

Model

Shiwen Xu

November 2, 2017

Decision Tree Model One (with data_unregistration)

```
model2=read.csv('model2.csv')

library(rpart)
library(rpart.plot)

## Warning: package 'rpart.plot' was built under R version 3.4.2
library(dplyr)

##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

model2=model2%>%
  mutate(drop=ifelse(final_result=='Withdrawn',1,0))
model2$date_unregistration=as.numeric(levels(model2$date_unregistration))[model2$date_unregistration]

## Warning: NAs introduced by coercion
```

Baseline Model

Baseline Accuracy - predict all students do not drop

```
mean(model2$drop==0)

## [1] 0.7547557
```

Split Training and Testing Dataset

```
trainrows=sample(29228,21921)
train=model2[trainrows,]
test=model2[-trainrows,]
```

Baseline Model

```
split=rpart(drop~date_unregistration+date_unregistration+
            sum_click+highest_education+num_of_prev_attempts+
            studied_credits, data=train,method='class',
            control=rpart.control(cp=0))
```

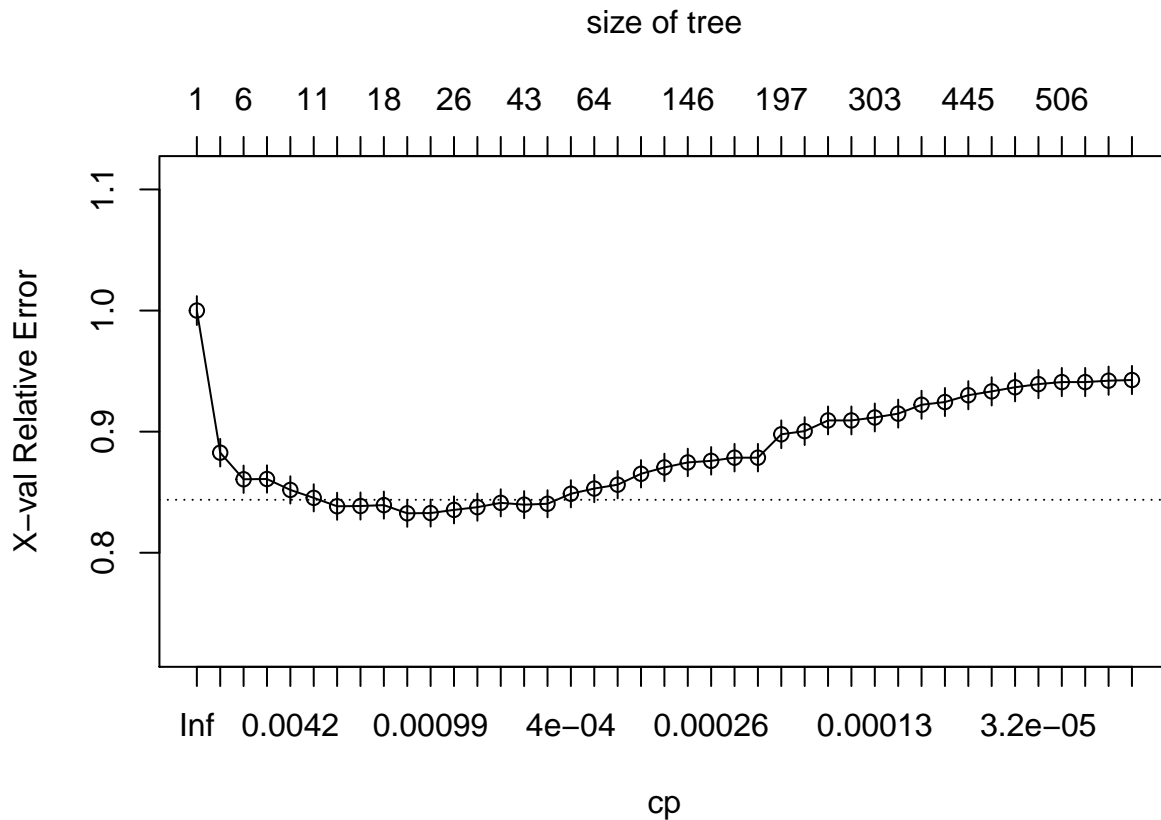
```
predict3=predict(split,test,type='class')
table(predict3,test$drop)
```

```
##
## predict3    0    1
##           0 4827 1119
##           1  646  715
mean(predict3==test$drop)
```

```
## [1] 0.7584508
```

Set the maximum split to 6

```
plotcp(split)
```

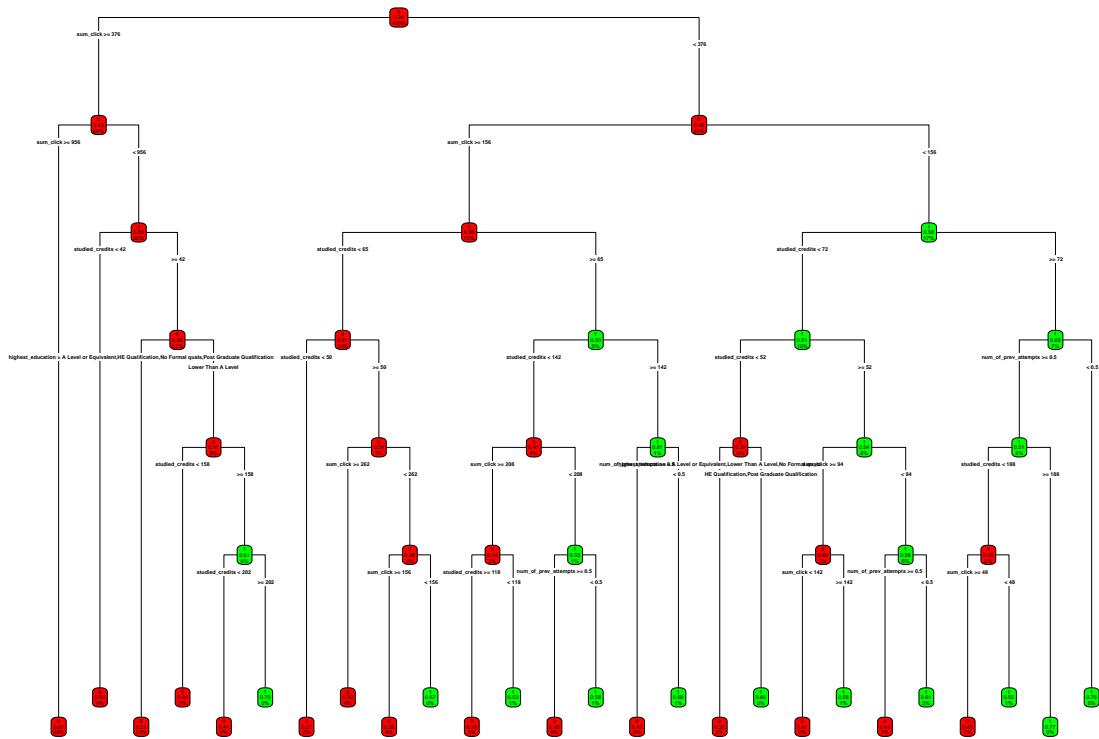


```
tree=rpart(drop~date_unregistration+date_unregistration+
            sum_click+highest_education+num_of_prev_attempts+
```

```
studied_credits, data=train,method='class',
control=rpart.control(maxdepth=6,cp=0))
```

```
predict=predict(tree,test,type='class')
rpart.plot(tree,type=4,box.palette = c('red','green'),fallen.leaves = TRUE)
```

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



```
summary(predict)
```

```
##      0      1
## 6190 1117
```

Confusion Matrix in Test

```
table(predict,test$drop)
```

```
##
## predict      0      1
##           0 5037 1153
##           1  436  681
```

```
mean(predict==test$drop)
```

```
## [1] 0.7825373
```

Confusion Matrix in Training

```
predict2=predict(tree,train,type='class')
table(predict2,train$drop)
```

```
##
## predict2      0      1
##           0 15510  3253
##           1  1077  2081
```

```
mean(predict2==train$drop)
```

```
## [1] 0.8024725
```

Decision Tree Model Two (Without date__unregistration)

Baseline Accuracy

```
mean(model2$drop==0)
```

```
## [1] 0.7547557
```

Baseline Model (exclude data__unregistration)

```
split2=rpart(drop~date_registration+
              sum_click+highest_education+num_of_prev_attempts+
              studied_credits,data=train,method='class',
              control=rpart.control(cp=0))
```

```
predict4=predict(split,test,type='class')
table(predict4,test$drop)
```

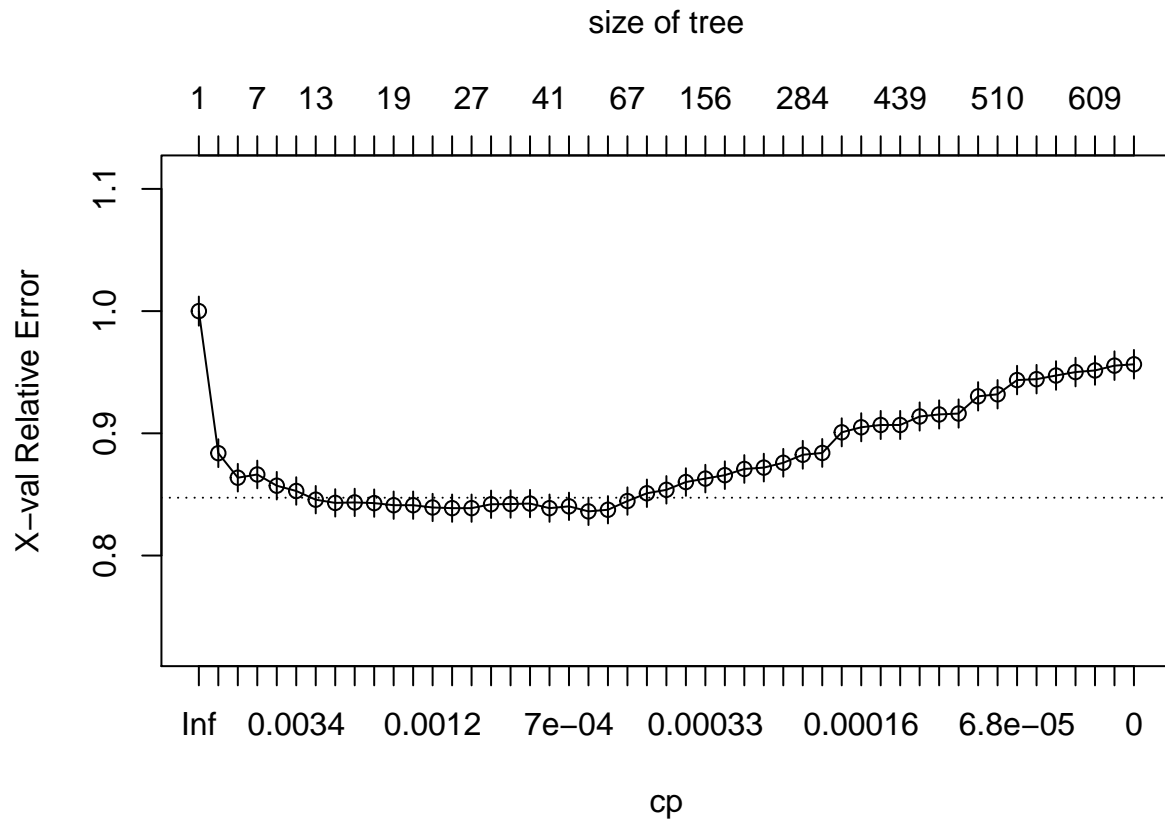
```
##
## predict4      0      1
##           0 4827 1119
##           1  646  715
```

```
mean(predict4==test$drop)
```

```
## [1] 0.7584508
```

Set Maximum Split at 7

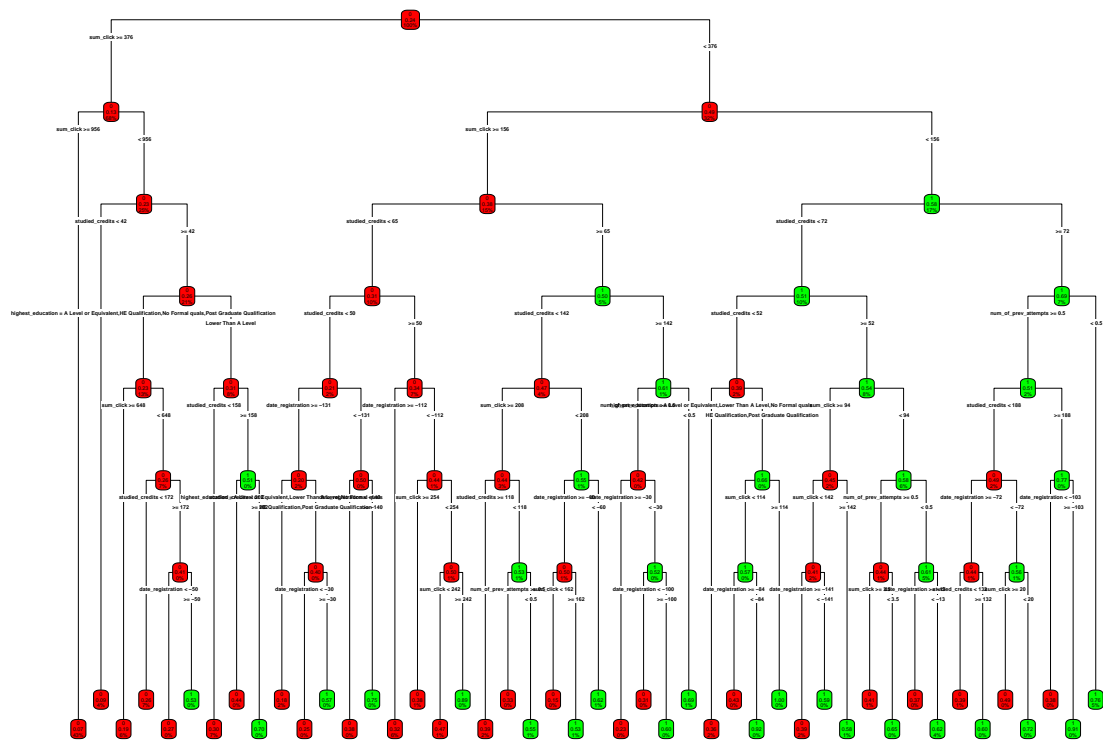
```
plotcp(split2)
```



```
tree2=rpart(drop~date_registration+
             sum_click+highest_education+num_of_prev_attempts+
             studied_credits, data=train,method='class',
             control=rpart.control(maxdepth=7,cp=0))

rpart.plot(tree2,type=4,box.palette = c('red','green'),fallen.leaves = TRUE)

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



Confusion Matrix in Test

```
predict5=predict(tree2,test,type='class')
table(predict5,test$drop)
```

```
##
## predict5    0    1
##           0 5048 1154
##           1  425  680
```

```
mean(predict5==test$drop)
```

```
## [1] 0.7839058
```

Confusion Matrix in Training

```
predict6=predict(tree2,train,type='class')
table(predict6,train$drop)
```

```
##
## predict6     0     1
##           0 15530 3205
##           1  1057 2129
```

```
mean(predict6==train$drop)
```

```
## [1] 0.8055746
```

Importance of Each Predictor

```
library(caret)
```

```
## Warning: package 'caret' was built under R version 3.4.2
```

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

```
## Loading required package: ggplot2
```

```
varImp(tree2)
```

```
##               Overall
## date_registration 207.5714
## highest_education 142.6575
## num_of_prev_attempts 158.4555
## studied_credits 525.7845
## sum_click 1754.1112
```

Multiple Regression (all)

```
regression = lm(drop~highest_education+code_module+code_presentation+num_of_prev_attempts+sum_click+studied_credits)
```

```
summary(regression)
```

```
##
## Call:
## lm(formula = drop ~ highest_education + code_module + code_presentation +
##     num_of_prev_attempts + sum_click + studied_credits, data = model2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.7762 -0.2761 -0.1529  0.3585  1.8561
##
## Coefficients:
##              Estimate Std. Error t value
## (Intercept)    1.966e-01  1.742e-02  11.288
## highest_educationHE Qualification -1.911e-04  6.939e-03  -0.028
## highest_educationLower Than A Level  6.876e-02  5.092e-03  13.502
## highest_educationNo Formal quals    1.288e-01  2.314e-02   5.565
## highest_educationPost Graduate Qualification -1.943e-02  2.293e-02  -0.847
## code_moduleBBB -5.737e-02  1.552e-02  -3.697
## code_moduleCCC  1.872e-01  1.610e-02  11.624
## code_moduleDDD  8.108e-02  1.557e-02   5.206
## code_moduleEEE  2.960e-02  1.645e-02   1.800
## code_moduleFFF  1.456e-01  1.544e-02   9.430
## code_moduleGGG -1.370e-01  1.717e-02  -7.977
## code_presentation2013J -1.984e-02  7.661e-03  -2.590
## code_presentation2014B -1.133e-02  8.060e-03  -1.406
```

```
## code_presentation2014J      1.551e-02  7.543e-03   2.056
## num_of_prev_attempts      -1.673e-02  5.006e-03  -3.341
## sum_click                  -8.846e-05  1.459e-06 -60.627
## studied_credits            1.157e-03  6.372e-05  18.159
##                             Pr(>|t|)
## (Intercept)                < 2e-16 ***
## highest_educationHE Qualification 0.978032
## highest_educationLower Than A Level < 2e-16 ***
## highest_educationNo Formal quals  2.65e-08 ***
## highest_educationPost Graduate Qualification 0.396813
## code_moduleBBB             0.000219 ***
## code_moduleCCC             < 2e-16 ***
## code_moduleDDD             1.94e-07 ***
## code_moduleEEE             0.071947 .
## code_moduleFFF             < 2e-16 ***
## code_moduleGGG             1.56e-15 ***
## code_presentation2013J      0.009604 **
## code_presentation2014B      0.159859
## code_presentation2014J      0.039809 *
## num_of_prev_attempts        0.000835 ***
## sum_click                   < 2e-16 ***
## studied_credits             < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3924 on 29211 degrees of freedom
## Multiple R-squared:  0.1686, Adjusted R-squared:  0.1682
## F-statistic: 370.3 on 16 and 29211 DF, p-value: < 2.2e-16
```

Multiple Regression

```
regression2 = lm(drop~num_of_prev_attempts+sum_click+studied_credits, data=model2)
summary(regression2)
```

```
##
## Call:
## lm(formula = drop ~ num_of_prev_attempts + sum_click + studied_credits,
##     data = model2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.8097 -0.2820 -0.2044  0.3463  1.6189
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.173e-01  5.578e-03  38.955 <2e-16 ***
## num_of_prev_attempts -1.540e-02  5.113e-03  -3.012  0.0026 **
## sum_click        -7.432e-05  1.371e-06 -54.219 <2e-16 ***
## studied_credits    1.681e-03  6.108e-05  27.522 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```



```
## Residual standard error: 0.4053 on 29224 degrees of freedom
## Multiple R-squared: 0.1127, Adjusted R-squared: 0.1126
## F-statistic: 1237 on 3 and 29224 DF, p-value: < 2.2e-16
```

```
confint(regression2, level=0.95)
```

```
##                2.5 %          97.5 %
## (Intercept)      2.063728e-01  2.282404e-01
## num_of_prev_attempts -2.542211e-02 -5.378531e-03
## sum_click         -7.701114e-05 -7.163741e-05
## studied_credits    1.561219e-03  1.800645e-03
```

Summary of the Decision Tree Model (exclude date_unregistration)

```
summary(tree2)
```

```
## Call:
## rpart(formula = drop ~ date_registration + sum_click + highest_education +
##       num_of_prev_attempts + studied_credits, data = train, method = "class",
##       control = rpart.control(maxdepth = 7, cp = 0))
## n= 21921
##
##           CP nsplit rel error   xerror   xstd
## 1  0.0590551181    0 1.0000000 1.0000000 0.01191042
## 2  0.0092800900    2 0.8818898 0.8835771 0.01140330
## 3  0.0056242970    5 0.8526434 0.8588301 0.01128552
## 4  0.0047806524    6 0.8470191 0.8575178 0.01127917
## 5  0.0037495313    9 0.8320210 0.8427072 0.01120677
## 6  0.0029996250   10 0.8282715 0.8352081 0.01116959
## 7  0.0022497188   12 0.8222722 0.8325834 0.01115650
## 8  0.0017497813   13 0.8200225 0.8350206 0.01116866
## 9  0.0016872891   16 0.8147732 0.8348331 0.01116773
## 10 0.0014998125   17 0.8130859 0.8310836 0.01114900
## 11 0.0013123360   18 0.8115861 0.8305212 0.01114618
## 12 0.0011248594   20 0.8089614 0.8322085 0.01115462
## 13 0.0008436445   21 0.8078365 0.8308961 0.01114806
## 14 0.0006561680   23 0.8061492 0.8299588 0.01114336
## 15 0.0003749531   25 0.8048369 0.8338958 0.01116305
## 16 0.0003437070   29 0.8033371 0.8353956 0.01117053
## 17 0.0002999625   35 0.8012748 0.8357705 0.01117239
## 18 0.0001874766   40 0.7997750 0.8378328 0.01118265
## 19 0.0001249844   42 0.7994001 0.8398950 0.01119287
## 20 0.0000000000   45 0.7990251 0.8421447 0.01120400
##
## Variable importance
##           sum_click      studied_credits num_of_prev_attempts
##              82              11              3
## date_registration highest_education
##              2              1
##
## Node number 1: 21921 observations, complexity param=0.05905512
## predicted class=0 expected loss=0.2433283 P(node) =1
## class counts: 16587 5334
## probabilities: 0.757 0.243
```

```

## left son=2 (14847 obs) right son=3 (7074 obs)
## Primary splits:
##   sum_click           < 376.5 to the right, improve=1259.02400, (0 missing)
##   studied_credits     < 82.5 to the left, improve= 127.58450, (0 missing)
##   highest_education   splits as LLRRL, improve= 36.02538, (0 missing)
##   date_registration   < -66.5 to the right, improve= 13.39262, (6 missing)
##   num_of_prev_attempts < 0.5 to the left, improve= 7.54207, (0 missing)
## Surrogate splits:
##   num_of_prev_attempts < 2.5 to the left, agree=0.678, adj=0.001, (0 split)
##   studied_credits     < 222.5 to the left, agree=0.677, adj=0.001, (0 split)
##   date_registration   < 46 to the left, agree=0.677, adj=0.000, (0 split)
##
## Node number 2: 14847 observations, complexity param=0.0002999625
## predicted class=0 expected loss=0.1263555 P(node) =0.6772957
## class counts: 12971 1876
## probabilities: 0.874 0.126
## left son=4 (9368 obs) right son=5 (5479 obs)
## Primary splits:
##   sum_click           < 955.5 to the right, improve=174.811200, (0 missing)
##   studied_credits     < 55 to the left, improve= 17.462840, (0 missing)
##   highest_education   splits as LLRRL, improve= 12.359080, (0 missing)
##   date_registration   < -146.5 to the right, improve= 4.860574, (2 missing)
##   num_of_prev_attempts < 0.5 to the left, improve= 3.765380, (0 missing)
## Surrogate splits:
##   studied_credits     < 35 to the right, agree=0.642, adj=0.030, (0 split)
##   date_registration   < -6 to the left, agree=0.633, adj=0.005, (0 split)
##   num_of_prev_attempts < 2.5 to the left, agree=0.631, adj=0.000, (0 split)
##
## Node number 3: 7074 observations, complexity param=0.05905512
## predicted class=0 expected loss=0.4888323 P(node) =0.3227043
## class counts: 3616 3458
## probabilities: 0.511 0.489
## left son=6 (3306 obs) right son=7 (3768 obs)
## Primary splits:
##   sum_click           < 155.5 to the right, improve=144.81440, (0 missing)
##   studied_credits     < 65 to the left, improve=129.37690, (0 missing)
##   date_registration   < -11.5 to the right, improve= 26.41009, (4 missing)
##   num_of_prev_attempts < 0.5 to the right, improve= 11.77175, (0 missing)
##   highest_education   splits as LLLRR, improve= 4.72818, (0 missing)
## Surrogate splits:
##   studied_credits     < 57.5 to the left, agree=0.535, adj=0.006, (0 split)
##   date_registration   < -185.5 to the left, agree=0.535, adj=0.006, (0 split)
##   num_of_prev_attempts < 5.5 to the right, agree=0.533, adj=0.001, (0 split)
##
## Node number 4: 9368 observations
## predicted class=0 expected loss=0.0676772 P(node) =0.4273528
## class counts: 8734 634
## probabilities: 0.932 0.068
##
## Node number 5: 5479 observations, complexity param=0.0002999625
## predicted class=0 expected loss=0.2266837 P(node) =0.249943
## class counts: 4237 1242
## probabilities: 0.773 0.227
## left son=10 (947 obs) right son=11 (4532 obs)

```

```

## Primary splits:
##   studied_credits    < 42.5   to the left,  improve=45.620060, (0 missing)
##   sum_click          < 659.5  to the right, improve=11.844930, (0 missing)
##   date_registration  < -146.5 to the right, improve= 9.779427, (2 missing)
##   highest_education  splits as LRRLR,      improve= 8.055828, (0 missing)
##   num_of_prev_attempts < 0.5   to the left,  improve= 1.002024, (0 missing)
##
## Node number 6: 3306 observations,    complexity param=0.004780652
##   predicted class=0 expected loss=0.3808227 P(node) =0.1508143
##   class counts:  2047  1259
##   probabilities: 0.619 0.381
##   left son=12 (2120 obs) right son=13 (1186 obs)
##   Primary splits:
##     studied_credits    < 65     to the left,  improve=53.2834100, (0 missing)
##     date_registration  < -8.5    to the right, improve=12.5672200, (2 missing)
##     sum_click          < 224.5   to the right, improve= 8.9991020, (0 missing)
##     highest_education  splits as LLLRL,      improve= 2.9557560, (0 missing)
##     num_of_prev_attempts < 3.5    to the left,  improve= 0.5097074, (0 missing)
##   Surrogate splits:
##     num_of_prev_attempts < 0.5    to the left,  agree=0.652, adj=0.029, (0 split)
##     date_registration  < -225.5  to the right, agree=0.644, adj=0.008, (0 split)
##
## Node number 7: 3768 observations,    complexity param=0.00928009
##   predicted class=1 expected loss=0.4164013 P(node) =0.17189
##   class counts:  1569  2199
##   probabilities: 0.416 0.584
##   left son=14 (2217 obs) right son=15 (1551 obs)
##   Primary splits:
##     studied_credits    < 72.5    to the left,  improve=62.474920, (0 missing)
##     num_of_prev_attempts < 0.5    to the right, improve=25.652860, (0 missing)
##     sum_click          < 17.5    to the right, improve=18.934020, (0 missing)
##     date_registration  < -12.5   to the right, improve=14.779610, (2 missing)
##     highest_education  splits as LRLLR,      improve= 5.745303, (0 missing)
##   Surrogate splits:
##     num_of_prev_attempts < 0.5    to the left,  agree=0.600, adj=0.027, (0 split)
##     date_registration  < -152.5  to the right, agree=0.593, adj=0.012, (0 split)
##
## Node number 10: 947 observations
##   predicted class=0 expected loss=0.08553326 P(node) =0.04320058
##   class counts:   866    81
##   probabilities: 0.914 0.086
##
## Node number 11: 4532 observations,    complexity param=0.0002999625
##   predicted class=0 expected loss=0.2561783 P(node) =0.2067424
##   class counts:  3371  1161
##   probabilities: 0.744 0.256
##   left son=22 (2869 obs) right son=23 (1663 obs)
##   Primary splits:
##     highest_education  splits as LLRLL,      improve=13.3968400, (0 missing)
##     sum_click          < 659.5   to the right, improve=13.0207800, (0 missing)
##     studied_credits    < 162.5   to the left,  improve= 9.6558540, (0 missing)
##     date_registration  < -146.5  to the right, improve= 5.3933530, (2 missing)
##     num_of_prev_attempts < 2.5   to the right, improve= 0.9114866, (0 missing)
##

```

```

## Node number 12: 2120 observations,    complexity param=0.000343707
## predicted class=0 expected loss=0.3136792 P(node) =0.09671092
## class counts: 1455 665
## probabilities: 0.686 0.314
## left son=24 (478 obs) right son=25 (1642 obs)
## Primary splits:
## studied_credits < 50 to the left, improve=12.93806, (0 missing)
## date_registration < -112.5 to the right, improve=10.44876, (1 missing)
## sum_click < 235.5 to the right, improve= 6.24567, (0 missing)
## highest_education splits as LRLRL, improve= 2.46301, (0 missing)
## num_of_prev_attempts < 1.5 to the left, improve= 0.28894, (0 missing)
## Surrogate splits:
## date_registration < -13.5 to the right, agree=0.776, adj=0.006, (0 split)
##
## Node number 13: 1186 observations,    complexity param=0.004780652
## predicted class=1 expected loss=0.4991568 P(node) =0.05410337
## class counts: 592 594
## probabilities: 0.499 0.501
## left son=26 (947 obs) right son=27 (239 obs)
## Primary splits:
## studied_credits < 142.5 to the left, improve=6.707424, (0 missing)
## num_of_prev_attempts < 0.5 to the right, improve=6.606907, (0 missing)
## date_registration < -8.5 to the right, improve=5.330634, (1 missing)
## highest_education splits as RLRRR, improve=3.835162, (0 missing)
## sum_click < 206.5 to the right, improve=3.481709, (0 missing)
## Surrogate splits:
## date_registration < 14 to the left, agree=0.8, adj=0.008, (0 split)
##
## Node number 14: 2217 observations,    complexity param=0.00928009
## predicted class=1 expected loss=0.4925575 P(node) =0.1011359
## class counts: 1092 1125
## probabilities: 0.493 0.507
## left son=28 (459 obs) right son=29 (1758 obs)
## Primary splits:
## studied_credits < 52.5 to the left, improve=15.386500, (0 missing)
## date_registration < -12.5 to the right, improve=10.762890, (0 missing)
## sum_click < 88.5 to the right, improve= 8.120225, (0 missing)
## num_of_prev_attempts < 0.5 to the right, improve= 6.094446, (0 missing)
## highest_education splits as LRLRR, improve= 2.655245, (0 missing)
## Surrogate splits:
## date_registration < -13.5 to the right, agree=0.806, adj=0.063, (0 split)
##
## Node number 15: 1551 observations,    complexity param=0.001749781
## predicted class=1 expected loss=0.3075435 P(node) =0.07075407
## class counts: 477 1074
## probabilities: 0.308 0.692
## left son=30 (397 obs) right son=31 (1154 obs)
## Primary splits:
## num_of_prev_attempts < 0.5 to the right, improve=36.982450, (0 missing)
## sum_click < 18.5 to the right, improve=10.729240, (0 missing)
## studied_credits < 187.5 to the left, improve= 7.314275, (0 missing)
## date_registration < 3.5 to the right, improve= 2.583983, (2 missing)
## highest_education splits as LRLLR, improve= 1.973756, (0 missing)
## Surrogate splits:

```

```

##      date_registration < 3.5    to the right, agree=0.750, adj=0.025, (0 split)
##      studied_credits  < 440    to the right, agree=0.745, adj=0.005, (0 split)
##
## Node number 22: 2869 observations,    complexity param=0.0001249844
## predicted class=0 expected loss=0.2269083 P(node) =0.1308791
##   class counts:  2218   651
##   probabilities: 0.773 0.227
## left son=44 (1318 obs) right son=45 (1551 obs)
## Primary splits:
##   sum_click          < 647.5 to the right, improve=8.204812, (0 missing)
##   studied_credits    < 87.5  to the left,  improve=3.816419, (0 missing)
##   date_registration  < -7.5  to the right, improve=2.924525, (1 missing)
##   highest_education  splits as LR-RL,      improve=1.441269, (0 missing)
##   num_of_prev_attempts < 2.5  to the right, improve=0.869814, (0 missing)
## Surrogate splits:
##   highest_education  splits as RL-RL,      agree=0.548, adj=0.015, (0 split)
##   date_registration  < 4.5  to the right, agree=0.541, adj=0.002, (0 split)
##   num_of_prev_attempts < 3.5  to the right, agree=0.541, adj=0.002, (0 split)
##
## Node number 23: 1663 observations,    complexity param=0.0002999625
## predicted class=0 expected loss=0.3066747 P(node) =0.07586333
##   class counts:  1153   510
##   probabilities: 0.693 0.307
## left son=46 (1589 obs) right son=47 (74 obs)
## Primary splits:
##   studied_credits    < 157.5 to the left,  improve=6.6266530, (0 missing)
##   sum_click          < 754.5 to the right, improve=6.4766310, (0 missing)
##   date_registration  < -146.5 to the right, improve=5.1241740, (1 missing)
##   num_of_prev_attempts < 0.5  to the right, improve=0.3709247, (0 missing)
##
## Node number 24: 478 observations,    complexity param=0.000343707
## predicted class=0 expected loss=0.2112971 P(node) =0.02180557
##   class counts:   377   101
##   probabilities: 0.789 0.211
## left son=48 (454 obs) right son=49 (24 obs)
## Primary splits:
##   date_registration  < -131  to the right, improve=4.2122650, (0 missing)
##   highest_education  splits as LRLRL,      improve=3.8686010, (0 missing)
##   sum_click          < 233   to the right, improve=1.0964730, (0 missing)
##   num_of_prev_attempts < 0.5  to the left,  improve=0.2418145, (0 missing)
##
## Node number 25: 1642 observations,    complexity param=0.000343707
## predicted class=0 expected loss=0.3434836 P(node) =0.07490534
##   class counts:  1078   564
##   probabilities: 0.657 0.343
## left son=50 (1351 obs) right son=51 (291 obs)
## Primary splits:
##   date_registration  < -112.5 to the right, improve=6.163933, (1 missing)
##   sum_click          < 261.5 to the right, improve=5.406893, (0 missing)
##   highest_education  splits as LLLRL,      improve=1.091257, (0 missing)
##   num_of_prev_attempts < 0.5  to the right, improve=0.287833, (0 missing)
##
## Node number 26: 947 observations,    complexity param=0.004780652
## predicted class=0 expected loss=0.4741288 P(node) =0.04320058

```

```

##      class counts:   498   449
##      probabilities: 0.526 0.474
##      left son=52 (678 obs) right son=53 (269 obs)
##      Primary splits:
##          sum_click          < 208.5   to the right, improve=4.782235, (0 missing)
##          date_registration   < -8.5    to the right, improve=4.233548, (1 missing)
##          studied_credits     < 117.5   to the right, improve=3.301139, (0 missing)
##          num_of_prev_attempts < 0.5     to the right, improve=2.999244, (0 missing)
##          highest_education   splits as LLRRR,      improve=2.788688, (0 missing)
##      Surrogate splits:
##          date_registration < -4.5    to the left,  agree=0.718, adj=0.007, (0 split)
##
## Node number 27: 239 observations,      complexity param=0.002249719
##      predicted class=1 expected loss=0.3933054 P(node) =0.01090279
##      class counts:    94   145
##      probabilities: 0.393 0.607
##      left son=54 (74 obs) right son=55 (165 obs)
##      Primary splits:
##          num_of_prev_attempts < 0.5     to the right, improve=7.5588230, (0 missing)
##          date_registration   < -141.5 to the right, improve=1.8943140, (0 missing)
##          highest_education   splits as RLRRR,      improve=1.4609370, (0 missing)
##          sum_click          < 304      to the right, improve=0.6839058, (0 missing)
##          studied_credits     < 155      to the right, improve=0.5433109, (0 missing)
##      Surrogate splits:
##          highest_education splits as RRRRL,      agree=0.699, adj=0.027, (0 split)
##          date_registration < -4.5    to the right, agree=0.695, adj=0.014, (0 split)
##
## Node number 28: 459 observations,      complexity param=0.002999625
##      predicted class=0 expected loss=0.3921569 P(node) =0.02093883
##      class counts:    279   180
##      probabilities: 0.608 0.392
##      left son=56 (409 obs) right son=57 (50 obs)
##      Primary splits:
##          highest_education   splits as LLLLL,      improve=8.051011, (0 missing)
##          date_registration   < -58.5   to the right, improve=4.074858, (0 missing)
##          sum_click          < 13.5     to the right, improve=2.026850, (0 missing)
##          num_of_prev_attempts < 0.5     to the left,  improve=1.075770, (0 missing)
##
## Node number 29: 1758 observations,      complexity param=0.00928009
##      predicted class=1 expected loss=0.4624573 P(node) =0.08019707
##      class counts:    813   945
##      probabilities: 0.462 0.538
##      left son=58 (535 obs) right son=59 (1223 obs)
##      Primary splits:
##          sum_click          < 93.5     to the right, improve=12.6846600, (0 missing)
##          num_of_prev_attempts < 0.5     to the right, improve=12.2203400, (0 missing)
##          date_registration   < -142.5 to the right, improve= 5.4591410, (0 missing)
##          highest_education   splits as LRRRR,      improve= 0.8431058, (0 missing)
##          studied_credits     < 65      to the right, improve= 0.1568138, (0 missing)
##      Surrogate splits:
##          studied_credits < 65          to the right, agree=0.697, adj=0.006, (0 split)
##
## Node number 30: 397 observations,      complexity param=0.001749781
##      predicted class=1 expected loss=0.4937028 P(node) =0.01811049

```

```

##      class counts:   196   201
##      probabilities: 0.494 0.506
##      left son=60 (367 obs) right son=61 (30 obs)
##      Primary splits:
##          studied_credits    < 187.5  to the left,  improve=4.400031, (0 missing)
##          sum_click          < 129.5  to the right, improve=2.587694, (0 missing)
##          date_registration  < -72.5  to the right, improve=2.153684, (1 missing)
##          highest_education  splits as RRLLR,      improve=1.714515, (0 missing)
##          num_of_prev_attempts < 2.5   to the right, improve=0.566612, (0 missing)
##
## Node number 31: 1154 observations
##      predicted class=1 expected loss=0.2435009 P(node) =0.05264358
##      class counts:   281   873
##      probabilities: 0.244 0.756
##
## Node number 44: 1318 observations
##      predicted class=0 expected loss=0.1858877 P(node) =0.06012499
##      class counts:  1073   245
##      probabilities: 0.814 0.186
##
## Node number 45: 1551 observations,      complexity param=0.0001249844
##      predicted class=0 expected loss=0.2617666 P(node) =0.07075407
##      class counts:   1145   406
##      probabilities: 0.738 0.262
##      left son=90 (1485 obs) right son=91 (66 obs)
##      Primary splits:
##          studied_credits    < 172.5  to the left,  improve=2.9923200, (0 missing)
##          date_registration  < -8.5    to the right, improve=2.3468390, (0 missing)
##          highest_education  splits as LR-RR,      improve=2.2954340, (0 missing)
##          sum_click          < 491.5  to the right, improve=1.3281800, (0 missing)
##          num_of_prev_attempts < 2.5   to the right, improve=0.3008333, (0 missing)
##
## Node number 46: 1589 observations
##      predicted class=0 expected loss=0.2970422 P(node) =0.07248757
##      class counts:   1117   472
##      probabilities: 0.703 0.297
##
## Node number 47: 74 observations,      complexity param=0.0002999625
##      predicted class=1 expected loss=0.4864865 P(node) =0.003375758
##      class counts:    36    38
##      probabilities: 0.486 0.514
##      left son=94 (54 obs) right son=95 (20 obs)
##      Primary splits:
##          studied_credits    < 202.5  to the left,  improve=1.906306, (0 missing)
##          date_registration  < -147.5 to the right, improve=1.825851, (0 missing)
##          num_of_prev_attempts < 0.5   to the right, improve=1.311482, (0 missing)
##          sum_click          < 846    to the left,  improve=1.003276, (0 missing)
##      Surrogate splits:
##          sum_click < 393.5  to the right, agree=0.757, adj=0.1, (0 split)
##
## Node number 48: 454 observations,      complexity param=0.0001874766
##      predicted class=0 expected loss=0.1960352 P(node) =0.02071073
##      class counts:   365    89
##      probabilities: 0.804 0.196

```

```

## left son=96 (424 obs) right son=97 (30 obs)
## Primary splits:
## highest_education splits as LRLLR, improve=2.67270800, (0 missing)
## date_registration < -115.5 to the left, improve=0.86644920, (0 missing)
## sum_click < 165 to the left, improve=0.86644920, (0 missing)
## num_of_prev_attempts < 0.5 to the left, improve=0.09659445, (0 missing)
##
## Node number 49: 24 observations, complexity param=0.000343707
## predicted class=0 expected loss=0.5 P(node) =0.001094841
## class counts: 12 12
## probabilities: 0.500 0.500
## left son=98 (16 obs) right son=99 (8 obs)
## Primary splits:
## date_registration < -139.5 to the left, improve=1.5, (0 missing)
## sum_click < 223 to the right, improve=0.8, (0 missing)
## highest_education splits as RRL--, improve=0.8, (0 missing)
## Surrogate splits:
## sum_click < 258 to the left, agree=0.875, adj=0.625, (0 split)
## highest_education splits as RLL--, agree=0.708, adj=0.125, (0 split)
##
## Node number 50: 1351 observations
## predicted class=0 expected loss=0.3234641 P(node) =0.0616304
## class counts: 914 437
## probabilities: 0.677 0.323
##
## Node number 51: 291 observations, complexity param=0.000343707
## predicted class=0 expected loss=0.4364261 P(node) =0.01327494
## class counts: 164 127
## probabilities: 0.564 0.436
## left son=102 (156 obs) right son=103 (135 obs)
## Primary splits:
## sum_click < 253.5 to the right, improve=1.80531600, (0 missing)
## highest_education splits as LRLRL, improve=1.78991200, (0 missing)
## date_registration < -113.5 to the left, improve=1.25349100, (0 missing)
## num_of_prev_attempts < 1.5 to the right, improve=0.06207728, (0 missing)
## Surrogate splits:
## date_registration < -128.5 to the left, agree=0.553, adj=0.037, (0 split)
## highest_education splits as LLLRR, agree=0.546, adj=0.022, (0 split)
## num_of_prev_attempts < 1.5 to the left, agree=0.543, adj=0.015, (0 split)
##
## Node number 52: 678 observations, complexity param=0.002999625
## predicted class=0 expected loss=0.4424779 P(node) =0.03092925
## class counts: 378 300
## probabilities: 0.558 0.442
## left son=104 (446 obs) right son=105 (232 obs)
## Primary splits:
## studied_credits < 117.5 to the right, improve=5.970828, (0 missing)
## highest_education splits as RLRRR, improve=4.278179, (0 missing)
## date_registration < -29.5 to the right, improve=3.123169, (0 missing)
## num_of_prev_attempts < 0.5 to the right, improve=1.831048, (0 missing)
## sum_click < 210.5 to the left, improve=1.496989, (0 missing)
## Surrogate splits:
## date_registration < -195 to the right, agree=0.665, adj=0.022, (0 split)
##

```



```

## Node number 53: 269 observations,      complexity param=0.0008436445
##   predicted class=1   expected loss=0.4460967   P(node) =0.01227134
##   class counts:      120   149
##   probabilities: 0.446 0.554
##   left son=106 (147 obs) right son=107 (122 obs)
##   Primary splits:
##       date_registration < -59.5   to the right, improve=2.2395370, (1 missing)
##       num_of_prev_attempts < 0.5   to the right, improve=1.6674250, (0 missing)
##       sum_click < 158.5   to the left, improve=1.4657480, (0 missing)
##       highest_education splits as LLRL-, improve=0.9540483, (0 missing)
##       studied_credits < 110   to the left, improve=0.1230389, (0 missing)
##   Surrogate splits:
##       sum_click < 164.5   to the right, agree=0.567, adj=0.049, (1 split)
##       highest_education splits as LLRL-, agree=0.567, adj=0.049, (0 split)
##
## Node number 54: 74 observations,      complexity param=0.000656168
##   predicted class=0   expected loss=0.4189189   P(node) =0.003375758
##   class counts:      43   31
##   probabilities: 0.581 0.419
##   left son=108 (26 obs) right son=109 (48 obs)
##   Primary splits:
##       date_registration < -30   to the right, improve=2.83792400, (0 missing)
##       sum_click < 347   to the right, improve=1.17841300, (0 missing)
##       num_of_prev_attempts < 1.5   to the right, improve=1.16495800, (0 missing)
##       highest_education splits as RLR-R, improve=1.10827700, (0 missing)
##       studied_credits < 172.5   to the left, improve=0.08452927, (0 missing)
##   Surrogate splits:
##       sum_click < 340   to the right, agree=0.662, adj=0.038, (0 split)
##
## Node number 55: 165 observations
##   predicted class=1   expected loss=0.3090909   P(node) =0.007527029
##   class counts:      51   114
##   probabilities: 0.309 0.691
##
## Node number 56: 409 observations
##   predicted class=0   expected loss=0.3594132   P(node) =0.01865791
##   class counts:      262   147
##   probabilities: 0.641 0.359
##
## Node number 57: 50 observations,      complexity param=0.0003749531
##   predicted class=1   expected loss=0.34   P(node) =0.002280918
##   class counts:      17   33
##   probabilities: 0.340 0.660
##   left son=114 (40 obs) right son=115 (10 obs)
##   Primary splits:
##       sum_click < 113.5   to the left, improve=2.890000, (0 missing)
##       date_registration < -60   to the right, improve=2.546952, (0 missing)
##       num_of_prev_attempts < 0.5   to the right, improve=0.090000, (0 missing)
##
## Node number 58: 535 observations,      complexity param=0.003749531
##   predicted class=0   expected loss=0.446729   P(node) =0.02440582
##   class counts:      296   239
##   probabilities: 0.553 0.447
##   left son=116 (409 obs) right son=117 (126 obs)

```

```

## Primary splits:
##   sum_click < 141.5 to the left, improve=5.7990190, (0 missing)
##   date_registration < -143.5 to the right, improve=3.5726420, (0 missing)
##   num_of_prev_attempts < 1.5 to the right, improve=1.8789360, (0 missing)
##   highest_education splits as RLRRR, improve=0.2833401, (0 missing)
##
## Node number 59: 1223 observations, complexity param=0.005624297
## predicted class=1 expected loss=0.422731 P(node) =0.05579125
## class counts: 517 706
## probabilities: 0.423 0.577
## left son=118 (232 obs) right son=119 (991 obs)
## Primary splits:
##   num_of_prev_attempts < 0.5 to the right, improve=11.534100, (0 missing)
##   sum_click < 3.5 to the right, improve= 4.098398, (0 missing)
##   date_registration < -13 to the right, improve= 3.894699, (0 missing)
##   highest_education splits as LRRRL, improve= 0.752320, (0 missing)
## Surrogate splits:
##   date_registration < 5.5 to the right, agree=0.812, adj=0.009, (0 split)
##
## Node number 60: 367 observations, complexity param=0.001749781
## predicted class=0 expected loss=0.4850136 P(node) =0.01674194
## class counts: 189 178
## probabilities: 0.515 0.485
## left son=120 (218 obs) right son=121 (149 obs)
## Primary splits:
##   date_registration < -72.5 to the right, improve=2.7108650, (1 missing)
##   sum_click < 47.5 to the right, improve=2.4485670, (0 missing)
##   highest_education splits as RRLRL, improve=2.3446100, (0 missing)
##   studied_credits < 132.5 to the left, improve=0.8764774, (0 missing)
##   num_of_prev_attempts < 2.5 to the right, improve=0.6215930, (0 missing)
## Surrogate splits:
##   num_of_prev_attempts < 2.5 to the left, agree=0.604, adj=0.027, (1 split)
##   studied_credits < 95 to the right, agree=0.604, adj=0.027, (0 split)
##   highest_education splits as LLLRL, agree=0.596, adj=0.007, (0 split)
##
## Node number 61: 30 observations, complexity param=0.0003749531
## predicted class=1 expected loss=0.2333333 P(node) =0.001368551
## class counts: 7 23
## probabilities: 0.233 0.767
## left son=122 (8 obs) right son=123 (22 obs)
## Primary splits:
##   date_registration < -103 to the left, improve=3.3469700, (0 missing)
##   sum_click < 16 to the right, improve=0.9942029, (0 missing)
##   studied_credits < 222.5 to the right, improve=0.9000000, (0 missing)
##   num_of_prev_attempts < 1.5 to the right, improve=0.6960663, (0 missing)
##   highest_education splits as LRRR-, improve=0.0539075, (0 missing)
## Surrogate splits:
##   sum_click < 119.5 to the right, agree=0.8, adj=0.25, (0 split)
##
## Node number 90: 1485 observations
## predicted class=0 expected loss=0.2552189 P(node) =0.06774326
## class counts: 1106 379
## probabilities: 0.745 0.255
##

```

```

## Node number 91: 66 observations,      complexity param=0.0001249844
##   predicted class=0   expected loss=0.4090909   P(node) =0.003010812
##   class counts:      39      27
##   probabilities: 0.591 0.409
##   left son=182 (30 obs) right son=183 (36 obs)
##   Primary splits:
##       date_registration < -50.5   to the left,   improve=2.231313, (0 missing)
##       highest_education   splits as LR-LL,       improve=1.363636, (0 missing)
##       num_of_prev_attempts < 0.5   to the right, improve=1.136415, (0 missing)
##       studied_credits < 215   to the left,   improve=0.733279, (0 missing)
##       sum_click < 436   to the right, improve=0.599287, (0 missing)
##   Surrogate splits:
##       num_of_prev_attempts < 0.5   to the right, agree=0.621, adj=0.167, (0 split)
##       studied_credits < 215   to the right, agree=0.621, adj=0.167, (0 split)
##       sum_click < 443.5   to the left,   agree=0.606, adj=0.133, (0 split)
##       highest_education   splits as RR-LL,       agree=0.576, adj=0.067, (0 split)
##
## Node number 94: 54 observations
##   predicted class=0   expected loss=0.4444444   P(node) =0.002463391
##   class counts:      30      24
##   probabilities: 0.556 0.444
##
## Node number 95: 20 observations
##   predicted class=1   expected loss=0.3   P(node) =0.0009123671
##   class counts:      6      14
##   probabilities: 0.300 0.700
##
## Node number 96: 424 observations
##   predicted class=0   expected loss=0.1816038   P(node) =0.01934218
##   class counts:      347      77
##   probabilities: 0.818 0.182
##
## Node number 97: 30 observations,      complexity param=0.0001874766
##   predicted class=0   expected loss=0.4   P(node) =0.001368551
##   class counts:      18      12
##   probabilities: 0.600 0.400
##   left son=194 (16 obs) right son=195 (14 obs)
##   Primary splits:
##       date_registration < -29.5   to the left,   improve=1.5428570, (0 missing)
##       sum_click < 247.5   to the left,   improve=0.5626794, (0 missing)
##   Surrogate splits:
##       sum_click < 258.5   to the right, agree=0.733, adj=0.429, (0 split)
##
## Node number 98: 16 observations
##   predicted class=0   expected loss=0.375   P(node) =0.0007298937
##   class counts:      10      6
##   probabilities: 0.625 0.375
##
## Node number 99: 8 observations
##   predicted class=1   expected loss=0.25   P(node) =0.0003649469
##   class counts:      2      6
##   probabilities: 0.250 0.750
##
## Node number 102: 156 observations

```

```

## predicted class=0 expected loss=0.3846154 P(node) =0.007116464
## class counts: 96 60
## probabilities: 0.615 0.385
##
## Node number 103: 135 observations, complexity param=0.000343707
## predicted class=0 expected loss=0.4962963 P(node) =0.006158478
## class counts: 68 67
## probabilities: 0.504 0.496
## left son=206 (126 obs) right son=207 (9 obs)
## Primary splits:
## sum_click < 242.5 to the left, improve=2.97248700, (0 missing)
## highest_education splits as LRLRL, improve=1.32717900, (0 missing)
## date_registration < -161.5 to the right, improve=0.91337460, (0 missing)
## num_of_prev_attempts < 0.5 to the right, improve=0.09163017, (0 missing)
##
## Node number 104: 446 observations
## predicted class=0 expected loss=0.3946188 P(node) =0.02034579
## class counts: 270 176
## probabilities: 0.605 0.395
##
## Node number 105: 232 observations, complexity param=0.001312336
## predicted class=1 expected loss=0.4655172 P(node) =0.01058346
## class counts: 108 124
## probabilities: 0.466 0.534
## left son=210 (21 obs) right son=211 (211 obs)
## Primary splits:
## num_of_prev_attempts < 0.5 to the right, improve=1.8684970, (0 missing)
## date_registration < -70.5 to the right, improve=1.8341780, (0 missing)
## sum_click < 265.5 to the right, improve=1.1324860, (0 missing)
## highest_education splits as RLLR-, improve=0.9924508, (0 missing)
## studied_credits < 72.5 to the right, improve=0.7697044, (0 missing)
##
## Node number 106: 147 observations, complexity param=0.0008436445
## predicted class=0 expected loss=0.4965986 P(node) =0.006705898
## class counts: 74 73
## probabilities: 0.503 0.497
## left son=212 (13 obs) right son=213 (134 obs)
## Primary splits:
## sum_click < 162.5 to the left, improve=3.3507890, (0 missing)
## num_of_prev_attempts < 0.5 to the right, improve=3.0537410, (0 missing)
## highest_education splits as LLRL-, improve=1.5542910, (0 missing)
## date_registration < -6.5 to the right, improve=1.0007940, (1 missing)
## studied_credits < 110 to the left, improve=0.1204806, (0 missing)
##
## Node number 107: 122 observations
## predicted class=1 expected loss=0.3770492 P(node) =0.00556544
## class counts: 46 76
## probabilities: 0.377 0.623
##
## Node number 108: 26 observations
## predicted class=0 expected loss=0.2307692 P(node) =0.001186077
## class counts: 20 6
## probabilities: 0.769 0.231
##

```

```

## Node number 109: 48 observations,    complexity param=0.000656168
##   predicted class=1   expected loss=0.4791667   P(node) =0.002189681
##   class counts:      23    25
##   probabilities: 0.479 0.521
##   left son=218 (13 obs) right son=219 (35 obs)
##   Primary splits:
##       date_registration < -100.5 to the left,   improve=1.6198720, (0 missing)
##       sum_click         < 329.5 to the right, improve=1.4083330, (0 missing)
##       studied_credits   < 232.5 to the right, improve=0.7788462, (0 missing)
##       num_of_prev_attempts < 1.5 to the right, improve=0.3688596, (0 missing)
##       highest_education splits as RLL-R,        improve=0.3569347, (0 missing)
##   Surrogate splits:
##       sum_click         < 362.5 to the right, agree=0.771, adj=0.154, (0 split)
##       highest_education splits as RLR-R,        agree=0.750, adj=0.077, (0 split)
##
## Node number 114: 40 observations,    complexity param=0.0003749531
##   predicted class=1   expected loss=0.425   P(node) =0.001824734
##   class counts:      17    23
##   probabilities: 0.425 0.575
##   left son=228 (28 obs) right son=229 (12 obs)
##   Primary splits:
##       date_registration < -84 to the right, improve=4.00238100, (0 missing)
##       sum_click         < 34 to the right, improve=1.01291600, (0 missing)
##       num_of_prev_attempts < 0.5 to the left, improve=0.01666667, (0 missing)
##
## Node number 115: 10 observations
##   predicted class=1   expected loss=0   P(node) =0.0004561836
##   class counts:      0    10
##   probabilities: 0.000 1.000
##
## Node number 116: 409 observations,    complexity param=0.001124859
##   predicted class=0   expected loss=0.405868   P(node) =0.01865791
##   class counts:      243   166
##   probabilities: 0.594 0.406
##   left son=232 (375 obs) right son=233 (34 obs)
##   Primary splits:
##       date_registration < -141 to the right, improve=2.466579, (0 missing)
##       sum_click         < 113.5 to the left, improve=2.175368, (0 missing)
##       num_of_prev_attempts < 0.5 to the right, improve=1.893132, (0 missing)
##       highest_education splits as RLRRR,        improve=0.232596, (0 missing)
##
## Node number 117: 126 observations
##   predicted class=1   expected loss=0.4206349   P(node) =0.005747913
##   class counts:      53    73
##   probabilities: 0.421 0.579
##
## Node number 118: 232 observations,    complexity param=0.001312336
##   predicted class=0   expected loss=0.4353448   P(node) =0.01058346
##   class counts:      131   101
##   probabilities: 0.565 0.435
##   left son=236 (209 obs) right son=237 (23 obs)
##   Primary splits:
##       sum_click         < 3.5 to the right, improve=2.4006820, (0 missing)
##       date_registration < -17.5 to the right, improve=1.2271350, (0 missing)

```

```

##      num_of_prev_attempts < 2.5      to the left,  improve=0.5557152, (0 missing)
##      highest_education      splits as  LRLRL,      improve=0.3614055, (0 missing)
##
## Node number 119: 991 observations,      complexity param=0.001499813
## predicted class=1 expected loss=0.3895055 P(node) =0.04520779
## class counts:      386      605
## probabilities: 0.390 0.610
## left son=238 (30 obs) right son=239 (961 obs)
## Primary splits:
##      date_registration < -13      to the right, improve=3.6784760, (0 missing)
##      sum_click          < 13.5    to the right, improve=2.7417510, (0 missing)
##      highest_education splits as  LRRRR,      improve=0.5239318, (0 missing)
##
## Node number 120: 218 observations,      complexity param=0.001687289
## predicted class=0 expected loss=0.4357798 P(node) =0.009944802
## class counts:      123      95
## probabilities: 0.564 0.436
## left son=240 (171 obs) right son=241 (47 obs)
## Primary splits:
##      studied_credits      < 132.5 to the left,  improve=3.0664610, (0 missing)
##      highest_education      splits as  RLLRR,      improve=2.9174600, (0 missing)
##      sum_click          < 11.5    to the left,  improve=2.3578450, (0 missing)
##      date_registration      < -4.5    to the right, improve=1.1331020, (1 missing)
##      num_of_prev_attempts < 2.5      to the right, improve=0.3864502, (0 missing)
##
## Node number 121: 149 observations,      complexity param=0.0003749531
## predicted class=1 expected loss=0.442953 P(node) =0.006797135
## class counts:      66      83
## probabilities: 0.443 0.557
## left son=242 (106 obs) right son=243 (43 obs)
## Primary splits:
##      sum_click          < 19.5    to the right, improve=3.2467440, (0 missing)
##      date_registration      < -156.5 to the left, improve=1.3063570, (0 missing)
##      studied_credits      < 165      to the right, improve=0.8131551, (0 missing)
##      num_of_prev_attempts < 2.5      to the right, improve=0.5100955, (0 missing)
##      highest_education      splits as  LRLLR,      improve=0.2419422, (0 missing)
## Surrogate splits:
##      highest_education splits as  LLLLR, agree=0.718, adj=0.023, (0 split)
##
## Node number 122: 8 observations
## predicted class=0 expected loss=0.375 P(node) =0.0003649469
## class counts:      5      3
## probabilities: 0.625 0.375
##
## Node number 123: 22 observations
## predicted class=1 expected loss=0.09090909 P(node) =0.001003604
## class counts:      2      20
## probabilities: 0.091 0.909
##
## Node number 182: 30 observations
## predicted class=0 expected loss=0.2666667 P(node) =0.001368551
## class counts:      22      8
## probabilities: 0.733 0.267
##

```

```

## Node number 183: 36 observations
##   predicted class=1   expected loss=0.4722222   P(node) =0.001642261
##   class counts:      17      19
##   probabilities: 0.472 0.528
##
## Node number 194: 16 observations
##   predicted class=0   expected loss=0.25   P(node) =0.0007298937
##   class counts:      12      4
##   probabilities: 0.750 0.250
##
## Node number 195: 14 observations
##   predicted class=1   expected loss=0.4285714   P(node) =0.000638657
##   class counts:       6      8
##   probabilities: 0.429 0.571
##
## Node number 206: 126 observations
##   predicted class=0   expected loss=0.468254   P(node) =0.005747913
##   class counts:      67      59
##   probabilities: 0.532 0.468
##
## Node number 207: 9 observations
##   predicted class=1   expected loss=0.1111111   P(node) =0.0004105652
##   class counts:       1      8
##   probabilities: 0.111 0.889
##
## Node number 210: 21 observations
##   predicted class=0   expected loss=0.3333333   P(node) =0.0009579855
##   class counts:      14      7
##   probabilities: 0.667 0.333
##
## Node number 211: 211 observations
##   predicted class=1   expected loss=0.4454976   P(node) =0.009625473
##   class counts:      94     117
##   probabilities: 0.445 0.555
##
## Node number 212: 13 observations
##   predicted class=0   expected loss=0.1538462   P(node) =0.0005930386
##   class counts:      11      2
##   probabilities: 0.846 0.154
##
## Node number 213: 134 observations
##   predicted class=1   expected loss=0.4701493   P(node) =0.00611286
##   class counts:      63      71
##   probabilities: 0.470 0.530
##
## Node number 218: 13 observations
##   predicted class=0   expected loss=0.3076923   P(node) =0.0005930386
##   class counts:       9      4
##   probabilities: 0.692 0.308
##
## Node number 219: 35 observations
##   predicted class=1   expected loss=0.4   P(node) =0.001596642
##   class counts:      14      21
##   probabilities: 0.400 0.600

```

```

##
## Node number 228: 28 observations
##   predicted class=0   expected loss=0.4285714   P(node) =0.001277314
##   class counts:      16      12
##   probabilities: 0.571 0.429
##
## Node number 229: 12 observations
##   predicted class=1   expected loss=0.08333333   P(node) =0.0005474203
##   class counts:       1      11
##   probabilities: 0.083 0.917
##
## Node number 232: 375 observations
##   predicted class=0   expected loss=0.3893333   P(node) =0.01710688
##   class counts:      229     146
##   probabilities: 0.611 0.389
##
## Node number 233: 34 observations
##   predicted class=1   expected loss=0.4117647   P(node) =0.001551024
##   class counts:       14      20
##   probabilities: 0.412 0.588
##
## Node number 236: 209 observations
##   predicted class=0   expected loss=0.4114833   P(node) =0.009534237
##   class counts:      123      86
##   probabilities: 0.589 0.411
##
## Node number 237: 23 observations
##   predicted class=1   expected loss=0.3478261   P(node) =0.001049222
##   class counts:        8      15
##   probabilities: 0.348 0.652
##
## Node number 238: 30 observations
##   predicted class=0   expected loss=0.3666667   P(node) =0.001368551
##   class counts:       19      11
##   probabilities: 0.633 0.367
##
## Node number 239: 961 observations
##   predicted class=1   expected loss=0.3818939   P(node) =0.04383924
##   class counts:      367     594
##   probabilities: 0.382 0.618
##
## Node number 240: 171 observations
##   predicted class=0   expected loss=0.3918129   P(node) =0.007800739
##   class counts:      104      67
##   probabilities: 0.608 0.392
##
## Node number 241: 47 observations
##   predicted class=1   expected loss=0.4042553   P(node) =0.002144063
##   class counts:       19      28
##   probabilities: 0.404 0.596
##
## Node number 242: 106 observations
##   predicted class=0   expected loss=0.490566   P(node) =0.004835546
##   class counts:       54      52

```



```
## probabilities: 0.509 0.491
##
## Node number 243: 43 observations
## predicted class=1 expected loss=0.2790698 P(node) =0.001961589
## class counts: 12 31
## probabilities: 0.279 0.721
```