Predicting complications of Chronic Heart Failure in Post-Myocardial Infarction Patients

Presented by

Nandhika Rajmanikandan & Ganesh Vannam

Goal of Study

This study aims to predict the chronic heart failure for patients using their clinical data to improve early intervention and enhance patient management in critical care settings.

Data Set Structure

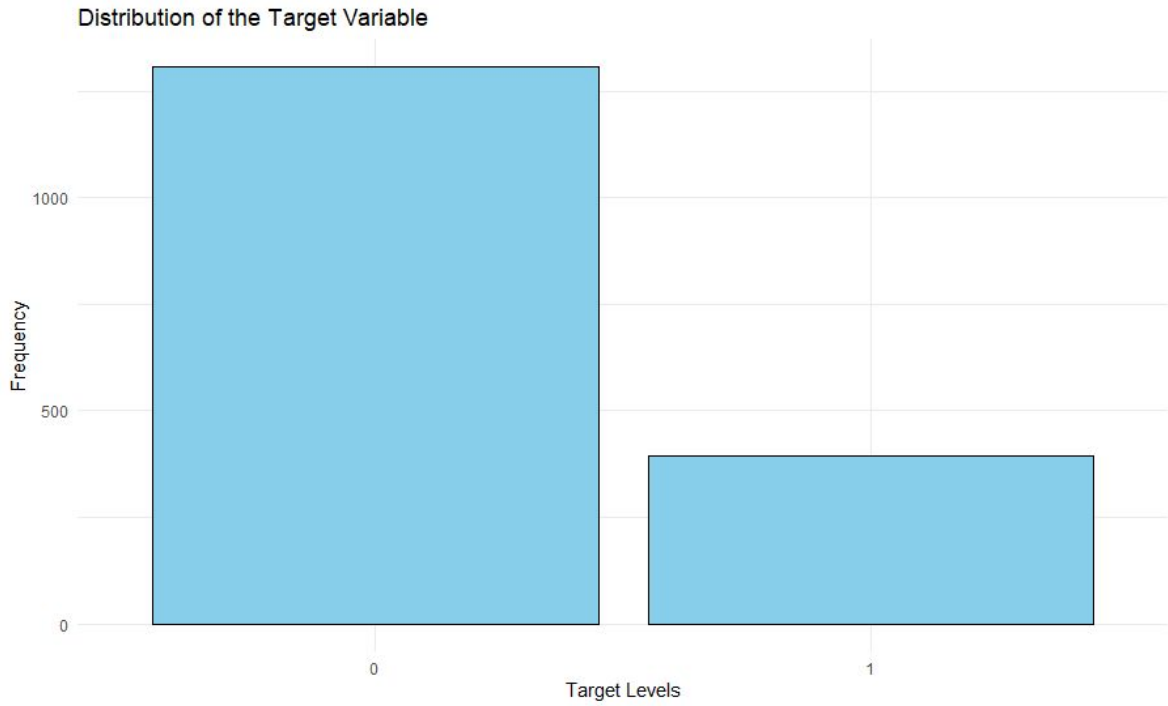
Sample Size: 1700

Number of Variables: 122

Continuous Variables: 14

Categorical Variables:108

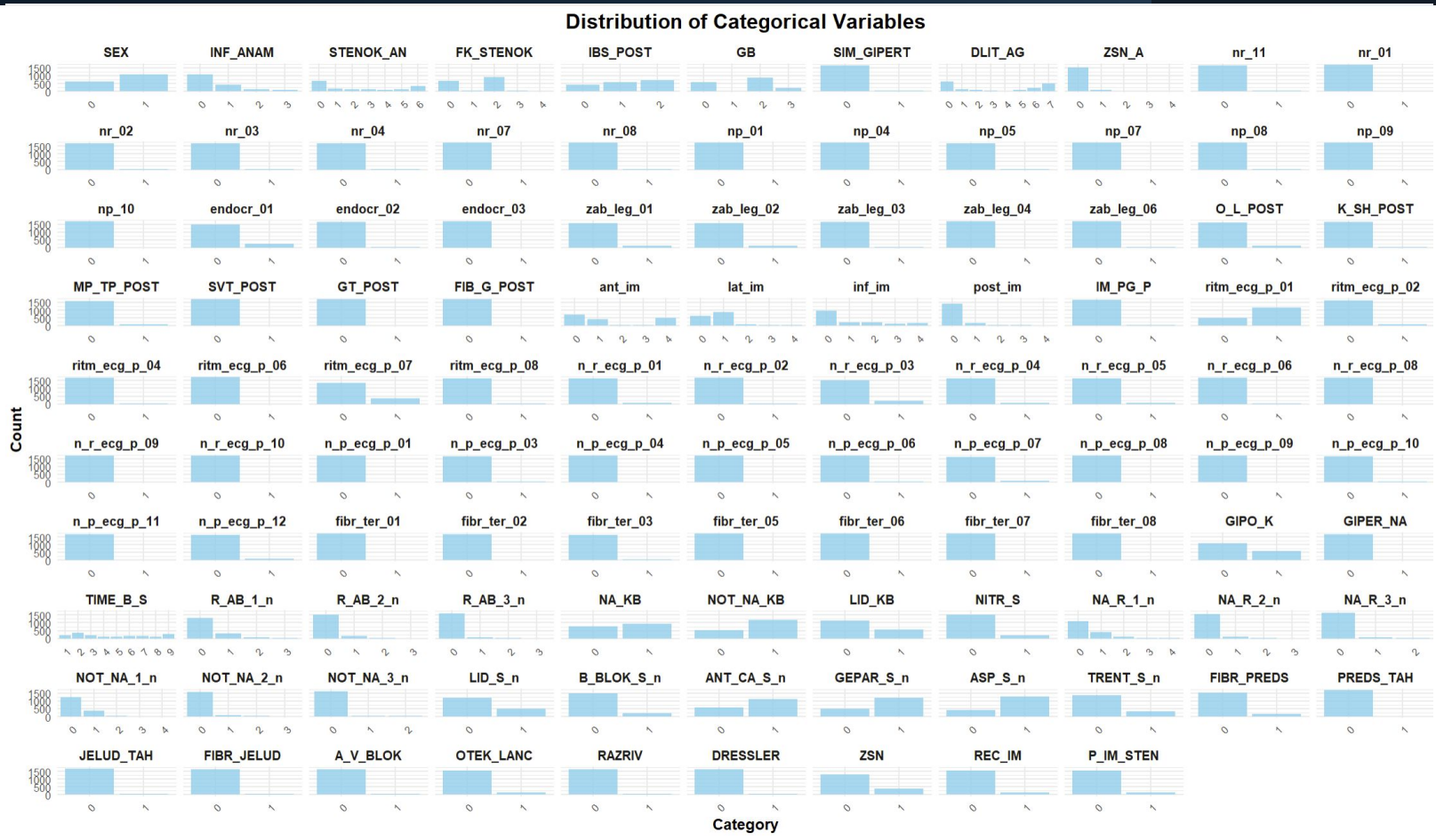
Response Variable: Chronic Heart Failure(ZSN)



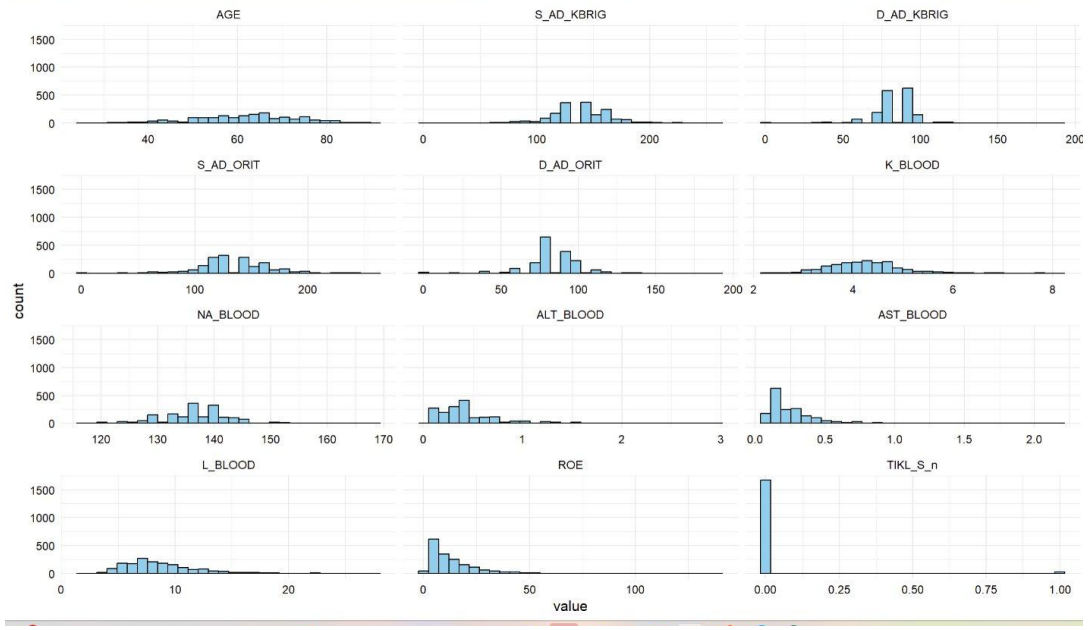
Data Preprocessing

- Converted categorical variables into factors.
- Target variable has no missing values.
- We have 15794 missing values, To handle those missing values we used KNN imputation with K=5.
- Number of feature after removing zero and near zero variance for categorical variables: 45
- Number of features after adding dummy variables: 115
- Number of features after removing zero and near zero variance predictors once again: 81
- To handle skewness and outliers we performed box-cox & Spatial sign transformation.

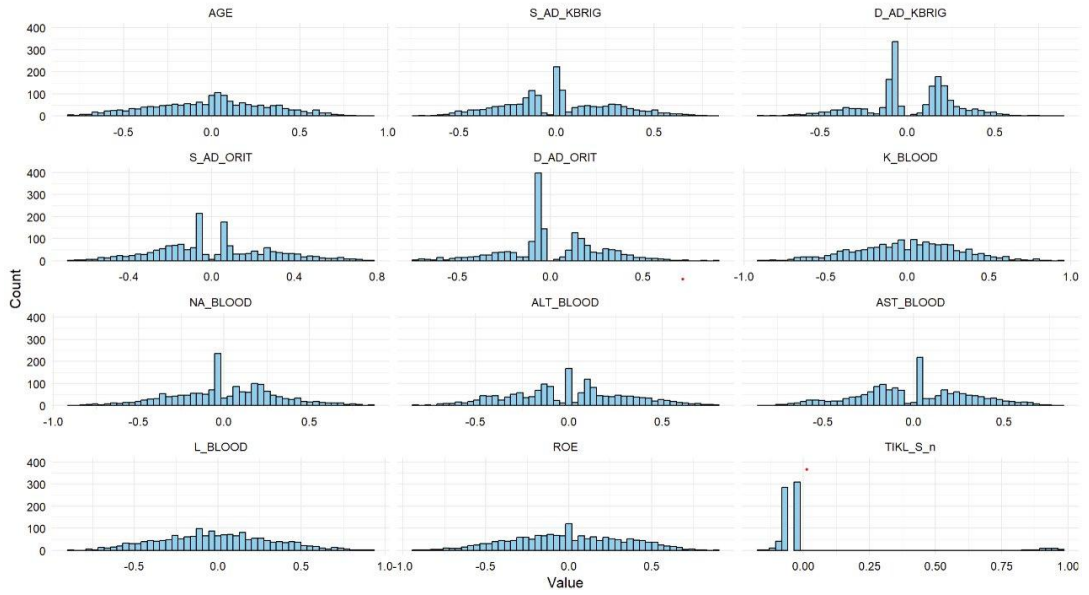
Distribution of Categorical Variables



Box-Cox Transformation

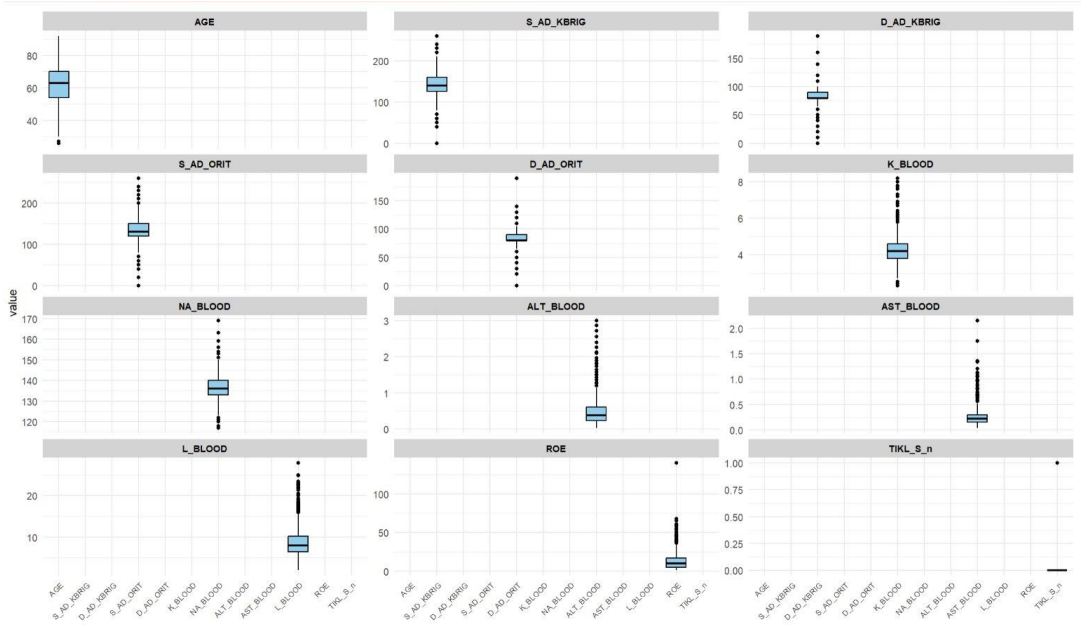


Before

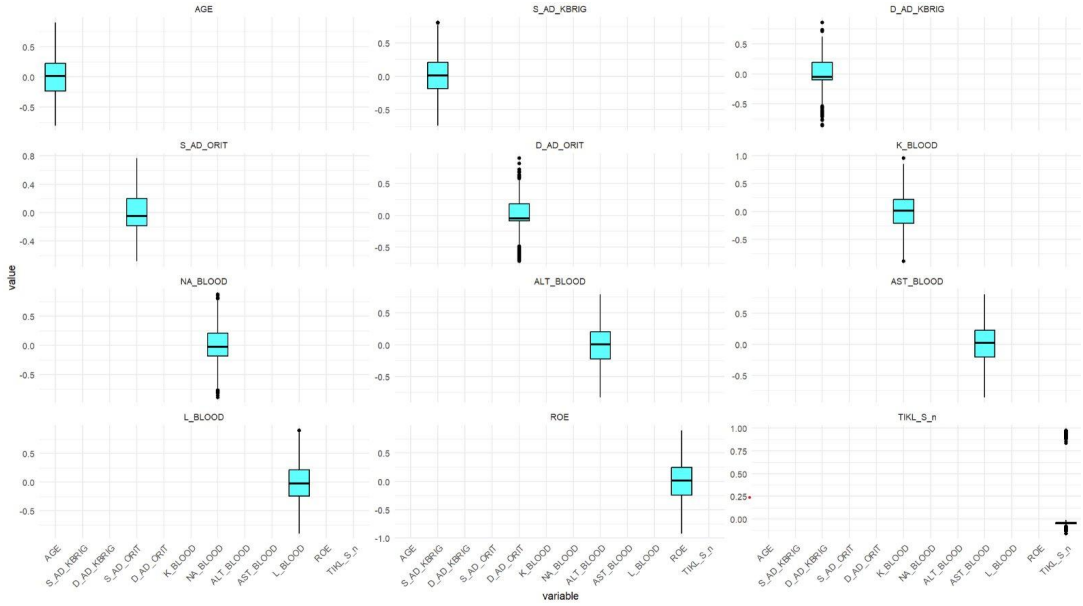


After

Spatial Sign Traansformation



Before



After

Model Building:

- We are using stratified random Sampling because dataset is imbalanced.
- Splitting 80% of data for training and 20% for testing set.
- 10-fold cross validation
- We are using Kappa metric to evaluate model performance, because our target variable is imbalanced.

Models

- **Linear Models:**
  - Logistic Regression
  - Linear Discriminant Analysis (LDA)
  - Partial Least Squares Discriminant Analysis (PLSDA)
  - Penalized Models
- **Non Linear Models:**
  - Quadratic Discriminant Analysis (QDA)
  - Regularized Discriminant Analysis (RDA)
  - Mixture Discriminant Analysis (MDA)
  - Flexible Discriminant Analysis (FDA)
  - Neural Network
  - K-Nearest Neighbors (KNN)
  - Support Vector Machines (SVM)
  - Naïve Bayes

Linear Models:

Logistic Regression Model

```
> print(logistic_model)
Generalized Linear Model

1361 samples
  76 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (76), scaled (76)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1224, 1225, 1225, 1226, 1225, 1225, ...
Resampling results:

  Accuracy   Kappa
  0.7950116  0.2996071
```

```
> print(confusion_matrix)
Confusion Matrix and Statistics

                Reference
Prediction       no_complication complication
  no_complication             250            52
  complication                 11            26

               Accuracy : 0.8142
                 95% CI : (0.7686, 0.8541)
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 0.02848

                  Kappa : 0.357

 Mcnemar's Test P-Value : 4.667e-07

            Sensitivity : 0.9579
            Specificity : 0.3333
         Pos Pred Value : 0.8278
         Neg Pred Value : 0.7027
             Prevalence : 0.7699
         Detection Rate : 0.7375
   Detection Prevalence : 0.8909
      Balanced Accuracy : 0.6456

       'Positive' Class : no_complication
```

Linear Discriminant Analysis Model

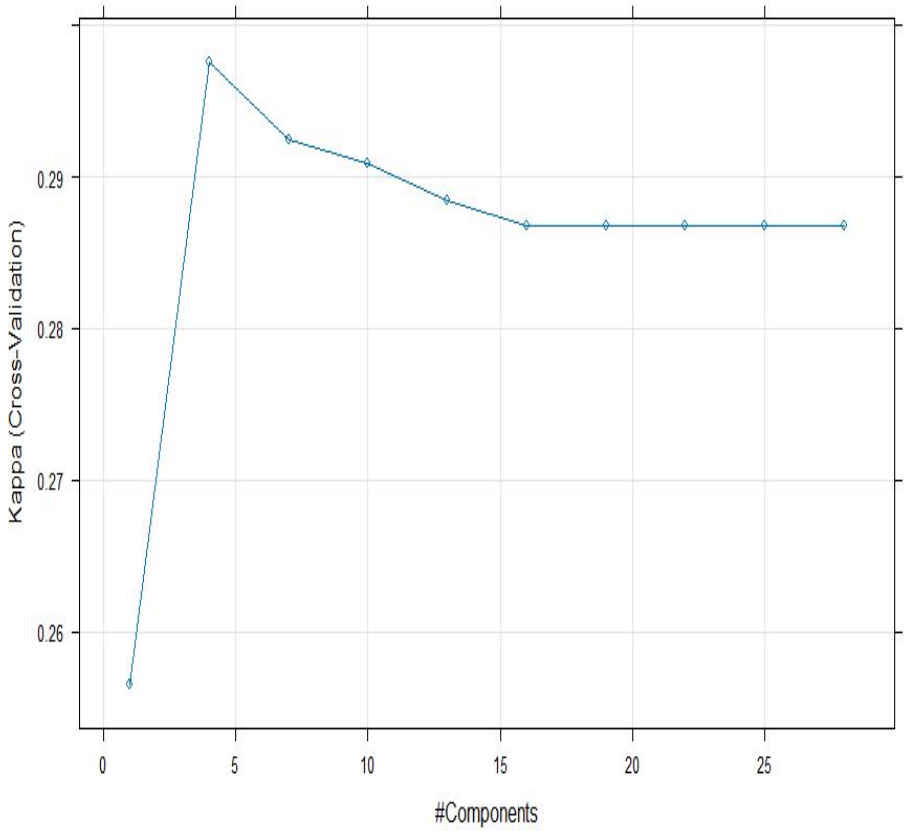
```
> print(lda_model)
Linear Discriminant Analysis

1361 samples
  76 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (76), scaled (76)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1224, 1225, 1225, 1226, 1225, 1225, ...
Resampling results:

  Accuracy   Kappa
  0.8001263  0.3028579
```

```
> print(confusion_matrix)
Confusion Matrix and Statistics

                Reference
Prediction       no_complication complication
  no_complication             253           54
  complication                  8           24

               Accuracy : 0.8171
                 95% CI : (0.7718, 0.8568)
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 0.02071

                  Kappa : 0.3492

 Mcnemar's Test P-Value : 1.097e-08

            Sensitivity : 0.9693
            Specificity : 0.3077
         Pos Pred Value : 0.8241
         Neg Pred Value : 0.7500
             Prevalence : 0.7699
         Detection Rate : 0.7463
   Detection Prevalence : 0.9056
      Balanced Accuracy : 0.6385

       'Positive' Class : no_complication
```

Partial Least Square Discriminative Analysis Model

```
print(plsda_model)
Partial Least Squares

1361 samples
  81 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (81), scaled (81)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1224, 1225, 1225, 1226, 1225, 1225, ...
Resampling results across tuning parameters:

  ncomp  Accuracy   Kappa
   1     0.7935520  0.2565033
   4     0.8045114  0.2975687
   7     0.8008403  0.2924411
  10     0.8001050  0.2908365
  13     0.8001050  0.2884129
  16     0.7993697  0.2868083
  19     0.7993697  0.2868083
  22     0.7993697  0.2868083
  25     0.7993697  0.2868083
  28     0.7993697  0.2868083

Kappa was used to select the optimal model using the largest value.
The final value used for the model was ncomp = 4.
```

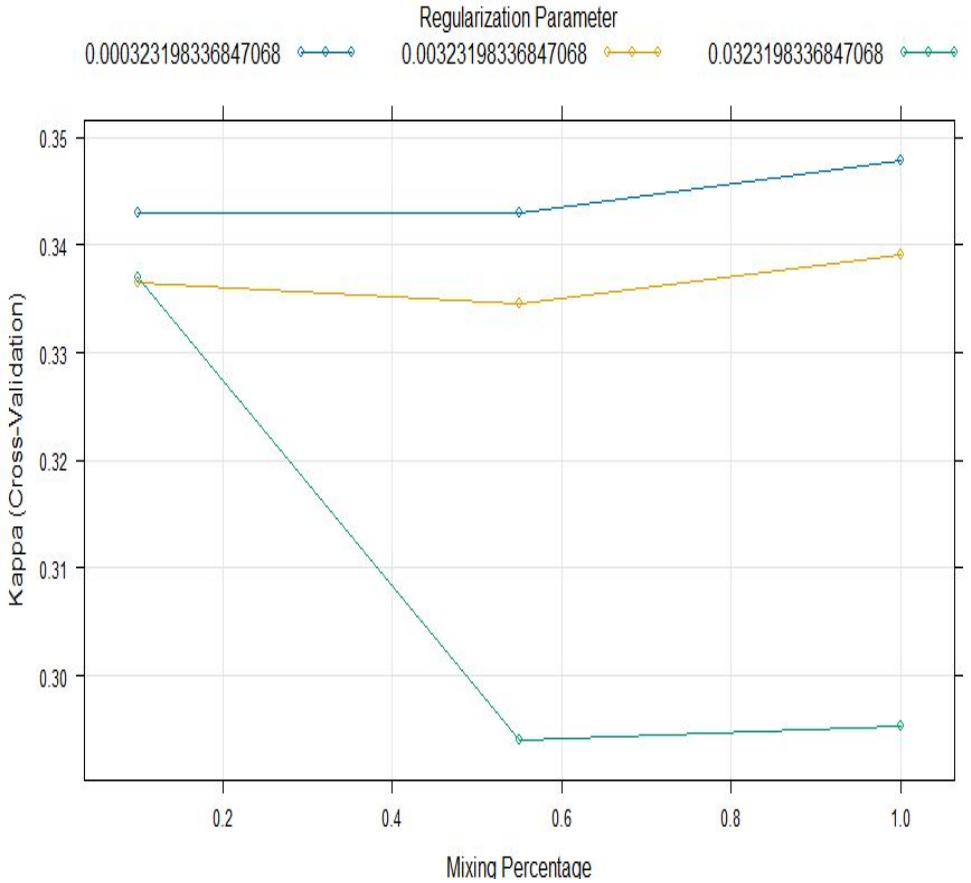
```
print(confusion_matrix_plsda)
Confusion Matrix and Statistics

                Reference
Prediction        no_complication complication
  no_complication             252            51
  complication                  9            27

               Accuracy : 0.823
                 95% CI : (0.7781, 0.8621)
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 0.01037

                  Kappa : 0.3842

 Mcnemar's Test P-Value : 1.203e-07

            Sensitivity : 0.9655
            Specificity : 0.3462
         Pos Pred Value : 0.8317
         Neg Pred Value : 0.7500
             Prevalence : 0.7699
         Detection Rate : 0.7434
   Detection Prevalence : 0.8938
      Balanced Accuracy : 0.6558

       'Positive' Class : no_complication
```



Penalized Model:

```
print(pm_model)
glmnet

1361 samples
  81 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (81), scaled (81)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1224, 1225, 1225, 1225, 1225, 1225, ...
Resampling results across tuning parameters:

  alpha  lambda        Accuracy   Kappa
  0.10   0.0003231983  0.8038462  0.3430345
  0.10   0.0032319834  0.8038355  0.3364745
  0.10   0.0323198337  0.8126593  0.3369634
  0.55   0.0003231983  0.8038462  0.3430063
  0.55   0.0032319834  0.8053115  0.3344751
  0.55   0.0323198337  0.8111618  0.2940654
  1.00   0.0003231983  0.8053115  0.3478082
  1.00   0.0032319834  0.8082365  0.3391610
  1.00   0.0323198337  0.8126323  0.2952853

Kappa was used to select the optimal model using the largest value.
The final values used for the model were alpha = 1 and lambda = 0.0003231983.
```

```
print(confusion_matrix_pm)
Confusion Matrix and Statistics

                Reference
Prediction       no_complication complication
  no_complication             251            60
  complication                 10            18

               Accuracy : 0.7935
                 95% CI : (0.7465, 0.8353)
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 0.1667

                  Kappa : 0.2482

 Mcnemar's Test P-Value : 4.724e-09

            Sensitivity : 0.9617
            Specificity : 0.2308
         Pos Pred Value : 0.8071
         Neg Pred Value : 0.6429
             Prevalence : 0.7699
         Detection Rate : 0.7404
   Detection Prevalence : 0.9174
      Balanced Accuracy : 0.5962
```



Non-Linear Models:
Quadratic Discriminant Analysis Model

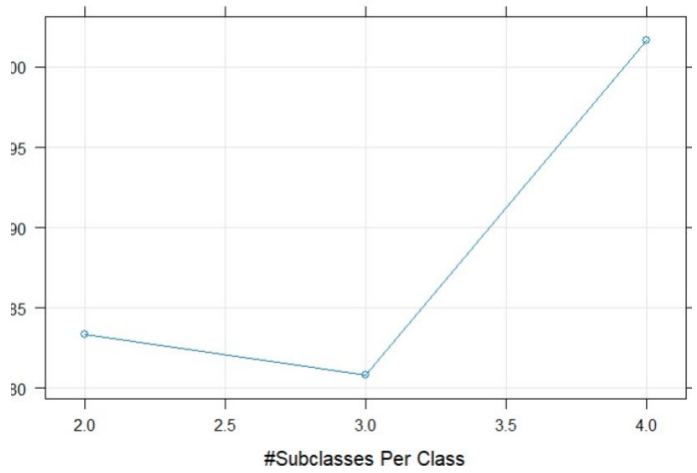
```
> print(qda_model)
Quadratic Discriminant Analysis

1361 samples
  76 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (76), scaled (76)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1224, 1225, 1225, 1226, 1225, 1225, ...
Resampling results:

  Accuracy   Kappa
  0.7758774  0.2841291
```

```
> print(confusion_matrix)
Confusion Matrix and Statistics

                Reference
Prediction       no_complication complication
  no_complication             238            54
  complication                 23            24

               Accuracy : 0.7729
                 95% CI : (0.7245, 0.8164)
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 0.4789007

                  Kappa : 0.2551

 Mcnemar's Test P-Value : 0.0006289

            Sensitivity : 0.9119
            Specificity : 0.3077
         Pos Pred Value : 0.8151
         Neg Pred Value : 0.5106
             Prevalence : 0.7699
         Detection Rate : 0.7021
   Detection Prevalence : 0.8614
      Balanced Accuracy : 0.6098

       'Positive' Class : no_complication
```

Mixture Discriminant Analysis Model

```
> print(mda_model)
Mixture Discriminant Analysis

1361 samples
  76 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (76), scaled (76)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1224, 1225, 1225, 1226, 1225
Resampling results across tuning parameters:

  subclasses  Accuracy   Kappa
  2           0.7803052  0.2833747
  3           0.7656427  0.2808036
  4           0.7751958  0.3016871

Kappa was used to select the optimal model using the
The final value used for the model was subclasses = 4
```

```
> print(confusion_matrix)
Confusion Matrix and Statistics

                Reference
Prediction       no_complication comp
  no_complication             245
  complication                 16

               Accuracy : 0.8053
                 95% CI : (0.7591, 0.8
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 0.06689

                  Kappa : 0.3514

 Mcnemar's Test P-Value : 4.865e-05

            Sensitivity : 0.9387
            Specificity : 0.3590
         Pos Pred Value : 0.8305
         Neg Pred Value : 0.6364
             Prevalence : 0.7699
         Detection Rate : 0.7227
   Detection Prevalence : 0.8702
      Balanced Accuracy : 0.6488

       'Positive' Class : no_complicat
```



Regularized Discriminant Analysis Model

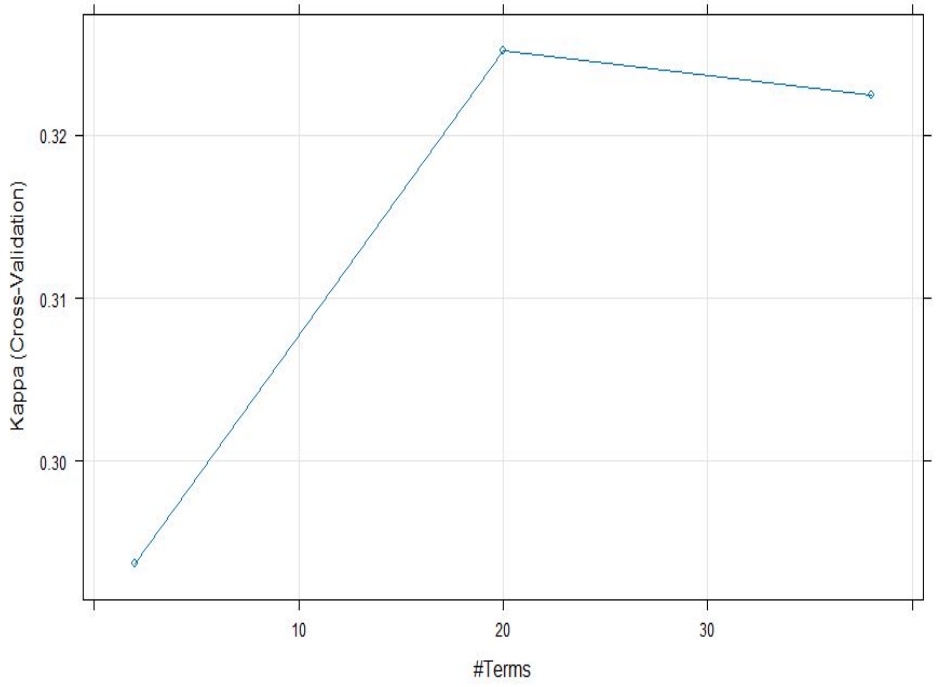
```
print(rda_model)
Regularized Discriminant Analysis

1361 samples
  81 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (81), scaled (81)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1224, 1226, 1225, 1225, 1226, 1225, ...
Resampling results across tuning parameters:

  gamma  lambda  Accuracy   Kappa
  0.1    0.1     0.7626532  0.2846853
  0.1    0.3     0.7582413  0.2938922
  0.1    0.5     0.7575006  0.2904087
  0.1    0.7     0.7670596  0.3104340
  0.1    0.9     0.7935473  0.3336677
  0.3    0.1     0.7575003  0.2894921
  0.3    0.3     0.7582194  0.2907668
  0.5    0.3     0.7641075  0.3142871
  0.5    0.5     0.7670379  0.3102348
  0.5    0.7     0.7788520  0.3209666
  0.5    0.9     0.7943479  0.3281710
  0.7    0.1     0.7641563  0.3207889
  0.7    0.3     0.7663352  0.3192700
  0.7    0.5     0.7722233  0.3279077
  0.7    0.7     0.7803173  0.3281934
  0.7    0.9     0.7980193  0.3457431
  0.9    0.1     0.7596960  0.3146620

Kappa was used to select the optimal model using the largest value.
The final values used for the model were gamma = 0.7 and lambda = 0.9.
```

```
print(confusion_matrix_rda)
Confusion Matrix and Statistics

                Reference
Prediction       no_complication complication
  no_complication             241            50
  complication                 20            28

               Accuracy : 0.7935
                 95% CI : (0.7465, 0.8353)
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 0.1667140

                  Kappa : 0.3263

 Mcnemar's Test P-Value : 0.0005279

            Sensitivity : 0.9234
            Specificity : 0.3590
         Pos Pred Value : 0.8282
         Neg Pred Value : 0.5833
             Prevalence : 0.7699
         Detection Rate : 0.7109
   Detection Prevalence : 0.8584
      Balanced Accuracy : 0.6412

       'Positive' Class : no_complication
```

Flexible Discriminant Analysis Model

```
print(fda_model)
Flexible Discriminant Analysis

1361 samples
  81 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (81), scaled (81)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1224, 1226, 1225, 1225, 1226, 1225, ...
Resampling results across tuning parameters:

  nprune  Accuracy   Kappa
   2      0.8134336  0.2937060
  20      0.8089892  0.3251865
  38      0.8075077  0.3224468

Tuning parameter 'degree' was held constant at a value of 1
Kappa was used to select the optimal model using the largest value.
The final values used for the model were degree = 1 and nprune = 20.
```

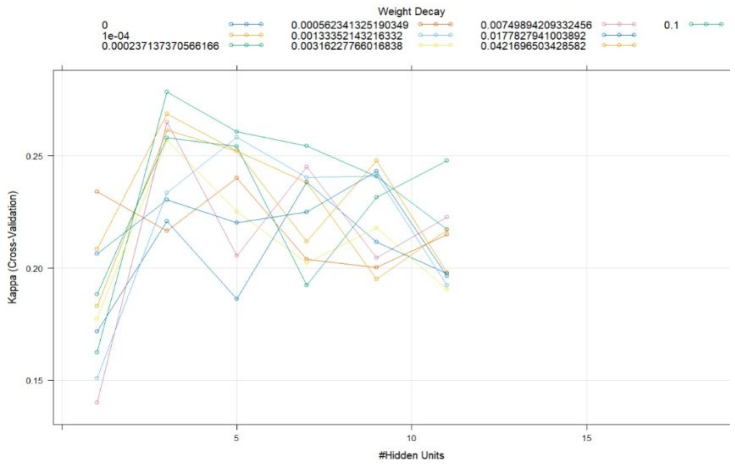
```
print(confusion_matrix_fda)
Confusion Matrix and Statistics

                Reference
Prediction        no_complication complication
  no_complication             254            57
  complication                  7            21

               Accuracy : 0.8112
                 95% CI : (0.7654, 0.8515)
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 0.0385

                  Kappa : 0.3127

 Mcnemar's Test P-Value : 9.068e-10

            Sensitivity : 0.9732
            Specificity : 0.2692
         Pos Pred Value : 0.8167
         Neg Pred Value : 0.7500
             Prevalence : 0.7699
         Detection Rate : 0.7493
   Detection Prevalence : 0.9174
      Balanced Accuracy : 0.6212

       'Positive' Class : no_complication
```



Neural Network Model

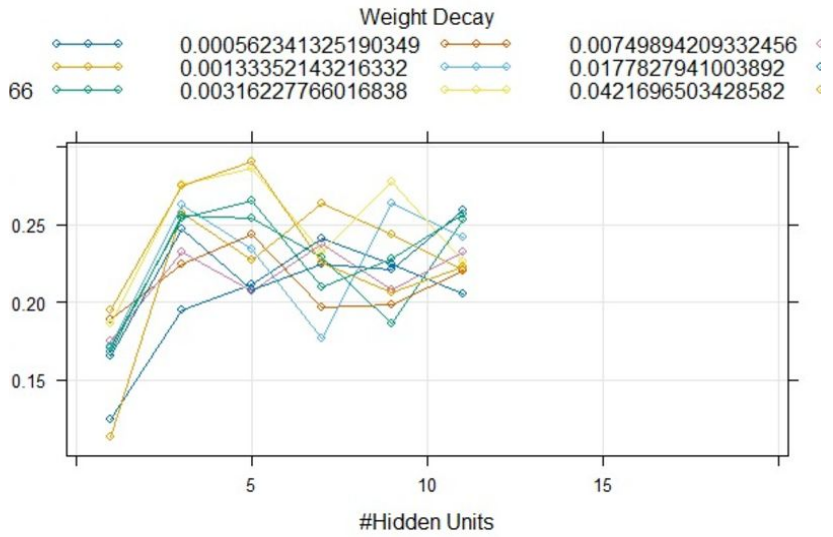
```
               Accuracy : 0.705
                 95% CI : (0.6533, 0.7531)
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 0.9977

                  Kappa : 0.2032

 Mcnemar's Test P-Value : 0.3681

            Sensitivity : 0.7893
            Specificity : 0.4231
         Pos Pred Value : 0.8207
         Neg Pred Value : 0.3750
             Prevalence : 0.7699
         Detection Rate : 0.6077
   Detection Prevalence : 0.7404
      Balanced Accuracy : 0.6062
```

```
Neural Network

1361 samples
  76 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (76), scaled (76)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1224, 1225, 1225, 1226, 1225
Resampling results across tuning parameters:

  size  decay         Accuracy   Kappa
   1    0.0000000000  0.7244113  0.1716757
   1    0.0001000000  0.7295803  0.1830693
   1    0.0002371374  0.6972916  0.1623249
   1    0.0005623413  0.7215620  0.2339528
   1    0.0013335214  0.7046932  0.1509042
   1    0.0031622777  0.6773520  0.1773054
   1    0.0074989421  0.7391177  0.1398685
   1    0.0177827941  0.7016870  0.2063797
   1    0.0421696503  0.7230001  0.2084366
   1    0.1000000000  0.7413234  0.1882287
   3    0.0000000000  0.7259690  0.2207703
   3    0.0001000000  0.7494173  0.2614507
   3    0.0002371374  0.7619283  0.2784134
   3    0.0005623413  0.7392529  0.2164454
   3    0.0013335214  0.7266709  0.2336487
   3    0.0031622777  0.7362357  0.2569384
   3    0.0074989421  0.7517097  0.2650375
   3    0.0177827941  0.7406258  0.2303367
   3    0.0421696503  0.7465242  0.2687773
   3    0.1000000000  0.7288828  0.2580100

Kappa was used to select the optimal model using the
The final values used for the model were size = 3 and
```



Neural Network with Spatial Sign

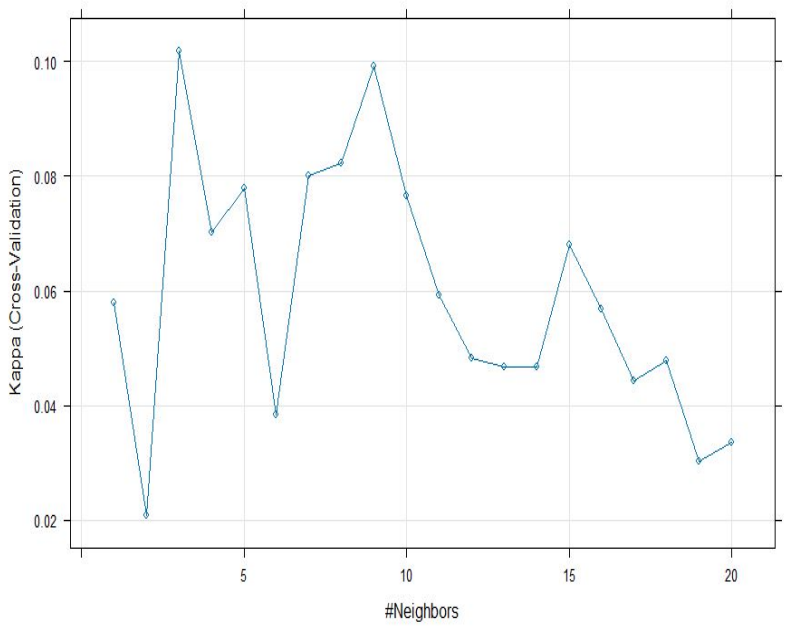
```
> print(nn_model)
Neural Network

1361 samples
  76 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (76), scaled (76)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1224, 1225, 1225, 1226, 1225, 12
Resampling results across tuning parameters:

  size  decay         Accuracy   Kappa
   1    0.0000000000  0.7170905  0.1249979
   1    0.0001000000  0.7325105  0.1139430
   1    0.0002371374  0.6986060  0.1682126
   1    0.0005623413  0.6980377  0.1889974
   1    0.0013335214  0.7105485  0.1713884
   1    0.0031622777  0.7075097  0.1866905
   1    0.0074989421  0.7229519  0.1751967
   1    0.0177827941  0.7001731  0.1655919
   1    0.0421696503  0.7016924  0.1947111
   1    0.1000000000  0.7472657  0.1711608
   3    0.0000000000  0.7141279  0.1951242
   3    0.0001000000  0.7538887  0.2573952
   3    0.0002371374  0.7523694  0.2555640
   3    0.0005623413  0.7495089  0.2241859
   3    0.0013335214  0.7384039  0.2622317
   3    0.0031622777  0.7376901  0.2756312
   3    0.0074989421  0.7465087  0.2327345
   3    0.0177827941  0.7413505  0.2472378
   3    0.0421696503  0.7428103  0.2744565
   3    0.1000000000  0.7201019  0.2540238
   5    0.0000000000  0.7230004  0.2114139
   5    0.0001000000  0.7346952  0.2274783
   5    0.0002371374  0.7443136  0.2540475
   5    0.0005623413  0.7274175  0.2434942
   5    0.0013335214  0.7296397  0.2343058
   5    0.0031622777  0.7385119  0.2857732
   5    0.0074989421  0.7046177  0.2071805
   5    0.0177827941  0.7178857  0.2077334
   5    0.0421696503  0.7457675  0.2904397
   5    0.1000000000  0.7355115  0.2651776

Kappa was used to select the optimal model using the largest value.
The final values used for the model were size = 5 and decay = 0.04216965.
> print(confusion_matrix)
Confusion Matrix and Statistics

                Reference
Prediction       no_complication complication
  no_complication             223            46
  complication                 38            32

               Accuracy : 0.7522
                 95% CI : (0.7027, 0.7973)
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 0.8002

                  Kappa : 0.2745

 Mcnemar's Test P-Value : 0.4450

            Sensitivity : 0.8544
            Specificity : 0.4103
         Pos Pred Value : 0.8290
         Neg Pred Value : 0.4571
             Prevalence : 0.7699
         Detection Rate : 0.6578
   Detection Prevalence : 0.7935
```



KNN Model

```
print(knn_model)
k-Nearest Neighbors

1361 samples
  81 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (81), scaled (81)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1226, 1226, 1225, 1225, 1226, 1225, ...
Resampling results across tuning parameters:

  k   Accuracy   Kappa
   1  0.6913320  0.05799704
   2  0.6804483  0.02083376
   3  0.7465245  0.10184021
   4  0.7384147  0.07012538
   5  0.7582684  0.07796487
   6  0.7502232  0.03846270
   7  0.7685845  0.08000863
   8  0.7671030  0.08230773
   9  0.7736993  0.09915661
  10  0.7700550  0.07660979
  11  0.7700010  0.05922164
  12  0.7692820  0.04831329
  13  0.7700173  0.04678443
  14  0.7700065  0.04678594
  15  0.7759053  0.06799653
  16  0.7751916  0.05691059
  17  0.7737102  0.04436763
  18  0.7737156  0.04774323
  19  0.7715042  0.03028619
  20  0.7714880  0.03359363

Kappa was used to select the optimal model using the largest value.
The final value used for the model was k = 3.
```

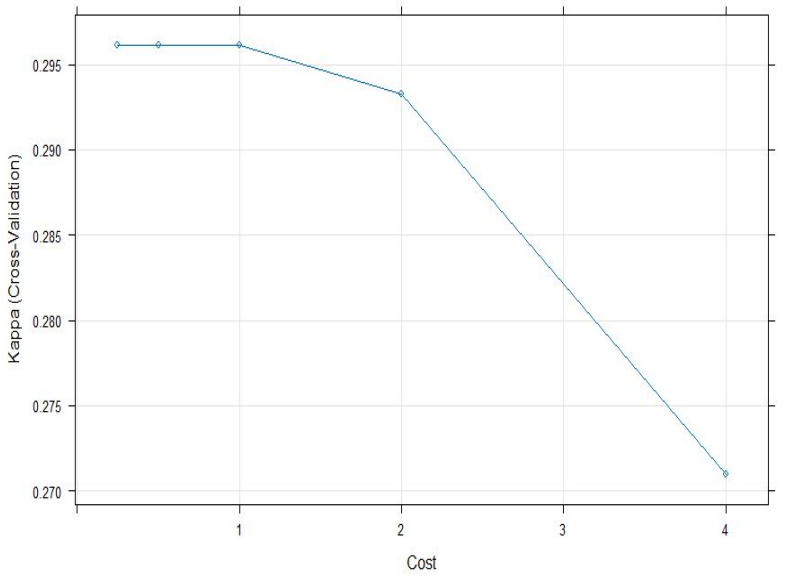
```
print(confusion_matrix_knn)
Confusion Matrix and Statistics

                Reference
Prediction       no_complication complication
  no_complication             238           61
  complication                 23           17

               Accuracy : 0.7522
                 95% CI : (0.7027, 0.7973)
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 0.8002

                  Kappa : 0.1566

 Mcnemar's Test P-Value : 5.413e-05

            Sensitivity : 0.9119
            Specificity : 0.2179
         Pos Pred Value : 0.7960
         Neg Pred Value : 0.4250
             Prevalence : 0.7699
         Detection Rate : 0.7021
   Detection Prevalence : 0.8820
      Balanced Accuracy : 0.5649

       'Positive' Class : no_complication
```



Support Vector Machine Model

```
print(svm_model)
Support Vector Machines with Radial Basis Function Kernel

1361 samples
  81 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (81), scaled (81)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1226, 1225, 1225, 1224, 1225, 1224, ...
Resampling results across tuning parameters:

  C     Accuracy   Kappa
  0.25  0.8126379  0.2961295
  0.50  0.8126379  0.2961295
  1.00  0.8126379  0.2961295
  2.00  0.8089612  0.2932952
  4.00  0.7986884  0.2709781

Tuning parameter 'sigma' was held constant at a value of 0.006848174
Kappa was used to select the optimal model using the largest value.
The final values used for the model were sigma = 0.006848174 and C = 0.25.
```

```
print(confusion_matrix_svm)
Confusion Matrix and Statistics

                Reference
Prediction        no_complication complication
  no_complication             260            62
  complication                  1            16

               Accuracy : 0.8142
                 95% CI : (0.7686, 0.8541)
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 0.02848

                  Kappa : 0.2773

 Mcnemar's Test P-Value : 4.053e-14

            Sensitivity : 0.9962
            Specificity : 0.2051
         Pos Pred Value : 0.8075
         Neg Pred Value : 0.9412
             Prevalence : 0.7699
         Detection Rate : 0.7670
   Detection Prevalence : 0.9499
      Balanced Accuracy : 0.6006

       'Positive' Class : no_complication
```



Naives Bayes Model

```
> print(nb_model)
Naive Bayes

1361 samples
  76 predictor
   2 classes: 'no_complication', 'complication'

Pre-processing: centered (76), scaled (76), Box-Cox transformation (7), spatial sign
 transformation (76)
Resampling: Cross-Validated (10 fold)
Summary of sample sizes: 1226, 1226, 1225, 1224, 1224, 1224, ...
Resampling results:

  Accuracy   Kappa
  0.6620008  0.1803054

Tuning parameter 'fL' was held constant at a value of 2
Tuning parameter 'usekernel' was
 held constant at a value of TRUE
Tuning parameter 'adjust' was held constant at a value of TRUE
>
```

```
> print(confusion_matrix_nb)
Confusion Matrix and Statistics

                Reference
Prediction        no_complication complication
  no_complication             198            46
  complication                 63            32

               Accuracy : 0.6785
                 95% CI : (0.6259, 0.7279)
    No Information Rate : 0.7699
    P-Value [Acc > NIR] : 1.0000

                  Kappa : 0.1569

 Mcnemar's Test P-Value : 0.1254

            Sensitivity : 0.7586
            Specificity : 0.4103
         Pos Pred Value : 0.8115
         Neg Pred Value : 0.3368
             Prevalence : 0.7699
         Detection Rate : 0.5841
   Detection Prevalence : 0.7198
      Balanced Accuracy : 0.5844

       'Positive' Class : no_complication
```

Summary of Linear Models

Models	Best Tuning Parameter	Training Kappa	Testing Kappa
LR	No	0.29960	0.357
LDA	No	0.302857	0.3492
PLSDA	ncomp=4	0.2976	0.3842
Penalized	Alpha=1 & lambda=0.00 03231	0.3478	0.2482

Summary of Non Linear Models

Models	Best Tuning Parameter	Training Kappa	Testing Kappa
QDA	No	0.2841	0.2551
RDA	Gamma=0.7 & Lambda=0.9	0.3457	0.3263
MDA	Subclasses=4	0.30168	0.3514
FDA	Degree=1 & nprune=20	0.3251	0.3127
NN	Size=3 & decay=0.00023	0.2784	0.2032
NN with Spatial Sign	Size = 5, decay = 0.04216	0.29043	0.2745
KNN	K=3	0.1018	0.1566
SVM	Sigma=0.0068 & c=0.25	0.2961	0.2773
NB	No	0.1803	0.1569

Best Models:

Linear Model: PLSDA

- Testing Kappa Value: 0.3842
- Optimal Components: 4

Non-Linear Model: MDA

- Testing Kappa Value: 0.3514
- Optimal Subclasses: 4

- The **PLSDA model** is chosen as the best model based on its superior kappa value .

Variable Importance

```
> varImp(plsda_model)
pls variable importance

  only 20 most important variables shown (out of 81)

                Overall
ZSN_A1           100.00
AGE               32.65
zab_leg_011       30.04
endocr_011        28.55
OTEK_LANC1        26.91
SEX1              26.37
SEX0              26.37
ritm_ecg_p_021    25.68
MP_TP_POST1       24.58
REC_IM1           23.58
FIBR_PREDS1       22.87
lat_im2           20.50
ritm_ecg_p_011    18.55
ALT_BLOOD         17.78
R_AB_3_n1         16.58
D_AD_KBRIG        16.29
GB3               15.64
TIME_B_S6         15.51
ANT_CA_S_n1       15.13
NOT_NA_1_n1       14.08
```

Thank You