> source('C:/Users/hp1/Desktop/R\_CODE\_AND\_OUTPUT/DECISION\_TREE.r', echo=TRUE)

> library(party)

Loading required package: grid

Loading required package: mvtnorm

Loading required package: modeltools

Loading required package: stats4

Loading required package: strucchange

Loading required package: zoo

Attaching package: ‘zoo’

The following objects are masked from ‘package:base’:

as.Date, as.Date.numeric

Loading required package: sandwich

> str(train)

'data.frame': 32769 obs. of 10 variables:

$ ACTION : int 1 1 1 1 1 0 1 1 1 1 ...

$ RESOURCE : int 39353 17183 36724 36135 42680 45333 25993 19666 31246 78766 ...

$ MGR\_ID : int 85475 1540 14457 5396 5905 14561 17227 4209 783 56683 ...

$ ROLE\_ROLLUP\_1 : int 117961 117961 118219 117961 117929 117951 117961 117961 117961 118079 ...

$ ROLE\_ROLLUP\_2 : int 118300 118343 118220 118343 117930 117952 118343 117969 118413 118080 ...

$ ROLE\_DEPTNAME : int 123472 123125 117884 119993 119569 118008 123476 118910 120584 117878 ...

$ ROLE\_TITLE : int 117905 118536 117879 118321 119323 118568 118980 126820 128230 117879 ...

$ ROLE\_FAMILY\_DESC: int 117906 118536 267952 240983 123932 118568 301534 269034 302830 304519 ...

$ ROLE\_FAMILY : int 290919 308574 19721 290919 19793 19721 118295 118638 4673 19721 ...

$ ROLE\_CODE : int 117908 118539 117880 118322 119325 118570 118982 126822 128231 117880 ...

> tra <-ctree(ACTION ~ROLE\_ROLLUP\_1+ROLE\_TITLE+ROLE\_CODE, data =train)

> print(tra)

Conditional inference tree with 5 terminal nodes

Response: ACTION

Inputs: ROLE\_ROLLUP\_1, ROLE\_TITLE, ROLE\_CODE

Number of observations: 32769

1) ROLE\_CODE <= 117880; criterion = 0.994, statistic = 9.635

2)\* weights = 1256

1) ROLE\_CODE > 117880

3) ROLE\_CODE <= 117908; criterion = 0.974, statistic = 6.846

4)\* weights = 4794

3) ROLE\_CODE > 117908

5) ROLE\_CODE <= 117948; criterion = 0.999, statistic = 12.456

6)\* weights = 329

5) ROLE\_CODE > 117948

7) ROLE\_CODE <= 119900; criterion = 0.997, statistic = 10.57

8)\* weights = 19046

7) ROLE\_CODE > 119900

9)\* weights = 7344

> plot(tra)

> predictedX <- predict(tra, train)

> plot(predictedX)

> table(train$ACTION, predictedX > 0.5)

TRUE

0 1897

1 30872

> ROCRpred = prediction(predictedX, train$ACTION)

> auc = as.numeric(performance(ROCRpred, "auc")@y.values)

> ROCRperf = performance(ROCRpred, "tpr", "fpr")

> plot(ROCRperf)

> auc

[1] 0.5798906

>