ECARTKNP	s	0.000	0.8790E+05		
433	ECARTKNC	s	0.000	0.5400E+05		
434	ECARTKUP	s	0.000	0.2643E+05		
435	ECARTKUC	s	0.000	0.2662E+05		
436	EOTHVEHP	s	0.000	0.1166E+05		
437	EOTHVEHC	s	0.000	6542.		
438	EENTRMTP	s	0.000	0.6318E+05		
439	EENTRMTC	s	0.000	0.1605E+05		

440	EOTHENTP	s	0.000	0.6300E+05		
441	EOTHENTC	s	0.000	7502.		
442	ENOMOTRP	s	0.000	7700.		
443	ENOMOTRC	s	0.000	1500.		
444	EMOTRVHP	s	0.000	0.6300E+05		
445	EMOTRVHC	s	0.000	6971.		
446	EENTMSCP	s	0.000	6000.		
447	EENTMSCC	s	0.000	5000.		
448	EMISCELP	s	0.000	0.2305E+05		
449	EMISCELC	s	0.000	0.1703E+05		
450	EMISCMTP	s	0.000	1096.		
451	EMISCMTC	s	0.000	2113.		
452	UNISTRQ	s	1.000	10.00		
453	UNISTRQ_	m			0	
455	WELF_EBX	c			2	
456	LUMPSUMB	s	2.000	12.00		4684
457	LUMP_UMB	m			2	
458	LMPSUMBX	s	1200.	0.8000E+05		4684
459	LMPS_MBX	m			2	
460	OTHRINCB	s	5.000	12.00		4687
461	OTHR_NCB	m			2	
462	OTRINCBX	s				4693
463	OTRI_CBX	m			1	
464	INCLASS2	s	1.000	7.000		
465	INCL_SS2	m			0	
467	HORREF1	c			6	4532
468	HORREF1_	m			1	
469	HORREF2	c			5	4577
470	HORREF2_	m			1	
471	ERANKHM	s	0.6205E-02	1.000		
472	ERANKHM_	m			0	
473	FGOVRETM	s	0.000	0.2509E+05		
474	FGOV_ETM	m			0	
475	FPRIPENM	s	0.000	0.5826E+05		
476	FPRI_ENM	m			0	
477	FRRDEDM	s	0.000	0.1043E+05		
478	FRRDEDM_	m			0	
479	PSU	c			21	2616
480	HISP_REF	c			2	
481	HISP2	c			2	1903
482	HIGH_EDU	s	10.00	16.00		
483	BUILT	s	1915.	2013.		592
484	BUILT_	m			2	
485	CREDFINX	s	0.000	6629.		4360
486	CRED_INX	m			2	
487	CREDITB	s	1.000	5.000		4667
488	CREDITB_	m			2	
489	CREDITBX	s	250.0	0.2250E+05		4667
490	CRED_TBX	m			2	
491	CREDITX	s	1.000	0.5132E+05		4311

492	CREDITX_	m			2	
493	CREDTYRX	s	0.000	0.5092E+05		4327
494	CRED_YRX	m			2	
495	CREDYRB	s	1.000	6.000		4656
496	CREDYRB_	m			2	
497	CREDYRBX	s	250.0	0.3500E+05		4656
498	CRED_RBX	m			2	
499	DEFBENRP	c			2	3551
500	DEFB_NRP	m			2	
501	EITC	c			2	1042
502	EITC_	m			2	
503	FMLPYYRX	s	4.000	4000.		4598
504	FMLP_YRX	m			2	
505	FS_MTHI	s	1.000	12.00		4644
506	FS_MTHI_	m			1	
507	FSMPFRMX	s	−0.4000E+06	0.1090E+07		
508	FSMP_RMX	m			0	
514	INTRDVX_	d			3	
522	IRAB	s	1.000	6.000		4514
523	IRAB_	m			2	
524	IRABX	s	1000.	0.7250E+06		4514
525	IRABX_	m			2	
526	IRAX	s	0.000	0.2635E+07		3917
527	IRAX_	m			2	
528	IRAYRB	s	1.000	6.000		4489
529	IRAYRB_	m			2	
530	IRAYRBX	s	1000.	0.7250E+06		4489
531	IRAYRBX_	m			2	
532	IRAYRX	s	0.000	0.2129E+07		3964
533	IRAYRX_	m			2	
534	JFS_AMT	s	0.000	4800.		
535	JFS_AMT_	m			0	
536	LIQDYRBX	s	250.0	0.3500E+05		4531
537	LIQD_RBX	m			2	
538	LIQUIDBX	s	250.0	0.3500E+05		4563
539	LIQU_DBX	m			2	
540	LIQUDYRB	s	1.000	6.000		4531
541	LIQU_YRB	m			2	
542	LIQUDYRX	s	0.000	0.5155E+06		3942
543	LIQU_YRX	m			2	
544	LIQUIDB	s	1.000	6.000		4563
545	LIQUIDB_	m			2	
546	LIQUIDX	s	0.000	0.4910E+06		3892
547	LIQUIDX_	m			2	
548	MEALSPAY	s	1.000	2.000		9
549	MEAL_PAY	m			2	
550	MLPAYWKX	s	2.000	300.0		4598
551	MLPA_WKX	m			2	
552	MLPYQWKS	s	1.000	52.00		4592
553	MLPY_WKS	m			2	

554	NETRENTB	s	0.000	12.00		4666
555	NETR_NTB	m			3	
556	NETRENTX	s	-0.5499E+05	0.1148E+06		4338
557	NETR_NTX	m			2	
558	NETRNTBX	s	-2400.	0.7130E+05		4666
559	NETR_TBX	m			2	
560	OTHASTBX	s	0.3000E+05	0.7250E+06		4673
561	OTHA_TBX	m			2	
562	OTHASTB	s	3.000	6.000		4673
563	OTHASTB_	m			3	
564	OTHASTX	s	2.000	0.2767E+07		4648
565	OTHASTX_	m			2	
566	OTHFINX	s	0.000	900.0		4654
567	OTHFINX_	m			2	
568	OTHLONBX	s	250.0	0.2250E+05		4690
569	OTHL_NBX	m			2	
570	OTHLYRBX	s	750.0	0.2250E+05		4689
571	OTHL_RBX	m			2	
572	OTHLNYRB	s	2.000	5.000		4689
573	OTHL_YRB	m			2	
574	OTHLNYRX	s	0.000	0.5500E+05		4642
575	OTHL_YRX	m			2	
576	OTHLOAN	s	1.000	2.000		3485
577	OTHLOAN_	m			1	
578	OTHLONB	s	1.000	5.000		4690
579	OTHLONB_	m			2	
580	OTHLONX	s	1.000	0.3106E+06		4638
581	OTHLONX_	m			2	
582	OTHREGBX	s	488.0	0.5000E+05		4678
583	OTHR_GBX	m			2	
584	OTHREGB	s	1.000	12.00		4678
585	OTHREGB_	m			2	
586	OTHREGX	s	36.00	0.6367E+05		4422
587	OTHREGX_	m			2	
588	OTHSYRBX	s	6000.	0.7250E+06		4669
589	OTHS_RBX	m			2	
590	OTHSTYRB	s	2.000	6.000		4669
591	OTHS_YRB	m			2	
592	OTHSTYRX	s	0.000	0.1533E+07		4656
593	OTHS_YRX	m			2	
594	RETSURVX	s	30.00	0.1269E+06		3571
595	RETS_RVX	m			2	
596	RETSRVBX	s	480.0	0.6200E+05		4626
597	RETS_VBX	m			2	
598	RETSURV	c			3	
599	RETSURV_	m			0	
600	RETSURVB	s	1.000	12.00		4626
601	RETS_RVB	m			2	
604	ROYESTBX	s	200.0	0.6000E+05		4654
605	ROYE_TBX	m			2	

606	ROYESTB	s	1.000	12.00		4654
607	ROYESTB_	m			2	
608	ROYESTX	s	1.000	0.1592E+06		4441
609	ROYESTX_	m			2	
610	STCKYRBX	s	1000.	0.7250E+06		4613
611	STCK_RBX	m			2	
612	STDNTYRB	s	3.000	6.000		4675
613	STDN_YRB	m			2	
614	STDNTYRX	s	0.000	0.4100E+06		4566
615	STDN_YRX	m			2	
616	STDTYRBX	s	1750.	0.3500E+05		4675
617	STDT_RBX	m			2	
618	STOCKYRB	s	1.000	6.000		4613
619	STOC_YRB	m			2	
620	STOCKYRX	s	0.000	0.5784E+07		4420
621	STOC_YRX	m			2	
622	STOCKB	s	1.000	6.000		4632
623	STOCKB_	m			3	
624	STOCKBX	s	1000.	0.7250E+06		4632
625	STOCKBX_	m			2	
626	STOCKX	s	25.00	0.6587E+07		4390
627	STOCKX_	m			2	
628	STUDFINX	s	0.000	9000.		4595
629	STUD_INX	m			2	
630	STUDNTBX	s	6250.	0.3500E+05		4682
631	STUD_TBX	m			2	
632	STUDNTB	s	4.000	6.000		4682
633	STUDNTB_	m			2	
634	STUDNTX	s	250.0	0.4200E+06		4556
635	STUDNTX_	m			2	
636	WHLFYRBX	s	250.0	0.3500E+05		4648
637	WHLF_RBX	m			2	
638	WHLFYRB	s	1.000	6.000		4648
639	WHLFYRB_	m			2	
640	WHLFYRX	s	0.000	0.7674E+06		4524
641	WHLFYRX_	m			3	
642	WHOLIFBX	s	250.0	0.3500E+05		4654
643	WHOL_FBX	m			2	
644	WHOLIFB	s	1.000	6.000		4654
645	WHOLIFB_	m			3	
646	WHOLIFX	s	1.000	0.7892E+06		4508
647	WHOLIFX_	m			3	
648	TOTXEST	s	−8990.	0.2938E+06		
649	FFTAXOWE	s	−8943.	0.2485E+06		
650	FFTA_OWE	m			0	
651	FSTAXOWE	s	−2505.	0.5991E+05		
652	FSTA_OWE	m			0	
653	ETOTA	s	1199.	0.2782E+06		

Total #cases w/   #missing

```
        #cases    miss. D  ord. vals   #X-var    #N-var    #F-var    #S-var
         4693         0       4693        16         0         0       422
        #P-var    #M-var    #B-var    #C-var    #I-var
            0       171         0        42         0
Number of cases used for training: 4693
Number of split variables: 464
Number of cases excluded due to 0 weight or missing D: 0

Number of trees in ensemble: 500
Number of variables used for splitting: 155
Warning: No linear splits; number of S variables must be < 225
Simple node models
Estimated priors
Unit misclassification costs
Warning: All positive weights treated as 1
Univariate split highest priority
No interaction splits
No linear splits
Fraction of cases used for splitting each node: .0213
Maximum number of split levels: 19
Minimum node sample size: 23
Mean number of terminal nodes:    139.9

Classification matrix for training sample:
Predicted       True class
class             C         D         T
C              1286        68         7
D               485      2770        77
T                 0         0         0
Total          1771      2838        84

Number of cases used for tree construction: 4693
Number misclassified: 637
Resubstitution estimate of mean misclassification cost: .1357

Number of OOB cases: 4693
Number OOB misclassified: 1041
OOB estimate of mean misclassification cost: .2218
Mean number of trees per OOB observation: 183.88

Predicted class probabilities are stored in gfpro.txt
Elapsed time in seconds: 1629.1
```

```
 A case goes into Node 4 if meanbp1 <= 68.500000 or NA
 meanbp1 mean = 72.674985
 Class       Number   Posterior
 NoRHC          774  0.4599E+00
 RHC            909  0.5401E+00
 Number of training cases misclassified = 774
 Predicted class is RHC
 _____
Node 4: Intermediate node
 A case goes into Node 8 if pafi1 <= 266.15625
 pafi1 mean = 241.37331
 Class       Number   Posterior
 NoRHC          424  0.3796E+00
 RHC            693  0.6204E+00
 Number of training cases misclassified = 424
 Predicted class is RHC
 _____
Node 8: Terminal node
 Class       Number   Posterior
 NoRHC          199  0.3038E+00
 RHC            456  0.6962E+00
 Number of training cases misclassified = 199
 Predicted class is RHC
 _____
Node 9: Intermediate node
 A case goes into Node 18 if ninsclas = "No insurance", "Private",
 "Private & Medicare"
 ninsclas mode = "Private"
 Class       Number   Posterior
 NoRHC          225  0.4870E+00
 RHC            237  0.5130E+00
 Number of training cases misclassified = 225
 Predicted class is RHC
 _____
Node 18: Terminal node
 Class       Number   Posterior
 NoRHC           91  0.3730E+00
 RHC            153  0.6270E+00
 Number of training cases misclassified = 91
 Predicted class is RHC
 _____
Node 19: Terminal node
 Class       Number   Posterior
 NoRHC          134  0.6147E+00
 RHC             84  0.3853E+00
 Number of training cases misclassified = 84
 Predicted class is NoRHC
 _____
Node 5: Intermediate node
 A case goes into Node 10 if alb1 <= 2.9499511
```

```
 alb1 mean = 3.1795715
 Class      Number  Posterior
 NoRHC         350  0.6184E+00
 RHC           216  0.3816E+00
 Number of training cases misclassified = 216
 Predicted class is NoRHC
 ──────────────────────────
Node 10: Intermediate node
 A case goes into Node 20 if pafi1 <= 216.25000
 pafi1 mean = 239.66683
 Class      Number  Posterior
 NoRHC          76  0.4810E+00
 RHC            82  0.5190E+00
 Number of training cases misclassified = 76
 Predicted class is RHC
 ──────────────────────────
Node 20: Terminal node
 Class      Number  Posterior
 NoRHC          23  0.3194E+00
 RHC            49  0.6806E+00
 Number of training cases misclassified = 23
 Predicted class is RHC
 ──────────────────────────
Node 21: Terminal node
 Class      Number  Posterior
 NoRHC          53  0.6163E+00
 RHC            33  0.3837E+00
 Number of training cases misclassified = 33
 Predicted class is NoRHC
 ──────────────────────────
Node 11: Terminal node
 Class      Number  Posterior
 NoRHC         274  0.6716E+00
 RHC           134  0.3284E+00
 Number of training cases misclassified = 134
 Predicted class is NoRHC
 ──────────────────────────
Node 3: Intermediate node
 A case goes into Node 6 if pafi1 <= 142.35938
 pafi1 mean = 211.08630
 Class      Number  Posterior
 NoRHC        2777  0.6853E+00
 RHC          1275  0.3147E+00
 Number of training cases misclassified = 1275
 Predicted class is NoRHC
 ──────────────────────────
Node 6: Intermediate node
 A case goes into Node 12 if resp = "No"
 resp mode = "Yes"
 Class      Number  Posterior
```

```
  NoRHC           667  0.5163E+00
  RHC             625  0.4837E+00
 Number of training cases misclassified = 625
 Predicted class is NoRHC
  _____
Node 12: Intermediate node
 A case goes into Node 24 if dnr1 = "No"
 dnr1 mode = "No"
 Class        Number   Posterior
 NoRHC           244  0.4200E+00
 RHC             337  0.5800E+00
 Number of training cases misclassified = 244
 Predicted class is RHC
  _____
Node 24: Terminal node
 Class        Number   Posterior
 NoRHC           201  0.3903E+00
 RHC             314  0.6097E+00
 Number of training cases misclassified = 201
 Predicted class is RHC
  _____
Node 25: Terminal node
 Class        Number   Posterior
 NoRHC            43  0.6515E+00
 RHC              23  0.3485E+00
 Number of training cases misclassified = 23
 Predicted class is NoRHC
  _____
Node 13: Intermediate node
 A case goes into Node 26 if seps = "Yes"
 seps mode = "No"
 Class        Number   Posterior
 NoRHC           423  0.5949E+00
 RHC             288  0.4051E+00
 Number of training cases misclassified = 288
 Predicted class is NoRHC
  _____
Node 26: Terminal node
 Class        Number   Posterior
 NoRHC            40  0.3636E+00
 RHC              70  0.6364E+00
 Number of training cases misclassified = 40
 Predicted class is RHC
  _____
Node 27: Terminal node
 Class        Number   Posterior
 NoRHC           383  0.6373E+00
 RHC             218  0.3627E+00
 Number of training cases misclassified = 218
 Predicted class is NoRHC
```

```
                _____
Node 7: Terminal node
 Class        Number    Posterior
 NoRHC         2110   0.7645E+00
 RHC            650   0.2355E+00
 Number of training cases misclassified = 650
 Predicted class is NoRHC
                _____
 *****************************************************
 Variables used for splitting:
 alb1
 cat1
 dnr1
 meanbp1
 ninsclas
 pafi1
 resp
 seps

 Number of terminal nodes: 11

 Scaled importance scores of predictor variables
        Score      Rank  Variable
  2.287E+01    1.00  cat1
  2.234E+01    2.00  aps1
  2.012E+01    3.00  crea1
  1.924E+01    4.00  pafi1
  1.855E+01    5.00  meanbp1
  1.219E+01    6.00  alb1
  1.196E+01    7.00  neuro
  1.066E+01    8.00  card
  1.021E+01    9.00  hema1
  1.010E+01   10.00  cat2
  9.315E+00   11.00  wtkilo1
  7.929E+00   12.00  seps
  6.409E+00   13.00  resp
  6.073E+00   14.00  dnr1
  6.029E+00   15.00  bili1
  5.410E+00   16.00  paco21
  4.260E+00   17.00  hrt1
  3.979E+00   18.00  transhx
  3.860E+00   19.00  chrpulhx
  3.860E+00   20.00  resp1
  3.311E+00   21.00  ninsclas
  3.183E+00   22.00  dementhx
  2.969E+00   23.00  ph1
  2.163E+00   24.00  psychhx
  2.029E+00   25.00  renal
  1.913E+00   26.00  gastr
  1.896E+00   27.00  income
```

```
 1.748E+00    28.00  cardiohx
 1.412E+00    29.00  urin1
 1.386E+00    30.00  trauma
 1.247E+00    31.00  age
 1.194E+00    32.00  sex
--------- variables above this line are highly important -----
 1.188E+00    33.00  edu
 1.176E+00    34.00  sod1
 1.056E+00    35.00  immunhx
--------- variables below this line are unimportant ----------
 9.441E-01    36.00  malighx
 9.116E-01    37.00  wblc1
 8.732E-01    38.00  ca
 8.616E-01    39.00  amihx
 8.071E-01    40.00  scoma1
 6.766E-01    41.00  chfhx
 5.859E-01    42.00  gibledhx
 4.112E-01    43.00  renalhx
 4.095E-01    44.00  pot1
 3.971E-01    45.00  ortho
 3.431E-01    46.00  liverhx
 3.412E-01    47.00  hema
 3.280E-01    48.00  meta
 2.586E-01    49.00  temp1
 1.296E-01    50.00  race


 Variables with scores above 1.19 are highly important
 Variables with scores between 1.0 and 1.19 are likely important
 Variables with scores below 1.0 are unimportant
 No. highly important, likely important, and unimportant split
variables: 32, 3, 15
 LaTeX code for tree is in imp.tex
 Importance scores are stored in imp.scr
 Elapsed time in seconds: 219.89
```