# Divide and Conquer

*Claudius Taylor*

*1/29/2019*

## Useful functions

```r
boolean_cover <- function(df,proportion=0.75){
#given a dataframe, return a boolean vector appropriate for selecting rows of the dataframe.
    n=dim(df)[1] #number of rows in the dataframe
    return(runif(n,0,1) > 1 - proportion)
}

min_max_normalize <- function(x) {
    return((x - min(x)) / (max(x) - min(x)))
}

discretize <- function(x){
#amplify and round
    round(x*10, 0)
}

entropy <- function(df) {
    df %>%
    filter(timestamp==x) %>%
    summarise( ch1_entropy=tryCatch(approx_entropy(ch1,edim=2,r=0.2*sd(ch1),elag=4),
                                    error = function(e) 0),
            ch2_entropy=tryCatch(approx_entropy(ch2,edim=2,r=0.2*sd(ch2),elag=4),
                                    error = function(e) 0),
            ch3_entropy=tryCatch(approx_entropy(ch3,edim=2,r=0.2*sd(ch3),elag=4),
                                    error = function(e) 0),
            ch4_entropy=tryCatch(approx_entropy(ch4,edim=2,r=0.2*sd(ch4),elag=4),
                                    error = function(e) 0)
        )
}

preprocess <- function(df) {
#given raw eeg data return a dataframe with normalized/discrete entropy values for each channel
    entropy(df) %>%
    select(matches('ch')) %>%
    apply(2, min_max_normalize) %>%
    apply(2, discretize) %>%
    as.data.frame()
}

read_eeg <- function(path){
#read a file with no header and give it specific column names
    column_names <- c('timestamp','ch1','ch2','ch3','ch4')
    fread(path, header=FALSE, col.names=column_names)
}
```

## load data

```
setwd("/Users/taylormade/Documents/Labs Data Mining")
S1_eo  <- read_eeg('./Subject1_eo.csv')
S2_eo  <- read_eeg('./Subject2_eo.csv')
S3_eo  <- read_eeg('./Subject3_eo.csv')
S1_ec  <- read_eeg('./Subject1_ec.csv')
S2_ec  <- read_eeg('./Subject2_ec.csv')
S3_ec  <- read_eeg('./Subject3_ec.csv')
S1_med <- read_eeg('./Subject1_med.csv')
S2_med <- read_eeg('./Subject2_med.csv')
S3_med <- read_eeg('./Subject3_med.csv')
S1_vid <- read_eeg('./Subject1_vid.csv')
S2_vid <- read_eeg('./Subject2_vid.csv')
S3_vid <- read_eeg('./Subject3_vid.csv')
```

## Entropy calculation

```
S1_eo_entropy  <- S1_eo %>% mutate(state='eyes-open')
S2_eo_entropy  <- S2_eo  %>% mutate(state='eyes-open')
S3_eo_entropy  <- S3_eo  %>% mutate(state='eyes-open')
S1_ec_entropy  <- S1_ec  %>% mutate(state='eyes-closed')
S2_ec_entropy  <- S2_ec  %>% mutate(state='eyes-closed')
S3_ec_entropy  <- S3_ec  %>% mutate(state='eyes-closed')
S1_med_entropy <- S1_med %>% mutate(state='meditate')
S2_med_entropy <- S2_med %>% mutate(state='meditate')
S3_med_entropy <- S3_med %>% mutate(state='meditate')
S1_vid_entropy <- S1_vid %>% mutate(state='video')
S2_vid_entropy <- S2_vid %>% mutate(state='video')
S3_vid_entropy <- S3_vid %>% mutate(state='video')
```

## Sample 75% of each dataset

```
S1_ec_sample     <- boolean_cover(S1_ec_entropy)
S2_ec_sample     <- boolean_cover(S2_ec_entropy)
S3_ec_sample     <- boolean_cover(S3_ec_entropy)
S1_eo_sample     <- boolean_cover(S1_eo_entropy)
S2_eo_sample     <- boolean_cover(S2_eo_entropy)
S3_eo_sample     <- boolean_cover(S3_eo_entropy)
S1_med_sample    <- boolean_cover(S1_med_entropy)
S2_med_sample    <- boolean_cover(S2_med_entropy)
S3_med_sample    <- boolean_cover(S3_med_entropy)
S1_vid_sample    <- boolean_cover(S1_vid_entropy)
S2_vid_sample    <- boolean_cover(S2_vid_entropy)
S3_vid_sample    <- boolean_cover(S3_vid_entropy)

S1_ec_train <- S1_ec_entropy[S1_ec_sample,]
S2_ec_train <- S2_ec_entropy[S2_ec_sample,]
S3_ec_train <- S3_ec_entropy[S3_ec_sample,]
S1_eo_train <- S1_eo_entropy[S1_eo_sample,]
```

```r
S2_eo_train <- S2_eo_entropy[S2_eo_sample,]
S3_eo_train <- S3_eo_entropy[S3_eo_sample,]
S1_med_train <- S1_med_entropy[S1_med_sample,]
S2_med_train <- S2_med_entropy[S2_med_sample,]
S3_med_train <- S3_med_entropy[S3_med_sample,]
S1_vid_train <- S1_vid_entropy[S1_vid_sample,]
S2_vid_train <- S2_vid_entropy[S2_vid_sample,]
S3_vid_train <- S3_vid_entropy[S3_vid_sample,]

S1_ec_test <- S1_ec_entropy[!S1_ec_sample,]
S2_ec_test <- S2_ec_entropy[!S2_ec_sample,]
S3_ec_test <- S3_ec_entropy[!S3_ec_sample,]
S1_eo_test <- S1_eo_entropy[!S1_eo_sample,]
S2_eo_test <- S2_eo_entropy[!S2_eo_sample,]
S3_eo_test <- S3_eo_entropy[!S3_eo_sample,]
S1_med_test <- S1_med_entropy[!S1_med_sample,]
S2_med_test <- S2_med_entropy[!S2_med_sample,]
S3_med_test <- S3_med_entropy[!S3_med_sample,]
S1_vid_test <- S1_vid_entropy[!S1_vid_sample,]
S2_vid_test <- S2_vid_entropy[!S2_vid_sample,]
S3_vid_test <- S3_vid_entropy[!S3_vid_sample,]
```

## Create necessary datasets

```r
S1_train <- rbind(S1_ec_train, S1_eo_train, S1_med_train, S1_vid_train) %>%
  mutate(state=as.factor(state))
S2_train <- rbind(S2_ec_train, S2_eo_train, S2_med_train, S2_vid_train) %>%
  mutate(state=as.factor(state))
S3_train <- rbind(S3_ec_train, S3_eo_train, S3_med_train, S3_vid_train) %>%
  mutate(state=as.factor(state))
S1S2_train <- rbind(S1_train, S2_train)
S1S3_train <- rbind(S1_train, S3_train)
S2S3_train <- rbind(S2_train, S3_train)
S1S2S3_train <- rbind(S1_train, S2_train, S3_train)

S1_test <- rbind(S1_ec_test, S1_eo_test, S1_med_test, S1_vid_test) %>%
  mutate(state=as.factor(state))
S2_test <- rbind(S2_ec_test, S2_eo_test, S2_med_test, S2_vid_test) %>%
  mutate(state=as.factor(state))
S3_test <- rbind(S3_ec_test, S3_eo_test, S3_med_test, S3_vid_test) %>%
  mutate(state=as.factor(state))
S1S2_test <- rbind(S1_test, S2_test)
S1S3_test <- rbind(S1_test, S3_test)
S2S3_test <- rbind(S2_test, S3_test)
S1S2S3_test <- rbind(S1_test, S2_test, S3_test)
```

## Naives Bayes, Decision Tree

```r
useNaiveBayes <- function(train, test){
#given two datafames, use naiveBayes with the correct arguments to return a prediction for test$state
```

```r
    model <- naiveBayes( state~., data=train, laplace = 1)
    predict( model, test )
}

S1_NaiveBayes <- useNaiveBayes(S1_train, S1_test)
S2_NaiveBayes <- useNaiveBayes(S2_train, S2_test)
S3_NaiveBayes <- useNaiveBayes(S3_train, S3_test)
S1S2_NaiveBayes <- useNaiveBayes(S1S2_train, S1S2_test)
S1S3_NaiveBayes <- useNaiveBayes(S1S3_train, S1S3_test)
S2S3_NaiveBayes <- useNaiveBayes(S2S3_train, S2S3_test)
S1S2S3_NaiveBayes <- useNaiveBayes(S1S2S3_train, S1S2S3_test)

useTree <- function(train, test){
#given two dataframes, use C5.0 with the correct arguments to return a prediction of test$state
    model <- C5.0( state~., data=train, trials = 10 )
    predict( model, test, type = 'class')
}

S1_Tree <- useTree(S1_train, S1_test)
S2_Tree <- useTree(S2_train, S2_test)
S3_Tree <- useTree(S3_train, S3_test)
S1S2_Tree <- useTree(S1S2_train, S1S2_test)
S1S3_Tree <- useTree(S1S3_train, S1S3_test)
S2S3_Tree <- useTree(S2S3_train, S2S3_test)
S1S2S3_Tree <- useTree(S1S2S3_train, S1S2S3_test)
```

## Crosstables

```r
#compare predicted to actual for every model used.
CrossTable(S1_NaiveBayes,S1_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |           N / Col Total |
## |-------------------------|
##
##
## Total Observations in Table:  36334
##
##
##               | S1_test$state
## S1_NaiveBayes | eyes-closed |    eyes-open |    meditate |       video |   Row Total |
## --------------|-------------|-------------|-------------|-------------|-------------|
##   eyes-closed |         932 |           1 |           0 |           0 |         933 |
##               |       0.974 |       0.000 |       0.000 |       0.000 |             |
## --------------|-------------|-------------|-------------|-------------|-------------|
##     eyes-open |           0 |        6523 |        5856 |         412 |       12791 |
##               |       0.000 |       0.599 |       0.326 |       0.063 |             |
## --------------|-------------|-------------|-------------|-------------|-------------|
```

4

```
##      meditate |           20 |        3227 |       12094 |         109 |       15450 |
##               |        0.021 |       0.296 |       0.674 |       0.017 |             |
## --------------|--------------|-------------|-------------|-------------|-------------|
##         video |            5 |        1143 |           0 |        6012 |        7160 |
##               |        0.005 |       0.105 |       0.000 |       0.920 |             |
## --------------|--------------|-------------|-------------|-------------|-------------|
##  Column Total |          957 |       10894 |       17950 |        6533 |       36334 |
##               |        0.026 |       0.300 |       0.494 |       0.180 |             |
## --------------|--------------|-------------|-------------|-------------|-------------|
##
##
```

**CrossTable**(S2_NaiveBayes,S2_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |          N / Col Total |
## |-------------------------|
##
##
## Total Observations in Table:  102851
##
##
##                | S2_test$state
## S2_NaiveBayes | eyes-closed |   eyes-open |    meditate |       video |   Row Total |
## --------------|-------------|-------------|-------------|-------------|-------------|
##   eyes-closed |        4633 |        1436 |           0 |           0 |        6069 |
##               |       0.930 |       0.027 |       0.000 |       0.000 |             |
## --------------|-------------|-------------|-------------|-------------|-------------|
##     eyes-open |         267 |       30529 |        1355 |          54 |       32205 |
##               |       0.054 |       0.575 |       0.034 |       0.010 |             |
## --------------|-------------|-------------|-------------|-------------|-------------|
##      meditate |          82 |           1 |       33243 |         459 |       33785 |
##               |       0.016 |       0.000 |       0.840 |       0.089 |             |
## --------------|-------------|-------------|-------------|-------------|-------------|
##         video |           0 |       21142 |        4992 |        4658 |       30792 |
##               |       0.000 |       0.398 |       0.126 |       0.901 |             |
## --------------|-------------|-------------|-------------|-------------|-------------|
##  Column Total |        4982 |       53108 |       39590 |        5171 |      102851 |
##               |       0.048 |       0.516 |       0.385 |       0.050 |             |
## --------------|-------------|-------------|-------------|-------------|-------------|
##
##
```

**CrossTable**(S3_NaiveBayes,S3_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |          N / Col Total |
```

```
## |-------------------------|
##
##
## Total Observations in Table:  40575
##
##
##               | S3_test$state
## S3_NaiveBayes | eyes-closed |   eyes-open |    meditate |       video |   Row Total |
## -------------|-------------|-------------|-------------|-------------|-------------|
##   eyes-closed |       19360 |        1226 |          16 |           0 |       20602 |
##               |       0.983 |       0.113 |       0.002 |       0.000 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##     eyes-open |         236 |        9568 |           0 |           0 |        9804 |
##               |       0.012 |       0.885 |       0.000 |       0.000 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##      meditate |           1 |           0 |        8220 |         152 |        8373 |
##               |       0.000 |       0.000 |       0.995 |       0.084 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##         video |          92 |          12 |          29 |        1663 |        1796 |
##               |       0.005 |       0.001 |       0.004 |       0.916 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##  Column Total |       19689 |       10806 |        8265 |        1815 |       40575 |
##               |       0.485 |       0.266 |       0.204 |       0.045 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##
##
```

```r
CrossTable(S1S2_NaiveBayes,S1S2_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |          N / Col Total |
## |-------------------------|
##
##
## Total Observations in Table:  139185
##
##
##                 | S1S2_test$state
## S1S2_NaiveBayes | eyes-closed |   eyes-open |    meditate |       video |   Row Total |
## ---------------|-------------|-------------|-------------|-------------|-------------|
##     eyes-closed |        3748 |       14961 |       24963 |        3062 |       46734 |
##                 |       0.631 |       0.234 |       0.434 |       0.262 |             |
## ---------------|-------------|-------------|-------------|-------------|-------------|
##       eyes-open |         223 |       14340 |        4316 |         471 |       19350 |
##                 |       0.038 |       0.224 |       0.075 |       0.040 |             |
## ---------------|-------------|-------------|-------------|-------------|-------------|
##        meditate |        1014 |       23791 |       10838 |        1644 |       37287 |
##                 |       0.171 |       0.372 |       0.188 |       0.140 |             |
## ---------------|-------------|-------------|-------------|-------------|-------------|
##           video |         954 |       10910 |       17423 |        6527 |       35814 |
##                 |       0.161 |       0.170 |       0.303 |       0.558 |             |
```

```
## ----------------|-------------|-------------|-------------|-------------|-------------|
##     Column Total |        5939 |       64002 |       57540 |       11704 |      139185 |
##                  |       0.043 |       0.460 |       0.413 |       0.084 |             |
## ----------------|-------------|-------------|-------------|-------------|-------------|
##
##
```

**CrossTable**(S1S3_NaiveBayes,S1S3_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)

```
##
##
##     Cell Contents
## |-------------------------|
## |                       N |
## |            N / Col Total |
## |-------------------------|
##
##
## Total Observations in Table:  76909
##
##
##                  | S1S3_test$state
## S1S3_NaiveBayes  | eyes-closed |    eyes-open |     meditate |        video |    Row Total |
## ----------------|-------------|-------------|-------------|-------------|-------------|
##      eyes-closed |       11258 |        5873 |        4662 |         592 |       22385 |
##                  |       0.545 |       0.271 |       0.178 |       0.071 |             |
## ----------------|-------------|-------------|-------------|-------------|-------------|
##        eyes-open |         396 |        1663 |         743 |          61 |        2863 |
##                  |       0.019 |       0.077 |       0.028 |       0.007 |             |
## ----------------|-------------|-------------|-------------|-------------|-------------|
##         meditate |        8467 |       10209 |       15369 |        5150 |       39195 |
##                  |       0.410 |       0.470 |       0.586 |       0.617 |             |
## ----------------|-------------|-------------|-------------|-------------|-------------|
##            video |         525 |        3955 |        5441 |        2545 |       12466 |
##                  |       0.025 |       0.182 |       0.208 |       0.305 |             |
## ----------------|-------------|-------------|-------------|-------------|-------------|
##     Column Total |       20646 |       21700 |       26215 |        8348 |       76909 |
##                  |       0.268 |       0.282 |       0.341 |       0.109 |             |
## ----------------|-------------|-------------|-------------|-------------|-------------|
##
##
```

**CrossTable**(S2S3_NaiveBayes,S2S3_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)

```
##
##
##     Cell Contents
## |-------------------------|
## |                       N |
## |            N / Col Total |
## |-------------------------|
##
##
## Total Observations in Table:  143426
##
```

```
## 
##                 | S2S3_test$state
## S2S3_NaiveBayes | eyes-closed |   eyes-open |    meditate |       video |   Row Total |
## ----------------|-------------|-------------|-------------|-------------|-------------|
##     eyes-closed |       17764 |        8076 |        8800 |        1985 |       36625 |
##                 |       0.720 |       0.126 |       0.184 |       0.284 |             |
## ----------------|-------------|-------------|-------------|-------------|-------------|
##       eyes-open |         886 |       22976 |        3725 |         348 |       27935 |
##                 |       0.036 |       0.359 |       0.078 |       0.050 |             |
## ----------------|-------------|-------------|-------------|-------------|-------------|
##        meditate |        6020 |       32862 |       35330 |        4653 |       78865 |
##                 |       0.244 |       0.514 |       0.738 |       0.666 |             |
## ----------------|-------------|-------------|-------------|-------------|-------------|
##           video |           1 |           0 |           0 |           0 |           1 |
##                 |       0.000 |       0.000 |       0.000 |       0.000 |             |
## ----------------|-------------|-------------|-------------|-------------|-------------|
##    Column Total |       24671 |       63914 |       47855 |        6986 |      143426 |
##                 |       0.172 |       0.446 |       0.334 |       0.049 |             |
## ----------------|-------------|-------------|-------------|-------------|-------------|
## 
## 
```

**CrossTable**(S1S2S3_NaiveBayes,S1S2S3_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)

```
## 
## 
##    Cell Contents
## |-------------------------|
## |                       N |
## |           N / Col Total |
## |-------------------------|
## 
## 
## Total Observations in Table:  179760 
## 
## 
##                   | S1S2S3_test$state
## S1S2S3_NaiveBayes | eyes-closed |   eyes-open |    meditate |       video |   Row Total |
## ------------------|-------------|-------------|-------------|-------------|-------------|
##       eyes-closed |       20138 |       26231 |       27551 |        4173 |       78093 |
##                   |       0.786 |       0.351 |       0.419 |       0.309 |             |
## ------------------|-------------|-------------|-------------|-------------|-------------|
##         eyes-open |        1947 |       15497 |        4207 |         418 |       22069 |
##                   |       0.076 |       0.207 |       0.064 |       0.031 |             |
## ------------------|-------------|-------------|-------------|-------------|-------------|
##          meditate |        2591 |       22257 |       16690 |        2404 |       43942 |
##                   |       0.101 |       0.298 |       0.254 |       0.178 |             |
## ------------------|-------------|-------------|-------------|-------------|-------------|
##             video |         952 |       10823 |       17357 |        6524 |       35656 |
##                   |       0.037 |       0.145 |       0.264 |       0.483 |             |
## ------------------|-------------|-------------|-------------|-------------|-------------|
##      Column Total |       25628 |       74808 |       65805 |       13519 |      179760 |
##                   |       0.143 |       0.416 |       0.366 |       0.075 |             |
## ------------------|-------------|-------------|-------------|-------------|-------------|
## 
```

```
##
```
```
CrossTable(S1_Tree,S1_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |           N / Col Total |
## |-------------------------|
##
##
## Total Observations in Table:  36334
##
##
##             | S1_test$state
##      S1_Tree | eyes-closed |   eyes-open |    meditate |       video |   Row Total |
## -------------|-------------|-------------|-------------|-------------|-------------|
##    eyes-open |           0 |        1077 |         794 |         303 |        2174 |
##              |       0.000 |       0.099 |       0.044 |       0.046 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##     meditate |           0 |        9526 |       17156 |        1018 |       27700 |
##              |       0.000 |       0.874 |       0.956 |       0.156 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##        video |         957 |         291 |           0 |        5212 |        6460 |
##              |       1.000 |       0.027 |       0.000 |       0.798 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
## Column Total |         957 |       10894 |       17950 |        6533 |       36334 |
##              |       0.026 |       0.300 |       0.494 |       0.180 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##
##
```
```
CrossTable(S2_Tree,S2_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)
```

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |           N / Col Total |
## |-------------------------|
##
##
## Total Observations in Table:  102851
##
##
##             | S2_test$state
##      S2_Tree | eyes-closed |   eyes-open |    meditate |       video |   Row Total |
## -------------|-------------|-------------|-------------|-------------|-------------|
##  eyes-closed |        4473 |           0 |           0 |           0 |        4473 |
##              |       0.898 |       0.000 |       0.000 |       0.000 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##    eyes-open |         509 |       53108 |           0 |           0 |       53617 |
```

```
##              |       0.102 |       1.000 |       0.000 |       0.000 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
##     meditate |           0 |           0 |       39571 |        5171 |       44742 |
##              |       0.000 |       0.000 |       1.000 |       1.000 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
##        video |           0 |           0 |          19 |           0 |          19 |
##              |       0.000 |       0.000 |       0.000 |       0.000 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
## Column Total |        4982 |       53108 |       39590 |        5171 |      102851 |
##              |       0.048 |       0.516 |       0.385 |       0.050 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
##
##
```

**CrossTable**(S3_Tree,S3_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |           N / Col Total |
## |-------------------------|
##
##
## Total Observations in Table:  40575
##
##
##              | S3_test$state
##      S3_Tree | eyes-closed |    eyes-open |    meditate |       video |   Row Total |
## ------------|-------------|-------------|-------------|-------------|-------------|
##  eyes-closed |       18895 |        1240 |           0 |           0 |       20135 |
##              |       0.960 |       0.115 |       0.000 |       0.000 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
##    eyes-open |         794 |        9566 |           0 |           0 |       10360 |
##              |       0.040 |       0.885 |       0.000 |       0.000 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
##     meditate |           0 |           0 |        7732 |         493 |        8225 |
##              |       0.000 |       0.000 |       0.936 |       0.272 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
##        video |           0 |           0 |         533 |        1322 |        1855 |
##              |       0.000 |       0.000 |       0.064 |       0.728 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
## Column Total |       19689 |       10806 |        8265 |        1815 |       40575 |
##              |       0.485 |       0.266 |       0.204 |       0.045 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
##
##
```

**CrossTable**(S1S2_Tree,S1S2_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)

```
##
##
##    Cell Contents
## |-------------------------|
```

```
## |                       N |
## |         N / Col Total |
## |-----------------------|
##
##
## Total Observations in Table:  139185
##
##
##                | S1S2_test$state
##     S1S2_Tree | eyes-closed |   eyes-open |    meditate |       video |   Row Total |
## -------------|-------------|-------------|-------------|-------------|-------------|
##   eyes-closed |        5430 |           0 |           0 |           0 |        5430 |
##              |       0.914 |       0.000 |       0.000 |       0.000 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##     eyes-open |         509 |       53108 |           0 |           0 |       53617 |
##              |       0.086 |       0.830 |       0.000 |       0.000 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##      meditate |           0 |       10603 |       57540 |        6492 |       74635 |
##              |       0.000 |       0.166 |       1.000 |       0.555 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##         video |           0 |         291 |           0 |        5212 |        5503 |
##              |       0.000 |       0.005 |       0.000 |       0.445 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
## Column Total |        5939 |       64002 |       57540 |       11704 |      139185 |
##              |       0.043 |       0.460 |       0.413 |       0.084 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##
##
```

**CrossTable**(S1S3_Tree,S1S3_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)

```
##
##
##    Cell Contents
## |-----------------------|
## |                     N |
## |         N / Col Total |
## |-----------------------|
##
##
## Total Observations in Table:  76909
##
##
##                | S1S3_test$state
##     S1S3_Tree | eyes-closed |   eyes-open |    meditate |       video |   Row Total |
## -------------|-------------|-------------|-------------|-------------|-------------|
##   eyes-closed |       20079 |        1434 |           0 |           0 |       21513 |
##              |       0.973 |       0.066 |       0.000 |       0.000 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##     eyes-open |         567 |       11286 |        1370 |         457 |       13680 |
##              |       0.027 |       0.520 |       0.052 |       0.055 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##      meditate |           0 |        8689 |       24235 |        1176 |       34100 |
##              |       0.000 |       0.400 |       0.924 |       0.141 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
```

```
##       video |           0 |         291 |         610 |        6715 |        7616 |
##             |       0.000 |       0.013 |       0.023 |       0.804 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
## Column Total |      20646 |       21700 |       26215 |        8348 |       76909 |
##             |       0.268 |       0.282 |       0.341 |       0.109 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
##
##
```

**CrossTable**(S2S3_Tree,S2S3_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |            N / Col Total |
## |-------------------------|
##
##
## Total Observations in Table:  143426
##
##
##              | S2S3_test$state
##    S2S3_Tree | eyes-closed |   eyes-open |    meditate |       video |   Row Total |
## ------------|-------------|-------------|-------------|-------------|-------------|
##  eyes-closed |       23808 |        1369 |           0 |           0 |       25177 |
##             |       0.965 |       0.021 |       0.000 |       0.000 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
##    eyes-open |         863 |       62545 |           0 |           0 |       63408 |
##             |       0.035 |       0.979 |       0.000 |       0.000 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
##     meditate |           0 |           0 |       47146 |        5428 |       52574 |
##             |       0.000 |       0.000 |       0.985 |       0.777 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
##        video |           0 |           0 |         709 |        1558 |        2267 |
##             |       0.000 |       0.000 |       0.015 |       0.223 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
## Column Total |      24671 |       63914 |       47855 |        6986 |      143426 |
##             |       0.172 |       0.446 |       0.334 |       0.049 |             |
## ------------|-------------|-------------|-------------|-------------|-------------|
##
##
```

**CrossTable**(S1S2S3_Tree,S1S2S3_test$state,prop.r = FALSE,prop.t = FALSE,prop.chisq = FALSE)

```
##
##
##    Cell Contents
## |-------------------------|
## |                       N |
## |            N / Col Total |
## |-------------------------|
##
##
```

```
## Total Observations in Table:  179760
##
##
##              | S1S2S3_test$state
##  S1S2S3_Tree | eyes-closed |  eyes-open  |   meditate  |      video  |  Row Total  |
## -------------|-------------|-------------|-------------|-------------|-------------|
##  eyes-closed |       24974 |        1608 |           0 |           0 |       26582 |
##              |       0.974 |       0.021 |       0.000 |       0.000 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##    eyes-open |         654 |       63044 |         464 |         267 |       64429 |
##              |       0.026 |       0.843 |       0.007 |       0.020 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##     meditate |           0 |        9865 |       64552 |        6376 |       80793 |
##              |       0.000 |       0.132 |       0.981 |       0.472 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##        video |           0 |         291 |         789 |        6876 |        7956 |
##              |       0.000 |       0.004 |       0.012 |       0.509 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
## Column Total |       25628 |       74808 |       65805 |       13519 |      179760 |
##              |       0.143 |       0.416 |       0.366 |       0.075 |             |
## -------------|-------------|-------------|-------------|-------------|-------------|
##
##
```

## Classification

```r
misclassification <- function(predicted, actual) {
    #given two vectors, return the complement of accuracy
    #accuracy is defined as TruePositives/All
    accuracy <- mean(predicted==actual)
    return(1-accuracy)
}

S1_NaiveBayes_mc <- misclassification(S1_NaiveBayes, S1_test$state)
S2_NaiveBayes_mc <- misclassification(S2_NaiveBayes, S2_test$state)
S3_NaiveBayes_mc <- misclassification(S3_NaiveBayes, S3_test$state)
S1S2_NaiveBayes_mc <- misclassification(S1S2_NaiveBayes, S1S2_test$state)
S1S3_NaiveBayes_mc <- misclassification(S1S3_NaiveBayes, S1S3_test$state)
S2S3_NaiveBayes_mc <- misclassification(S2S3_NaiveBayes, S2S3_test$state)
S1S2S3_NaiveBayes_mc <- misclassification(S1S2S3_NaiveBayes, S1S2S3_test$state)

S1_Tree_mc <- misclassification(S1_Tree, S1_test$state)
S2_Tree_mc <- misclassification(S2_Tree, S2_test$state)
S3_Tree_mc <- misclassification(S3_Tree, S3_test$state)
S1S2_Tree_mc <- misclassification(S1S2_Tree, S1S2_test$state)
S1S3_Tree_mc <- misclassification(S1S3_Tree, S1S3_test$state)
S2S3_Tree_mc <- misclassification(S2S3_Tree, S2S3_test$state)
S1S2S3_Tree_mc <- misclassification(S1S2S3_Tree, S1S2S3_test$state)

summary_data <- data.frame( nsubjects=c(1,1,1,2,2,2,3,1,1,1,2,2,2,3),
                            algorithm=c(rep('naiveBayes',7), rep('decisionTree',7)),
                            miscl_rate=c(S1_NaiveBayes_mc, S2_NaiveBayes_mc, S3_NaiveBayes_mc,
```

```
                                      S1S2_NaiveBayes_mc, S1S3_NaiveBayes_mc, S2S3_NaiveBayes_mc,
                                      S1S2S3_NaiveBayes_mc, S1_Tree_mc, S2_Tree_mc, S3_Tree_mc,
                                      S1S2_Tree_mc, S1S3_Tree_mc, S2S3_Tree_mc, S1S2S3_Tree_mc)
                        )

ggplot(summary_data, aes( x = nsubjects, y = miscl_rate, color = algorithm )) +
  geom_point() +
  geom_smooth()
```

## `geom_smooth()` using method = 'loess' and formula 'y ~ x'