Adaboost

Adaboost

Prepared from https://qizeresearch.wordpress.com/2013/12/05/short-example-for-adaboost/

AdaBoost (adaptive boost) algorithm is another ensemble classification technology in data mining. Let???s suppose the y value (yes or not) is (+1, -1)

1. first find a week classifier with N observations (assign each observation equal weight $1/N$)

2. calculate errors between predicted y values and real y values use the errors as information to reweigh each observation (the total weight is set to be 1);

3. repeat this process T times;

4. finally output the sign of a combination of all classifier outputs.

The reason for using AdaBoost algorithm is that: we could not find a strong classifier some time at the beginning. Normally, in this case, the most important things we need to do is to brainstorm some new features, which really captures some insight of (people???s) behaviors. However, in case we cannot figure out new features, we had to use AdaBoost to somehow leverage our current knowledge, and build a strong classifier.

Let's look at the algorithm of AdaBoost: (suppose $Y = \{-1, +1\}$)

- **Step 1**: suppose we have N observations, we weight each observation with $1/N$ weight (W_i).

$D_i = \frac{W_i}{\sum W_i}$. In other word, D_i is the normalized weight.

- **Step 2**: To the following sub-steps for $t = 1$ to T, repeat T times:
    - **Sub-step2.1**:
    select a base classifier h (can be CART, Logit, probit, etc), and train data with the given weight
    - **Sub-step2.2**: Typically, we recognize the err of the classifier:
    $Err(t) = \sum D_i * I(y_i! = h_t(x_i))$. In other words, we sum D_i for all the observations with the predicted value not equal to the true value. $a_t = 1/2 * ln(\frac{1 - Err(t)}{Err(t)})$
    - **Sub-step2.3**:
    update Weight $W_i(t + 1) = D_i(t) * exp(-a_t * Y_i * h_t(X_i))$
    - **Sub-step2.4** : normalized weight for each observation: $Di(t + 1) = Wi(t + 1)/Sum(Wi(t + 1))$. In other word, Di(t+1) is normalized Wi(t+1) in t+t step.

Step 3: H(x)=sign(sum(a(t)*ht(x))). t from 1 to T.

In R, the adaboost package (named adabag) can be downloaded and install. This package implements the Freund and Schapire???s Adaboost.M1 algorithm. The ???weak??? classifier used in this algorithm is CART. Notice that before install adabag package, you first need: rpart, mlbench, caret, lattice, reshape2, plyr, cluster, foreach packages. Some of them are pre-install by default in R. The Lack of system integration is a drawback for open source software.

```
library(adabag)
```

```
## Loading required package: rpart

## Warning: package 'rpart' was built under R version 3.5.2

## Loading required package: caret

## Warning: package 'caret' was built under R version 3.5.2

## Loading required package: lattice
```

```
## Loading required package: ggplot2

## Loading required package: foreach

## Loading required package: doParallel

## Warning: package 'doParallel' was built under R version 3.5.2

## Loading required package: iterators

## Loading required package: parallel
```

```r
adadata <- read.csv("/Users/jvsingh/work/github/R-codes/data/bank/bank-full.csv", header=TRUE,sep=";")
adaboost<-boosting(y ~ ., data=adadata, boos=TRUE, mfinal=20, coeflearn='Breiman')
```

Here,

- **mfinal**=20 indicates the times of repeated process is 20.

- **coeflearn**="Breiman" indicates using M1 algorithm proposed by Breiman. Other options are "Freund" and "Zhu". The differences are in the calculation of a(t) .

- **errorevol** function can be used to show the evolution of error during the boosting process.

- **predict** function can be used to predict new data based on the existing adaboost trees.

In reality, we can use adaboost$trees to get all 20 trees, and also use adaboost$weights to get all weights, then calculate the predict result for new data by ourselves (saving time).

```
##            Length Class    Mode
## formula        3  formula  call
## trees         20  -none-   list
## weights       20  -none-   numeric
## votes      90422  -none-   numeric
## prob       90422  -none-   numeric
## class      45211  -none-   character
## importance    16  -none-   numeric
## terms          3  terms    call
## call           6  -none-   call

## [[1]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##  1) root 45211 5260 no (0.88365663 0.11634337)
##     2) duration< 503.5 39944 3039 no (0.92391849 0.07608151)
##       4) poutcome=failure,other,unknown 38694 2284 no (0.94097276 0.05902724) *
##       5) poutcome=success 1250  495 yes (0.39600000 0.60400000)
##        10) duration< 146.5 276   67 no (0.75724638 0.24275362) *
##        11) duration>=146.5 974  286 yes (0.29363450 0.70636550) *
##     3) duration>=503.5 5267 2221 no (0.57831783 0.42168217)
##       6) duration< 648.5 2127  611 no (0.71274095 0.28725905) *
##       7) duration>=648.5 3140 1530 yes (0.48726115 0.51273885)
##        14) contact=telephone,unknown 1062  411 no (0.61299435 0.38700565) *
##        15) contact=cellular 2078  879 yes (0.42300289 0.57699711) *
##
## [[2]]
## n= 45211
##
```

```
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##  1) root 45211 9939 no (0.7801641 0.2198359)
##    2) duration< 365.5 33188 4571 no (0.8622695 0.1377305)
##      4) month=apr,aug,feb,jan,jul,jun,may,nov 31014 3249 no (0.8952409 0.1047591) *
##      5) month=dec,mar,oct,sep 2174  852 yes (0.3919043 0.6080957)
##       10) duration< 99.5 324   65 no (0.7993827 0.2006173) *
##       11) duration>=99.5 1850  593 yes (0.3205405 0.6794595) *
##    3) duration>=365.5 12023 5368 no (0.5535224 0.4464776)
##      6) housing=yes 6277 2330 no (0.6288036 0.3711964) *
##      7) housing=no 5746 2708 yes (0.4712844 0.5287156)
##       14) month=aug,feb,jan,jul,jun,may,nov 4775 2308 no (0.5166492 0.4833508)
##         28) pdays< 32.5 4108 1866 no (0.5457644 0.4542356)
##           56) duration< 972.5 3456 1485 no (0.5703125 0.4296875) *
##           57) duration>=972.5 652  271 yes (0.4156442 0.5843558) *
##         29) pdays>=32.5 667  225 yes (0.3373313 0.6626687) *
##       15) month=apr,dec,mar,oct,sep 971  241 yes (0.2481977 0.7518023) *
##
## [[3]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##  1) root 45211 14754 no (0.67366349 0.32633651)
##    2) duration< 225.5 22402  3642 no (0.83742523 0.16257477) *
##    3) duration>=225.5 22809 11112 no (0.51282389 0.48717611)
##      6) duration< 408.5 10540  4089 no (0.61204934 0.38795066)
##       12) contact=unknown 2052    91 no (0.95565302 0.04434698) *
##       13) contact=cellular,telephone 8488  3998 no (0.52898209 0.47101791)
##         26) month=aug,jan,jul,may,nov 5234  1942 no (0.62896446 0.37103554) *
##         27) month=apr,dec,feb,jun,mar,oct,sep 3254  1198 yes (0.36816226 0.63183774) *
##      7) duration>=408.5 12269  5246 yes (0.42758171 0.57241829)
##       14) month=apr,feb,jan,jul,nov,sep 5259  2623 no (0.50123598 0.49876402)
##         28) duration>=650.5 2549   992 no (0.61082778 0.38917222) *
##         29) duration< 650.5 2710  1079 yes (0.39815498 0.60184502) *
##       15) month=aug,dec,jun,mar,may,oct 7010  2610 yes (0.37232525 0.62767475) *
##
## [[4]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##   1) root 45211 16365 no (0.63803057 0.36196943)
##    2) duration< 129.5 10305  1429 no (0.86132945 0.13867055) *
##    3) duration>=129.5 34906 14936 no (0.57210795 0.42789205)
##      6) contact=unknown 7187  1779 no (0.75246974 0.24753026)
##       12) duration< 536.5 4873   372 no (0.92366099 0.07633901) *
##       13) duration>=536.5 2314   907 yes (0.39196197 0.60803803) *
##      7) contact=cellular,telephone 27719 13157 no (0.52534363 0.47465637)
##       14) housing=yes 11506  4363 no (0.62080654 0.37919346) *
##       15) housing=no 16213  7419 yes (0.45759576 0.54240424)
```

```
##             30) month=aug,dec,feb,jan,jul,nov,oct,sep 12214   5913 no (0.51588341 0.48411659)
##               60) pdays< 37.5 9161   4026 no (0.56052833 0.43947167)
##                120) job=admin.,blue-collar,entrepreneur,housemaid,management,self-employed,services,tec
##                121) job=retired,student 1558    581 yes (0.37291399 0.62708601) *
##               61) pdays>=37.5 3053   1166 yes (0.38191942 0.61808058) *
##             31) month=apr,jun,mar,may 3999   1118 yes (0.27956989 0.72043011) *
##
## [[5]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##    1) root 45211 18583 no (0.5889717 0.4110283)
##      2) duration< 116.5 8109   1427 no (0.8240227 0.1759773) *
##      3) duration>=116.5 37102 17156 no (0.5375991 0.4624009)
##        6) contact=unknown 7201   1813 no (0.7482294 0.2517706) *
##        7) contact=cellular,telephone 29901 14558 yes (0.4868733 0.5131267)
##         14) month=apr,aug,dec,jan,jul,mar,may,nov,oct,sep 25621 12616 no (0.5075914 0.4924086)
##           28) duration< 204.5 6638   2468 no (0.6282013 0.3717987)
##             56) month=aug,dec,jan,jul,may,nov 4771   1390 no (0.7086565 0.2913435) *
##             57) month=apr,mar,oct,sep 1867    789 yes (0.4226031 0.5773969) *
##           29) duration>=204.5 18983   8835 yes (0.4654164 0.5345836)
##             58) pdays< 411 18594   8778 yes (0.4720878 0.5279122)
##              116) balance< 1490.5 12903   6442 no (0.5007363 0.4992637)
##                232) duration< 518.5 8037   3735 no (0.5352744 0.4647256)
##                  464) age>=25.5 7484   3346 no (0.5529129 0.4470871) *
##                  465) age< 25.5 553    164 yes (0.2965642 0.7034358) *
##                233) duration>=518.5 4866   2159 yes (0.4436909 0.5563091) *
##              117) balance>=1490.5 5691   2317 yes (0.4071341 0.5928659) *
##             59) pdays>=411 389     57 yes (0.1465296 0.8534704) *
##         15) month=feb,jun 4280   1553 yes (0.3628505 0.6371495) *
##
## [[6]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##    1) root 45211 19354 no (0.57191834 0.42808166)
##      2) duration< 77.5 3495    296 no (0.91530758 0.08469242) *
##      3) duration>=77.5 41716 19058 no (0.54314891 0.45685109)
##        6) contact=unknown 7639   2202 no (0.71174237 0.28825763)
##         12) duration< 637.5 5728    976 no (0.82960894 0.17039106)
##           24) month=jul,jun,may 5368    640 no (0.88077496 0.11922504) *
##           25) month=apr,aug,feb,jan,mar,nov,oct,sep 360     24 yes (0.06666667 0.93333333) *
##         13) duration>=637.5 1911    685 yes (0.35845107 0.64154893) *
##        7) contact=cellular,telephone 34077 16856 no (0.50535552 0.49464448)
##         14) campaign>=1.5 18682   8376 no (0.55165400 0.44834600)
##           28) month=apr,aug,dec,feb,jan,jul,may,nov,oct,sep 17380   7496 no (0.56869965 0.43130035)
##             56) duration< 220.5 6036   1999 no (0.66882041 0.33117959)
##              112) month=aug,jan,jul,may,nov,oct 4307   1022 no (0.76271186 0.23728814) *
##              113) month=apr,dec,feb,sep 1729    752 yes (0.43493349 0.56506651) *
##             57) duration>=220.5 11344   5497 no (0.51542666 0.48457334)
```

4

```
##             114) education=primary,secondary 6847   3068 no (0.55192055 0.44807945) *
##              115) education=tertiary,unknown 4497   2068 yes (0.45986213 0.54013787) *
##         29) month=jun,mar 1302    422 yes (0.32411674 0.67588326) *
##       15) campaign< 1.5 15395  6915 yes (0.44917181 0.55082819)
##         30) month=apr,dec,feb,jan,jul,may,nov,oct,sep 12493  6037 yes (0.48323061 0.51676939)
##           60) job=blue-collar,entrepreneur,housemaid 2603   1013 no (0.61083365 0.38916635) *
##           61) job=admin.,management,retired,self-employed,services,student,technician,unemployed,unl
##         31) month=aug,jun,mar 2902    878 yes (0.30254997 0.69745003) *
##
## [[7]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##   1) root 45211 19289 no (0.5733560 0.4266440)
##     2) duration< 77.5 3074    351 no (0.8858165 0.1141835) *
##     3) duration>=77.5 42137 18938 no (0.5505613 0.4494387)
##       6) contact=unknown 7271  2282 no (0.6861505 0.3138495) *
##       7) contact=cellular,telephone 34866 16656 no (0.5222853 0.4777147)
##        14) loan=yes 4191   1518 no (0.6377953 0.3622047) *
##        15) loan=no 30675 15138 no (0.5065037 0.4934963)
##          30) marital=married 17079  7852 no (0.5402541 0.4597459) *
##          31) marital=divorced,single 13596  6310 yes (0.4641071 0.5358929)
##            62) age>=30.5 8754  4373 yes (0.4995431 0.5004569)
##             124) month=apr,dec,feb,jan,mar,may,nov 4978  2253 no (0.5474086 0.4525914) *
##             125) month=aug,jul,jun,oct,sep 3776  1648 yes (0.4364407 0.5635593) *
##            63) age< 30.5 4842  1937 yes (0.4000413 0.5999587) *
##
## [[8]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##  1) root 45211 20688 no (0.5424122 0.4575878)
##    2) duration< 168.5 13036  4145 no (0.6820344 0.3179656)
##      4) month=aug,jan,jul,jun,may 7775  1462 no (0.8119614 0.1880386) *
##      5) month=apr,dec,feb,mar,nov,oct,sep 5261  2578 yes (0.4900209 0.5099791)
##       10) duration< 60.5 477     50 no (0.8951782 0.1048218) *
##       11) duration>=60.5 4784  2151 yes (0.4496237 0.5503763)
##         22) poutcome=failure,other 1196    483 no (0.5961538 0.4038462) *
##         23) poutcome=success,unknown 3588  1438 yes (0.4007804 0.5992196) *
##    3) duration>=168.5 32175 15632 yes (0.4858430 0.5141570)
##      6) month=dec,feb,jan,mar,may,oct,sep 13930  6517 no (0.5321608 0.4678392)
##       12) duration< 325.5 5616  2192 no (0.6096866 0.3903134) *
##       13) duration>=325.5 8314  3989 yes (0.4797931 0.5202069)
##         26) housing=no 3180  1413 no (0.5556604 0.4443396) *
##         27) housing=yes 5134  2222 yes (0.4328009 0.5671991) *
##      7) month=apr,aug,jul,jun,nov 18245  8219 yes (0.4504796 0.5495204)
##       14) day>=16.5 8792  4339 no (0.5064832 0.4935168)
##         28) day< 24.5 6181  2770 no (0.5518525 0.4481475) *
##         29) day>=24.5 2611  1042 yes (0.3990808 0.6009192) *
##       15) day< 16.5 9453  3766 yes (0.3983920 0.6016080) *
```

```
## 
## [[9]]
## n= 45211
## 
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
## 
##  1) root 45211 20658 no (0.54307580 0.45692420)
##    2) duration< 88.5 3731   653 no (0.82497990 0.17502010) *
##    3) duration>=88.5 41480 20005 no (0.51771938 0.48228062)
##      6) contact=unknown 6925  2500 no (0.63898917 0.36101083)
##       12) duration< 368.5 3091   458 no (0.85182789 0.14817211)
##         24) month=jul,jun,may 2782   167 no (0.93997124 0.06002876) *
##         25) month=apr,aug,feb,jan,mar,nov,oct,sep 309    18 yes (0.05825243 0.94174757) *
##       13) duration>=368.5 3834  1792 yes (0.46739697 0.53260303) *
##      7) contact=cellular,telephone 34555 17050 yes (0.49341629 0.50658371)
##       14) duration< 154.5 6197  2670 no (0.56914636 0.43085364)
##         28) poutcome=failure,other,unknown 5466  2066 no (0.62202708 0.37797292)
##           56) month=aug,dec,jan,jul,may,nov,sep 3081   717 no (0.76728335 0.23271665) *
##           57) month=apr,feb,jun,mar,oct 2385  1036 yes (0.43438155 0.56561845) *
##         29) poutcome=success 731   127 yes (0.17373461 0.82626539) *
##       15) duration>=154.5 28358 13523 yes (0.47686720 0.52313280)
##         30) pdays>=99.5 6744  3112 no (0.53855279 0.46144721)
##           60) month=apr,dec,feb,jan,jun,mar,may,nov,oct,sep 5659  2303 no (0.59303764 0.40696236) *
##           61) month=aug,jul 1085   276 yes (0.25437788 0.74562212) *
##         31) pdays< 99.5 21614  9891 yes (0.45762006 0.54237994) *
## 
## [[10]]
## n= 45211
## 
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
## 
##   1) root 45211 19855 no (0.5608370 0.4391630)
##    2) duration< 77.5 2607   461 no (0.8231684 0.1768316) *
##    3) duration>=77.5 42604 19394 no (0.5447845 0.4552155)
##      6) contact=unknown 6887  2254 no (0.6727167 0.3272833) *
##      7) contact=cellular,telephone 35717 17140 no (0.5201165 0.4798835)
##       14) poutcome=failure,unknown 29980 13843 no (0.5382588 0.4617412)
##         28) duration< 157.5 5965  2233 no (0.6256496 0.3743504) *
##         29) duration>=157.5 24015 11610 no (0.5165522 0.4834478)
##           58) age< 59.5 21883 10299 no (0.5293607 0.4706393)
##            116) month=apr,aug,feb,jan,jul 13390  5858 no (0.5625093 0.4374907) *
##            117) month=dec,jun,mar,may,nov,oct,sep 8493  4052 yes (0.4770988 0.5229012) *
##           59) age>=59.5 2132   821 yes (0.3850844 0.6149156) *
##       15) poutcome=other,success 5737  2440 yes (0.4253094 0.5746906) *
## 
## [[11]]
## n= 45211
## 
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
## 
##   1) root 45211 20308 no (0.55081728 0.44918272)
```

```
##      2) duration< 62.5 1496    208 no (0.86096257 0.13903743) *
##      3) duration>=62.5 43715 20100 no (0.54020359 0.45979641)
##        6) contact=unknown 7178  2655 no (0.63011981 0.36988019)
##         12) duration< 459.5 3724    765 no (0.79457573 0.20542427)
##           24) month=jul,jun,may 3351    411 no (0.87735004 0.12264996) *
##           25) month=apr,aug,feb,jan,mar,nov,oct,sep 373    19 yes (0.05093834 0.94906166) *
##         13) duration>=459.5 3454  1564 yes (0.45280834 0.54719166) *
##        7) contact=cellular,telephone 36537 17445 no (0.52253880 0.47746120)
##         14) pdays< 384.5 35862 16948 no (0.52741063 0.47258937)
##           28) loan=yes 4342  1680 no (0.61308153 0.38691847) *
##           29) loan=no 31520 15268 no (0.51560914 0.48439086)
##             58) duration>=659.5 4875  1969 no (0.59610256 0.40389744) *
##             59) duration< 659.5 26645 13299 no (0.50088197 0.49911803)
##              118) housing=yes 9705  4318 no (0.55507470 0.44492530)
##                236) month=apr,feb,jan,jul,may,nov 8006  3247 no (0.59442918 0.40557082)
##                  472) duration< 363.5 5358  1815 no (0.66125420 0.33874580) *
##                  473) duration>=363.5 2648  1216 yes (0.45921450 0.54078550) *
##                237) month=aug,dec,jun,mar,oct,sep 1699    628 yes (0.36962919 0.63037081) *
##              119) housing=no 16940  7959 yes (0.46983471 0.53016529)
##                238) month=aug,dec,mar,oct,sep 6634  3142 no (0.52637926 0.47362074) *
##                239) month=apr,feb,jan,jul,jun,may,nov 10306  4467 yes (0.43343683 0.56656317)
##                  478) day>=28.5 1048    377 no (0.64026718 0.35973282) *
##                  479) day< 28.5 9258  3796 yes (0.41002376 0.58997624)
##                    958) balance< 99.5 1386    591 no (0.57359307 0.42640693) *
##                    959) balance>=99.5 7872  3001 yes (0.38122459 0.61877541) *
##         15) pdays>=384.5 675    178 yes (0.26370370 0.73629630) *
##
## [[12]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##  1) root 45211 21200 no (0.5310876 0.4689124)
##    2) duration< 62.5 1441    218 no (0.8487162 0.1512838) *
##    3) duration>=62.5 43770 20982 no (0.5206306 0.4793694)
##      6) contact=unknown 6904  2476 no (0.6413673 0.3586327)
##       12) month=jul,jun,may 6502  2115 no (0.6747155 0.3252845) *
##       13) month=apr,aug,feb,jan,mar,nov,oct,sep 402    41 yes (0.1019900 0.8980100) *
##      7) contact=cellular,telephone 36866 18360 yes (0.4980199 0.5019801)
##       14) month=dec,jul,jun,may,nov,oct,sep 19981  9507 no (0.5241980 0.4758020)
##         28) duration< 246.5 7402  3051 no (0.5878141 0.4121859) *
##         29) duration>=246.5 12579  6123 yes (0.4867637 0.5132363)
##           58) balance>=665.5 6107  2860 no (0.5316850 0.4683150) *
##           59) balance< 665.5 6472  2876 yes (0.4443758 0.5556242) *
##       15) month=apr,aug,feb,jan,mar 16885  7886 yes (0.4670418 0.5329582) *
##
## [[13]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##  1) root 45211 20907 no (0.5375683 0.4624317)
```

```
##      2) duration< 60.5 1299   241 no (0.8144727 0.1855273) *
##      3) duration>=60.5 43912 20666 no (0.5293769 0.4706231)
##        6) contact=unknown 6855  2618 no (0.6180890 0.3819110) *
##        7) contact=cellular,telephone 37057 18048 no (0.5129665 0.4870335)
##         14) balance< -41 1439   467 no (0.6754691 0.3245309) *
##         15) balance>=-41 35618 17581 no (0.5064013 0.4935987)
##           30) day< 8.5 10021  4549 no (0.5460533 0.4539467)
##             60) month=apr,dec,feb,may,oct,sep 5399  1888 no (0.6503056 0.3496944) *
##             61) month=aug,jan,jul,jun,mar,nov 4622  1961 yes (0.4242752 0.5757248) *
##           31) day>=8.5 25597 12565 yes (0.4908778 0.5091222)
##             62) month=apr,aug,dec,jan,jul,jun,mar,may,nov,oct,sep 24196 11913 no (0.5076459 0.4923541)
##              124) duration< 121.5 2570   979 no (0.6190661 0.3809339)
##                248) month=aug,jan,jul,may,nov 1590   349 no (0.7805031 0.2194969) *
##                249) month=apr,dec,jun,mar,oct,sep 980   350 yes (0.3571429 0.6428571) *
##              125) duration>=121.5 21626 10692 yes (0.4944049 0.5055951)
##                250) job=blue-collar,housemaid,management,retired,self-employed,technician,unknown 1542
##                251) job=admin.,entrepreneur,services,student,unemployed 6200  2801 yes (0.4517742 0.54
##             63) month=feb 1401   282 yes (0.2012848 0.7987152) *
##
## [[14]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##   1) root 45211 21932 no (0.5148968 0.4851032)
##     2) duration< 62.5 1341   291 no (0.7829978 0.2170022) *
##     3) duration>=62.5 43870 21641 no (0.5067016 0.4932984)
##       6) balance< 60.5 7866  3284 no (0.5825070 0.4174930) *
##       7) balance>=60.5 36004 17647 yes (0.4901400 0.5098600)
##        14) day< 9.5 11237  5220 no (0.5354632 0.4645368)
##          28) month=apr,aug,dec,feb,jun,may,oct,sep 9245  3979 no (0.5696052 0.4303948) *
##          29) month=jan,jul,mar,nov 1992   751 yes (0.3770080 0.6229920) *
##        15) day>=9.5 24767 11630 yes (0.4695765 0.5304235)
##          30) month=apr,aug,dec,jan,jul,jun,mar,may,nov,oct,sep 23677 11434 yes (0.4829159 0.5170841)
##            60) pdays< 3 17083  8417 no (0.5072879 0.4927121)
##             120) duration< 211.5 5021  1948 no (0.6120295 0.3879705) *
##             121) duration>=211.5 12062  5593 yes (0.4636876 0.5363124)
##               242) duration< 800.5 9438  4635 yes (0.4910998 0.5089002)
##                 484) housing=yes 4533  2024 no (0.5534966 0.4465034) *
##                 485) housing=no 4905  2126 yes (0.4334353 0.5665647) *
##               243) duration>=800.5 2624   958 yes (0.3650915 0.6349085) *
##            61) pdays>=3 6594  2768 yes (0.4197756 0.5802244) *
##          31) month=feb 1090   196 yes (0.1798165 0.8201835) *
##
## [[15]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##   1) root 45211 22422 no (0.50405875 0.49594125)
##     2) duration< 77.5 2191   692 no (0.68416248 0.31583752) *
##     3) duration>=77.5 43020 21290 yes (0.49488610 0.50511390)
```

```
##        6) duration< 503.5 30092 14719 no (0.51086668 0.48913332)
##          12) contact=unknown 3264  1078 no (0.66973039 0.33026961)
##            24) month=dec,jul,jun,may 2937   769 no (0.73816820 0.26183180) *
##            25) month=apr,aug,feb,jan,mar,nov,oct,sep 327    18 yes (0.05504587 0.94495413) *
##          13) contact=cellular,telephone 26828 13187 yes (0.49153869 0.50846131)
##            26) balance< -0.5 902   265 no (0.70620843 0.29379157) *
##            27) balance>=-0.5 25926 12550 yes (0.48407005 0.51592995)
##              54) education=primary 2650  1092 no (0.58792453 0.41207547) *
##              55) education=secondary,tertiary,unknown 23276 10992 yes (0.47224609 0.52775391)
##               110) month=feb,jul,nov 7464  3586 no (0.51956056 0.48043944)
##                 220) pdays< 33.5 5185  2205 no (0.57473481 0.42526519) *
##                 221) pdays>=33.5 2279   898 yes (0.39403247 0.60596753) *
##               111) month=apr,aug,dec,jan,jun,mar,may,oct,sep 15812  7114 yes (0.44991146 0.55008854) *
##        7) duration>=503.5 12928  5917 yes (0.45768874 0.54231126)
##          14) housing=no 5807  2724 no (0.53091097 0.46908903) *
##          15) housing=yes 7121  2834 yes (0.39797781 0.60202219) *
##
## [[16]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##   1) root 45211 22101 no (0.5111588 0.4888412)
##     2) duration< 60.5 1237   355 no (0.7130154 0.2869846) *
##     3) duration>=60.5 43974 21746 no (0.5054805 0.4945195)
##       6) month=may 9162  4006 no (0.5627592 0.4372408) *
##       7) month=apr,aug,dec,feb,jan,jul,jun,mar,nov,oct,sep 34812 17072 yes (0.4904056 0.5095944)
##        14) duration< 475.5 24794 12209 no (0.5075825 0.4924175)
##          28) campaign>=3.5 3524  1412 no (0.5993190 0.4006810) *
##          29) campaign< 3.5 21270 10473 yes (0.4923836 0.5076164)
##            58) balance< -0.5 558   149 no (0.7329749 0.2670251) *
##            59) balance>=-0.5 20712 10064 yes (0.4859019 0.5140981)
##             118) age< 59.5 18340  9158 yes (0.4993457 0.5006543)
##               236) day>=15.5 8123  3678 no (0.5472116 0.4527884)
##                 472) month=apr,jan,jul,jun,nov 5324  1929 no (0.6376784 0.3623216) *
##                 473) month=aug,dec,feb,mar,oct,sep 2799  1050 yes (0.3751340 0.6248660) *
##               237) day< 15.5 10217  4713 yes (0.4612900 0.5387100) *
##             119) age>=59.5 2372   906 yes (0.3819562 0.6180438) *
##        15) duration>=475.5 10018  4487 yes (0.4478938 0.5521062) *
##
## [[17]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##  1) root 45211 21817 no (0.5174404 0.4825596)
##    2) duration< 287.5 21283  9587 no (0.5495466 0.4504534) *
##    3) duration>=287.5 23928 11698 yes (0.4888833 0.5111167)
##      6) job=housemaid,management,retired,services,student,unknown 10365  4914 no (0.5259045 0.4740955
##       12) duration>=406.5 7187  3171 no (0.5587867 0.4412133) *
##       13) duration< 406.5 3178  1435 yes (0.4515419 0.5484581) *
##      7) job=admin.,blue-collar,entrepreneur,self-employed,technician,unemployed 13563  6247 yes (0.4(
```

9

```
## 
## [[18]]
## n= 45211
## 
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
## 
##    1) root 45211 22103 no (0.5111146 0.4888854)
##      2) pdays< 377.5 44365 21548 no (0.5143018 0.4856982)
##        4) duration< 62.5 1106   345 no (0.6880651 0.3119349) *
##        5) duration>=62.5 43259 21203 no (0.5098592 0.4901408)
##         10) duration< 473.5 29108 13840 no (0.5245293 0.4754707)
##           20) contact=unknown 3028  1022 no (0.6624835 0.3375165)
##             40) month=dec,jun,may 2571   662 no (0.7425126 0.2574874) *
##             41) month=apr,aug,feb,jan,jul,mar,nov,oct,sep 457    97 yes (0.2122538 0.7877462) *
##           21) contact=cellular,telephone 26080 12818 no (0.5085123 0.4914877)
##             42) age< 60.5 23806 11418 no (0.5203730 0.4796270)
##               84) job=blue-collar,entrepreneur,retired,self-employed,services,unknown 7044  2973 no (
##               85) job=admin.,housemaid,management,student,technician,unemployed 16762  8317 yes (0.49
##                170) month=aug,dec,jan,jul,mar,nov,oct,sep 9536  4456 no (0.5327181 0.4672819)
##                  340) day>=16.5 4566  1827 no (0.5998686 0.4001314) *
##                  341) day< 16.5 4970  2341 yes (0.4710262 0.5289738)
##                    682) month=aug,dec,jul,mar,oct,sep 4181  1953 no (0.5328869 0.4671131) *
##                    683) month=jan,nov 789   113 yes (0.1432193 0.8567807) *
##                171) month=apr,feb,jun,may 7226  3237 yes (0.4479657 0.5520343) *
##             43) age>=60.5 2274   874 yes (0.3843448 0.6156552) *
##         11) duration>=473.5 14151  6788 yes (0.4796834 0.5203166)
##           22) contact=cellular 9695  4709 no (0.5142857 0.4857143)
##             44) job=housemaid,retired,student,unemployed 1603   576 no (0.6406737 0.3593263) *
##             45) job=admin.,blue-collar,entrepreneur,management,self-employed,services,technician,unkno
##               90) duration>=648.5 4492  2097 no (0.5331701 0.4668299) *
##               91) duration< 648.5 3600  1564 yes (0.4344444 0.5655556) *
##           23) contact=telephone,unknown 4456  1802 yes (0.4043986 0.5956014) *
##      3) pdays>=377.5 846   291 yes (0.3439716 0.6560284) *
## 
## [[19]]
## n= 45211
## 
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
## 
##    1) root 45211 22209 no (0.5087700 0.4912300)
##      2) housing=yes 20121  9386 no (0.5335222 0.4664778)
##        4) duration< 892.5 17789  7895 no (0.5561864 0.4438136)
##          8) pdays< 9.5 12445  5016 no (0.5969466 0.4030534)
##           16) month=apr,aug,jan,jul,may,nov 9959  3662 no (0.6322924 0.3677076) *
##           17) month=dec,feb,jun,mar,oct,sep 2486  1132 yes (0.4553500 0.5446500)
##             34) day< 9.5 1440   603 no (0.5812500 0.4187500) *
##             35) day>=9.5 1046   295 yes (0.2820268 0.7179732) *
##          9) pdays>=9.5 5344  2465 yes (0.4612650 0.5387350) *
##        5) duration>=892.5 2332   841 yes (0.3606346 0.6393654) *
##      3) housing=no 25090 12267 yes (0.4889199 0.5110801)
##        6) duration>=1008.5 1220   404 no (0.6688525 0.3311475) *
##        7) duration< 1008.5 23870 11451 yes (0.4797235 0.5202765)
```

```
##        14) pdays>=99.5 4223  1887 no (0.5531613 0.4468387) *
##        15) pdays< 99.5 19647  9115 yes (0.4639385 0.5360615)
##          30) month=aug,feb,jul,jun,sep 12234  6107 no (0.5008174 0.4991826)
##            60) age>=27.5 10722  5169 no (0.5179071 0.4820929)
##             120) duration< 329.5 5908  2525 no (0.5726134 0.4273866) *
##             121) duration>=329.5 4814  2170 yes (0.4507686 0.5492314) *
##            61) age< 27.5 1512   574 yes (0.3796296 0.6203704) *
##          31) month=apr,dec,jan,mar,may,nov,oct 7413  2988 yes (0.4030757 0.5969243) *
##
## [[20]]
## n= 45211
##
## node), split, n, loss, yval, (yprob)
##       * denotes terminal node
##
##   1) root 45211 22415 no (0.5042136 0.4957864)
##     2) duration< 22.5 260    55 no (0.7884615 0.2115385) *
##     3) duration>=22.5 44951 22360 no (0.5025695 0.4974305)
##       6) pdays< 373.5 44119 21819 no (0.5054512 0.4945488)
##        12) pdays>=202.5 3758  1505 no (0.5995210 0.4004790) *
##        13) pdays< 202.5 40361 20047 yes (0.4966924 0.5033076)
##          26) duration< 62.5 658   217 no (0.6702128 0.3297872) *
##          27) duration>=62.5 39703 19606 yes (0.4938166 0.5061834)
##            54) job=retired,self-employed,services,student,unemployed,unknown 10798  5075 no (0.530005
##            55) job=admin.,blue-collar,entrepreneur,housemaid,management,technician 28905 13883 yes (
##             110) month=apr,aug,jun,may,nov 18919  9437 yes (0.4988107 0.5011893)
##               220) age>=29.5 16420  7973 no (0.5144336 0.4855664) *
##               221) age< 29.5 2499   990 yes (0.3961585 0.6038415) *
##             111) month=dec,feb,jan,jul,mar,oct,sep 9986  4446 yes (0.4452233 0.5547767) *
##       7) pdays>=373.5 832   291 yes (0.3497596 0.6502404) *
##
##  [1] 1.09080908 0.72374993 0.53112591 0.45202319 0.34504673 0.32242773
##  [7] 0.21536392 0.24721016 0.23773348 0.19448808 0.25674829 0.14951176
## [13] 0.17566721 0.17996303 0.17179451 0.15215157 0.08343577 0.18179705
## [19] 0.16699761 0.08373943
##
##      age    balance   campaign    contact        day    default
## 0.6357254  0.8029164  0.4851111 10.2384278  0.9715816  0.0000000
##  duration  education    housing        job       loan    marital
## 55.2693328  0.1847255  2.4839408  0.9350582  0.3296736  0.1355105
##     month      pdays   poutcome   previous
## 19.3717456  1.7985532  6.3576974  0.0000000
##
## $error
##  [1] 0.10141337 0.10141337 0.10019685 0.09747628 0.09880339 0.09813983
##  [7] 0.09727721 0.09690120 0.09606069 0.09497689 0.09539714 0.09504324
## [13] 0.09462299 0.09411426 0.09311893 0.09325164 0.09362766 0.09307469
## [19] 0.09322952 0.09303046
##
## attr(,"class")
## [1] "errorevol"
```