C4T3

Code

Objective:

Your job is to evaluate multiple machine learning models to see which produces the best result, enabling us to make a recommendation to the client.

Hide

```
require(pacman)

pacman:: p_load(pacman, dplyr, GGally, ggplot2, ggrepel, patchwork, gifski, ggforce, ggthemes, m
aps, sf, concaveman, remotes, readxl, ggthemes, ggvis, httr, plotly, rmarkdown, extrafont, shin
y, isoband, stringr, rio, tidyr, labeling, caret, jquerylib, farver, corrgram, caTools, cowplot,
randomForest, RMariaDB, lubridate, zoo, scales, ggfortify, forecast, doParallel,e1071)
```

Parallel computing using multiple cores:

Hide

```


#cl <- makeCluster(5)
#registerDoParallel(cl)
## Here you put the processes
#stopCluster(cl)
```

Hide

```
df <- import("trainingData.csv")

df <- within(df, MasterID <- paste("R", BUILDINGID, FLOOR, SPACEID, RELATIVEPOSITION, sep='_'))

df <- select(df, -c(LONGITUDE, LATITUDE, USERID, PHONEID, TIMESTAMP))

df$MasterID <- as.factor(as.character(df$MasterID))

#group_indices_
#df$MasterID
```

Loaded the data, created independent variable called ID that contains “FLOOR”, “BUILDINGID”, “SPACEID”, “RELATIVEPOSITION” information.
Dropped useless columns like Longitude and Latitude.

Hide

```
#Random Forest
df_modelling <- select(df, -c("BUILDINGID", "FLOOR", "SPACEID", "RELATIVEPOSITION"))
ctrl <- trainControl(method = "repeatedcv", number = 10, repeats = 3, classProbs = TRUE)
str(df_modelling$MasterID)
```

```
 Factor w/ 905 levels "R_0_0_102_2",..: 400 400 394 392 16 398 394 390 407 393 ...
```

Hide

```
#cl <- makeCluster(6)
#registerDoParallel(cl)
#RF_Fit1 <- train(MasterID~., data = df_modelling, method = "rf", tuneLength = 1, trControl=ctr
l)
#stopCluster(cl)
#names(df_modelling)
```

Even with parallel processing, the computer takes around an hour to run only one model, therefore i will only focus on one builing instead of all of them.

Hide

```
#summary(as.factor(df$BUILDINGID))

df_modelling2 <- filter(df, df$BUILDINGID == 0)
df_modelling2 <- select(df_modelling2, -c("FLOOR", "SPACEID", "RELATIVEPOSITION","BUILDINGID"))

str(df_modelling2$MasterID)
```

```
 Factor w/ 905 levels "R_0_0_102_2",..: 16 1 4 5 3 2 9 8 7 6 ...
```

Hide

```
#str(df_modelling2$MasterID)
#str(df_modelling$MasterID)
#str(df_modelling$BUILDINGID)
#str(df_modelling2)
```

##Starting from Scratch

Hide

```
df3 <- import("trainingData.csv")
df3 <- filter(df3, df$BUILDINGID == 0)
df3 <- within(df3, MasterID <- paste("R", BUILDINGID, FLOOR, SPACEID, RELATIVEPOSITION, sep='_'
))
df3 <- select(df3, -c(LONGITUDE, LATITUDE, USERID, PHONEID, TIMESTAMP, FLOOR, SPACEID, RELATIVEP
OSITION,BUILDINGID))
df3$MasterID <- as.factor(as.character(df3$MasterID))

ctrl <- trainControl(method = "repeatedcv", number = 10, repeats = 3, classProbs = TRUE)
```

Hide

```
cl <- makeCluster(7)
registerDoParallel(cl)
RF_Fit3 <- train(MasterID~., data = df3, method = "rf", tuneLength = 1, trControl=ctrl)
```

```
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 15 (<-NZXT-Mar:11357)
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 14 (<-NZXT-Mar:11357)
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 13 (<-NZXT-Mar:11357)
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 12 (<-NZXT-Mar:11357)
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 11 (<-NZXT-Mar:11357)
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 10 (<-NZXT-Mar:11357)
```

Hide

```
stopCluster(cl)


#fitControl <- trainControl(method = "repeatedcv", number = 10, repeats = 3, returnResamp="all")
#cl <- makeCluster(6)
#registerDoParallel(cl)
#C50Fit1 <- train(MasterID~., data = df3,trControl=fitControl,method="C5.0",verbose=FALSE)
#stopCluster(cl)

#cl <- makeCluster(6)
#registerDoParallel(cl)
#svmFit1 = svm( MasterID~., data = df3, scale = FALSE, kernel = "radial", cost = 5)
#stopCluster(cl)

#ModelData <- resamples(list(RF = RF_Fit3, SVM = svmFit1, C50 = c50Fit1))
```

Hide

```
RF_Fit3
```

```
Random Forest

5249 samples
 520 predictor
 259 classes: 'R_0_0_102_2', 'R_0_0_106_2', 'R_0_0_107_2', 'R_0_0_110_2', 'R_0_0_111_2', 'R_0_0_
112_2', 'R_0_0_113_2', 'R_0_0_114_2', 'R_0_0_115_2', 'R_0_0_116_2', 'R_0_0_117_2', 'R_0_0_118_
2', 'R_0_0_119_2', 'R_0_0_120_2', 'R_0_0_121_2', 'R_0_0_122_2', 'R_0_0_123_2', 'R_0_0_125_2', 'R
_0_0_126_2', 'R_0_0_127_2', 'R_0_0_128_2', 'R_0_0_129_2', 'R_0_0_130_2', 'R_0_0_131_2', 'R_0_0_1
32_2', 'R_0_0_133_2', 'R_0_0_134_2', 'R_0_0_201_2', 'R_0_0_202_2', 'R_0_0_208_2', 'R_0_0_209_2',
'R_0_0_211_2', 'R_0_0_212_2', 'R_0_0_213_2', 'R_0_0_214_2', 'R_0_0_215_2', 'R_0_0_216_2', 'R_0_0
_218_2', 'R_0_0_219_2', 'R_0_0_220_2', 'R_0_0_222_2', 'R_0_0_224_2', 'R_0_0_225_2', 'R_0_0_226_
2', 'R_0_0_227_2', 'R_0_0_229_2', 'R_0_0_230_2', 'R_0_0_231_2', 'R_0_0_232_2', 'R_0_0_233_2', 'R
_0_0_234_2', 'R_0_0_235_2', 'R_0_0_236_2', 'R_0_0_237_2', 'R_0_1_101_2', 'R_0_1_102_2', 'R_0_1_1
03_2', 'R_0_1_104_2', 'R_0_1_105_2', 'R_0_1_106_2', 'R_0_1_107_2', 'R_0_1_108_2', 'R_0_1_109_2',
'R_0_1_110_2', 'R_0_1_111_2', 'R_0_1_112_2', 'R_0_1_113_2', 'R_0_1_114_2', 'R_0_1_115_2', 'R_0_1
_116_2', 'R_0_1_117_2', 'R_0_1_118_2', 'R_0_1_119_2', 'R_0_1_120_2', 'R_0_1_121_2', 'R_0_1_122_
2', 'R_0_1_123_2', 'R_0_1_124_2', 'R_0_1_125_2', 'R_0_1_126_2', 'R_0_1_127_2', 'R_0_1_128_2', 'R
_0_1_129_2', 'R_0_1_130_2', 'R_0_1_136_2', 'R_0_1_137_2', 'R_0_1_138_2', 'R_0_1_201_2', 'R_0_1_2
02_2', 'R_0_1_203_2', 'R_0_1_204_2', 'R_0_1_205_1', 'R_0_1_205_2', 'R_0_1_206_2', 'R_0_1_207_2',
'R_0_1_208_2', 'R_0_1_209_2', 'R_0_1_210_2', 'R_0_1_211_2', 'R_0_1_212_2', 'R_0_1_213_2', 'R_0_1
_214_2', 'R_0_1_215_2', 'R_0_1_216_2', 'R_0_1_217_2', 'R_0_1_218_2', 'R_0_1_219_2', 'R_0_1_220_
2', 'R_0_1_221_2', 'R_0_1_222_2', 'R_0_1_223_2', 'R_0_1_224_2', 'R_0_1_225_2', 'R_0_1_226_2', 'R
_0_1_227_2', 'R_0_1_228_2', 'R_0_1_229_2', 'R_0_1_230_2', 'R_0_1_233_2', 'R_0_1_234_2', 'R_0_1_2
35_2', 'R_0_2_101_2', 'R_0_2_102_2', 'R_0_2_103_2', 'R_0_2_104_2', 'R_0_2_105_2', 'R_0_2_106_2',
'R_0_2_107_2', 'R_0_2_108_2', 'R_0_2_109_2', 'R_0_2_110_2', 'R_0_2_111_2', 'R_0_2_112_2', 'R_0_2
_113_2', 'R_0_2_114_2', 'R_0_2_115_2', 'R_0_2_117_2', 'R_0_2_118_2', 'R_0_2_119_2', 'R_0_2_120_
2', 'R_0_2_121_2', 'R_0_2_122_2', 'R_0_2_123_2', 'R_0_2_124_2', 'R_0_2_125_2', 'R_0_2_126_2', 'R
_0_2_127_2', 'R_0_2_128_1', 'R_0_2_128_2', 'R_0_2_129_2', 'R_0_2_130_2', 'R_0_2_132_2', 'R_0_2_1
33_2', 'R_0_2_134_2', 'R_0_2_138_2', 'R_0_2_139_2', 'R_0_2_140_2', 'R_0_2_201_2', 'R_0_2_202_2',
'R_0_2_203_2', 'R_0_2_204_2', 'R_0_2_205_2', 'R_0_2_206_2', 'R_0_2_207_2', 'R_0_2_208_2', 'R_0_2
_209_2', 'R_0_2_210_2', 'R_0_2_211_2', 'R_0_2_212_2', 'R_0_2_213_2', 'R_0_2_214_1', 'R_0_2_214_
2', 'R_0_2_216_2', 'R_0_2_217_2', 'R_0_2_218_2', 'R_0_2_219_2', 'R_0_2_220_2', 'R_0_2_221_2', 'R
_0_2_222_2', 'R_0_2_223_2', 'R_0_2_224_2', 'R_0_2_225_2', 'R_0_2_226_2', 'R_0_2_227_2', 'R_0_2_2
28_2', 'R_0_2_229_2', 'R_0_2_230_2', 'R_0_2_231_2', 'R_0_2_234_2', 'R_0_2_235_2', 'R_0_2_241_2',
'R_0_3_101_2', 'R_0_3_102_2', 'R_0_3_103_2', 'R_0_3_104_2', 'R_0_3_105_2', 'R_0_3_106_2', 'R_0_3
_107_2', 'R_0_3_108_2', 'R_0_3_109_2', 'R_0_3_110_2', 'R_0_3_111_2', 'R_0_3_112_2', 'R_0_3_113_
2', 'R_0_3_114_2', 'R_0_3_115_2', 'R_0_3_116_2', 'R_0_3_117_2', 'R_0_3_118_2', 'R_0_3_119_2', 'R
_0_3_120_2', 'R_0_3_121_2', 'R_0_3_122_2', 'R_0_3_123_2', 'R_0_3_124_2', 'R_0_3_125_2', 'R_0_3_1
26_2', 'R_0_3_127_2', 'R_0_3_128_2', 'R_0_3_129_2', 'R_0_3_130_2', 'R_0_3_131_2', 'R_0_3_135_2',
'R_0_3_136_2', 'R_0_3_137_2', 'R_0_3_201_2', 'R_0_3_202_2', 'R_0_3_203_2', 'R_0_3_204_2', 'R_0_3
_205_2', 'R_0_3_206_2', 'R_0_3_207_2', 'R_0_3_208_2', 'R_0_3_209_2', 'R_0_3_210_2', 'R_0_3_211_
2', 'R_0_3_212_2', 'R_0_3_213_2', 'R_0_3_214_2', 'R_0_3_215_2', 'R_0_3_216_2', 'R_0_3_217_2', 'R
_0_3_218_2', 'R_0_3_219_2', 'R_0_3_220_2', 'R_0_3_221_2', 'R_0_3_222_2', 'R_0_3_223_2', 'R_0_3_2
24_2', 'R_0_3_225_2', 'R_0_3_226_2', 'R_0_3_227_2', 'R_0_3_228_2', 'R_0_3_229_2', 'R_0_3_230_2',
'R_0_3_231_2', 'R_0_3_234_2', 'R_0_3_235_2', 'R_0_3_236_2'

No pre-processing
Resampling: Cross-Validated (10 fold, repeated 3 times)
Summary of sample sizes: 4724, 4727, 4718, 4720, 4723, 4730, ...
Resampling results:

  Accuracy   Kappa
  0.7347186  0.7336129
```

```
Tuning parameter 'mtry' was held constant at a value of 22
```

Kappa Score is a metric that compares an Observed Accuracy with an Expected Accuracy and it is used not only to evaluate a single classifier, but also to evaluate multiple classifiers when they have been used on the same problem. In general it is less misleading than simply using accuracy as a metric; computation of Observed Accuracy and Expected Accuracy is integral to comprehension of the Kappa Score, and is most easily seen in the use of a confusion matrix.

Observed Accuracy is simply the number of instances that were classified correctly throughout the entire confusion matrix. Expected Accuracy is defined as the accuracy that any random classifier would be expected to achieve based on the confusion matrix. The Expected Accuracy is directly related to the number of instances of each class combined with the number of instances that the machine learning classifier agreed with as being ground truth.

Hide

```
cl <- makeCluster(7)
registerDoParallel(cl)
C50Fit1 <- train(MasterID~., data = df3,trControl=ctrl,method="C5.0", tuneLength = 5)
stopCluster(cl)
```

Hide

```
C50Fit1
```

```
C5.0

5249 samples
 520 predictor
 259 classes: 'R_0_0_102_2', 'R_0_0_106_2', 'R_0_0_107_2', 'R_0_0_110_2', 'R_0_0_111_2', 'R_0_0_
112_2', 'R_0_0_113_2', 'R_0_0_114_2', 'R_0_0_115_2', 'R_0_0_116_2', 'R_0_0_117_2', 'R_0_0_118_
2', 'R_0_0_119_2', 'R_0_0_120_2', 'R_0_0_121_2', 'R_0_0_122_2', 'R_0_0_123_2', 'R_0_0_125_2', 'R
_0_0_126_2', 'R_0_0_127_2', 'R_0_0_128_2', 'R_0_0_129_2', 'R_0_0_130_2', 'R_0_0_131_2', 'R_0_0_1
32_2', 'R_0_0_133_2', 'R_0_0_134_2', 'R_0_0_201_2', 'R_0_0_202_2', 'R_0_0_208_2', 'R_0_0_209_2',
'R_0_0_211_2', 'R_0_0_212_2', 'R_0_0_213_2', 'R_0_0_214_2', 'R_0_0_215_2', 'R_0_0_216_2', 'R_0_0
_218_2', 'R_0_0_219_2', 'R_0_0_220_2', 'R_0_0_222_2', 'R_0_0_224_2', 'R_0_0_225_2', 'R_0_0_226_
2', 'R_0_0_227_2', 'R_0_0_229_2', 'R_0_0_230_2', 'R_0_0_231_2', 'R_0_0_232_2', 'R_0_0_233_2', 'R
_0_0_234_2', 'R_0_0_235_2', 'R_0_0_236_2', 'R_0_0_237_2', 'R_0_1_101_2', 'R_0_1_102_2', 'R_0_1_1
03_2', 'R_0_1_104_2', 'R_0_1_105_2', 'R_0_1_106_2', 'R_0_1_107_2', 'R_0_1_108_2', 'R_0_1_109_2',
'R_0_1_110_2', 'R_0_1_111_2', 'R_0_1_112_2', 'R_0_1_113_2', 'R_0_1_114_2', 'R_0_1_115_2', 'R_0_1
_116_2', 'R_0_1_117_2', 'R_0_1_118_2', 'R_0_1_119_2', 'R_0_1_120_2', 'R_0_1_121_2', 'R_0_1_122_
2', 'R_0_1_123_2', 'R_0_1_124_2', 'R_0_1_125_2', 'R_0_1_126_2', 'R_0_1_127_2', 'R_0_1_128_2', 'R
_0_1_129_2', 'R_0_1_130_2', 'R_0_1_136_2', 'R_0_1_137_2', 'R_0_1_138_2', 'R_0_1_201_2', 'R_0_1_2
02_2', 'R_0_1_203_2', 'R_0_1_204_2', 'R_0_1_205_1', 'R_0_1_205_2', 'R_0_1_206_2', 'R_0_1_207_2',
'R_0_1_208_2', 'R_0_1_209_2', 'R_0_1_210_2', 'R_0_1_211_2', 'R_0_1_212_2', 'R_0_1_213_2', 'R_0_1
_214_2', 'R_0_1_215_2', 'R_0_1_216_2', 'R_0_1_217_2', 'R_0_1_218_2', 'R_0_1_219_2', 'R_0_1_220_
2', 'R_0_1_221_2', 'R_0_1_222_2', 'R_0_1_223_2', 'R_0_1_224_2', 'R_0_1_225_2', 'R_0_1_226_2', 'R
_0_1_227_2', 'R_0_1_228_2', 'R_0_1_229_2', 'R_0_1_230_2', 'R_0_1_233_2', 'R_0_1_234_2', 'R_0_1_2
35_2', 'R_0_2_101_2', 'R_0_2_102_2', 'R_0_2_103_2', 'R_0_2_104_2', 'R_0_2_105_2', 'R_0_2_106_2',
'R_0_2_107_2', 'R_0_2_108_2', 'R_0_2_109_2', 'R_0_2_110_2', 'R_0_2_111_2', 'R_0_2_112_2', 'R_0_2
_113_2', 'R_0_2_114_2', 'R_0_2_115_2', 'R_0_2_117_2', 'R_0_2_118_2', 'R_0_2_119_2', 'R_0_2_120_
2', 'R_0_2_121_2', 'R_0_2_122_2', 'R_0_2_123_2', 'R_0_2_124_2', 'R_0_2_125_2', 'R_0_2_126_2', 'R
_0_2_127_2', 'R_0_2_128_1', 'R_0_2_128_2', 'R_0_2_129_2', 'R_0_2_130_2', 'R_0_2_132_2', 'R_0_2_1
33_2', 'R_0_2_134_2', 'R_0_2_138_2', 'R_0_2_139_2', 'R_0_2_140_2', 'R_0_2_201_2', 'R_0_2_202_2',
'R_0_2_203_2', 'R_0_2_204_2', 'R_0_2_205_2', 'R_0_2_206_2', 'R_0_2_207_2', 'R_0_2_208_2', 'R_0_2
_209_2', 'R_0_2_210_2', 'R_0_2_211_2', 'R_0_2_212_2', 'R_0_2_213_2', 'R_0_2_214_1', 'R_0_2_214_
2', 'R_0_2_216_2', 'R_0_2_217_2', 'R_0_2_218_2', 'R_0_2_219_2', 'R_0_2_220_2', 'R_0_2_221_2', 'R
_0_2_222_2', 'R_0_2_223_2', 'R_0_2_224_2', 'R_0_2_225_2', 'R_0_2_226_2', 'R_0_2_227_2', 'R_0_2_2
28_2', 'R_0_2_229_2', 'R_0_2_230_2', 'R_0_2_231_2', 'R_0_2_234_2', 'R_0_2_235_2', 'R_0_2_241_2',
'R_0_3_101_2', 'R_0_3_102_2', 'R_0_3_103_2', 'R_0_3_104_2', 'R_0_3_105_2', 'R_0_3_106_2', 'R_0_3
_107_2', 'R_0_3_108_2', 'R_0_3_109_2', 'R_0_3_110_2', 'R_0_3_111_2', 'R_0_3_112_2', 'R_0_3_113_
2', 'R_0_3_114_2', 'R_0_3_115_2', 'R_0_3_116_2', 'R_0_3_117_2', 'R_0_3_118_2', 'R_0_3_119_2', 'R
_0_3_120_2', 'R_0_3_121_2', 'R_0_3_122_2', 'R_0_3_123_2', 'R_0_3_124_2', 'R_0_3_125_2', 'R_0_3_1
26_2', 'R_0_3_127_2', 'R_0_3_128_2', 'R_0_3_129_2', 'R_0_3_130_2', 'R_0_3_131_2', 'R_0_3_135_2',
'R_0_3_136_2', 'R_0_3_137_2', 'R_0_3_201_2', 'R_0_3_202_2', 'R_0_3_203_2', 'R_0_3_204_2', 'R_0_3
_205_2', 'R_0_3_206_2', 'R_0_3_207_2', 'R_0_3_208_2', 'R_0_3_209_2', 'R_0_3_210_2', 'R_0_3_211_
2', 'R_0_3_212_2', 'R_0_3_213_2', 'R_0_3_214_2', 'R_0_3_215_2', 'R_0_3_216_2', 'R_0_3_217_2', 'R
_0_3_218_2', 'R_0_3_219_2', 'R_0_3_220_2', 'R_0_3_221_2', 'R_0_3_222_2', 'R_0_3_223_2', 'R_0_3_2
24_2', 'R_0_3_225_2', 'R_0_3_226_2', 'R_0_3_227_2', 'R_0_3_228_2', 'R_0_3_229_2', 'R_0_3_230_2',
'R_0_3_231_2', 'R_0_3_234_2', 'R_0_3_235_2', 'R_0_3_236_2'

No pre-processing
Resampling: Cross-Validated (10 fold, repeated 3 times)
Summary of sample sizes: 4719, 4726, 4726, 4724, 4724, 4727, ...
Resampling results across tuning parameters:

  model  winnow  trials  Accuracy   Kappa
  rules  FALSE    1      0.6000210  0.5983891
```

```
  rules  FALSE   10      0.7177421  0.7165906
  rules  FALSE   20      0.7394604  0.7383983
  rules  FALSE   30      0.7485460  0.7475215
  rules  FALSE   40      0.7529214  0.7519144
  rules   TRUE    1      0.5994339  0.5977954
  rules   TRUE   10      0.7231910  0.7220623
  rules   TRUE   20      0.7401948  0.7391352
  rules   TRUE   30      0.7462928  0.7452590
  rules   TRUE   40      0.7499187  0.7488995
  tree   FALSE    1      0.6087787  0.6071864
  tree   FALSE   10      0.7177178  0.7165661
  tree   FALSE   20      0.7328938  0.7318042
  tree   FALSE   30      0.7393134  0.7382500
  tree   FALSE   40      0.7443261  0.7432838
  tree    TRUE    1      0.6054638  0.6038552
  tree    TRUE   10      0.7194301  0.7182856
  tree    TRUE   20      0.7346084  0.7335260
  tree    TRUE   30      0.7402004  0.7391408
  tree    TRUE   40      0.7426793  0.7416295

Accuracy was used to select the optimal model using the largest value.
The final values used for the model were trials = 40, model = rules and winnow = FALSE.
```

Hide

```
ModelData <- resamples(list(RF = RF_Fit3, SVM = svmFit1, C50 = C50Fit1))
```

```
Error in vapply(x, function(x) x$control$method, character(1)) :
  values must be length 1,
 but FUN(X[[2]]) result is length 0
```

Hide

```
svmFit1
```

```


Call:
svm(formula = MasterID ~ ., data = df3, kernel = "radial", cost = 5, scale = FALSE)


Parameters:
   SVM-Type:  C-classification
 SVM-Kernel:  radial
       cost:  5

Number of Support Vectors:  5051
```

Hide

```
cl <- makeCluster(7)
registerDoParallel(cl)
#svmFit2 <- train(MasterID~., data = df3, method = "svmLinear", trControl=ctrl, tuneLength = 5)
knnFit1 <- train(MasterID~., data = df3, method = "knn", trControl = ctrl, preProcess = c("cente
r","scale"), tuneLength = 20)
```

```
Warning in numInClass[i] :
  closing unused connection 17 (<-NZXT-Mar:11357)
Warning in numInClass[i] :
  closing unused connection 16 (<-NZXT-Mar:11357)
Warning in numInClass[i] :
  closing unused connection 15 (<-NZXT-Mar:11357)
Warning in numInClass[i] :
  closing unused connection 14 (<-NZXT-Mar:11357)
Warning in numInClass[i] :
  closing unused connection 13 (<-NZXT-Mar:11357)
Warning in numInClass[i] :
  closing unused connection 12 (<-NZXT-Mar:11357)
Warning in numInClass[i] :
  closing unused connection 11 (<-NZXT-Mar:11357)
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 10 (<-NZXT-Mar:11357)
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 9 (<-NZXT-Mar:11357)
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 8 (<-NZXT-Mar:11357)
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 7 (<-NZXT-Mar:11357)
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 6 (<-NZXT-Mar:11357)
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 5 (<-NZXT-Mar:11357)
Warning in .Internal(gc(verbose, reset, full)) :
  closing unused connection 4 (<-NZXT-Mar:11357)
These variables have zero variances: WAP003, WAP004, WAP005, WAP006, WAP010, WAP011, WAP012, WAP
015, WAP016, WAP021, WAP022, WAP037, WAP038, WAP055, WAP056, WAP059, WAP060, WAP061, WAP062, WAP
063, WAP064, WAP065, WAP066, WAP067, WAP068, WAP069, WAP070, WAP073, WAP074, WAP077, WAP078, WAP
082, WAP083, WAP084, WAP085, WAP086, WAP087, WAP088, WAP089, WAP090, WAP091, WAP092, WAP093, WAP
094, WAP095, WAP096, WAP097, WAP098, WAP099, WAP101, WAP102, WAP105, WAP106, WAP107, WAP108, WAP
109, WAP110, WAP111, WAP112, WAP113, WAP114, WAP115, WAP116, WAP117, WAP118, WAP119, WAP120, WAP
121, WAP122, WAP125, WAP126, WAP127, WAP128, WAP129, WAP130, WAP131, WAP132, WAP133, WAP134, WAP
135, WAP136, WAP137, WAP138, WAP139, WAP140, WAP141, WAP144, WAP145, WAP146, WAP147, WAP148, WAP
149, WAP152, WAP157, WAP158, WAP159, WAP160, WAP164, WAP165, WAP174, WAP175, WAP176, WAP177, WAP
178, WAP179, WAP180, WAP181, WAP186, WAP187, WAP188, WAP189, WAP190, WAP193, WAP194, WAP195, WAP
196, WAP197, WAP198, WAP199, WAP200, WAP203, WAP204, WAP206, WAP208, WAP209, WAP213, WAP214, WAP
215, WAP216, WAP217, WAP218, WAP219, WAP220, WAP222, WAP223, WAP226, WAP227, WAP228, WAP229, WAP
230, WAP231, WAP232, WAP233, WAP234, WAP235, WAP236, WAP237, WAP238, WAP239, WAP240, WAP241, WAP
242, WAP243, WAP244, WAP245, WAP246, WAP247, WAP249, WAP250, WAP254, WAP255, WAP256, WAP257, WAP
258, WAP259, WAP260, WAP261, WAP262, WAP263, WAP274, WAP277, WAP278, WAP279, WAP282, WAP284, WAP
286, WAP288, WAP293, WAP295, WAP296, WAP301, WAP303, WAP304, WAP307, WAP310, WAP311, WAP312, WAP
313, WAP314, WAP315, WAP316, WAP317, WAP318, WAP329, WAP332, WAP333, WAP334, WAP335, WAP338, WAP
340, WAP342, WAP344, WAP349, WAP351, WAP353, WAP360, WAP365, WAP366, WAP367, WAP368, WAP369, WAP
370, WAP371, WAP372, WAP373, WAP374, WAP375, WAP386, WAP389, WAP390, WAP391, WAP394, WAP396, WAP
398, WAP400, WAP405, WAP407, WAP416, WAP417, WAP418, WAP419, WAP420, WAP421, WAP422, WAP423, WAP
424, WAP425, WAP427, WAP428, WAP429, WAP430, WAP431, WAP432, WAP433, WAP435, WAP436, WAP437, WAP
438, WAP439, WAP440, WAP441, WAP442, WAP444, WAP445, WAP446, WAP448, WAP449, WAP450, WAP451, WAP
453, WAP454, WAP455, WAP456, WAP457, WAP458, WAP460, WAP461, WAP462, WAP463, WAP464, WAP465, WAP
466, WAP467, WAP468, WAP469, WAP470, WAP471, WAP472, WAP473, WAP474, WAP476, WAP477, WAP478, WAP
479, WAP480, WAP481, WAP482, WAP483, WAP484, WAP485, WAP486, WAP487, WAP488, WAP489, WAP490, WAP
```

```
491, WAP492, WAP493, WAP495, WAP496, WAP497, WAP498, WAP499, WAP501, WAP502, WAP503, WAP504, WAP
505, WAP506, WAP507, WAP509, WAP510, WAP511, WAP512, WAP513, WAP514, WAP516, WAP517, WAP518, WAP
520
```

Hide

```
stopCluster(cl)
```

Hide

```
knnFit1
```

```
k-Nearest Neighbors

5249 samples
 520 predictor
 259 classes: 'R_0_0_102_2', 'R_0_0_106_2', 'R_0_0_107_2', 'R_0_0_110_2', 'R_0_0_111_2', 'R_0_0_
112_2', 'R_0_0_113_2', 'R_0_0_114_2', 'R_0_0_115_2', 'R_0_0_116_2', 'R_0_0_117_2', 'R_0_0_118_
2', 'R_0_0_119_2', 'R_0_0_120_2', 'R_0_0_121_2', 'R_0_0_122_2', 'R_0_0_123_2', 'R_0_0_125_2', 'R
_0_0_126_2', 'R_0_0_127_2', 'R_0_0_128_2', 'R_0_0_129_2', 'R_0_0_130_2', 'R_0_0_131_2', 'R_0_0_1
32_2', 'R_0_0_133_2', 'R_0_0_134_2', 'R_0_0_201_2', 'R_0_0_202_2', 'R_0_0_208_2', 'R_0_0_209_2',
'R_0_0_211_2', 'R_0_0_212_2', 'R_0_0_213_2', 'R_0_0_214_2', 'R_0_0_215_2', 'R_0_0_216_2', 'R_0_0
_218_2', 'R_0_0_219_2', 'R_0_0_220_2', 'R_0_0_222_2', 'R_0_0_224_2', 'R_0_0_225_2', 'R_0_0_226_
2', 'R_0_0_227_2', 'R_0_0_229_2', 'R_0_0_230_2', 'R_0_0_231_2', 'R_0_0_232_2', 'R_0_0_233_2', 'R
_0_0_234_2', 'R_0_0_235_2', 'R_0_0_236_2', 'R_0_0_237_2', 'R_0_1_101_2', 'R_0_1_102_2', 'R_0_1_1
03_2', 'R_0_1_104_2', 'R_0_1_105_2', 'R_0_1_106_2', 'R_0_1_107_2', 'R_0_1_108_2', 'R_0_1_109_2',
'R_0_1_110_2', 'R_0_1_111_2', 'R_0_1_112_2', 'R_0_1_113_2', 'R_0_1_114_2', 'R_0_1_115_2', 'R_0_1
_116_2', 'R_0_1_117_2', 'R_0_1_118_2', 'R_0_1_119_2', 'R_0_1_120_2', 'R_0_1_121_2', 'R_0_1_122_
2', 'R_0_1_123_2', 'R_0_1_124_2', 'R_0_1_125_2', 'R_0_1_126_2', 'R_0_1_127_2', 'R_0_1_128_2', 'R
_0_1_129_2', 'R_0_1_130_2', 'R_0_1_136_2', 'R_0_1_137_2', 'R_0_1_138_2', 'R_0_1_201_2', 'R_0_1_2
02_2', 'R_0_1_203_2', 'R_0_1_204_2', 'R_0_1_205_1', 'R_0_1_205_2', 'R_0_1_206_2', 'R_0_1_207_2',
'R_0_1_208_2', 'R_0_1_209_2', 'R_0_1_210_2', 'R_0_1_211_2', 'R_0_1_212_2', 'R_0_1_213_2', 'R_0_1
_214_2', 'R_0_1_215_2', 'R_0_1_216_2', 'R_0_1_217_2', 'R_0_1_218_2', 'R_0_1_219_2', 'R_0_1_220_
2', 'R_0_1_221_2', 'R_0_1_222_2', 'R_0_1_223_2', 'R_0_1_224_2', 'R_0_1_225_2', 'R_0_1_226_2', 'R
_0_1_227_2', 'R_0_1_228_2', 'R_0_1_229_2', 'R_0_1_230_2', 'R_0_1_233_2', 'R_0_1_234_2', 'R_0_1_2
35_2', 'R_0_2_101_2', 'R_0_2_102_2', 'R_0_2_103_2', 'R_0_2_104_2', 'R_0_2_105_2', 'R_0_2_106_2',
'R_0_2_107_2', 'R_0_2_108_2', 'R_0_2_109_2', 'R_0_2_110_2', 'R_0_2_111_2', 'R_0_2_112_2', 'R_0_2
_113_2', 'R_0_2_114_2', 'R_0_2_115_2', 'R_0_2_117_2', 'R_0_2_118_2', 'R_0_2_119_2', 'R_0_2_120_
2', 'R_0_2_121_2', 'R_0_2_122_2', 'R_0_2_123_2', 'R_0_2_124_2', 'R_0_2_125_2', 'R_0_2_126_2', 'R
_0_2_127_2', 'R_0_2_128_1', 'R_0_2_128_2', 'R_0_2_129_2', 'R_0_2_130_2', 'R_0_2_132_2', 'R_0_2_1
33_2', 'R_0_2_134_2', 'R_0_2_138_2', 'R_0_2_139_2', 'R_0_2_140_2', 'R_0_2_201_2', 'R_0_2_202_2',
'R_0_2_203_2', 'R_0_2_204_2', 'R_0_2_205_2', 'R_0_2_206_2', 'R_0_2_207_2', 'R_0_2_208_2', 'R_0_2
_209_2', 'R_0_2_210_2', 'R_0_2_211_2', 'R_0_2_212_2', 'R_0_2_213_2', 'R_0_2_214_1', 'R_0_2_214_
2', 'R_0_2_216_2', 'R_0_2_217_2', 'R_0_2_218_2', 'R_0_2_219_2', 'R_0_2_220_2', 'R_0_2_221_2', 'R
_0_2_222_2', 'R_0_2_223_2', 'R_0_2_224_2', 'R_0_2_225_2', 'R_0_2_226_2', 'R_0_2_227_2', 'R_0_2_2
28_2', 'R_0_2_229_2', 'R_0_2_230_2', 'R_0_2_231_2', 'R_0_2_234_2', 'R_0_2_235_2', 'R_0_2_241_2',
'R_0_3_101_2', 'R_0_3_102_2', 'R_0_3_103_2', 'R_0_3_104_2', 'R_0_3_105_2', 'R_0_3_106_2', 'R_0_3
_107_2', 'R_0_3_108_2', 'R_0_3_109_2', 'R_0_3_110_2', 'R_0_3_111_2', 'R_0_3_112_2', 'R_0_3_113_
2', 'R_0_3_114_2', 'R_0_3_115_2', 'R_0_3_116_2', 'R_0_3_117_2', 'R_0_3_118_2', 'R_0_3_119_2', 'R
_0_3_120_2', 'R_0_3_121_2', 'R_0_3_122_2', 'R_0_3_123_2', 'R_0_3_124_2', 'R_0_3_125_2', 'R_0_3_1
26_2', 'R_0_3_127_2', 'R_0_3_128_2', 'R_0_3_129_2', 'R_0_3_130_2', 'R_0_3_131_2', 'R_0_3_135_2',
'R_0_3_136_2', 'R_0_3_137_2', 'R_0_3_201_2', 'R_0_3_202_2', 'R_0_3_203_2', 'R_0_3_204_2', 'R_0_3
_205_2', 'R_0_3_206_2', 'R_0_3_207_2', 'R_0_3_208_2', 'R_0_3_209_2', 'R_0_3_210_2', 'R_0_3_211_
2', 'R_0_3_212_2', 'R_0_3_213_2', 'R_0_3_214_2', 'R_0_3_215_2', 'R_0_3_216_2', 'R_0_3_217_2', 'R
_0_3_218_2', 'R_0_3_219_2', 'R_0_3_220_2', 'R_0_3_221_2', 'R_0_3_222_2', 'R_0_3_223_2', 'R_0_3_2
24_2', 'R_0_3_225_2', 'R_0_3_226_2', 'R_0_3_227_2', 'R_0_3_228_2', 'R_0_3_229_2', 'R_0_3_230_2',
'R_0_3_231_2', 'R_0_3_234_2', 'R_0_3_235_2', 'R_0_3_236_2'

Pre-processing: centered (520), scaled (520)
Resampling: Cross-Validated (10 fold, repeated 3 times)
Summary of sample sizes: 4721, 4723, 4724, 4719, 4728, 4728, ...
Resampling results across tuning parameters:

  k   Accuracy   Kappa
   5  0.5331114  0.5312129
```

```
   7  0.5137396  0.5117584
   9  0.4979818  0.4959351
  11  0.4886605  0.4865756
  13  0.4728360  0.4706896
  15  0.4611507  0.4589529
  17  0.4443897  0.4421246
  19  0.4239463  0.4215928
  21  0.4065500  0.4041257
  23  0.3902922  0.3877968
  25  0.3755103  0.3729502
  27  0.3605153  0.3578814
  29  0.3444666  0.3417561
  31  0.3309418  0.3281674
  33  0.3143163  0.3114632
  35  0.2982416  0.2953115
  37  0.2898809  0.2869075
  39  0.2773618  0.2743245
  41  0.2668011  0.2637148
  43  0.2566549  0.2535136

Accuracy was used to select the optimal model using the largest value.
The final value used for the model was k = 5.
```

Hide

```
ModelData <- resamples(list(RF = RF_Fit3, KNN = knnFit1, C50 = C50Fit1))
summary(ModelData)
```

```

Call:
summary.resamples(object = ModelData)

Models: RF, KNN, C50
Number of resamples: 30

Accuracy
         Min.   1st Qu.    Median      Mean   3rd Qu.      Max. NA's
RF  0.7005650 0.7234963 0.7326427 0.7347186 0.7458378 0.7740113    0
KNN 0.4971209 0.5208729 0.5301417 0.5331114 0.5472822 0.5727969    0
C50 0.7142857 0.7436695 0.7557172 0.7529214 0.7624909 0.7817837    0

Kappa
         Min.   1st Qu.    Median      Mean   3rd Qu.      Max. NA's
RF  0.6993248 0.7223434 0.7315311 0.7336129 0.7447809 0.7730688    0
KNN 0.4950449 0.5189198 0.5282289 0.5312129 0.5454369 0.5710685    0
C50 0.7131304 0.7426229 0.7547239 0.7519144 0.7615170 0.7808938    0
```