recode

Code

Hide

```
# create a new dataset that will be used for recoding sentiment
iphone_smallMatrix_RC <- iphone_smallMatrix
# recode sentiment to combine factor levels 0 & 1 and 4 & 5
iphone_smallMatrix_RC$iphonesentiment <- recode(iphone_smallMatrix_RC$iphonesentiment,
'0' = 1, '1' = 1, '2' = 2, '3' = 3, '4' = 4, '5' = 4)
# inspect results
summary(iphone_smallMatrix_RC)
```

```
     iphone       samsunggalaxy       sonyxperia       nokialumina          htcphone
 Min.   : 0.000   Min.   :0.00000   Min.   :0.00000   Min.   :0.000000   Min.   :  0.000
0
 1st Qu.: 1.000   1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.000000   1st Qu.:  0.000
0
 Median : 1.000   Median :0.00000   Median :0.00000   Median :0.000000   Median :  0.000
0
 Mean   : 2.148   Mean   :0.07115   Mean   :0.02405   Mean   :0.002312   Mean   :  0.137
1
 3rd Qu.: 1.000   3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.000000   3rd Qu.:  0.000
0
 Max.   :46.000   Max.   :8.00000   Max.   :8.00000   Max.   :2.000000   Max.   :479.000
0
      ios          googleandroid      iphonecampos      samsungcampos        sonycampos
 Min.   :0.0000   Min.   :0.00000   Min.   :  0.0000   Min.   : 0.00000   Min.   :0.0000
00
 1st Qu.:0.0000   1st Qu.:0.00000   1st Qu.:  0.0000   1st Qu.: 0.00000   1st Qu.:0.0000
00
 Median :0.0000   Median :0.00000   Median :  0.0000   Median : 0.00000   Median :0.0000
00
 Mean   :0.1523   Mean   :0.03962   Mean   :  0.2896   Mean   : 0.05373   Mean   :0.0099
44
 3rd Qu.:0.0000   3rd Qu.:0.00000   3rd Qu.:  0.0000   3rd Qu.: 0.00000   3rd Qu.:0.0000
00
 Max.   :6.0000   Max.   :6.00000   Max.   :156.0000   Max.   :65.00000   Max.   :8.0000
00
  nokiacampos          htccampos          iphonecamneg      samsungcamneg         sonycamneg
 Min.   : 0.00000   Min.   :  0.0000   Min.   : 0.0000   Min.   : 0.00000   Min.   :0.00
0000
 1st Qu.: 0.00000   1st Qu.:  0.0000   1st Qu.: 0.0000   1st Qu.: 0.00000   1st Qu.:0.00
0000
 Median : 0.00000   Median :  0.0000   Median : 0.0000   Median : 0.00000   Median :0.00
0000
 Mean   : 0.00686   Mean   :  0.1132   Mean   : 0.2346   Mean   : 0.05473   Mean   :0.00
2312
 3rd Qu.: 0.00000   3rd Qu.:  0.0000   3rd Qu.: 0.0000   3rd Qu.: 0.00000   3rd Qu.:0.00
0000
 Max.   :17.00000   Max.   :156.0000   Max.   :31.0000   Max.   :17.00000   Max.   :5.00
0000
  nokiacamneg           htccamneg          iphonecamunc      samsungcamunc         sonycamun
c
 Min.   : 0.000000   Min.   : 0.00000   Min.   : 0.0000   Min.   : 0.00000   Min.   :0.0
00000
 1st Qu.: 0.000000   1st Qu.: 0.00000   1st Qu.: 0.0000   1st Qu.: 0.00000   1st Qu.:0.0
00000
 Median : 0.000000   Median : 0.00000   Median : 0.0000   Median : 0.00000   Median :0.0
00000
 Mean   : 0.006167   Mean   : 0.09296   Mean   : 0.2535   Mean   : 0.01873   Mean   :0.0
03237
 3rd Qu.: 0.000000   3rd Qu.: 0.00000   3rd Qu.: 0.0000   3rd Qu.: 0.00000   3rd Qu.:0.0
00000
 Max.   :12.000000   Max.   :21.00000   Max.   :45.0000   Max.   :11.00000   Max.   :4.0
00000
```

```
  nokiacamunc          htccamunc        iphonedispos     samsungdispos        sonydispos
 Min.   : 0.000000   Min.   : 0.0000   Min.   : 0.0000   Min.   : 0.00000   Min.   : 0.0
0000
 1st Qu.: 0.000000   1st Qu.: 0.0000   1st Qu.: 0.0000   1st Qu.: 0.00000   1st Qu.: 0.0
0000
 Median : 0.000000   Median : 0.0000   Median : 0.0000   Median : 0.00000   Median : 0.0
0000
 Mean   : 0.003623   Mean   : 0.0471   Mean   : 0.7172   Mean   : 0.06313   Mean   : 0.0
1141
 3rd Qu.: 0.000000   3rd Qu.: 0.0000   3rd Qu.: 0.0000   3rd Qu.: 0.00000   3rd Qu.: 0.0
0000
 Max.   :10.000000   Max.   :20.0000   Max.   :66.0000   Max.   :51.00000   Max.   :32.0
0000
  nokiadispos          htcdispos         iphonedisneg      samsungdisneg        sonydisn
eg
 Min.   : 0.000000   Min.   :  0.0000   Min.   :  0.0000   Min.   : 0.00000   Min.   :
0.00000
 1st Qu.: 0.000000   1st Qu.:  0.0000   1st Qu.:  0.0000   1st Qu.: 0.00000   1st Qu.:
0.00000
 Median : 0.000000   Median :  0.0000   Median :  0.0000   Median : 0.00000   Median :
0.00000
 Mean   : 0.009944   Mean   :  0.1581   Mean   :  0.6085   Mean   : 0.06598   Mean   :
0.01241
 3rd Qu.: 0.000000   3rd Qu.:  0.0000   3rd Qu.:  0.0000   3rd Qu.: 0.00000   3rd Qu.:
0.00000
 Max.   :23.000000   Max.   :479.0000   Max.   :120.0000   Max.   :48.00000   Max.   :9
7.00000
  nokiadisneg          htcdisneg        iphonedisunc     samsungdisunc        sonydisunc
 Min.   : 0.00000   Min.   : 0.0000   Min.   : 0.0000   Min.   : 0.00000   Min.   : 0.00
0000
 1st Qu.: 0.00000   1st Qu.: 0.0000   1st Qu.: 0.0000   1st Qu.: 0.00000   1st Qu.: 0.00
0000
 Median : 0.00000   Median : 0.0000   Median : 0.0000   Median : 0.00000   Median : 0.00
0000
 Mean   : 0.00794   Mean   : 0.1163   Mean   : 0.5049   Mean   : 0.02713   Mean   : 0.00
4548
 3rd Qu.: 0.00000   3rd Qu.: 0.0000   3rd Qu.: 0.0000   3rd Qu.: 0.00000   3rd Qu.: 0.00
0000
 Max.   :18.00000   Max.   :48.0000   Max.   :77.0000   Max.   :42.00000   Max.   :11.00
0000
  nokiadisunc          htcdisunc         iphoneperpos     samsungperpos        sonyperpos
 Min.   : 0.000000   Min.   : 0.00000   Min.   : 0.0000   Min.   : 0.0000   Min.   : 0.0
00000
 1st Qu.: 0.000000   1st Qu.: 0.00000   1st Qu.: 0.0000   1st Qu.: 0.0000   1st Qu.: 0.0
00000
 Median : 0.000000   Median : 0.00000   Median : 0.0000   Median : 0.0000   Median : 0.0
00000
 Mean   : 0.004933   Mean   : 0.06259   Mean   : 0.6054   Mean   : 0.0619   Mean   : 0.0
08633
 3rd Qu.: 0.000000   3rd Qu.: 0.00000   3rd Qu.: 0.0000   3rd Qu.: 0.0000   3rd Qu.: 0.0
00000
 Max.   :16.000000   Max.   :42.00000   Max.   :46.0000   Max.   :57.0000   Max.   :14.0
00000
  nokiaperpos          htcperpos        iphoneperneg     samsungperneg        sonyperneg
```

```
 Min.   : 0.00000   Min.   : 0.000   Min.   : 0.0000   Min.   : 0.00000   Min.   : 0.000
000
 1st Qu.: 0.00000   1st Qu.: 0.000   1st Qu.: 0.0000   1st Qu.: 0.00000   1st Qu.: 0.000
000
 Median : 0.00000   Median : 0.000   Median : 0.0000   Median : 0.00000   Median : 0.000
000
 Mean   : 0.01064   Mean   : 0.128   Mean   : 0.4866   Mean   : 0.06413   Mean   : 0.006
398
 3rd Qu.: 0.00000   3rd Qu.: 0.000   3rd Qu.: 0.0000   3rd Qu.: 0.00000   3rd Qu.: 0.000
000
 Max.   :21.00000   Max.   :65.000   Max.   :33.0000   Max.   :33.00000   Max.   :28.000
000
  nokiaperneg         htcperneg        iphoneperunc     samsungperunc        sonyperunc
 Min.   : 0.00000   Min.   : 0.0000   Min.   : 0.0000   Min.   : 0.00000   Min.   :0.000
000
 1st Qu.: 0.00000   1st Qu.: 0.0000   1st Qu.: 0.0000   1st Qu.: 0.00000   1st Qu.:0.000
000
 Median : 0.00000   Median : 0.0000   Median : 0.0000   Median : 0.00000   Median :0.000
000
 Mean   : 0.01025   Mean   : 0.1204   Mean   : 0.2991   Mean   : 0.02212   Mean   :0.002
312
 3rd Qu.: 0.00000   3rd Qu.: 0.0000   3rd Qu.: 0.0000   3rd Qu.: 0.00000   3rd Qu.:0.000
000
 Max.   :18.00000   Max.   :33.0000   Max.   :31.0000   Max.   :31.00000   Max.   :5.000
000
  nokiaperunc         htcperunc          iosperpos        googleperpos         iosperneg
 Min.   : 0.00000   Min.   : 0.00000   Min.   : 0.0000   Min.   : 0.00000   Min.   : 0.0
0000
 1st Qu.: 0.00000   1st Qu.: 0.00000   1st Qu.: 0.0000   1st Qu.: 0.00000   1st Qu.: 0.0
0000
 Median : 0.00000   Median : 0.00000   Median : 0.0000   Median : 0.00000   Median : 0.0
0000
 Mean   : 0.00555   Mean   : 0.06413   Mean   : 0.0592   Mean   : 0.03854   Mean   : 0.0
5758
 3rd Qu.: 0.00000   3rd Qu.: 0.00000   3rd Qu.: 0.0000   3rd Qu.: 0.00000   3rd Qu.: 0.0
0000
 Max.   :13.00000   Max.   :59.00000   Max.   :26.0000   Max.   :24.00000   Max.   :20.0
0000
  googleperneg       iosperunc          googleperunc      iphonesentiment
 Min.   : 0.0000   Min.   : 0.00000   Min.   : 0.00000   Min.   :1.000
 1st Qu.: 0.0000   1st Qu.: 0.00000   1st Qu.: 0.00000   1st Qu.:3.000
 Median : 0.0000   Median : 0.00000   Median : 0.00000   Median :4.000
 Mean   : 0.0582   Mean   : 0.03307   Mean   : 0.01611   Mean   :3.295
 3rd Qu.: 0.0000   3rd Qu.: 0.00000   3rd Qu.: 0.00000   3rd Qu.:4.000
 Max.   :25.0000   Max.   :10.00000   Max.   :24.00000   Max.   :4.000
```

Hide

```
str(iphone_smallMatrix_RC)
```

```
'data.frame':   12973 obs. of  59 variables:
 $ iphone          : int  1 1 1 1 1 41 1 1 1 1 ...
 $ samsunggalaxy   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ sonyxperia      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ nokialumina     : int  0 0 0 0 0 0 0 0 0 0 ...
 $ htcphone        : int  0 0 0 0 0 0 0 0 0 0 ...
 $ ios             : int  0 0 0 0 0 6 0 0 0 0 ...
 $ googleandroid   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ iphonecampos    : int  0 0 0 0 0 1 1 0 0 0 ...
 $ samsungcampos   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ sonycampos      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ nokiacampos     : int  0 0 0 0 0 0 0 0 0 0 ...
 $ htccampos       : int  0 0 0 0 0 0 0 0 0 0 ...
 $ iphonecamneg    : int  0 0 0 0 0 3 1 0 0 0 ...
 $ samsungcamneg   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ sonycamneg      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ nokiacamneg     : int  0 0 0 0 0 0 0 0 0 0 ...
 $ htccamneg       : int  0 0 0 0 0 0 0 0 0 0 ...
 $ iphonecamunc    : int  0 0 0 0 0 7 1 0 0 0 ...
 $ samsungcamunc   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ sonycamunc      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ nokiacamunc     : int  0 0 0 0 0 0 0 0 0 0 ...
 $ htccamunc       : int  0 0 0 0 0 0 0 0 0 0 ...
 $ iphonedispos    : int  0 0 0 0 0 1 13 0 0 0 ...
 $ samsungdispos   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ sonydispos      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ nokiadispos     : int  0 0 0 0 0 0 0 0 0 0 ...
 $ htcdispos       : int  0 0 0 0 0 0 0 0 0 0 ...
 $ iphonedisneg    : int  0 0 0 0 0 3 10 0 0 0 ...
 $ samsungdisneg   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ sonydisneg      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ nokiadisneg     : int  0 0 0 0 0 0 0 0 0 0 ...
 $ htcdisneg       : int  0 0 0 0 0 0 0 0 0 0 ...
 $ iphonedisunc    : int  0 0 0 0 0 4 9 0 0 0 ...
 $ samsungdisunc   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ sonydisunc      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ nokiadisunc     : int  0 0 0 0 0 0 0 0 0 0 ...
 $ htcdisunc       : int  0 0 0 0 0 0 0 0 0 0 ...
 $ iphoneperpos    : int  0 1 0 1 1 0 5 3 0 0 ...
 $ samsungperpos   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ sonyperpos      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ nokiaperpos     : int  0 0 0 0 0 0 0 0 0 0 ...
 $ htcperpos       : int  0 0 0 0 0 0 0 0 0 0 ...
 $ iphoneperneg    : int  0 0 0 0 0 0 4 1 0 0 ...
 $ samsungperneg   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ sonyperneg      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ nokiaperneg     : int  0 0 0 0 0 0 0 0 0 0 ...
 $ htcperneg       : int  0 0 0 0 0 0 0 0 0 0 ...
 $ iphoneperunc    : int  0 0 0 1 0 0 5 0 0 0 ...
 $ samsungperunc   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ sonyperunc      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ nokiaperunc     : int  0 0 0 0 0 0 0 0 0 0 ...
 $ htcperunc       : int  0 0 0 0 0 0 0 0 0 0 ...
```

```
 $ iosperpos      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ googleperpos   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ iosperneg      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ googleperneg   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ iosperunc      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ googleperunc   : int  0 0 0 0 0 0 0 0 0 0 ...
 $ iphonesentiment: num  1 1 1 1 1 4 4 1 1 1 ...
```

Hide

```
# make iphonesentiment a factor
iphone_smallMatrix_RC$iphonesentiment <- as.factor(iphone_smallMatrix_RC$iphonesentimen
t)
```

iphone_smallMatrix_RC

Hide

```
# convert variable types, categorical
#iphone_smallMatrix_RCCOR$iphonesentiment <- as.factor(iphone_smallMatrix_RCCOR$iphonese
ntiment)
```

Hide

```


# create 10-fold cross validation fitcontrol
fitControl <- trainControl(method = "cv", number = 10)
```

Train and Test Set:

Hide

```
# Create Train and Test Set for iphone_smallMatrix_RC
# create 75% sample of row indices
in_training <-createDataPartition(iphone_smallMatrix_RC$iphonesentiment, p = .7, list =
FALSE)
# create 75% sample of data and save it to trainData
trainData_iphone_smallMatrix_RC <- iphone_smallMatrix_RC[in_training, ]
 # create 25% sample of data and save it to test_data
testData_iphone_smallMatrix_RC <- iphone_smallMatrix_RC[-in_training, ]
# verify split percentages
nrow(trainData_iphone_smallMatrix_RC) / nrow(iphone_smallMatrix_RC)
```

```
[1] 0.7001465
```

Hide

```
#c5
c5_iphone_smallMatrix_RC <- train(iphonesentiment ~., data = trainData_iphone_smallMatri
x_RC, method = "C5.0",trControl = fitControl)
```

Compare Accuracy on Prediction Results:

Hide

```
#c5
prediction_c5_iphone_smallMatrix_RC <- predict(c5_iphone_smallMatrix_RC, testData_iphone
_smallMatrix_RC)
postResample(prediction_c5_iphone_smallMatrix_RC, testData_iphone_smallMatrix_RC$iphones
entiment)
```

Hide

```
summary(c5_iphone_smallMatrix_RC)
```

```
Call:
(function (x, y, trials = 1, rules = FALSE, weights = NULL, control = C5.0Control(), cos
ts
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
0, 0,
 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 6, 0, 0, 0, 0, 0, 0,
0,


C5.0 [Release 2.07 GPL Edition]     Mon Mar  9 17:18:10 2020
-------------------------------

Class specified by attribute `outcome'

Read 9083 cases (59 attributes) from undefined.data

Decision tree:

iphone <= 0:
:...googleandroid > 0: 2 (40)
:   googleandroid <= 0:
:   :...samsungdisneg <= 0: 1 (596/1)
:       samsungdisneg > 0:
:       :...samsungperneg <= 1: 1 (23/8)
:           samsungperneg > 1: 4 (10/4)
iphone > 0:
:...samsungcamneg > 4: 1 (88/2)
    samsungcamneg <= 4:
    :...googleandroid > 0:
        :...iphoneperpos <= 0: 3 (102/2)
        :   iphoneperpos > 0: 4 (13/3)
        googleandroid <= 0:
        :...sonyxperia > 0:
            :...iphoneperpos <= 0: 1 (148/4)
            :   iphoneperpos > 0: 4 (28/8)
            sonyxperia <= 0:
            :...samsunggalaxy > 0:
                :...iphoneperpos > 0: 4 (15/1)
                :   iphoneperpos <= 0:
                :   :...iphonecamunc > 0:
                :       :...iphonedisunc <= 0: 1 (5)
                :       :   iphonedisunc > 0: 4 (3/1)
                :       iphonecamunc <= 0:
                :       :...iphonedispos <= 0: 3 (137/2)
                :           iphonedispos > 0:
                :           :...iphonedisunc <= 0: 4 (8/3)
                :               iphonedisunc > 0: 3 (3)
                samsunggalaxy <= 0:
                :...htccampos > 0:
                    :...iphone > 2:
                    :   :...iphonecamunc <= 2: 4 (12/1)
                    :   :   iphonecamunc > 2: 3 (12/2)
```

```
:   iphone <= 2:
:   :...iphonedisunc <= 0: 1 (24/3)
:       iphonedisunc > 0:
:       :...iphonecampos <= 1: 3 (12/3)
:           iphonecampos > 1: 1 (8)
htccampos <= 0:
:...iphonedispos <= 0:
    :...iphonedisunc <= 0: 4 (5887/1084)
    :   iphonedisunc > 0:
    :   :...iphoneperunc > 0: 4 (46/2)
    :       iphoneperunc <= 0:
    :       :...iphonedisneg > 0: 4 (35/7)
    :           iphonedisneg <= 0:
    :           :...iphoneperpos > 0: 4 (16/4)
    :               iphoneperpos <= 0:
    :               :...iphone > 2: 4 (6/1)
    :                   iphone <= 2:
    :                   :...iphonecamneg <= 0: 3 (62/7)
    :                       iphonecamneg > 0: 4 (2)
    iphonedispos > 0:
    :...iphoneperpos > 0: 4 (1030/162)
        iphoneperpos <= 0:
        :...iphonedisneg <= 0:
            :...iphone <= 1:
            :   :...iphonecampos <= 0:
            :   :   :...iphoneperunc <= 0: 3 (179/7)
            :   :   :   iphoneperunc > 0: 4 (5/2)
            :   :   iphonecampos > 0:
            :   :   :...iphonedisunc <= 0: 4 (27/5)
            :   :       iphonedisunc > 0: 3 (12/2)
            :   iphone > 1:
            :   :...iphonedisunc <= 0: 4 (51/10)
            :       iphonedisunc > 0:
            :       :...iphonedispos <= 2: 3 (18/4)
            :           iphonedispos > 2: 4 (9/1)
            iphonedisneg > 0:
            :...iphonedispos > 1: 4 (105/14)
                iphonedispos <= 1:
                :...iphonedisunc <= 0: 4 (60/8)
                    iphonedisunc > 0:
                    :...iphone > 1: 4 (200/2)
                        iphone <= 1:
                        :...iphoneperneg > 0: 4 (3/1)
                            iphoneperneg <= 0:
                            :...iphonecamneg > 0: 4 (3)
                                iphonecamneg <= 0: [S1]

SubTree [S1]

iphoneperunc <= 0: 3 (38)
iphoneperunc > 0: 4 (2)


Evaluation on training data (9083 cases):
```

```
        Decision Tree
      ----------------
      Size      Errors

        42 1371(15.1%)   <<


       (a)   (b)   (c)   (d)    <-classified as
      ----  ----  ----  ----
       874           5   768    (a): class 1
              40     4   274    (b): class 2
         4         546   282    (c): class 3
        14          20  6252    (d): class 4


    Attribute usage:

    100.00% iphone
     99.03% googleandroid
     92.63% samsungcamneg
     90.40% sonyxperia
     88.46% samsunggalaxy
     87.46% iphonedispos
     86.58% htccampos
     72.00% iphonedisunc
     25.21% iphoneperpos
      9.17% iphonedisneg
      6.93% samsungdisneg
      4.30% iphoneperunc
      2.68% iphonecampos
      1.98% iphonecamunc
      1.18% iphonecamneg
      0.51% iphoneperneg
      0.36% samsungperneg


Time: 0.1 secs
```