

Case Study 3 Model

```
library(randomForest)

## randomForest 4.6-14
## Type rfNews() to see new features/changes/bug fixes.

library(MASS)
library(cvAUC)

## Loading required package: ROCR
## Warning: package 'ROCR' was built under R version 3.6.2
## Loading required package: data.table
##
## cvAUC version: 1.1.0
## Notice to cvAUC users: Major speed improvements in version 1.1.0
##

load("final_data.rda")
data <- data.frame(new_data)

summary(data)
```

##	id	label	Age	Height
##	Min. : 1.000	Min. :0.0000	Min. :24.00	Min. :165.0
##	1st Qu.: 4.000	1st Qu.:0.0000	1st Qu.:26.00	1st Qu.:172.0
##	Median : 8.000	Median :0.0000	Median :27.00	Median :178.0
##	Mean : 8.041	Mean :0.2217	Mean :27.47	Mean :177.6
##	3rd Qu.:12.000	3rd Qu.:0.0000	3rd Qu.:28.00	3rd Qu.:184.0
##	Max. :15.000	Max. :1.0000	Max. :35.00	Max. :189.0
##	Weight	Gender	ACC_wrist_mean	ACC_wrist_sd
##	Min. :54.00	Min. :0.000	Min. :62.18	Min. : 0.01898
##	1st Qu.:66.00	1st Qu.:1.000	1st Qu.:62.93	1st Qu.: 0.42623
##	Median :75.00	Median :1.000	Median :63.23	Median : 1.18674
##	Mean :73.12	Mean :0.866	Mean :63.64	Mean : 2.06887
##	3rd Qu.:80.00	3rd Qu.:1.000	3rd Qu.:64.44	3rd Qu.: 2.68943
##	Max. :90.00	Max. :1.000	Max. :67.96	Max. :17.68095
##	BVP_mean	BVP_sd	BVP_HR	EDA_wrist_mean
##	Min. : -10.794000	Min. : 2.383	Min. : 47.0	Min. : 0.09245
##	1st Qu.: -0.335750	1st Qu.: 17.895	1st Qu.: 69.0	1st Qu.: 0.31399
##	Median : 0.005333	Median : 33.912	Median : 75.0	Median : 0.77401
##	Mean : 0.008307	Mean : 43.700	Mean : 75.6	Mean : 1.81387
##	3rd Qu.: 0.351833	3rd Qu.: 55.378	3rd Qu.: 81.0	3rd Qu.: 2.50610
##	Max. : 11.532292	Max. :311.408	Max. :138.0	Max. :15.63060
##	EDA_wrist_sd	EDA_wrist_min	EDA_wrist_max	EDA_wrist_range
##	Min. :0.000889	Min. : 0.05527	Min. : 0.1015	Min. :0.005125
##	1st Qu.:0.004387	1st Qu.: 0.30089	1st Qu.: 0.3239	1st Qu.:0.023029

```

## Median :0.014692 Median : 0.71665 Median : 0.8319 Median :0.071646
## Mean :0.051516 Mean : 1.70082 Mean : 1.9368 Mean :0.236023
## 3rd Qu.:0.055648 3rd Qu.: 2.36744 3rd Qu.: 2.6814 3rd Qu.:0.250763
## Max. :1.353600 Max. :14.67037 Max. :15.9215 Max. :4.653134
## EDA_wrist_slope Temp_wrist_mean Temp_wrist_sd Temp_wrist_min
## Min. :-1.825027 Min. :29.01 Min. :0.008575 Min. :28.97
## 1st Qu.: -0.003838 1st Qu.:31.70 1st Qu.:0.016053 1st Qu.:31.65
## Median : -0.001279 Median :32.95 Median :0.022193 Median :32.91
## Mean : -0.000508 Mean :32.75 Mean :0.032356 Mean :32.68
## 3rd Qu.: 0.001280 3rd Qu.:33.85 3rd Qu.:0.035804 3rd Qu.:33.75
## Max. : 2.261773 Max. :35.93 Max. :0.302920 Max. :35.91
## Temp_wrist_max Temp_wrist_range Temp_wrist_slope ACC_chest_mean
## Min. :29.05 Min. :0.0200 Min. : -1.00e-01 Min. :0.9019
## 1st Qu.:31.75 1st Qu.:0.0700 1st Qu.: -2.00e-02 1st Qu.:0.9200
## Median :32.99 Median :0.1000 Median : 0.00e+00 Median :0.9285
## Mean :32.81 Mean :0.1234 Mean : -9.06e-06 Mean :0.9333
## 3rd Qu.:33.95 3rd Qu.:0.1400 3rd Qu.: 2.00e-02 3rd Qu.:0.9434
## Max. :35.97 Max. :0.9600 Max. : 8.00e-02 Max. :0.9833
## ACC_chest_sd ECG_mean ECG_sd ECG_HR
## Min. :0.002264 Min. : -0.060135 Min. :0.07933 Min. : 47.00
## 1st Qu.:0.004257 1st Qu.: -0.008511 1st Qu.:0.21325 1st Qu.: 66.00
## Median :0.006064 Median : 0.002120 Median :0.26234 Median : 75.00
## Mean :0.008955 Mean : 0.002572 Mean :0.25821 Mean : 80.82
## 3rd Qu.:0.010841 3rd Qu.: 0.012479 3rd Qu.:0.30286 3rd Qu.: 88.00
## Max. :0.093678 Max. : 0.079029 Max. :0.51159 Max. :149.00
## EDA_chest_mean EDA_chest_sd EDA_chest_min EDA_chest_max
## Min. : 0.4691 Min. :0.00595 Min. : 0.285 Min. : 0.5604
## 1st Qu.: 2.0327 1st Qu.:0.01359 1st Qu.: 1.954 1st Qu.: 2.1214
## Median : 3.6993 Median :0.01903 Median : 3.592 Median : 3.7609
## Mean : 4.6070 Mean :0.04767 Mean : 4.517 Mean : 4.7390
## 3rd Qu.: 6.4996 3rd Qu.:0.04072 3rd Qu.: 6.422 3rd Qu.: 6.5853
## Max. :20.2740 Max. :1.80241 Max. :19.970 Max. :21.1349
## EDA_chest_range EDA_chest_slope EMG_mean EMG_sd
## Min. :0.03471 Min. : -0.6877899 Min. : -0.8151398 Min. :0.005057
## 1st Qu.:0.08774 1st Qu.: -0.0099182 1st Qu.: -0.0093842 1st Qu.:0.008527
## Median :0.12398 Median : -0.0007629 Median : -0.0000458 Median :0.010563
## Mean :0.22175 Mean : -0.0001620 Mean : -0.0000412 Mean :0.012033
## 3rd Qu.:0.19150 3rd Qu.: 0.0080109 3rd Qu.: 0.0093384 3rd Qu.:0.013822
## Max. :5.04379 Max. : 0.8102417 Max. : 1.0171051 Max. :0.108135
## EMG_range Resp_Volume Resp_range breath_rate
## Min. :0.02820 Min. : 48.01 Min. : 1.376 Min. : 6.00
## 1st Qu.:0.05777 1st Qu.: 444.60 1st Qu.:10.791 1st Qu.:13.00
## Median :0.07425 Median : 584.53 Median :16.168 Median :15.00
## Mean :0.09356 Mean : 635.47 Mean :18.768 Mean :14.73
## 3rd Qu.:0.10936 3rd Qu.: 758.93 3rd Qu.:24.377 3rd Qu.:17.00
## Max. :1.83792 Max. :2681.30 Max. :72.710 Max. :28.00
## Temp_chest_mean Temp_chest_sd Temp_chest_min Temp_chest_max
## Min. :28.20 Min. :0.009757 Min. :28.13 Min. :28.34
## 1st Qu.:33.60 1st Qu.:0.027085 1st Qu.:33.50 1st Qu.:33.70
## Median :34.18 Median :0.031548 Median :34.08 Median :34.28
## Mean :33.88 Mean :0.042550 Mean :33.77 Mean :33.99
## 3rd Qu.:34.59 3rd Qu.:0.039311 3rd Qu.:34.49 3rd Qu.:34.70
## Max. :35.63 Max. :1.844897 Max. :35.52 Max. :35.74
## Temp_chest_range Temp_chest_slope

```

```

## Min. :0.06625 Min. :-0.3031310
## 1st Qu.:0.16806 1st Qu.: -0.0155620
## Median :0.19464 Median : 0.0000000
## Mean :0.22308 Mean : 0.0003405
## 3rd Qu.:0.22629 3rd Qu.: 0.0168150
## Max. :5.47601 Max. : 0.3354500

nrow(subset(data, label == 1))/nrow(data)

## [1] 0.2216921

nrow(subset(data, label == 0))/nrow(data)

## [1] 0.7783079

data$label = as.factor(data$label)

test = subset(data, id == 14)
train = subset(data, id != 14)
nrow(data)

## [1] 179817

nrow(test)

## [1] 12052

nrow(train)

## [1] 167765

colnames(train)

## [1] "id" "label" "Age" "Height"
## [5] "Weight" "Gender" "ACC_wrist_mean" "ACC_wrist_sd"
## [9] "BVP_mean" "BVP_sd" "BVP_HR" "EDA_wrist_mean"
## [13] "EDA_wrist_sd" "EDA_wrist_min" "EDA_wrist_max" "EDA_wrist_range"
## [17] "EDA_wrist_slope" "Temp_wrist_mean" "Temp_wrist_sd" "Temp_wrist_min"
## [21] "Temp_wrist_max" "Temp_wrist_range" "Temp_wrist_slope" "ACC_chest_mean"
## [25] "ACC_chest_sd" "ECG_mean" "ECG_sd" "ECG_HR"
## [29] "EDA_chest_mean" "EDA_chest_sd" "EDA_chest_min" "EDA_chest_max"
## [33] "EDA_chest_range" "EDA_chest_slope" "EMG_mean" "EMG_sd"
## [37] "EMG_range" "Resp_Volume" "Resp_range" "breath_rate"
## [41] "Temp_chest_mean" "Temp_chest_sd" "Temp_chest_min" "Temp_chest_max"
## [45] "Temp_chest_range" "Temp_chest_slope"

personal = colnames(train)[3:6]
wrist_acc = colnames(train)[7:8]
chest_acc = colnames(train)[24:25]
wrist_bvp = colnames(train)[9:11]
wrist_eda = colnames(train)[12:17]
wrist_temp = colnames(train)[18:23]
wrist_physio = colnames(train)[9:23]
chest_ecg = colnames(train)[26:28]
chest_eda = colnames(train)[29:34]
chest_emg = colnames(train)[35:37]
chest_resp = colnames(train)[38:40]
chest_temp = colnames(train)[41:46]
chest_physio = colnames(train)[26:46]
all_wrist = colnames(train)[7:23]

```

```

all_chest = colnames(train)[24:46]
all_physio = colnames(train)[c(9:23,26:46)]
all_modalities = colnames(train)[c(7:46)]

predictor_vars <- c("personal", "wrist_acc", "chest_acc", "wrist_bvp", "wrist_eda", "wrist_temp", "wrist_hr")

test_sample = test
train_sample = train

set.seed(1)
train_indices = sample(nrow(train), 1000)
train_sample = train[train_indices,]
# Run this instead to train on the full train set
# test_sample = train

rf <- function(train_sample, test_sample, predictors){
  set.seed(1)
  model_rf <- randomForest(as.formula(paste("label ~ ", paste(predictors, collapse = ' + '))), ntree = 1000)
  predict_rf <- predict(model_rf, test_sample)
  cat("Accuracy is", mean(test_sample$label == predict_rf)*100, "% \n")
  cat("AUROC is", AUC(as.numeric(as.character(predict_rf))), as.numeric(as.character(test_sample$label))), "\n")
  if (mean(test_sample$label == predict_rf) == 1){
    df <- data.frame(importance(model_rf, type = 1))
    print(df)
    cat('\n')
  }
}

print_baseline <- function(test_sample){
  predict_rf <- rep(0,nrow(test_sample))
  cat("Accuracy is", mean(test_sample$label == predict_rf)*100, "% \n")
  cat("AUROC is", AUC(as.numeric(as.character(predict_rf))), as.numeric(as.character(test_sample$label))), "\n")
}

print_baseline(test_sample)

## Accuracy is 77.66346 %
## AUROC is 0.5
##

for (i in 1:length(predictor_vars)){
  cat("Predictors: ", predictor_vars[i], "\n")
  rf(train_sample, test_sample, eval(parse(text = predictor_vars[i])))
}

## Predictors: personal
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: wrist_acc
## Accuracy is 72.01294 %
## AUROC is 0.514696
##
## Predictors: chest_acc
## Accuracy is 91.32094 %
## AUROC is 0.8467388

```

```

##
## Predictors:  wrist_bvp
## Accuracy is 89.39595 %
## AUROC is 0.7755971
##
## Predictors:  wrist_eda
## Accuracy is 85.84467 %
## AUROC is 0.6831352
##
## Predictors:  wrist_temp
## Accuracy is 76.00398 %
## AUROC is 0.4893162
##
## Predictors:  wrist_physio
## Accuracy is 98.48988 %
## AUROC is 0.9661961
##
## Predictors:  chest_ecg
## Accuracy is 82.04447 %
## AUROC is 0.8844017
##
## Predictors:  chest_eda
## Accuracy is 87.58712 %
## AUROC is 0.7823438
##
## Predictors:  chest_emg
## Accuracy is 75.307 %
## AUROC is 0.5258472
##
## Predictors:  chest_resp
## Accuracy is 86.89844 %
## AUROC is 0.7839966
##
## Predictors:  chest_temp
## Accuracy is 72.4693 %
## AUROC is 0.4893183
##
## Predictors:  chest_physio
## Accuracy is 100 %
## AUROC is 1
##
##                               MeanDecreaseAccuracy
## ECG_mean                      9.1944086
## ECG_sd                       25.0320995
## ECG_HR                       38.7903753
## EDA_chest_mean               17.8834158
## EDA_chest_sd                 28.2825712
## EDA_chest_min                16.4386864
## EDA_chest_max                17.5592515
## EDA_chest_range              26.9618053
## EDA_chest_slope              5.8585576
## EMG_mean                     3.8458007
## EMG_sd                      16.1981758
## EMG_range                   13.2845247

```

```

## Resp_Volume                21.2544550
## Resp_range                 23.2212202
## breath_rate                18.2416750
## Temp_chest_mean            20.4689125
## Temp_chest_sd              11.5026323
## Temp_chest_min             19.7646250
## Temp_chest_max             23.5620817
## Temp_chest_range           9.8265898
## Temp_chest_slope           0.7674039
##
## Predictors:  all_wrist
## Accuracy is 98.07501 %
## AUROC is 0.9569094
##
## Predictors:  all_chest
## Accuracy is 100 %
## AUROC is 1
##
##                               MeanDecreaseAccuracy
## ACC_chest_mean                18.505801
## ACC_chest_sd                  35.536270
## ECG_mean                      7.156738
## ECG_sd                       23.654543
## ECG_HR                       33.345066
## EDA_chest_mean                16.370751
## EDA_chest_sd                  24.659602
## EDA_chest_min                 16.738178
## EDA_chest_max                 16.022553
## EDA_chest_range               25.706254
## EDA_chest_slope               4.080738
## EMG_mean                     2.278767
## EMG_sd                       13.921465
## EMG_range                     10.607432
## Resp_Volume                  16.883980
## Resp_range                    16.961546
## breath_rate                   12.709862
## Temp_chest_mean              18.018281
## Temp_chest_sd                 9.543670
## Temp_chest_min               18.337576
## Temp_chest_max               17.299264
## Temp_chest_range              9.176240
## Temp_chest_slope             1.655409
##
## Predictors:  all_physio
## Accuracy is 100 %
## AUROC is 1
##
##                               MeanDecreaseAccuracy
## BVP_mean                      5.1902428
## BVP_sd                       12.9703019
## BVP_HR                       21.6696975
## EDA_wrist_mean                13.3486438
## EDA_wrist_sd                  16.6281973
## EDA_wrist_min                 12.5954115

```

```

## EDA_wrist_max          12.5525290
## EDA_wrist_range        21.9972161
## EDA_wrist_slope        9.6959078
## Temp_wrist_mean        16.5087362
## Temp_wrist_sd          10.2185320
## Temp_wrist_min         15.7221794
## Temp_wrist_max         15.8384320
## Temp_wrist_range        8.8685467
## Temp_wrist_slope        0.3267863
## ECG_mean               5.2190745
## ECG_sd                 19.2078886
## ECG_HR                 32.0590820
## EDA_chest_mean         11.9365614
## EDA_chest_sd           17.1995642
## EDA_chest_min          13.3069963
## EDA_chest_max          12.7629753
## EDA_chest_range        15.2868472
## EDA_chest_slope        2.3578981
## EMG_mean               2.0624674
## EMG_sd                 11.1473068
## EMG_range              8.4404319
## Resp_Volume            15.4262581
## Resp_range             13.3920321
## breath_rate            11.5785588
## Temp_chest_mean        14.7093078
## Temp_chest_sd           7.0202151
## Temp_chest_min         13.0545724
## Temp_chest_max         15.0042640
## Temp_chest_range        7.6417845
## Temp_chest_slope        0.2047934
##
## Predictors:  all_modalities
## Accuracy is 100 %
## AUROC is 1
##
##                               MeanDecreaseAccuracy
## ACC_wrist_mean          11.5297374
## ACC_wrist_sd            18.6395392
## BVP_mean                4.8201706
## BVP_sd                  10.9503515
## BVP_HR                  18.1121031
## EDA_wrist_mean          12.8646034
## EDA_wrist_sd            15.2604242
## EDA_wrist_min           12.9207128
## EDA_wrist_max           12.6665262
## EDA_wrist_range         17.5800369
## EDA_wrist_slope         8.8611327
## Temp_wrist_mean         15.0648210
## Temp_wrist_sd           10.3008997
## Temp_wrist_min          15.0049072
## Temp_wrist_max          15.5572699
## Temp_wrist_range        8.7126559
## Temp_wrist_slope        -0.9031157
## ACC_chest_mean          14.6436046

```

```
## ACC_chest_sd          22.2688727
## ECG_mean             5.2207103
## ECG_sd               17.5521278
## ECG_HR               25.6752359
## EDA_chest_mean       12.6535387
## EDA_chest_sd         17.6136293
## EDA_chest_min        13.0799608
## EDA_chest_max        12.4364134
## EDA_chest_range      15.6900874
## EDA_chest_slope      4.5227018
## EMG_mean             0.8656493
## EMG_sd               8.3112467
## EMG_range            6.4201386
## Resp_Volume          12.1790078
## Resp_range           10.5988051
## breath_rate          10.7947002
## Temp_chest_mean      13.0184214
## Temp_chest_sd        7.7159645
## Temp_chest_min       13.3668074
## Temp_chest_max       13.6918899
## Temp_chest_range     6.9245280
## Temp_chest_slope     -0.3812264
```

```
for (i in 2:length(predictor_vars)){
  cat("Predictors: personal +", predictor_vars[i], "\n")
  rf(train_sample, test_sample, c(eval(parse(text = predictor_vars[1])), eval(parse(text = predictor_var
})
```

```
## Predictors: personal + wrist_acc
## Accuracy is 78.36044 %
## AUROC is 0.5158664
##
## Predictors: personal + chest_acc
## Accuracy is 81.76236 %
## AUROC is 0.5953259
##
## Predictors: personal + wrist_bvp
## Accuracy is 95.9011 %
## AUROC is 0.9736111
##
## Predictors: personal + wrist_eda
## Accuracy is 87.6452 %
## AUROC is 0.7234398
##
## Predictors: personal + wrist_temp
## Accuracy is 77.64686 %
## AUROC is 0.4998932
##
## Predictors: personal + wrist_physio
## Accuracy is 99.07899 %
## AUROC is 0.9793834
##
## Predictors: personal + chest_ecg
## Accuracy is 80.01991 %
## AUROC is 0.8713675
```



```

##
## Predictors: personal + chest_eda
## Accuracy is 93.27912 %
## AUROC is 0.8498189
##
## Predictors: personal + chest_emg
## Accuracy is 77.58048 %
## AUROC is 0.5011859
##
## Predictors: personal + chest_resp
## Accuracy is 93.81016 %
## AUROC is 0.9119863
##
## Predictors: personal + chest_temp
## Accuracy is 77.35646 %
## AUROC is 0.5092704
##
## Predictors: personal + chest_physio
## Accuracy is 100 %
## AUROC is 1
##
##
##           MeanDecreaseAccuracy
## Age                10.1842887
## Height             13.4742560
## Weight             16.2952322
## Gender              5.9430760
## ECG_mean            9.9096055
## ECG_sd              26.3261286
## ECG_HR              41.1159521
## EDA_chest_mean      16.3722350
## EDA_chest_sd        28.8708497
## EDA_chest_min       16.1055584
## EDA_chest_max       15.5656027
## EDA_chest_range     27.4060284
## EDA_chest_slope      5.7664941
## EMG_mean            3.2901841
## EMG_sd              14.8999984
## EMG_range           12.5665622
## Resp_Volume         19.4347022
## Resp_range          21.5662945
## breath_rate         16.6137066
## Temp_chest_mean     21.4057921
## Temp_chest_sd       11.3477263
## Temp_chest_min      19.2522584
## Temp_chest_max      22.1404247
## Temp_chest_range     9.5071240
## Temp_chest_slope    0.4247978
##
## Predictors: personal + all_wrist
## Accuracy is 98.78029 %
## AUROC is 0.9726969
##
## Predictors: personal + all_chest
## Accuracy is 100 %

```

```

## AUROC is 1
##
##                               MeanDecreaseAccuracy
## Age                           10.0655628
## Height                        12.5650463
## Weight                        12.9337844
## Gender                        5.4902022
## ACC_chest_mean                19.9669458
## ACC_chest_sd                  32.6726314
## ECG_mean                      7.0881633
## ECG_sd                        24.7299534
## ECG_HR                        31.8263929
## EDA_chest_mean                16.6494455
## EDA_chest_sd                  27.0082500
## EDA_chest_min                 16.1246031
## EDA_chest_max                 15.7406402
## EDA_chest_range               25.5083904
## EDA_chest_slope               6.8338910
## EMG_mean                      -1.1948827
## EMG_sd                        12.1666646
## EMG_range                     11.5572864
## Resp_Volume                   16.9370236
## Resp_range                    16.6410147
## breath_rate                   13.1641246
## Temp_chest_mean               20.1403430
## Temp_chest_sd                  8.2085096
## Temp_chest_min                17.1108865
## Temp_chest_max                18.7743842
## Temp_chest_range              8.7151822
## Temp_chest_slope              0.9414631
##
## Predictors: personal + all_physio
## Accuracy is 100 %
## AUROC is 1
##
##                               MeanDecreaseAccuracy
## Age                           8.4779440
## Height                        9.0921047
## Weight                        9.9320793
## Gender                        4.6431506
## BVP_mean                      6.4744976
## BVP_sd                        12.6151169
## BVP_HR                        18.8952805
## EDA_wrist_mean                12.6208762
## EDA_wrist_sd                  14.9356950
## EDA_wrist_min                 12.3449662
## EDA_wrist_max                 12.9172447
## EDA_wrist_range               20.4316439
## EDA_wrist_slope               8.0766320
## Temp_wrist_mean               15.9757508
## Temp_wrist_sd                  10.2613390
## Temp_wrist_min                15.7765025
## Temp_wrist_max                15.4282995
## Temp_wrist_range              8.0115150

```

```

## Temp_wrist_slope          0.6057994
## ECG_mean                  4.2731884
## ECG_sd                    20.8140657
## ECG_HR                    32.8912375
## EDA_chest_mean            12.2021209
## EDA_chest_sd              17.3834783
## EDA_chest_min             12.3317824
## EDA_chest_max             12.1374888
## EDA_chest_range           15.9742233
## EDA_chest_slope           5.0479395
## EMG_mean                  2.1504473
## EMG_sd                    9.9017597
## EMG_range                 6.5314745
## Resp_Volume               14.9979715
## Resp_range                12.9124687
## breath_rate               12.1769804
## Temp_chest_mean           13.4629781
## Temp_chest_sd             7.8009873
## Temp_chest_min            12.5381746
## Temp_chest_max            13.7517824
## Temp_chest_range           7.4150509
## Temp_chest_slope          -1.3612837
##
## Predictors: personal + all_modalities
## Accuracy is 100 %
## AUROC is 1
##
##                               MeanDecreaseAccuracy
## Age                          9.7521556
## Height                       8.2137147
## Weight                      10.0052626
## Gender                      3.5823850
## ACC_wrist_mean              11.0690527
## ACC_wrist_sd                16.9938969
## BVP_mean                    4.8204221
## BVP_sd                     11.2240390
## BVP_HR                     16.9881589
## EDA_wrist_mean              11.7056152
## EDA_wrist_sd                14.8945730
## EDA_wrist_min               13.7271223
## EDA_wrist_max               12.6292392
## EDA_wrist_range             16.5418665
## EDA_wrist_slope             8.9030749
## Temp_wrist_mean             14.3668687
## Temp_wrist_sd               9.1321051
## Temp_wrist_min              14.2649729
## Temp_wrist_max              16.0602563
## Temp_wrist_range            7.1668184
## Temp_wrist_slope            1.2855250
## ACC_chest_mean              13.6029956
## ACC_chest_sd                23.3657412
## ECG_mean                    2.9865623
## ECG_sd                     17.7184739
## ECG_HR                     26.1565724

```

```
## EDA_chest_mean          12.1934769
## EDA_chest_sd            16.3688510
## EDA_chest_min          11.5781451
## EDA_chest_max          11.3053344
## EDA_chest_range        16.6042314
## EDA_chest_slope        3.8386206
## EMG_mean               4.7092855
## EMG_sd                 7.8627211
## EMG_range              7.1031113
## Resp_Volume            12.3879513
## Resp_range             11.2621574
## breath_rate            8.9101815
## Temp_chest_mean        13.7259716
## Temp_chest_sd          8.1659235
## Temp_chest_min         12.6297490
## Temp_chest_max         11.5809801
## Temp_chest_range       6.1703317
## Temp_chest_slope       -0.8928669
```

```
train_sample = train
```

LDA

```
LDA <- function(train_sample, test_sample, predictors){
  model_lda <- lda(as.formula(paste("label ~ ", paste(predictors, collapse = ' + '))), data = train_sample)
  predict_lda <- predict(model_lda, test_sample)[[1]]
  cat("Accuracy is", mean(test_sample$label == predict_lda)*100, "% \n")
  cat("AUROC is", AUC(as.numeric(as.character(predict_lda)), as.numeric(as.character(test_sample$label))))
}
```

```
for (i in 1:length(predictor_vars)){
  cat("Predictors: ", predictor_vars[i], "\n")
  LDA(train_sample, test_sample, eval(parse(text = predictor_vars[i])))
}
```

```
## Predictors:  personal
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors:  wrist_acc
## Accuracy is 76.13674 %
## AUROC is 0.5018148
##
## Predictors:  chest_acc
## Accuracy is 84.47561 %
## AUROC is 0.6524889
##
## Predictors:  wrist_bvp
## Accuracy is 97.02954 %
## AUROC is 0.9463414
##
## Predictors:  wrist_eda
## Accuracy is 77.66346 %
## AUROC is 0.5
```

```

##
## Predictors:  wrist_temp
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors:  wrist_physio
## Accuracy is 98.53966 %
## AUROC is 0.990466
##
## Predictors:  chest_ecg
## Accuracy is 100 %
## AUROC is 1
##
## Predictors:  chest_eda
## Accuracy is 80.43478 %
## AUROC is 0.5620357
##
## Predictors:  chest_emg
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors:  chest_resp
## Accuracy is 91.7607 %
## AUROC is 0.8168878
##
## Predictors:  chest_temp
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors:  chest_physio
## Accuracy is 100 %
## AUROC is 1
##
## Predictors:  all_wrist
## Accuracy is 98.57285 %
## AUROC is 0.9683181
##
## Predictors:  all_chest
## Accuracy is 100 %
## AUROC is 1
##
## Predictors:  all_physio
## Accuracy is 99.15367 %
## AUROC is 0.9945513
##
## Predictors:  all_modalities
## Accuracy is 100 %
## AUROC is 1
##
for (i in 2:length(predictor_vars)){
  cat("Predictors: personal +", predictor_vars[i], "\n")
  LDA(train_sample, test_sample, c(eval(parse(text = predictor_vars[1])), eval(parse(text = predictor_v
}

```

```

## Predictors: personal + wrist_acc
## Accuracy is 78.36044 %
## AUROC is 0.5156018
##
## Predictors: personal + chest_acc
## Accuracy is 78.74212 %
## AUROC is 0.5241456
##
## Predictors: personal + wrist_bvp
## Accuracy is 97.44441 %
## AUROC is 0.964758
##
## Predictors: personal + wrist_eda
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: personal + wrist_temp
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: personal + wrist_physio
## Accuracy is 98.84666 %
## AUROC is 0.9741828
##
## Predictors: personal + chest_ecg
## Accuracy is 100 %
## AUROC is 1
##
## Predictors: personal + chest_eda
## Accuracy is 84.16031 %
## AUROC is 0.6454309
##
## Predictors: personal + chest_emg
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: personal + chest_resp
## Accuracy is 89.48722 %
## AUROC is 0.7657316
##
## Predictors: personal + chest_temp
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: personal + chest_physio
## Accuracy is 99.90873 %
## AUROC is 0.9979569
##
## Predictors: personal + all_wrist
## Accuracy is 94.83903 %
## AUROC is 0.8844725
##
## Predictors: personal + all_chest
## Accuracy is 97.74311 %

```

```
## AUROC is 0.9494799
##
## Predictors: personal + all_physio
## Accuracy is 100 %
## AUROC is 1
##
## Predictors: personal + all_modalities
## Accuracy is 99.73448 %
## AUROC is 0.9940565
##
```

Logistic Regression

```
logistic <- function(train_sample, test_sample, predictors){
  model_logistic <- glm(as.formula(paste("label ~ ", paste(predictors, collapse = ' + '))), family=binomial)
  predict_logistic <- predict(model_logistic, test_sample)
  predict_logistic <- ifelse(predict_logistic > 0.5, 1, 0)
  cat("Accuracy is", mean(test_sample$label == predict_logistic)*100, "% \n")
  cat("AUROC is", AUC(as.numeric(as.character(predict_logistic)), as.numeric(as.character(test_sample$label))), "% \n")
}
```

```
for (i in 1:length(predictor_vars)){
  cat("Predictors: ", predictor_vars[i], "\n")
  logistic(train_sample, test_sample, eval(parse(text = predictor_vars[i])))
}
```

```
## Predictors: personal
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: wrist_acc
## Accuracy is 78.36044 %
## AUROC is 0.5156018
##
## Predictors: chest_acc
## Accuracy is 83.2227 %
## AUROC is 0.6244428
##
## Predictors: wrist_bvp
## Accuracy is 94.40757 %
## AUROC is 0.8748143
##
## Predictors: wrist_eda
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: wrist_temp
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: wrist_physio
## Accuracy is 98.57285 %
## AUROC is 0.9861809
```

```

##
## Predictors: chest_ecg
## Accuracy is 100 %
## AUROC is 1
##
## Predictors: chest_eda
## Accuracy is 79.65483 %
## AUROC is 0.5445765
##
## Predictors: chest_emg
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: chest_resp
## Accuracy is 85.67873 %
## AUROC is 0.6794205
##
## Predictors: chest_temp
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: chest_physio
## Accuracy is 100 %
## AUROC is 1
##
## Predictors: all_wrist
## Accuracy is 98.58945 %
## AUROC is 0.9685573
##
## Predictors: all_chest
## Accuracy is 100 %
## AUROC is 1
##
## Predictors: all_physio
## Accuracy is 99.3694 %
## AUROC is 0.9959402
##
## Predictors: all_modalities
## Accuracy is 99.85065 %
## AUROC is 0.9990385
##
for (i in 2:length(predictor_vars)){
  cat("Predictors: personal +", predictor_vars[i], "\n")
  logistic(train_sample, test_sample, c(eval(parse(text = predictor_vars[1])), eval(parse(text = predic
})

## Predictors: personal + wrist_acc
## Accuracy is 78.35214 %
## AUROC is 0.515416
##
## Predictors: personal + chest_acc
## Accuracy is 77.93727 %
## AUROC is 0.5061293
##

```



```
## Predictors: personal + wrist_bvp
## Accuracy is 97.71822 %
## AUROC is 0.9737983
##
## Predictors: personal + wrist_eda
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: personal + wrist_temp
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: personal + wrist_physio
## Accuracy is 96.27448 %
## AUROC is 0.976015
##
## Predictors: personal + chest_ecg
## Accuracy is 100 %
## AUROC is 1
##
## Predictors: personal + chest_eda
## Accuracy is 88.26751 %
## AUROC is 0.73737
##
## Predictors: personal + chest_emg
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: personal + chest_resp
## Accuracy is 83.48822 %
## AUROC is 0.6303863
##
## Predictors: personal + chest_temp
## Accuracy is 77.66346 %
## AUROC is 0.5
##
## Predictors: personal + chest_physio
## Accuracy is 100 %
## AUROC is 1
##
## Predictors: personal + all_wrist
## Accuracy is 98.26585 %
## AUROC is 0.9884385
##
## Predictors: personal + all_chest
## Accuracy is 100 %
## AUROC is 1
##
## Predictors: personal + all_physio
## Accuracy is 91.96814 %
## AUROC is 0.9482906
##
## Predictors: personal + all_modalities
## Accuracy is 98.77199 %
```

```
## AUROC is 0.992094
##
```

We choose the Random Forest

Cross-Validation

- We got 100 percent accuracy for predictors wrist_physio, chest_ecg, chest_physio, all_wrist, all_chest, all_physio, all_modalities
- Let us do cross validation.

```
# cv <- data.frame(matrix(ncol = 16, nrow = 14))
# rownames(cv) <- c("wrist_physio acc", "wrist_physio auc", "chest_ecg acc", "chest_ecg auc", "chest_ph
# colnames(cv) <- c("predictor", c(1:15))
# cv$predictor <- c("wrist_physio", "wrist_physio", "chest_ecg", "chest_ecg", "chest_physio", "chest_ph

# cv

# for(i in 1:7){
#   for (j in 1:15){
#     set.seed(1)
#     test = subset(data, id == j)
#     train = subset(data, id != j)
#     test_sample = test
#
#     # Run this instead to train on the full train set
#     # test_sample = train
#     train_indices = sample(nrow(train), 100)
#     train_sample = train[train_indices,]
#
#     predictor = cv$predictor[(i-1)*2+1]
#     predictors = eval(parse(text = predictor))
#     model_rf <- randomForest(as.formula(paste("label ~ ", paste(predictors, collapse = ' + '))), ntre
#     predict_rf <- predict(model_rf, test_sample)
#     acc = mean(test_sample$label == predict_rf)*100
#     auc = AUC(as.numeric(as.character(predict_rf)), as.numeric(as.character(test_sample$label)))
#
#     cv[(i-1)*2+1,j+1] <- acc
#     cv[(i-1)*2+2,j+1] <- auc
#   }
# }
```

Let us use fewer training samples

1000 Training Samples

```
# for(i in 1:7){
#   for (j in 1:15){
#     set.seed(1)
#     test = subset(data, id == j)
```

```

#   train = subset(data, id != j)
#   test_sample = test
#
#   train_indices = sample(nrow(train), 1000)
#   train_sample = train[train_indices,]
#
#   predictor = cv$predictor[(i-1)*2+1]
#   predictors = eval(parse(text = predictor))
#   model_rf <- randomForest(as.formula(paste("label ~ ", paste(predictors, collapse = ' + '))), ntree = 500)
#   predict_rf <- predict(model_rf, test_sample)
#   acc = mean(test_sample$label == predict_rf)*100
#   auc = AUC(as.numeric(predict_rf), as.numeric(test_sample$label))
#
#   cv[(i-1)*2+1,j+1] <- acc
#   cv[(i-1)*2+2,j+1] <- auc
# }
# }

# cv

```

100 Training Samples

```

# for(i in 1:7){
#   for (j in 1:15){
#     set.seed(1)
#     test = subset(data, id == j)
#     train = subset(data, id != j)
#     test_sample = test
#
#     train_indices = sample(nrow(train), 100)
#     train_sample = train[train_indices,]
#
#     predictor = cv$predictor[(i-1)*2+1]
#     predictors = eval(parse(text = predictor))
#     model_rf <- randomForest(as.formula(paste("label ~ ", paste(predictors, collapse = ' + '))), ntree = 500)
#     predict_rf <- predict(model_rf, test_sample)
#     acc = mean(test_sample$label == predict_rf)*100
#     auc = AUC(as.numeric(predict_rf), as.numeric(test_sample$label))
#
#     cv[(i-1)*2+1,j+1] <- acc
#     cv[(i-1)*2+2,j+1] <- auc
#   }
# }

# cv

```