

shredCVL ML Notebook

Using the template of the machine learning module in datacamp.

```
## [1] "forcats"      "stringr"      "dplyr"        "purrr"
## [5] "readr"        "tidyr"        "tibble"       "tidyverse"
## [9] "rpart"        "caretEnsemble" "C50"          "caTools"
## [13] "mlbench"      "caret"        "ggplot2"      "lattice"
## [17] "e1071"        "ranger"       "stats"        "graphics"
## [21] "grDevices"    "utils"        "datasets"     "methods"
## [25] "base"

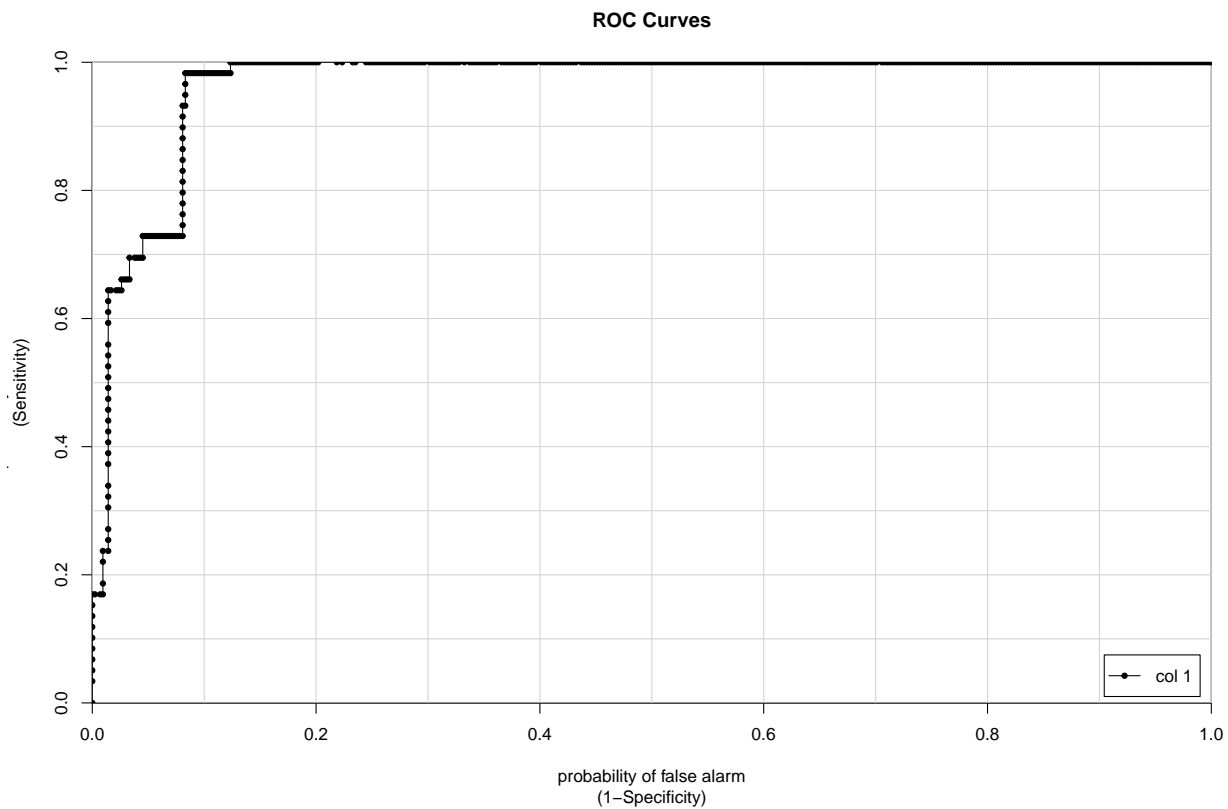
##
##      0      1
## 2107  294

## 'data.frame':  1802 obs. of  9 variables:
## $ statusesCount      : num  2646 7941 61939 61939 15338 ...
## $ friendsCount       : num  392 104 395 395 432 95 378 608 608 883 ...
## $ followersCount     : num  209 144 6937 6937 537 ...
## $ listedCount        : num  0 0 13 13 3 351 0 21 21 4 ...
## $ acct_age           : num  499010 2314876 1293856 1293856 1917603 ...
## $ langDiv            : num  0.872 0.878 0.851 0.851 0.915 ...
## $ mean_time_betwn_tweets: num  188.6 291.5 20.9 20.9 125 ...
## $ mCount             : int   1 1 2 2 1 2 2 2 2 2 ...
## $ bot                : Factor w/ 2 levels "0","1": 2 2 2 2 2 2 2 2 2 2 ...

##
##      0      1
## 526  73

##
##      0      1
## 1581  221

## [1] 0.2523
```



```
##          [,1]
## 0 vs. 1 0.9677
## [1] 0.2585

## + Fold01: parameter=none
## - Fold01: parameter=none
## + Fold02: parameter=none
## - Fold02: parameter=none
## + Fold03: parameter=none
## - Fold03: parameter=none
## + Fold04: parameter=none
## - Fold04: parameter=none
## + Fold05: parameter=none
## - Fold05: parameter=none
## + Fold06: parameter=none
## - Fold06: parameter=none
## + Fold07: parameter=none
## - Fold07: parameter=none
## + Fold08: parameter=none
## - Fold08: parameter=none
## + Fold09: parameter=none
## - Fold09: parameter=none
## + Fold10: parameter=none
## - Fold10: parameter=none
## Aggregating results
## Fitting final model on full training set
## Generalized Linear Model
```

```

##
## 2401 samples
##      8 predictor
##      2 classes: '0', '1'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 2160, 2160, 2162, 2161, 2161, 2161, ...
## Resampling results:
##
##   Accuracy   Kappa
##   0.9967     0.9847

## + Fold1: parameter=none
## - Fold1: parameter=none
## + Fold2: parameter=none
## - Fold2: parameter=none
## + Fold3: parameter=none
## - Fold3: parameter=none
## + Fold4: parameter=none
## - Fold4: parameter=none
## + Fold5: parameter=none
## - Fold5: parameter=none
## Aggregating results
## Fitting final model on full training set

## Generalized Linear Model
##
## 2401 samples
##      8 predictor
##      2 classes: '0', '1'
##
## No pre-processing
## Resampling: Cross-Validated (5 fold)
## Summary of sample sizes: 1921, 1921, 1920, 1921, 1921
## Resampling results:
##
##   Accuracy   Kappa
##   0.995     0.977

## + Fold1: parameter=none
## - Fold1: parameter=none
## + Fold2: parameter=none
## - Fold2: parameter=none
## + Fold3: parameter=none
## - Fold3: parameter=none
## + Fold4: parameter=none
## - Fold4: parameter=none
## + Fold5: parameter=none
## - Fold5: parameter=none
## Aggregating results
## Fitting final model on full training set

## Generalized Linear Model
##
## 2401 samples

```



```

## [1463] 1 0 1 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [1497] 0 0 0 0 0 0 1 0 0 1 0 1 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
## [1531] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [1565] 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [1599] 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0 0 0 0 1
## [1633] 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
## [1667] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0
## [1701] 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [1735] 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0
## [1769] 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 0 0 0 0 0 1 0
## [1803] 0 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [1837] 0 0 1 0 1 0 0 0 1 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0
## [1871] 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [1905] 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0
## [1939] 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0
## [1973] 0 0 0 0 0 1 0 0 0 0 0 0 1 1 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0 0 0 0 1 0 0
## [2007] 0 0 0 0 0 0 0 1 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0
## [2041] 0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0
## [2075] 0 0 1 0 0 0 1 0 1 0 0 0 0 0 0 0 1 1 0 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0 0
## [2109] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [2143] 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 1 0 1
## [2177] 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 1
## [2211] 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [2245] 0 0 0 0 1 1 1 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [2279] 1 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
## [2313] 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0
## [2347] 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 1 0 0 0 0 0 0
## [2381] 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
## Levels: 0 1

##           [,1]
## 0 vs. 1      1

## Confusion Matrix and Statistics
##
##           Reference
## Prediction  0    1
##           0 420    0
##           1    1  59
##
##           Accuracy : 0.998
##           95% CI : (0.988, 1)
##           No Information Rate : 0.877
##           P-Value [Acc > NIR] : <2e-16
##
##           Kappa : 0.99
##           McNemar's Test P-Value : 1
##
##           Sensitivity : 0.998
##           Specificity : 1.000
##           Pos Pred Value : 1.000
##           Neg Pred Value : 0.983
##           Prevalence : 0.877
##           Detection Rate : 0.875
##           Detection Prevalence : 0.875

```

```

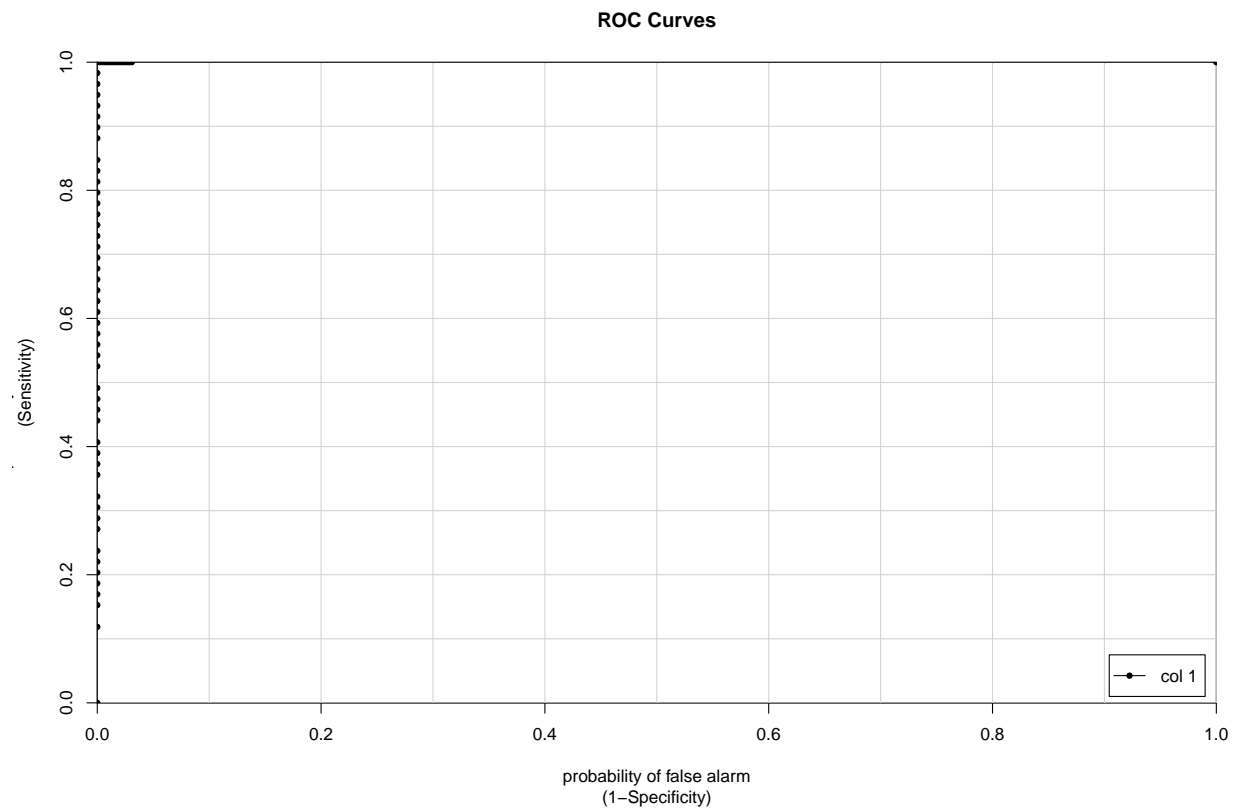
##          Balanced Accuracy : 0.999
##
##          'Positive' Class : 0
##

## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##           0 420    0
##           1    1  59
##
##              Accuracy : 0.998
##              95% CI : (0.988, 1)
##          No Information Rate : 0.877
##          P-Value [Acc > NIR] : <2e-16
##
##              Kappa : 0.99
##  McNemar's Test P-Value : 1
##
##          Sensitivity : 0.998
##          Specificity : 1.000
##          Pos Pred Value : 1.000
##          Neg Pred Value : 0.983
##          Prevalence : 0.877
##          Detection Rate : 0.875
##          Detection Prevalence : 0.875
##          Balanced Accuracy : 0.999
##
##          'Positive' Class : 0
##

## Confusion Matrix and Statistics
##
##              Reference
## Prediction    0    1
##           0 420    0
##           1    1  59
##
##              Accuracy : 0.998
##              95% CI : (0.988, 1)
##          No Information Rate : 0.877
##          P-Value [Acc > NIR] : <2e-16
##
##              Kappa : 0.99
##  McNemar's Test P-Value : 1
##
##          Sensitivity : 0.998
##          Specificity : 1.000
##          Pos Pred Value : 1.000
##          Neg Pred Value : 0.983
##          Prevalence : 0.877
##          Detection Rate : 0.875
##          Detection Prevalence : 0.875
##          Balanced Accuracy : 0.999

```

```
##
##      'Positive' Class : 0
##
```



```
##      [,1]
## 0 vs. 1    1

## + Fold01: parameter=none
## - Fold01: parameter=none
## + Fold02: parameter=none
## - Fold02: parameter=none
## + Fold03: parameter=none
## - Fold03: parameter=none
## + Fold04: parameter=none
## - Fold04: parameter=none
## + Fold05: parameter=none
## - Fold05: parameter=none
## + Fold06: parameter=none
## - Fold06: parameter=none
## + Fold07: parameter=none
## - Fold07: parameter=none
## + Fold08: parameter=none
## - Fold08: parameter=none
## + Fold09: parameter=none
## - Fold09: parameter=none
## + Fold10: parameter=none
## - Fold10: parameter=none
## Aggregating results
```

```

## Fitting final model on full training set

## Generalized Linear Model
##
## 2401 samples
##    8 predictor
##    2 classes: 'X0', 'X1'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 2162, 2161, 2161, 2160, 2160, 2161, ...
## Resampling results:
##
##      ROC      Sens      Spec
## 0.9988 0.9957 0.9864

## + Fold1: mtry=2, min.node.size=1, splitrule=gini
## - Fold1: mtry=2, min.node.size=1, splitrule=gini
## + Fold1: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold1: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold2: mtry=2, min.node.size=1, splitrule=gini
## - Fold2: mtry=2, min.node.size=1, splitrule=gini
## + Fold2: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold2: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold3: mtry=2, min.node.size=1, splitrule=gini
## - Fold3: mtry=2, min.node.size=1, splitrule=gini
## + Fold3: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold3: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold4: mtry=2, min.node.size=1, splitrule=gini
## - Fold4: mtry=2, min.node.size=1, splitrule=gini
## + Fold4: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold4: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold5: mtry=2, min.node.size=1, splitrule=gini
## - Fold5: mtry=2, min.node.size=1, splitrule=gini
## + Fold5: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold5: mtry=2, min.node.size=1, splitrule=extratrees
## Aggregating results
## Selecting tuning parameters
## Fitting mtry = 2, splitrule = gini, min.node.size = 1 on full training set

## Random Forest
##
## 2401 samples
##    8 predictor
##    2 classes: '0', '1'
##
## No pre-processing
## Resampling: Cross-Validated (5 fold)
## Summary of sample sizes: 1921, 1921, 1920, 1921, 1921
## Resampling results across tuning parameters:
##
##      splitrule      Accuracy      Kappa
##      gini          0.9979      0.9905
##      extratrees    0.9950      0.9765
##

```



```

## Tuning parameter 'mtry' was held constant at a value of 2
## Tuning
## parameter 'min.node.size' was held constant at a value of 1
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were mtry = 2, splitrule = gini
## and min.node.size = 1.

## + Fold1: mtry=2, min.node.size=1, splitrule=gini
## - Fold1: mtry=2, min.node.size=1, splitrule=gini
## + Fold1: mtry=5, min.node.size=1, splitrule=gini
## - Fold1: mtry=5, min.node.size=1, splitrule=gini
## + Fold1: mtry=8, min.node.size=1, splitrule=gini
## - Fold1: mtry=8, min.node.size=1, splitrule=gini
## + Fold1: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold1: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold1: mtry=5, min.node.size=1, splitrule=extratrees
## - Fold1: mtry=5, min.node.size=1, splitrule=extratrees
## + Fold1: mtry=8, min.node.size=1, splitrule=extratrees
## - Fold1: mtry=8, min.node.size=1, splitrule=extratrees
## + Fold2: mtry=2, min.node.size=1, splitrule=gini
## - Fold2: mtry=2, min.node.size=1, splitrule=gini
## + Fold2: mtry=5, min.node.size=1, splitrule=gini
## - Fold2: mtry=5, min.node.size=1, splitrule=gini
## + Fold2: mtry=8, min.node.size=1, splitrule=gini
## - Fold2: mtry=8, min.node.size=1, splitrule=gini
## + Fold2: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold2: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold2: mtry=5, min.node.size=1, splitrule=extratrees
## - Fold2: mtry=5, min.node.size=1, splitrule=extratrees
## + Fold2: mtry=8, min.node.size=1, splitrule=extratrees
## - Fold2: mtry=8, min.node.size=1, splitrule=extratrees
## + Fold3: mtry=2, min.node.size=1, splitrule=gini
## - Fold3: mtry=2, min.node.size=1, splitrule=gini
## + Fold3: mtry=5, min.node.size=1, splitrule=gini
## - Fold3: mtry=5, min.node.size=1, splitrule=gini
## + Fold3: mtry=8, min.node.size=1, splitrule=gini
## - Fold3: mtry=8, min.node.size=1, splitrule=gini
## + Fold3: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold3: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold3: mtry=5, min.node.size=1, splitrule=extratrees
## - Fold3: mtry=5, min.node.size=1, splitrule=extratrees
## + Fold3: mtry=8, min.node.size=1, splitrule=extratrees
## - Fold3: mtry=8, min.node.size=1, splitrule=extratrees
## + Fold4: mtry=2, min.node.size=1, splitrule=gini
## - Fold4: mtry=2, min.node.size=1, splitrule=gini
## + Fold4: mtry=5, min.node.size=1, splitrule=gini
## - Fold4: mtry=5, min.node.size=1, splitrule=gini
## + Fold4: mtry=8, min.node.size=1, splitrule=gini
## - Fold4: mtry=8, min.node.size=1, splitrule=gini
## + Fold4: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold4: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold4: mtry=5, min.node.size=1, splitrule=extratrees
## - Fold4: mtry=5, min.node.size=1, splitrule=extratrees
## + Fold4: mtry=8, min.node.size=1, splitrule=extratrees

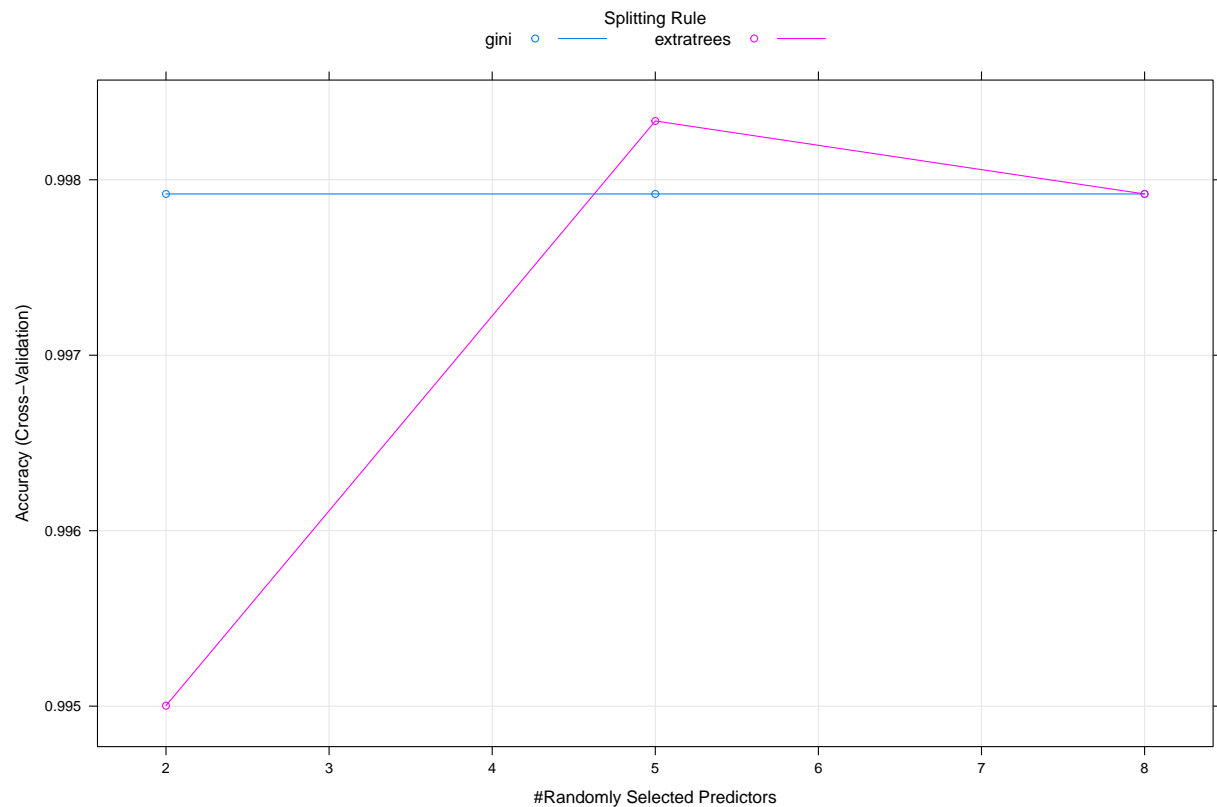
```

```

## - Fold4: mtry=8, min.node.size=1, splitrule=extratrees
## + Fold5: mtry=2, min.node.size=1, splitrule=gini
## - Fold5: mtry=2, min.node.size=1, splitrule=gini
## + Fold5: mtry=5, min.node.size=1, splitrule=gini
## - Fold5: mtry=5, min.node.size=1, splitrule=gini
## + Fold5: mtry=8, min.node.size=1, splitrule=gini
## - Fold5: mtry=8, min.node.size=1, splitrule=gini
## + Fold5: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold5: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold5: mtry=5, min.node.size=1, splitrule=extratrees
## - Fold5: mtry=5, min.node.size=1, splitrule=extratrees
## + Fold5: mtry=8, min.node.size=1, splitrule=extratrees
## - Fold5: mtry=8, min.node.size=1, splitrule=extratrees
## Aggregating results
## Selecting tuning parameters
## Fitting mtry = 5, splitrule = extratrees, min.node.size = 1 on full training set

## Random Forest
##
## 2401 samples
##      8 predictor
##      2 classes: '0', '1'
##
## No pre-processing
## Resampling: Cross-Validated (5 fold)
## Summary of sample sizes: 1921, 1922, 1920, 1921, 1920
## Resampling results across tuning parameters:
##
##      mtry  splitrule  Accuracy  Kappa
##      2     gini      0.9979    0.9904
##      2     extratrees 0.9950    0.9762
##      5     gini      0.9979    0.9904
##      5     extratrees 0.9983    0.9923
##      8     gini      0.9979    0.9904
##      8     extratrees 0.9979    0.9904
##
## Tuning parameter 'min.node.size' was held constant at a value of 1
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were mtry = 5, splitrule =
## extratrees and min.node.size = 1.

```



```
## + Fold01: alpha=0.10, lambda=0.03425
## - Fold01: alpha=0.10, lambda=0.03425
## + Fold01: alpha=0.55, lambda=0.03425
## - Fold01: alpha=0.55, lambda=0.03425
## + Fold01: alpha=1.00, lambda=0.03425
## - Fold01: alpha=1.00, lambda=0.03425
## + Fold02: alpha=0.10, lambda=0.03425
## - Fold02: alpha=0.10, lambda=0.03425
## + Fold02: alpha=0.55, lambda=0.03425
## - Fold02: alpha=0.55, lambda=0.03425
## + Fold02: alpha=1.00, lambda=0.03425
## - Fold02: alpha=1.00, lambda=0.03425
## + Fold03: alpha=0.10, lambda=0.03425
## - Fold03: alpha=0.10, lambda=0.03425
## + Fold03: alpha=0.55, lambda=0.03425
## - Fold03: alpha=0.55, lambda=0.03425
## + Fold03: alpha=1.00, lambda=0.03425
## - Fold03: alpha=1.00, lambda=0.03425
## + Fold04: alpha=0.10, lambda=0.03425
## - Fold04: alpha=0.10, lambda=0.03425
## + Fold04: alpha=0.55, lambda=0.03425
## - Fold04: alpha=0.55, lambda=0.03425
## + Fold04: alpha=1.00, lambda=0.03425
## - Fold04: alpha=1.00, lambda=0.03425
## + Fold05: alpha=0.10, lambda=0.03425
## - Fold05: alpha=0.10, lambda=0.03425
## + Fold05: alpha=0.55, lambda=0.03425
```

```

## - Fold05: alpha=0.55, lambda=0.03425
## + Fold05: alpha=1.00, lambda=0.03425
## - Fold05: alpha=1.00, lambda=0.03425
## + Fold06: alpha=0.10, lambda=0.03425
## - Fold06: alpha=0.10, lambda=0.03425
## + Fold06: alpha=0.55, lambda=0.03425
## - Fold06: alpha=0.55, lambda=0.03425
## + Fold06: alpha=1.00, lambda=0.03425
## - Fold06: alpha=1.00, lambda=0.03425
## + Fold07: alpha=0.10, lambda=0.03425
## - Fold07: alpha=0.10, lambda=0.03425
## + Fold07: alpha=0.55, lambda=0.03425
## - Fold07: alpha=0.55, lambda=0.03425
## + Fold07: alpha=1.00, lambda=0.03425
## - Fold07: alpha=1.00, lambda=0.03425
## + Fold08: alpha=0.10, lambda=0.03425
## - Fold08: alpha=0.10, lambda=0.03425
## + Fold08: alpha=0.55, lambda=0.03425
## - Fold08: alpha=0.55, lambda=0.03425
## + Fold08: alpha=1.00, lambda=0.03425
## - Fold08: alpha=1.00, lambda=0.03425
## + Fold09: alpha=0.10, lambda=0.03425
## - Fold09: alpha=0.10, lambda=0.03425
## + Fold09: alpha=0.55, lambda=0.03425
## - Fold09: alpha=0.55, lambda=0.03425
## + Fold09: alpha=1.00, lambda=0.03425
## - Fold09: alpha=1.00, lambda=0.03425
## + Fold10: alpha=0.10, lambda=0.03425
## - Fold10: alpha=0.10, lambda=0.03425
## + Fold10: alpha=0.55, lambda=0.03425
## - Fold10: alpha=0.55, lambda=0.03425
## + Fold10: alpha=1.00, lambda=0.03425
## - Fold10: alpha=1.00, lambda=0.03425
## Aggregating results
## Selecting tuning parameters
## Fitting alpha = 1, lambda = 0.000343 on full training set

## glmnet
##
## 2401 samples
##      8 predictor
##      2 classes: 'no', 'yes'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 2161, 2162, 2160, 2160, 2162, 2160, ...
## Resampling results across tuning parameters:
##
##      alpha  lambda      ROC      Sens      Spec
##      0.10   0.0003425  0.9994  0.9943  0.9933
##      0.10   0.0034251  0.9983  0.9953  0.8034
##      0.10   0.0342512  0.9921  0.9991  0.7049
##      0.55   0.0003425  0.9995  0.9948  0.9933
##      0.55   0.0034251  0.9987  0.9943  0.9497

```

```

## 0.55 0.0342512 0.9965 0.9991 0.7185
## 1.00 0.0003425 0.9999 0.9962 1.0000
## 1.00 0.0034251 0.9994 0.9929 1.0000
## 1.00 0.0342512 0.9987 0.9981 0.7185
##
## ROC was used to select the optimal model using the largest value.
## The final values used for the model were alpha = 1 and lambda = 0.0003425.

## [1] 1

## + Fold01: alpha=0, lambda=1
## - Fold01: alpha=0, lambda=1
## + Fold01: alpha=1, lambda=1
## - Fold01: alpha=1, lambda=1
## + Fold02: alpha=0, lambda=1
## - Fold02: alpha=0, lambda=1
## + Fold02: alpha=1, lambda=1
## - Fold02: alpha=1, lambda=1
## + Fold03: alpha=0, lambda=1
## - Fold03: alpha=0, lambda=1
## + Fold03: alpha=1, lambda=1
## - Fold03: alpha=1, lambda=1
## + Fold04: alpha=0, lambda=1
## - Fold04: alpha=0, lambda=1
## + Fold04: alpha=1, lambda=1
## - Fold04: alpha=1, lambda=1
## + Fold05: alpha=0, lambda=1
## - Fold05: alpha=0, lambda=1
## + Fold05: alpha=1, lambda=1
## - Fold05: alpha=1, lambda=1
## + Fold06: alpha=0, lambda=1
## - Fold06: alpha=0, lambda=1
## + Fold06: alpha=1, lambda=1
## - Fold06: alpha=1, lambda=1
## + Fold07: alpha=0, lambda=1
## - Fold07: alpha=0, lambda=1
## + Fold07: alpha=1, lambda=1
## - Fold07: alpha=1, lambda=1
## + Fold08: alpha=0, lambda=1
## - Fold08: alpha=0, lambda=1
## + Fold08: alpha=1, lambda=1
## - Fold08: alpha=1, lambda=1
## + Fold09: alpha=0, lambda=1
## - Fold09: alpha=0, lambda=1
## + Fold09: alpha=1, lambda=1
## - Fold09: alpha=1, lambda=1
## + Fold10: alpha=0, lambda=1
## - Fold10: alpha=0, lambda=1
## + Fold10: alpha=1, lambda=1
## - Fold10: alpha=1, lambda=1
## Aggregating results
## Selecting tuning parameters
## Fitting alpha = 1, lambda = 1e-04 on full training set
## glmnet

```

```

##
## 2401 samples
##      8 predictor
##      2 classes: 'no', 'yes'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 2162, 2162, 2160, 2161, 2161, 2160, ...
## Resampling results across tuning parameters:
##
##   alpha  lambda  ROC      Sens   Spec
##   0      0.00010 0.9947  0.9976  0.71425
##   0      0.05273 0.9881  0.9991  0.27517
##   0      0.10535 0.9827  1.0000  0.27517
##   0      0.15798 0.9796  1.0000  0.08161
##   0      0.21061 0.9775  1.0000  0.08161
##   0      0.26323 0.9755  1.0000  0.03747
##   0      0.31586 0.9741  1.0000  0.00000
##   0      0.36848 0.9730  1.0000  0.00000
##   0      0.42111 0.9720  1.0000  0.00000
##   0      0.47374 0.9712  1.0000  0.00000
##   0      0.52636 0.9707  1.0000  0.00000
##   0      0.57899 0.9702  1.0000  0.00000
##   0      0.63162 0.9698  1.0000  0.00000
##   0      0.68424 0.9694  1.0000  0.00000
##   0      0.73687 0.9692  1.0000  0.00000
##   0      0.78949 0.9690  1.0000  0.00000
##   0      0.84212 0.9688  1.0000  0.00000
##   0      0.89475 0.9686  1.0000  0.00000
##   0      0.94737 0.9685  1.0000  0.00000
##   0      1.00000 0.9684  1.0000  0.00000
##   1      0.00010 0.9999  0.9976  0.99667
##   1      0.05273 0.9982  0.9995  0.40437
##   1      0.10535 0.9980  1.0000  0.00000
##   1      0.15798 0.9999  1.0000  0.00000
##   1      0.21061 0.5000  1.0000  0.00000
##   1      0.26323 0.5000  1.0000  0.00000
##   1      0.31586 0.5000  1.0000  0.00000
##   1      0.36848 0.5000  1.0000  0.00000
##   1      0.42111 0.5000  1.0000  0.00000
##   1      0.47374 0.5000  1.0000  0.00000
##   1      0.52636 0.5000  1.0000  0.00000
##   1      0.57899 0.5000  1.0000  0.00000
##   1      0.63162 0.5000  1.0000  0.00000
##   1      0.68424 0.5000  1.0000  0.00000
##   1      0.73687 0.5000  1.0000  0.00000
##   1      0.78949 0.5000  1.0000  0.00000
##   1      0.84212 0.5000  1.0000  0.00000
##   1      0.89475 0.5000  1.0000  0.00000
##   1      0.94737 0.5000  1.0000  0.00000
##   1      1.00000 0.5000  1.0000  0.00000
##
## ROC was used to select the optimal model using the largest value.
## The final values used for the model were alpha = 1 and lambda = 1e-04.

```

```

## [1] 0.9999

## + Fold01: parameter=None
## - Fold01: parameter=None
## + Fold02: parameter=None
## - Fold02: parameter=None
## + Fold03: parameter=None
## - Fold03: parameter=None
## + Fold04: parameter=None
## - Fold04: parameter=None
## + Fold05: parameter=None
## - Fold05: parameter=None
## + Fold06: parameter=None
## - Fold06: parameter=None
## + Fold07: parameter=None
## - Fold07: parameter=None
## + Fold08: parameter=None
## - Fold08: parameter=None
## + Fold09: parameter=None
## - Fold09: parameter=None
## + Fold10: parameter=None
## - Fold10: parameter=None
## Aggregating results
## Fitting final model on full training set

## Generalized Linear Model
##
## 2401 samples
##      8 predictor
##      2 classes: 'no', 'yes'
##
## Pre-processing: median imputation (8)
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 2161, 2160, 2161, 2161, 2161, 2162, ...
## Resampling results:
##
##      ROC      Sens      Spec
##      0.9964  0.9919  0.9898

## + Fold01: parameter=None
## - Fold01: parameter=None
## + Fold02: parameter=None
## - Fold02: parameter=None
## + Fold03: parameter=None
## - Fold03: parameter=None
## + Fold04: parameter=None
## - Fold04: parameter=None
## + Fold05: parameter=None
## - Fold05: parameter=None
## + Fold06: parameter=None
## - Fold06: parameter=None
## + Fold07: parameter=None
## - Fold07: parameter=None
## + Fold08: parameter=None
## - Fold08: parameter=None

```

```

## + Fold09: parameter=None
## - Fold09: parameter=None
## + Fold10: parameter=None
## - Fold10: parameter=None
## Aggregating results
## Fitting final model on full training set

## Generalized Linear Model
##
## 2401 samples
##      8 predictor
##      2 classes: 'no', 'yes'
##
## Pre-processing: nearest neighbor imputation (8), centered (8), scaled (8)
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 2161, 2162, 2161, 2161, 2161, 2160, ...
## Resampling results:
##
##      ROC      Sens      Spec
##      0.9991   0.9967   0.99

## + Fold01: parameter=None
## - Fold01: parameter=None
## + Fold02: parameter=None
## - Fold02: parameter=None
## + Fold03: parameter=None
## - Fold03: parameter=None
## + Fold04: parameter=None
## - Fold04: parameter=None
## + Fold05: parameter=None
## - Fold05: parameter=None
## + Fold06: parameter=None
## - Fold06: parameter=None
## + Fold07: parameter=None
## - Fold07: parameter=None
## + Fold08: parameter=None
## - Fold08: parameter=None
## + Fold09: parameter=None
## - Fold09: parameter=None
## + Fold10: parameter=None
## - Fold10: parameter=None
## Aggregating results
## Fitting final model on full training set

## Generalized Linear Model
##
## 2401 samples
##      8 predictor
##      2 classes: 'no', 'yes'
##
## Pre-processing: median imputation (8)
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 2160, 2162, 2162, 2161, 2160, 2161, ...
## Resampling results:
##

```



```

##      ROC      Sens   Spec
##    0.9939  0.991  0.9866

## + Fold01: parameter=none
## - Fold01: parameter=none
## + Fold02: parameter=none
## - Fold02: parameter=none
## + Fold03: parameter=none
## - Fold03: parameter=none
## + Fold04: parameter=none
## - Fold04: parameter=none
## + Fold05: parameter=none
## - Fold05: parameter=none
## + Fold06: parameter=none
## - Fold06: parameter=none
## + Fold07: parameter=none
## - Fold07: parameter=none
## + Fold08: parameter=none
## - Fold08: parameter=none
## + Fold09: parameter=none
## - Fold09: parameter=none
## + Fold10: parameter=none
## - Fold10: parameter=none
## Aggregating results
## Fitting final model on full training set

## Generalized Linear Model
##
## 2401 samples
##      8 predictor
##      2 classes: 'no', 'yes'
##
## Pre-processing: median imputation (8), centered (8), scaled (8)
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 2162, 2161, 2161, 2161, 2161, 2161, ...
## Resampling results:
##
##      ROC      Sens   Spec
##    0.9994  0.9976  0.9898

## Generalized Linear Model
##
## 2401 samples
##      7 predictor
##      2 classes: 'no', 'yes'
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 2401, 2401, 2401, 2401, 2401, 2401, ...
## Resampling results:
##
##      Accuracy  Kappa
##    0.996      0.9812

## Generalized Linear Model
##

```

```

## 2401 samples
##    7 predictor
##    2 classes: 'no', 'yes'
##
## No pre-processing
## Resampling: Bootstrapped (25 reps)
## Summary of sample sizes: 2401, 2401, 2401, 2401, 2401, 2401, ...
## Resampling results:
##
##   Accuracy   Kappa
##   0.9779     0.8689

## + Fold1: alpha=0.10, lambda=0.03425
## - Fold1: alpha=0.10, lambda=0.03425
## + Fold1: alpha=0.55, lambda=0.03425
## - Fold1: alpha=0.55, lambda=0.03425
## + Fold1: alpha=1.00, lambda=0.03425
## - Fold1: alpha=1.00, lambda=0.03425
## + Fold2: alpha=0.10, lambda=0.03425
## - Fold2: alpha=0.10, lambda=0.03425
## + Fold2: alpha=0.55, lambda=0.03425
## - Fold2: alpha=0.55, lambda=0.03425
## + Fold2: alpha=1.00, lambda=0.03425
## - Fold2: alpha=1.00, lambda=0.03425
## + Fold3: alpha=0.10, lambda=0.03425
## - Fold3: alpha=0.10, lambda=0.03425
## + Fold3: alpha=0.55, lambda=0.03425
## - Fold3: alpha=0.55, lambda=0.03425
## + Fold3: alpha=1.00, lambda=0.03425
## - Fold3: alpha=1.00, lambda=0.03425
## + Fold4: alpha=0.10, lambda=0.03425
## - Fold4: alpha=0.10, lambda=0.03425
## + Fold4: alpha=0.55, lambda=0.03425
## - Fold4: alpha=0.55, lambda=0.03425
## + Fold4: alpha=1.00, lambda=0.03425
## - Fold4: alpha=1.00, lambda=0.03425
## + Fold5: alpha=0.10, lambda=0.03425
## - Fold5: alpha=0.10, lambda=0.03425
## + Fold5: alpha=0.55, lambda=0.03425
## - Fold5: alpha=0.55, lambda=0.03425
## + Fold5: alpha=1.00, lambda=0.03425
## - Fold5: alpha=1.00, lambda=0.03425
## Aggregating results
## Selecting tuning parameters
## Fitting alpha = 1, lambda = 0.000343 on full training set

## + Fold1: mtry=2, min.node.size=1, splitrule=gini
## - Fold1: mtry=2, min.node.size=1, splitrule=gini
## + Fold1: mtry=5, min.node.size=1, splitrule=gini
## - Fold1: mtry=5, min.node.size=1, splitrule=gini
## + Fold1: mtry=8, min.node.size=1, splitrule=gini
## - Fold1: mtry=8, min.node.size=1, splitrule=gini
## + Fold1: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold1: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold1: mtry=5, min.node.size=1, splitrule=extratrees

```

```

## - Fold1: mtry=5, min.node.size=1, splitrule=extratrees
## + Fold1: mtry=8, min.node.size=1, splitrule=extratrees
## - Fold1: mtry=8, min.node.size=1, splitrule=extratrees
## + Fold2: mtry=2, min.node.size=1, splitrule=gini
## - Fold2: mtry=2, min.node.size=1, splitrule=gini
## + Fold2: mtry=5, min.node.size=1, splitrule=gini
## - Fold2: mtry=5, min.node.size=1, splitrule=gini
## + Fold2: mtry=8, min.node.size=1, splitrule=gini
## - Fold2: mtry=8, min.node.size=1, splitrule=gini
## + Fold2: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold2: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold2: mtry=5, min.node.size=1, splitrule=extratrees
## - Fold2: mtry=5, min.node.size=1, splitrule=extratrees
## + Fold2: mtry=8, min.node.size=1, splitrule=extratrees
## - Fold2: mtry=8, min.node.size=1, splitrule=extratrees
## + Fold3: mtry=2, min.node.size=1, splitrule=gini
## - Fold3: mtry=2, min.node.size=1, splitrule=gini
## + Fold3: mtry=5, min.node.size=1, splitrule=gini
## - Fold3: mtry=5, min.node.size=1, splitrule=gini
## + Fold3: mtry=8, min.node.size=1, splitrule=gini
## - Fold3: mtry=8, min.node.size=1, splitrule=gini
## + Fold3: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold3: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold3: mtry=5, min.node.size=1, splitrule=extratrees
## - Fold3: mtry=5, min.node.size=1, splitrule=extratrees
## + Fold3: mtry=8, min.node.size=1, splitrule=extratrees
## - Fold3: mtry=8, min.node.size=1, splitrule=extratrees
## + Fold4: mtry=2, min.node.size=1, splitrule=gini
## - Fold4: mtry=2, min.node.size=1, splitrule=gini
## + Fold4: mtry=5, min.node.size=1, splitrule=gini
## - Fold4: mtry=5, min.node.size=1, splitrule=gini
## + Fold4: mtry=8, min.node.size=1, splitrule=gini
## - Fold4: mtry=8, min.node.size=1, splitrule=gini
## + Fold4: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold4: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold4: mtry=5, min.node.size=1, splitrule=extratrees
## - Fold4: mtry=5, min.node.size=1, splitrule=extratrees
## + Fold4: mtry=8, min.node.size=1, splitrule=extratrees
## - Fold4: mtry=8, min.node.size=1, splitrule=extratrees
## + Fold5: mtry=2, min.node.size=1, splitrule=gini
## - Fold5: mtry=2, min.node.size=1, splitrule=gini
## + Fold5: mtry=5, min.node.size=1, splitrule=gini
## - Fold5: mtry=5, min.node.size=1, splitrule=gini
## + Fold5: mtry=8, min.node.size=1, splitrule=gini
## - Fold5: mtry=8, min.node.size=1, splitrule=gini
## + Fold5: mtry=2, min.node.size=1, splitrule=extratrees
## - Fold5: mtry=2, min.node.size=1, splitrule=extratrees
## + Fold5: mtry=5, min.node.size=1, splitrule=extratrees
## - Fold5: mtry=5, min.node.size=1, splitrule=extratrees
## + Fold5: mtry=8, min.node.size=1, splitrule=extratrees
## - Fold5: mtry=8, min.node.size=1, splitrule=extratrees
## Aggregating results
## Selecting tuning parameters
## Fitting mtry = 2, splitrule = gini, min.node.size = 1 on full training set

```

```
##
## Call:
## summary.resamples(object = resamples)
##
## Models: item1, item2
## Number of resamples: 5
##
## ROC
##      Min. 1st Qu. Median   Mean 3rd Qu. Max. NA's
## item1 0.9987  0.9998 0.9999 0.9996  0.9999    1    0
## item2 0.9999  0.9999 0.9999 0.9999  1.0000    1    0
##
## Sens
##      Min. 1st Qu. Median   Mean 3rd Qu.   Max. NA's
## item1 0.9852  0.9953 0.9953 0.9938  0.9964 0.9970    0
## item2 0.9976  0.9976 0.9982 0.9985  0.9994 0.9994    0
##
## Spec
##      Min. 1st Qu. Median   Mean 3rd Qu. Max. NA's
## item1 0.9957  1.0000 1.0000 0.9991    1    1    0
## item2 0.9830  0.9957 0.9957 0.9949    1    1    0
```

