Predicting sepsis from heart rate variability analysis

Erin Lopez

2022-05-01

Introduction Sepsis is a rapidly developing, life-threatening response to infection characterized by tachycardia, confusion, extreme pain, fever, shortness of breath, and clammy skin. In hospitals, 1 in 3 patients who die have sepsis (CDC 2021). Quickly identifying the symptoms and administering treatment can greatly increase the odds of survival (CDC 2021, Huang 2019).

Heart rate variability (HRV) is the milliseconds long fluctuations in time between consecutive heartbeats. HRV parameters including time, frequency, and non-linear domain variables, can all be affected by bodily stress and infection. Previous studies have shown differences in heart rate variability parameters in response to strenuous exercise (Gronwald 2020) or COVID19 infection (Mol 2021).

This study aims to predict the onset of sepsis from heart rate variability analysis.

Data Source The dataset from Kuan-Fu Chen at Chang Gung University College of Medicine was accessed via Mendeley Data (Chen, 2020). The data contains 4314 observations of 58 variables. The binary Sepsis3 variable denotes sepsis as '1'. HRV time, frequency, and non-linear domain measures are included. Frequency domain parameters are high frequency (HF), low frequency (LF), and very low frequency (VLF) power. Non-linear domain variables include poincare plot SD1, poincare plot SD2, detrended fluctuation analysis alpha 1 (DFAa1) and DFAa2.

Libraries The following libraries were utilized: rmarkdown, knitr, dplyr, DataExplorer, caret, rpart, randomForest, readxl, psych, ggfortify, rpart.plot.

Methods *Data Preparation*

Data was split into training and testing sets at ratios of 40/60, 50/50, 60/40, and 70/30. After initial testing, the 70/30 ratio was chosen to proceed with. 3048 of the 4314 observations were split into the training set. The remaining 1266 observations were split into the testing set.

Exploratory Analysis

The data was grouped by sepsis status to calculate the mean of each of the 57 predictors. Student's T-test for each predictor was calculated to identify, with 95% confidence, statistically significant differences between sepsis positive and negative groups. A correlation plot and boxplots of each predictor were created to visualize predictor differences and interactions.

Classification Trees

rpart

Classification trees were generated using the rpart library. Each model was trained using the training dataset of predictors. Sepsis predictions were produced from the predictors for both the training and testing set. A table stored the predicted and actual sepsis outcomes for validation.

Random Forest

Random forest classification was generated using the training dataset to predict the sepsis outcome. The additional parameters of importance and mtry were included to assess the importance of predictors and

determine the number of randomly sampled variables at each node, respectively. The random forest classification predicted the sepsis outcome based on training or testing data and stored the predicted versus actual values in a table for validation.

Model Evaluation

Each model was evaluated with a confusion matrix to display the predicted and actual sepsis outcomes and report the sensitivity, specificity, accuracy, kappa, and F1 values.

Results *Exploratory Analysis*

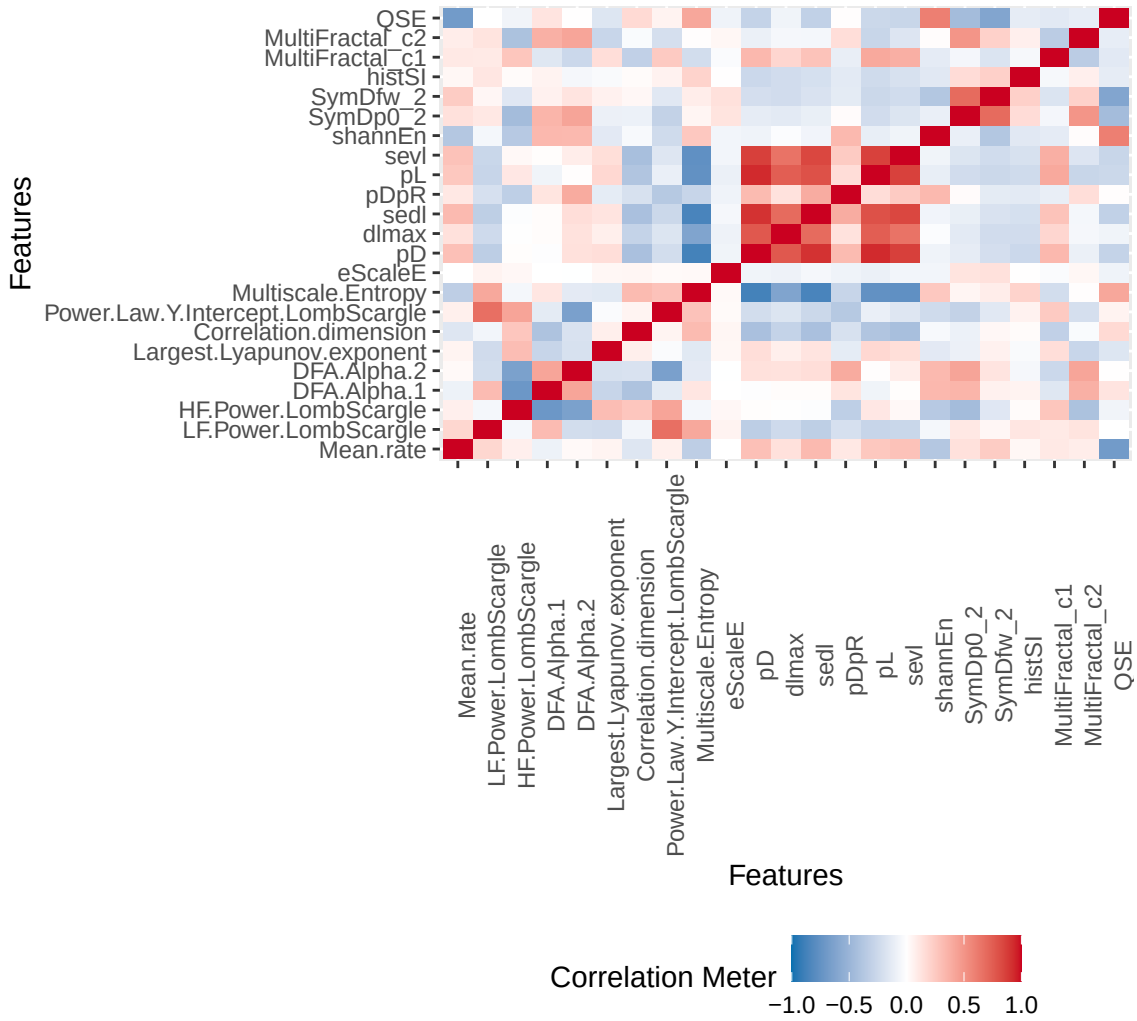
The mean value of each predictor calculated for the sepsis positive and sepsis negative groups, identifies 23 statistically significant variables with p-values below 0.05 (Fig. 1).

Figure 1. Predictors with statistically significant differences in mean between sepsis positive and sepsis negative groups.

Significant Predictors
Mean.rate
LF.Power.LombScargle
HF.Power.LombScargle
DFA.Alpha.1
DFA.Alpha.2
Largest.Lyapunov.exponent
Correlation.dimension
Power.Law.Y.Intercept.LombScargle
Multiscale.Entropy
eScaleE
pD
dlmax
sedl
pDpR
pL
sevl
shannEn
SymDp0_2
SymDfw_2
histSI
MultiFractal_c1
MultiFractal_c2
QSE

A correlation plot of the 23 significant variables shows a significant positive correlation between sevl, pL, sedl, dlmax, and pD. This subset of variables shows a significant negative correlation with Multiscale.Entropy.

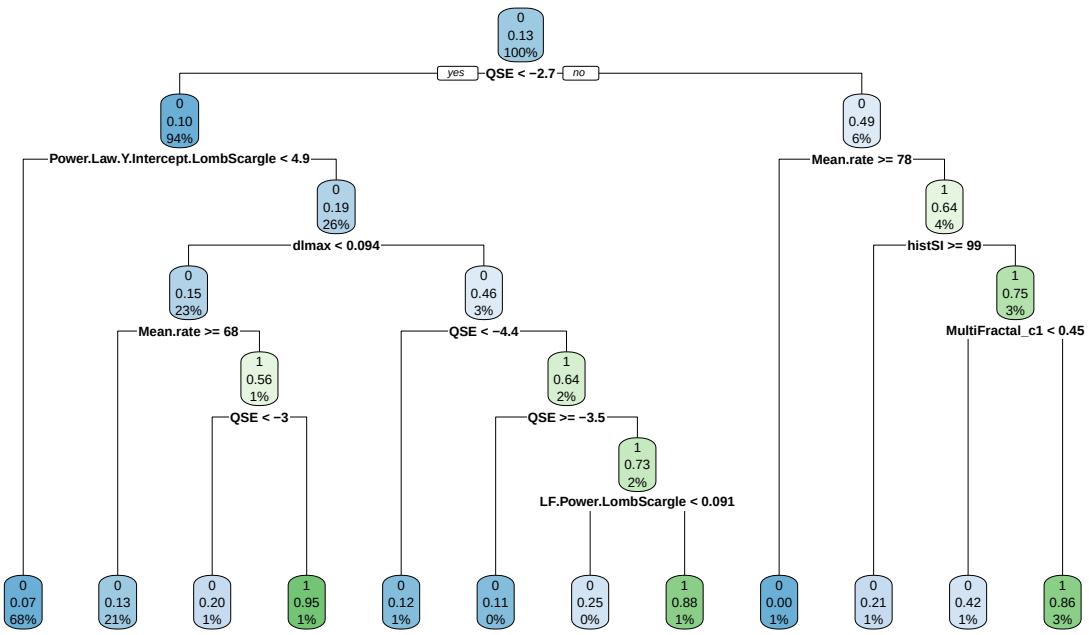
Figure 2. Correlation plot between each of the 23 predictors identified in Figure 1.



Classification Trees

The rpart classification tree using the previously identified 23 significant variables as sepsis predictors determined QSE, Mean.rate, and dlmax to be the most important variables.

Figure 3. a) Rpart classification tree using Table 1 predictors to classify sepsis outcome. b) Variable importance of each predictor considered in 3a.



	Importance
QSE	77.806036
Mean.rate	46.735879
dlmax	22.114698
Power.Law.Y.Intercept.LombScargle	17.543943
LF.Power.LombScargle	14.681248
histSI	13.667105
DFA.Alpha.1	13.148247
HF.Power.LombScargle	12.779813
MultiFractal_c1	11.911267
SymDp0_2	9.489226
Multiscale.Entropy	9.380369
Correlation.dimension	6.826715
sevl	6.352908
pL	6.105720
DFA.Alpha.2	5.986567
Largest.Lyapunov.exponent	5.977190
SymDfw_2	5.728816
shannEn	5.155934
pD	4.727407
sedl	4.238365
pDpR	2.723329
eScaleE	2.407347
MultiFractal_c2	1.004782

The predictions for the training set correctly classified 2646 sepsis negative instances and 122 sepsis positive instances. The model resulted in 263 false negatives and 17 false positives for the training set. The confusion matrix showed 0.9081 accuracy, 0.4273 kappa, and 0.46565 F1.

Figure 4. Statistical evaluation of classification tree (Fig. 3) training set predictions.

```
## Confusion Matrix and Statistics
##
##          Actual
## Predicted    0    1
##         0 2646  263
##         1   17  122
##
##                Accuracy : 0.9081
##                  95% CI : (0.8973, 0.9182)
##     No Information Rate : 0.8737
##     P-Value [Acc > NIR] : 1.406e-09
##
##                   Kappa : 0.4273
##
##  Mcnemar's Test P-Value : < 2.2e-16
##
##             Sensitivity : 0.31688
##             Specificity : 0.99362
##          Pos Pred Value : 0.87770
##          Neg Pred Value : 0.90959
##               Precision : 0.87770
##                  Recall : 0.31688
##                      F1 : 0.46565
##              Prevalence : 0.12631
##          Detection Rate : 0.04003
##    Detection Prevalence : 0.04560
##       Balanced Accuracy : 0.65525
##
##        'Positive' Class : 1
##
```

The predictions for the testing set correctly classified 1088 sepsis negative instances and 47 sepsis positive instances. The model resulted in 106 false negatives and 25 false positives for the testing set. The confusion matrix showed 0.8965 accuracy, 0.369 kappa, and 0.41778 F1.

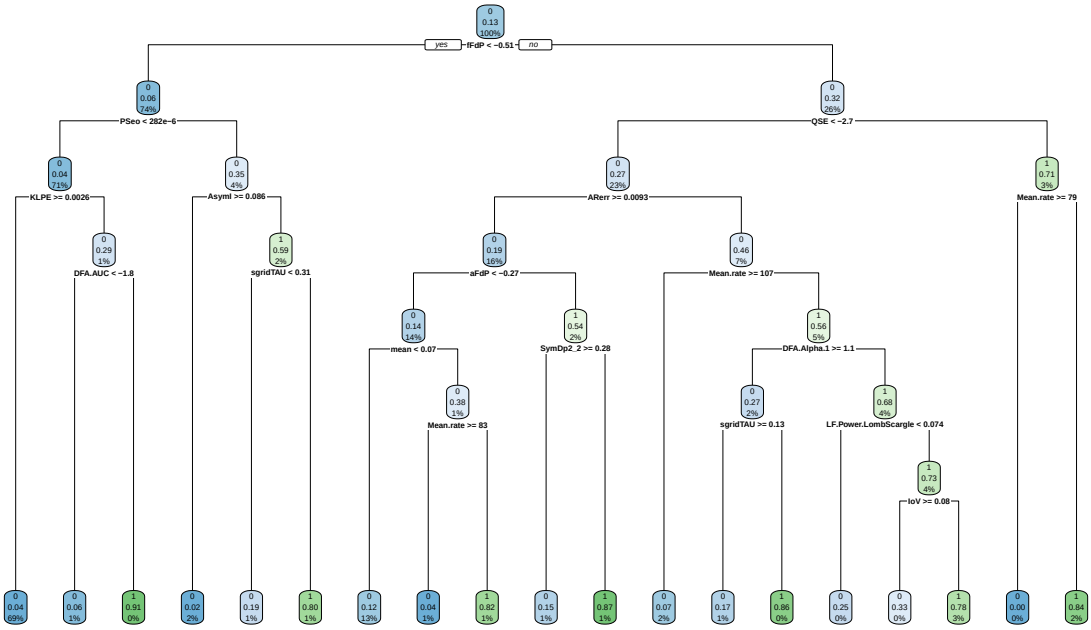
Figure. 5 Statistical evaluation of classification tree (Fig. 3) testing set predictions.

```
## Confusion Matrix and Statistics
##
##          Actual
## Predicted    0    1
##         0 1088  106
##         1   25   47
##
##                Accuracy : 0.8965
##                  95% CI : (0.8784, 0.9128)
##     No Information Rate : 0.8791
##     P-Value [Acc > NIR] : 0.02989
##
##                   Kappa : 0.369
##
##  Mcnemar's Test P-Value : 2.756e-12
##
##             Sensitivity : 0.30719
```

```
##             Specificity : 0.97754
##          Pos Pred Value : 0.65278
##          Neg Pred Value : 0.91122
##               Precision : 0.65278
##                  Recall : 0.30719
##                      F1 : 0.41778
##              Prevalence : 0.12085
##          Detection Rate : 0.03712
##    Detection Prevalence : 0.05687
##       Balanced Accuracy : 0.64236
##
##        'Positive' Class : 1
##
```

The second rpart classification tree was modeled based on each of the 57 predictors to evaluate importance of variables that were originally excluded. In this tree, fFdP, IoV, and aFdP were found to be significantly more important than QSE. A total of 13 predictors were ranked with importance greater than 20.

Figure 6. a) Rpart classification tree based on all 57 possible predictors. b) Variable importance of predictors used in 6a.



	Importance
fFdP	87.9756752
IoV	76.2361690
aFdP	57.1028612
Mean.rate	52.6718503
QSE	46.2280196
ARerr	41.7636652
DFA.AUC	40.1015973
sgridAND	32.8538020
Complexity	31.6276273

	Importance
CVI	29.8404995
gcount	28.9691757
PSeo	26.2142695
SymDp2_2	20.1006462
SymDp0_2	18.6682250
sgridTAU	18.5847700
DFA.Alpha.1	17.8080730
AsymI	16.9464965
Coefficient.of.variation	15.5849345
SymDse_2	14.6559258
sgridWGT	14.5741606
Poincar..SD2	13.4396888
KLPE	12.7040239
SymDce_2	11.0331880
Teo	10.6304830
Poincar..SD1	10.5329557
MultiFractal_c1	10.3437306
mean	9.9711095
shannEn	8.3413523
median	8.2062741
SymDfw_2	7.3078007
CSI	6.2665062
LF.Power.LombScargle	4.9507186
formF	3.9240781
Correlation.dimension	2.6221074
LF.HF.ratio.LombScargle	2.0028438
Power.Law.Slope.LombScargle	1.7791096
Power.Law.Y.Intercept.LombScargle	1.2376797
VLF.Power.LombScargle	1.2376797
MultiFractal_c2	0.8121449
SDLEmean	0.7052085
eScaleE	0.2353114

The predictions for the training set correctly classified 2613 sepsis negative instances and 225 sepsis positive instances. The model resulted in 160 false negatives and 50 false positives for the training set. The confusion matrix showed 0.9311 accuracy, 0.6444 kappa, and 0.68182 F1.

Figure 7. Statistical evaluation of classification tree (Fig. 6) training set predictions.

```
## Confusion Matrix and Statistics
##
##          Actual
## Predicted    0    1
##         0 2613  160
##         1   50  225
##
##                Accuracy : 0.9311
##                  95% CI : (0.9215, 0.9398)
##     No Information Rate : 0.8737
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.6444
##
```

```
##  Mcnemar's Test P-Value : 5.406e-14
##
##             Sensitivity : 0.58442
##             Specificity : 0.98122
##          Pos Pred Value : 0.81818
##          Neg Pred Value : 0.94230
##               Precision : 0.81818
##                  Recall : 0.58442
##                      F1 : 0.68182
##              Prevalence : 0.12631
##          Detection Rate : 0.07382
##    Detection Prevalence : 0.09022
##       Balanced Accuracy : 0.78282
##
##        'Positive' Class : 1
##
```

The predictions for the testing set correctly classified 1069 sepsis negative instances and 75 sepsis positive instances. The model resulted in 78 false negatives and 44 false positives for the testing set. The confusion matrix showed 0.9036 accuracy, 0.4984 kappa, and 0.55147 F1.

Figure 8. Statistical evaluation of classification tree (Fig. 6) testing set predictions.

```
## Confusion Matrix and Statistics
##
##          Actual
## Predicted    0    1
##         0 1069   78
##         1   44   75
##
##                Accuracy : 0.9036
##                  95% CI : (0.886, 0.9193)
##     No Information Rate : 0.8791
##     P-Value [Acc > NIR] : 0.003455
##
##                   Kappa : 0.4984
##
##  Mcnemar's Test P-Value : 0.002811
##
##             Sensitivity : 0.49020
##             Specificity : 0.96047
##          Pos Pred Value : 0.63025
##          Neg Pred Value : 0.93200
##               Precision : 0.63025
##                  Recall : 0.49020
##                      F1 : 0.55147
##              Prevalence : 0.12085
##          Detection Rate : 0.05924
##    Detection Prevalence : 0.09400
##       Balanced Accuracy : 0.72533
##
##        'Positive' Class : 1
##
```

The complexity parameter with the lowest xerror was determined to be 0.01, so the parameter did not need to be further adjusted in this model.

The random forest, considering importance and testing 25 random predictors at each node, correctly classified 2632 sepsis negative cases and 183 sepsis positive cases in the training set. The training predictions resulted in 202 false negatives and 31 false positives. The model showed 0.9236 accuracy, 0.5724 kappa, and 0.61102 F1.

Figure 9. Evaluation of random forest classification model predictions using the training dataset.

```
## Confusion Matrix and Statistics
##
##          actual
## predicted    0    1
##         0 2630  203
##         1   33  182
##
##                Accuracy : 0.9226
##                  95% CI : (0.9125, 0.9318)
##     No Information Rate : 0.8737
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.5675
##
##  Mcnemar's Test P-Value : < 2.2e-16
##
##             Sensitivity : 0.47273
##             Specificity : 0.98761
##          Pos Pred Value : 0.84651
##          Neg Pred Value : 0.92834
##               Precision : 0.84651
##                  Recall : 0.47273
##                      F1 : 0.60667
##              Prevalence : 0.12631
##          Detection Rate : 0.05971
##    Detection Prevalence : 0.07054
##       Balanced Accuracy : 0.73017
##
##        'Positive' Class : 1
##
```

The predictions for the testing set correctly classified 1093 sepsis negative instances and 88 sepsis positive instances. The model resulted in 65 false negatives and 20 false positives for the testing set. The confusion matrix showed 0.9329 accuracy, 0.6381 kappa, and 0.67433 F1.

Figure 10. Evaluation of random forest classification model using testing data.

```
## Confusion Matrix and Statistics
##
##          actual
## predicted    0    1
##         0 1094   64
##         1   19   89
##
##                Accuracy : 0.9344
##                  95% CI : (0.9194, 0.9474)
##     No Information Rate : 0.8791
##     P-Value [Acc > NIR] : 4.661e-11
```

```
##
##                   Kappa : 0.6467
##
##  Mcnemar's Test P-Value : 1.368e-06
##
##             Sensitivity : 0.58170
##             Specificity : 0.98293
##          Pos Pred Value : 0.82407
##          Neg Pred Value : 0.94473
##               Precision : 0.82407
##                  Recall : 0.58170
##                      F1 : 0.68199
##              Prevalence : 0.12085
##          Detection Rate : 0.07030
##    Detection Prevalence : 0.08531
##       Balanced Accuracy : 0.78231
##
##        'Positive' Class : 1
##
```

The random forest model determined the 9 predictors that would result in the greatest decrease in accuracy if removed were Mean.rate, LF.Power.LombScargle, DFA.AUC, fFdP, IoV, KLPE, ARerr, QSE, and mean.

Figure 11. Random forest predictor importance based on mean decrease in accuracy and mean decrease in gini score if predictor is removed.

```
##                          X0 X1 MeanDecreaseAccuracy MeanDecreaseGini
## Mean.rate                14 23                   23               31
## LF.HF.ratio.LombScargle 15  6                   18               13
## LF.Power.LombScargle    15 14                   19               19
## DFA.AUC                 15 14                   18               27
## fFdP                    31 21                   36               57
## IoV                     29 14                   32               31
## KLPE                    21 16                   24               22
## ARerr                   19  7                   21               17
## QSE                     17 22                   28               33
## mean                    20 15                   24               23
```

In combination, the rpart and random forest classification models determined the most important predictors to be fFdP, IoV, Mean.rate, QSE, ARerr, and DFA.AUC.

Discussion

The first classification tree resulted in 0.9 accuracy for both the training and testing datasets. This is highly accurate and the equivalency between training and testing suggests that overfitting is not a concern. The kappa coefficient, which measures the agreement between classification and actual values, showed moderate agreement. The F1 score, which gives the weighted average of precision and recall on a range from 0 to 1, is also moderate. These metrics combined indicate that the model can correctly predict sepsis; however, there is room for improvement.

The second classification tree, considering all predictors, marginally increased the accuracy for both the training and testing data compared to the initial tree. However, the kappa coefficient and F1 score significantly increased to show moderate-to-substantial agreement, indicating an improved classification model.

The random forest classification tree resulted in the highest accuracy out of all of the models for both the training and testing datasets. Additionally, the testing predictions resulted in higher accuracy than the

training predictions for the random forest model, indicating greater prediction ability. The kappa and F1 scores showed substantial agreement between the predictions and actual outcomes.

In conclusion, the random forest model resulted in greater predictive accuracy of sepsis compared to the rpart classification models. Each of the developed models were able to categorize the outcome as sepsis positive or sepsis negative with high accuracy. This improvement by the random forest method is likely due to the amount of random trees generated and analyzed. While the rpart classification model built one effective tree, the random forest model built hundreds of trees to identify the most accurate model.

Limitations

Building a random forest model from an large dataset can require considerable time and computer power. Additionally, the application of a model to predict sepsis is limited by the breadth of continuous variables included and the lack of clear, binary sepsis diagnostic criteria. While the model predictions were highly accurate here, it may be difficult to collect all of the parameters simultaneously and issue an alert for sepsis earlier than the symptoms may be recognized in a medical facility.

Future Directions

The ability to continuously monitor non-invasive bodily parameters and detect health events in realtime holds great opportunity for the future of healthcare technology. With further development, a sepsis alert could be triggered at an early stage in hospitals, greatly improving the odds of survival. Outside of the hospital setting, similar health prediction models may be translated to wearable technology to detect and alert the wearer of possible illness or stress.

Citations

Chen, Kuan-Fu (2020), "HRV sepsis dataset", Mendeley Data, V1, doi: 10.17632/5tgk53r4v5.1

Centers for Disease Control and Prevention. (2021, August 17). What is sepsis? Centers for Disease Control and Prevention. Retrieved May 2, 2022, from https://www.cdc.gov/sepsis/what-is-sepsis.html

Gronwald T, Hoos O. Correlation properties of heart rate variability during endurance exercise: A systematic review. Ann Noninvasive Electrocardiol. 2020 Jan;25(1):e12697. doi: 10.1111/anec.12697. Epub 2019 Sep 9. PMID: 31498541; PMCID: PMC7358842.

Huang M, Cai S, Su J. The Pathogenesis of Sepsis and Potential Therapeutic Targets. Int J Mol Sci. 2019 Oct 29;20(21):5376. doi: 10.3390/ijms20215376. PMID: 31671729; PMCID: PMC6862039.

Mol, M., Strous, M., van Osch, F., Vogelaar, F. J., Barten, D. G., Farchi, M., Foudraine, N. A., & Gidron, Y. (2021). Heart-rate-variability (HRV), predicts outcomes in COVID-19. PloS one, 16(10), e0258841. https://doi.org/10.1371/journal.pone.0258841

Appendix Raw Code

```
######### libraries
library(rmarkdown)
library(dplyr)
library(DataExplorer)
library(caret)
library(rpart)
library(randomForest)
library(readxl)
library(psych)
library(ggfortify)
```

```
library(pROC)
library(rpart.plot)
######### datasets
data <- read_excel('HRV data 20201209.xlsx')
set.seed(6006)
ind <- sample(2, nrow(data), replace = T, prob = c(0.7, 0.3)) #tested 70/30, 60/40, 50/50, 40/60 conc: 70/30 best
train <- data[ind == 1,]
test <- data[ind == 2,]
xtrain <- train %>% select(-Sepsis3)
xtest <- test %>% select(-Sepsis3)
ytrain <- as.factor(train$Sepsis3)
ytest <- as.factor(test$Sepsis3)
######### exploratory
qq <- plot_qq(data, by = 'Sepsis3')
cor <- plot_correlation(data)
box <- plot_boxplot(data, by = 'Sepsis3')
#sig variables
st <- data %>% group_by(Sepsis3) %>% summarise_all(mean, na.rm = T)
sig <- c()
for (x in 1:58){
t <- t.test(st[x])
p <- t$p.value
if (p < 0.05){
print(st[0, x]);print(p)
sig <- c(sig, st[0,x])
}
}
names(sig)
sigdata <- data %>% select(names(sig))
plot_correlation(sigdata) #correlation between significatn variabels
#########rpart sig
sigrp <- rpart(ytrain ~ Mean.rate + LF.Power.LombScargle + HF.Power.LombScargle + DFA.Alpha.1 + DFA.Alpha.2 + Largest.Lyapunov.exponent + Correlation.dimension + Power.Law.Y.Intercept.LombScargle + Multiscale.Entropy + eScaleE + pD + dlmax + sedl + pDpR + pL + sevl + shannEn + SymDp0_2 + SymDfw_2 + histSI + MultiFractal_c1 + MultiFractal_c2 + QSE, data = xtrain, method = 'class' )
######### rpart accuracy
```

```
ptrain.sigrp <- predict(sigrp, type = 'class')
ptest.sigrp <- predict(sigrp, xtest, type = 'class')
traintablesigrp <- table(Predicted = ptrain.sigrp, Actual = ytrain)
cftrainsigrp <- confusionMatrix(traintablesigrp, mode = 'everything', positive = '1')
cftrainsigrp #0.9081 accuracy #0.46565 f1
testtablesigrp <- table(Predicted = ptest.sigrp, Actual = ytest)
cftestsigrp <- confusionMatrix(testtablesigrp, mode = 'everything', positive = '1')
cftestsigrp #0.8965 acc # 0.41778 F1
######### rpart
rp <- rpart(ytrain ~ ., data = xtrain, method = 'class')
rpart.plot(rp, main = 'Sepsis3 Classification Tree 2')
######### rpart accuracy
ptrain.rp <- predict(rp, type = 'class')
ptest.rp <- predict(rp, xtest, type = 'class')
traintable <- table(Predicted = ptrain.rp, Actual = ytrain)
cftrain <- confusionMatrix(traintable, mode = 'everything', positive = '1')
cftrain #0.9311 accuracy (70/30); 0.9378 (60/40); 0.9332 (40/60); 0.9463 (50/50)
#0.6444 kappa #0.68182
testtable <- table(Predicted = ptest.rp, Actual = ytest)
cftest <- confusionMatrix(testtable, mode = 'everything', positive = '1')
cftest #0.9036 accuracy; 0.897 (60/40); 0.9034 (40/60); 0.8922 (50/50)
#0.9036 kappa #0.55147 f1
######### rpart2 tuning
printcp(rp); plotcp(rp)
######### randomforest
rf <- randomForest(ytrain ~ ., data = xtrain, type = 'classification', importance = T, mtry = 25)
######### randomforest evaluation
ptrain.rf <- predict(rf, type = 'class')
ptest.rf <- predict(rf, newdata = xtest, type = 'class')
traintable.rf <- table(predicted = ptrain.rf, actual = ytrain)
cftrain.rf <- confusionMatrix(traintable.rf, mode = 'everything', positive = '1')
cftrain.rf #0.917 ;1 accuracy (70/30); 1 (60/40); 1 (40/60); 1 (50/50)
#1 kappa #acc 0.9268 mtry = 25 #0.61307 f1
testtable.rf <- table(predicted = ptest.rf, actual = ytest)
cftest.rf <- confusionMatrix(testtable.rf, mode = 'everything',positive = '1')
cftest.rf #0.9305 accuracy; 0.9087 (60/40); 0.9242 (40/60); 0.9064 (50/50)
```

```
#0.5864 kappa #acc 0.9384 mtry = 25 #0.68726
######### randomforest2 tuning
imp <- data.frame(round(importance(rf)))
imp <- imp %>% filter(imp[,3] > 17)
impgini <- imp %>% filter(imp[,4] > 20 )
sort(imp[,4], decreasing = T)
####################
m <- rpart(ytrain ~ fFdP + IoV + KLPE + aFdP + Mean.rate + QSE, data = xtrain, method = 'class')
rpart.plot(m, main = 'Sepsis3 Classification Tree 3')
mtrain.rp <- predict(m, type = 'class')
mtest.rp <- predict(m, xtest, type = 'class')
traintablem <- table(Predicted = mtrain.rp, Actual = ytrain)
cftrainm <- confusionMatrix(traintablem, mode = 'everything', positive = '1')
cftrainm #0.9236
testtablem <- table(Predicted = mtest.rp, Actual = ytest)
cftestm <- confusionMatrix(testtablem, mode = 'everything', positive = '1')
cftestm #0.8989
#####
```