MIR_Discrimination_Analysis_R

Zhijun

21/04/2021

Introduction

The objective of this Markdown page is to summary data analysis manuscript for food's data.

Data set

The NIR data was from Quadram Institute. Example Datasets for Download Copyright: In accordance with BBSRC policy on data sharing, many of our datasets which have been collected via publicly funded projects are freely available for further research and analysis. Below is a list of datasets, including concise metadata, ready for download. All datasets are in zipped '.csv' format. The accompanying citations give full experimental details relating to the original data acquisition. These datasets are distributed under the terms of **the Creative Commons Attribution License**, which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

Project 1 - Geographical origin of Extra Virgin Olive Oils (Quadram open dataset,2003)

Load dataset

```
##                V1          V2          V3          V4          V5
## 1  Sample Number:           1           1           2           2
## 2     Group Code:           1           1           1           1
## 3     Wavenumbers      Greece      Greece      Greece      Greece
## 4         798.892 0.127523009 0.126498181 0.130411785 0.130022227
## 5        800.8215 0.127949615 0.127130974 0.130675401 0.130406662
## 6         802.751 0.129282219 0.128510777  0.13201661 0.132018029
## 7        804.6805 0.131174169  0.13033991 0.133824061 0.134007275
## 8          806.61 0.133590328 0.132527221 0.136095296 0.136270568
## 9        808.5395 0.136425525 0.135308508 0.138943757  0.13887477
## 10        810.469 0.139357827  0.13835292 0.141722779 0.141481132
```

Data wrangling

The raw data form are not suitable for R programming, therefore we need to transpose it ti suitable form.

```
setwd('C:/Users/wang298/OneDrive - WageningenUR/Rworkspace/github/R_Foods/Dataset/FTIRSpectraOliveOils')
## data transpose
oil_transpose <- as.data.frame(t(oil))

print(oil_transpose[1:10,1:5])
```

```
##                 V1          V2          V3          V4          V5
## V1  Sample Number: Group Code: Wavenumbers     798.892    800.8215
## V2               1           1      Greece 0.127523009 0.127949615
```

```
## V3               1           1      Greece 0.126498181 0.127130974
## V4               2           1      Greece 0.130411785 0.130675401
## V5               2           1      Greece 0.130022227 0.130406662
## V6               3           1      Greece 0.128601989 0.128789565
## V7               3           1      Greece 0.128217254 0.128282253
## V8               4           1      Greece 0.126174933 0.126732773
## V9               4           1      Greece 0.126466053 0.126915413
## V10              5           1      Greece 0.127060105 0.127551128
```

```
## rename some col names to make data analysis easier

colnames(oil_transpose)=oil_transpose[1,]
oil_transpose <- oil_transpose[-1,]
library(dplyr)
oil_transpose <- rename(oil_transpose,c("Group" ="Group Code:",
                                        "Number" ="Sample Number:",
                                        "Countries" ="Wavenumbers"))
## Get transposed dataset

oil.data <- oil_transpose

print(oil.data[1:10,1:5])
```

```
##     Number Group Countries     798.892    800.8215
## V2       1     1    Greece 0.127523009 0.127949615
## V3       1     1    Greece 0.126498181 0.127130974
## V4       2     1    Greece 0.130411785 0.130675401
## V5       2     1    Greece 0.130022227 0.130406662
## V6       3     1    Greece 0.128601989 0.128789565
## V7       3     1    Greece 0.128217254 0.128282253
## V8       4     1    Greece 0.126174933 0.126732773
## V9       4     1    Greece 0.126466053 0.126915413
## V10      5     1    Greece 0.127060105 0.127551128
## V11      5     1    Greece 0.126812707 0.127460743
```

Principal Component Analysis

The PCA is used to reduce the dimensionality of the spectral value, and initially explore the distribution of the sample.

The original data and related transformated data (MSC, SNV and Savitzky-Golay filtering) are used to obtain the PCA plots.

Raw data

```
df <- oil.data[,4:573]
df <- as.data.frame(apply(df, 2, as.numeric))
print(df[1:10,1:5])
```

```
##      798.892  800.8215   802.751  804.6805    806.61
## 1  0.1275230 0.1279496 0.1292822 0.1311742 0.1335903
## 2  0.1264982 0.1271310 0.1285108 0.1303399 0.1325272
## 3  0.1304118 0.1306754 0.1320166 0.1338241 0.1360953
## 4  0.1300222 0.1304067 0.1320180 0.1340073 0.1362706
## 5  0.1286020 0.1287896 0.1300223 0.1320119 0.1344266
## 6  0.1282173 0.1282823 0.1296366 0.1317986 0.1340615
```

```
## 7  0.1261749 0.1267328 0.1282438 0.1298927 0.1317546
## 8  0.1264661 0.1269154 0.1282541 0.1299583 0.1320672
## 9  0.1270601 0.1275511 0.1289000 0.1306090 0.1329558
## 10 0.1268127 0.1274607 0.1287653 0.1306390 0.1331313
```

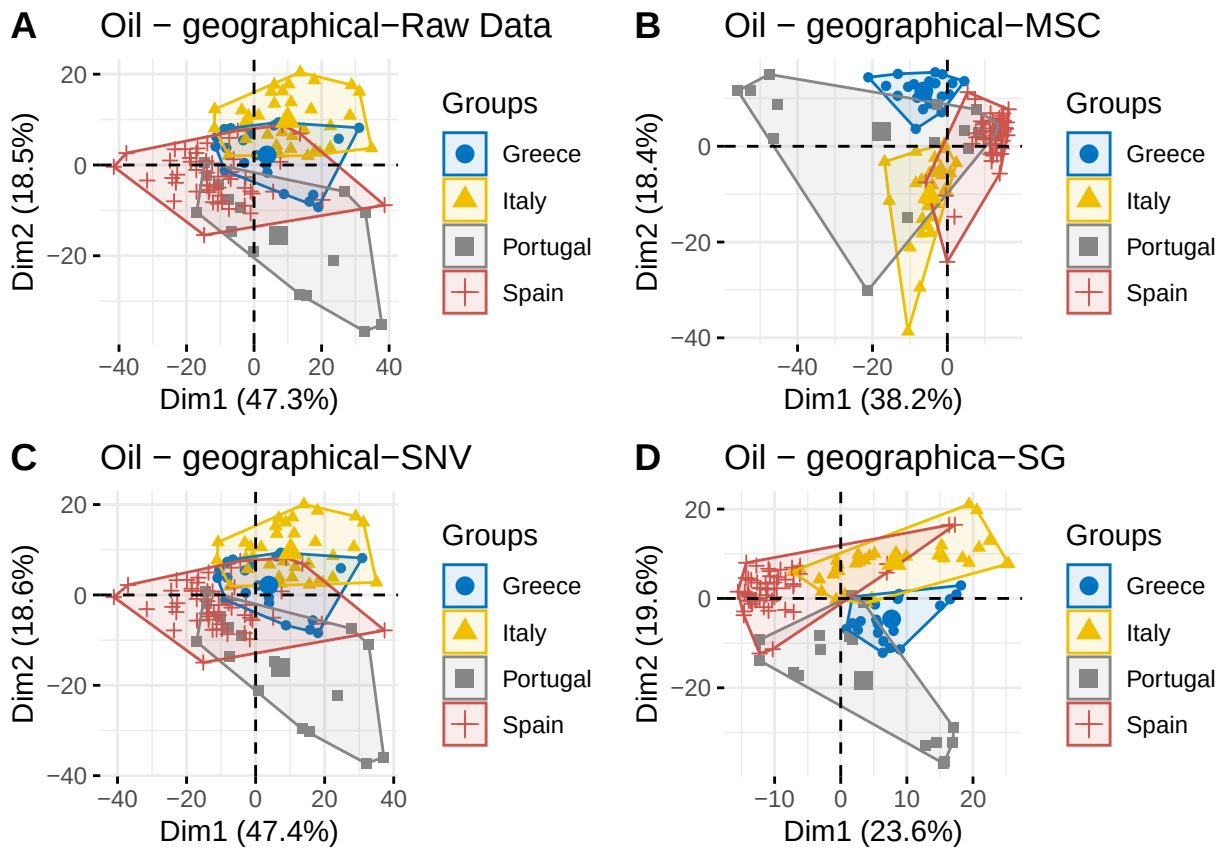
```
oil.pca <- PCA(df,scale.unit = TRUE,ncp = 5,graph = TRUE)
```

Multiplicative Scatter Correction

Standard Normal Variate

Savitzky-Golay filtering

PCA plots



According to PCA plots, the best option is use RAW-SNV data set, therefore following data analysis only RAW-snv transformated data will be used.

PLS-DA

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece        8     0        0     0
##   Italy         0    10        2     2
##   Portugal      0     0        5     0
##   Spain         1     0        0    12
```

```
##
## Overall Statistics
##
##                Accuracy : 0.875
##                  95% CI : (0.732, 0.9581)
##     No Information Rate : 0.35
##     P-Value [Acc > NIR] : 9.085e-12
##
##                   Kappa : 0.8292
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 0.8889       1.0000          0.7143       0.8571
## Specificity                 1.0000       0.8667          1.0000       0.9615
## Pos Pred Value              1.0000       0.7143          1.0000       0.9231
## Neg Pred Value              0.9688       1.0000          0.9429       0.9259
## Prevalence                  0.2250       0.2500          0.1750       0.3500
## Detection Rate              0.2000       0.2500          0.1250       0.3000
## Detection Prevalence        0.2000       0.3500          0.1250       0.3250
## Balanced Accuracy           0.9444       0.9333          0.8571       0.9093
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece        9     0        0     0
##   Italy         0    10        2     2
##   Portugal      0     0        5     0
##   Spain         0     0        0    12
##
## Overall Statistics
##
##                Accuracy : 0.9
##                  95% CI : (0.7634, 0.9721)
##     No Information Rate : 0.35
##     P-Value [Acc > NIR] : 6.678e-13
##
##                   Kappa : 0.8639
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 1.000       1.0000          0.7143       0.8571
## Specificity                 1.000       0.8667          1.0000       1.0000
## Pos Pred Value              1.000       0.7143          1.0000       1.0000
## Neg Pred Value              1.000       1.0000          0.9429       0.9286
## Prevalence                  0.225       0.2500          0.1750       0.3500
## Detection Rate              0.225       0.2500          0.1250       0.3000
## Detection Prevalence        0.225       0.3500          0.1250       0.3000
```

```
## Balanced Accuracy            1.000       0.9333          0.8571       0.9286
```

Other models

Create the training and test datasets

pls

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece       12     0        0     0
##   Italy         3    27        1     4
##   Portugal      1     0        5     0
##   Spain         0     1        7    36
##
## Overall Statistics
##
##                Accuracy : 0.8247
##                  95% CI : (0.7343, 0.8945)
##     No Information Rate : 0.4124
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.7423
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 0.7500       0.9643         0.38462       0.9000
## Specificity                 1.0000       0.8841         0.98810       0.8596
## Pos Pred Value              1.0000       0.7714         0.83333       0.8182
## Neg Pred Value              0.9529       0.9839         0.91209       0.9245
## Precision                   1.0000       0.7714         0.83333       0.8182
## Recall                      0.7500       0.9643         0.38462       0.9000
## F1                          0.8571       0.8571         0.52632       0.8571
## Prevalence                  0.1649       0.2887         0.13402       0.4124
## Detection Rate              0.1237       0.2784         0.05155       0.3711
## Detection Prevalence        0.1237       0.3608         0.06186       0.4536
## Balanced Accuracy           0.8750       0.9242         0.68636       0.8798
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece        1     0        0     0
##   Italy         1     6        1     0
##   Portugal      1     0        1     0
##   Spain         1     0        1    10
##
## Overall Statistics
##
##                Accuracy : 0.7826
```

```
##                  95% CI : (0.563, 0.9254)
##     No Information Rate : 0.4348
##     P-Value [Acc > NIR] : 0.0007403
##
##                   Kappa : 0.6724
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                0.25000       1.0000         0.33333       1.0000
## Specificity                1.00000       0.8824         0.95000       0.8462
## Pos Pred Value             1.00000       0.7500         0.50000       0.8333
## Neg Pred Value             0.86364       1.0000         0.90476       1.0000
## Precision                  1.00000       0.7500         0.50000       0.8333
## Recall                     0.25000       1.0000         0.33333       1.0000
## F1                         0.40000       0.8571         0.40000       0.9091
## Prevalence                 0.17391       0.2609         0.13043       0.4348
## Detection Rate             0.04348       0.2609         0.04348       0.4348
## Detection Prevalence       0.04348       0.3478         0.08696       0.5217
## Balanced Accuracy          0.62500       0.9412         0.64167       0.9231
```

Neural Networks with Feature Extraction

```
model_pcaNNet = train(Countries ~ ., data=trainData, method='pcaNNet')
```

```
pcaNNet_predicted_1 <- predict(model_pcaNNet, trainData[,2:561])
pcaNNet_predicted_3 <- predict(model_pcaNNet, testData[,2:561])

# Compute the confusion matrix
pcaNNet_reference1 <- as.factor(trainData$Countries)
pcaNNet_reference3 <- as.factor(testData$Countries)

confusionMatrix(pcaNNet_reference1, data = pcaNNet_predicted_1,
                mode='everything', positive='Spain')#Calibration
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece       16     0        0     0
##   Italy         0    28        0     0
##   Portugal      0     0       13     0
##   Spain         0     0        0    40
##
## Overall Statistics
##
##                Accuracy : 1
##                  95% CI : (0.9627, 1)
##     No Information Rate : 0.4124
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 1
```

```
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 1.0000       1.0000           1.000       1.0000
## Specificity                 1.0000       1.0000           1.000       1.0000
## Pos Pred Value              1.0000       1.0000           1.000       1.0000
## Neg Pred Value              1.0000       1.0000           1.000       1.0000
## Precision                   1.0000       1.0000           1.000       1.0000
## Recall                      1.0000       1.0000           1.000       1.0000
## F1                          1.0000       1.0000           1.000       1.0000
## Prevalence                  0.1649       0.2887           0.134       0.4124
## Detection Rate              0.1649       0.2887           0.134       0.4124
## Detection Prevalence        0.1649       0.2887           0.134       0.4124
## Balanced Accuracy           1.0000       1.0000           1.000       1.0000
```

```
confusionMatrix(pcaNNet_reference3, data = pcaNNet_predicted_3,
                mode='everything', positive='Spain')#Predict
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece        4     0        0     0
##   Italy         0     6        0     0
##   Portugal      0     0        3     0
##   Spain         0     0        0    10
##
## Overall Statistics
##
##                Accuracy : 1
##                  95% CI : (0.8518, 1)
##     No Information Rate : 0.4348
##     P-Value [Acc > NIR] : 4.789e-09
##
##                   Kappa : 1
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 1.0000       1.0000           1.0000       1.0000
## Specificity                 1.0000       1.0000           1.0000       1.0000
## Pos Pred Value              1.0000       1.0000           1.0000       1.0000
## Neg Pred Value              1.0000       1.0000           1.0000       1.0000
## Precision                   1.0000       1.0000           1.0000       1.0000
## Recall                      1.0000       1.0000           1.0000       1.0000
## F1                          1.0000       1.0000           1.0000       1.0000
## Prevalence                  0.1739       0.2609           0.1304       0.4348
## Detection Rate              0.1739       0.2609           0.1304       0.4348
## Detection Prevalence        0.1739       0.2609           0.1304       0.4348
## Balanced Accuracy           1.0000       1.0000           1.0000       1.0000
```

LDA

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece       16     0        0     0
##   Italy         0    28        0     0
##   Portugal      0     0       13     0
##   Spain         0     0        0    40
##
## Overall Statistics
##
##                Accuracy : 1
##                  95% CI : (0.9627, 1)
##     No Information Rate : 0.4124
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 1
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 1.0000       1.0000           1.000       1.0000
## Specificity                 1.0000       1.0000           1.000       1.0000
## Pos Pred Value              1.0000       1.0000           1.000       1.0000
## Neg Pred Value              1.0000       1.0000           1.000       1.0000
## Precision                   1.0000       1.0000           1.000       1.0000
## Recall                      1.0000       1.0000           1.000       1.0000
## F1                          1.0000       1.0000           1.000       1.0000
## Prevalence                  0.1649       0.2887           0.134       0.4124
## Detection Rate              0.1649       0.2887           0.134       0.4124
## Detection Prevalence        0.1649       0.2887           0.134       0.4124
## Balanced Accuracy           1.0000       1.0000           1.000       1.0000
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece        4     0        0     0
##   Italy         0     6        0     0
##   Portugal      0     0        3     0
##   Spain         0     0        0    10
##
## Overall Statistics
##
##                Accuracy : 1
##                  95% CI : (0.8518, 1)
##     No Information Rate : 0.4348
##     P-Value [Acc > NIR] : 4.789e-09
##
##                   Kappa : 1
##
```

```
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 1.0000       1.0000          1.0000       1.0000
## Specificity                 1.0000       1.0000          1.0000       1.0000
## Pos Pred Value              1.0000       1.0000          1.0000       1.0000
## Neg Pred Value              1.0000       1.0000          1.0000       1.0000
## Precision                   1.0000       1.0000          1.0000       1.0000
## Recall                      1.0000       1.0000          1.0000       1.0000
## F1                          1.0000       1.0000          1.0000       1.0000
## Prevalence                  0.1739       0.2609          0.1304       0.4348
## Detection Rate              0.1739       0.2609          0.1304       0.4348
## Detection Prevalence        0.1739       0.2609          0.1304       0.4348
## Balanced Accuracy           1.0000       1.0000          1.0000       1.0000
```

SVM

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece       16     0        0     0
##   Italy         0    28        0     3
##   Portugal      0     0       12     0
##   Spain         0     0        1    37
##
## Overall Statistics
##
##                Accuracy : 0.9588
##                  95% CI : (0.8978, 0.9887)
##     No Information Rate : 0.4124
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.9413
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 1.0000       1.0000          0.9231       0.9250
## Specificity                 1.0000       0.9565          1.0000       0.9825
## Pos Pred Value              1.0000       0.9032          1.0000       0.9737
## Neg Pred Value              1.0000       1.0000          0.9882       0.9492
## Precision                   1.0000       0.9032          1.0000       0.9737
## Recall                      1.0000       1.0000          0.9231       0.9250
## F1                          1.0000       0.9492          0.9600       0.9487
## Prevalence                  0.1649       0.2887          0.1340       0.4124
## Detection Rate              0.1649       0.2887          0.1237       0.3814
## Detection Prevalence        0.1649       0.3196          0.1237       0.3918
## Balanced Accuracy           1.0000       0.9783          0.9615       0.9537
```

```
## Confusion Matrix and Statistics
```

```
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece        4     0        0     0
##   Italy         0     6        0     0
##   Portugal      0     0        2     1
##   Spain         0     0        1     9
##
## Overall Statistics
##
##                Accuracy : 0.913
##                  95% CI : (0.7196, 0.9893)
##     No Information Rate : 0.4348
##     P-Value [Acc > NIR] : 2.196e-06
##
##                   Kappa : 0.875
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 1.0000       1.0000         0.66667       0.9000
## Specificity                 1.0000       1.0000         0.95000       0.9231
## Pos Pred Value              1.0000       1.0000         0.66667       0.9000
## Neg Pred Value              1.0000       1.0000         0.95000       0.9231
## Precision                   1.0000       1.0000         0.66667       0.9000
## Recall                      1.0000       1.0000         0.66667       0.9000
## F1                          1.0000       1.0000         0.66667       0.9000
## Prevalence                  0.1739       0.2609         0.13043       0.4348
## Detection Rate              0.1739       0.2609         0.08696       0.3913
## Detection Prevalence        0.1739       0.2609         0.13043       0.4348
## Balanced Accuracy           1.0000       1.0000         0.80833       0.9115
```

ELM

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece        0     0        0     1
##   Italy         3    27        2     4
##   Portugal      2     0        5     0
##   Spain        11     1        6    35
##
## Overall Statistics
##
##                Accuracy : 0.6907
##                  95% CI : (0.5888, 0.7807)
##     No Information Rate : 0.4124
##     P-Value [Acc > NIR] : 2.831e-08
##
##                   Kappa : 0.5287
##
```

```
##  Mcnemar's Test P-Value : 0.000753
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                0.00000       0.9643         0.38462       0.8750
## Specificity                0.98765       0.8696         0.97619       0.6842
## Pos Pred Value             0.00000       0.7500         0.71429       0.6604
## Neg Pred Value             0.83333       0.9836         0.91111       0.8864
## Precision                  0.00000       0.7500         0.71429       0.6604
## Recall                     0.00000       0.9643         0.38462       0.8750
## F1                             NaN       0.8437         0.50000       0.7527
## Prevalence                 0.16495       0.2887         0.13402       0.4124
## Detection Rate             0.00000       0.2784         0.05155       0.3608
## Detection Prevalence       0.01031       0.3711         0.07216       0.5464
## Balanced Accuracy          0.49383       0.9169         0.68040       0.7796
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece        0     0        0     0
##   Italy         2     6        1     0
##   Portugal      0     0        1     0
##   Spain         2     0        1    10
##
## Overall Statistics
##
##                Accuracy : 0.7391
##                  95% CI : (0.5159, 0.8977)
##     No Information Rate : 0.4348
##     P-Value [Acc > NIR] : 0.003074
##
##                   Kappa : 0.5965
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 0.0000       1.0000         0.33333       1.0000
## Specificity                 1.0000       0.8235         1.00000       0.7692
## Pos Pred Value                 NaN       0.6667         1.00000       0.7692
## Neg Pred Value              0.8261       1.0000         0.90909       1.0000
## Precision                       NA       0.6667         1.00000       0.7692
## Recall                      0.0000       1.0000         0.33333       1.0000
## F1                              NA       0.8000         0.50000       0.8696
## Prevalence                  0.1739       0.2609         0.13043       0.4348
## Detection Rate              0.0000       0.2609         0.04348       0.4348
## Detection Prevalence        0.0000       0.3913         0.04348       0.5652
## Balanced Accuracy           0.5000       0.9118         0.66667       0.8846
```

KNN

```
## Confusion Matrix and Statistics
```

```
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece       16     0        0     0
##   Italy         0    28        1     4
##   Portugal      0     0       12     0
##   Spain         0     0        0    36
##
## Overall Statistics
##
##                Accuracy : 0.9485
##                  95% CI : (0.8838, 0.9831)
##     No Information Rate : 0.4124
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.9269
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 1.0000       1.0000          0.9231       0.9000
## Specificity                 1.0000       0.9275          1.0000       1.0000
## Pos Pred Value              1.0000       0.8485          1.0000       1.0000
## Neg Pred Value              1.0000       1.0000          0.9882       0.9344
## Precision                   1.0000       0.8485          1.0000       1.0000
## Recall                      1.0000       1.0000          0.9231       0.9000
## F1                          1.0000       0.9180          0.9600       0.9474
## Prevalence                  0.1649       0.2887          0.1340       0.4124
## Detection Rate              0.1649       0.2887          0.1237       0.3711
## Detection Prevalence        0.1649       0.3402          0.1237       0.3711
## Balanced Accuracy           1.0000       0.9638          0.9615       0.9500
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece        4     0        1     0
##   Italy         0     6        1     0
##   Portugal      0     0        1     1
##   Spain         0     0        0     9
##
## Overall Statistics
##
##                Accuracy : 0.8696
##                  95% CI : (0.6641, 0.9722)
##     No Information Rate : 0.4348
##     P-Value [Acc > NIR] : 2.083e-05
##
##                   Kappa : 0.814
##
##  Mcnemar's Test P-Value : NA
##
```

```
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 1.0000       1.0000         0.33333       0.9000
## Specificity                 0.9474       0.9412         0.95000       1.0000
## Pos Pred Value              0.8000       0.8571         0.50000       1.0000
## Neg Pred Value              1.0000       1.0000         0.90476       0.9286
## Precision                   0.8000       0.8571         0.50000       1.0000
## Recall                      1.0000       1.0000         0.33333       0.9000
## F1                          0.8889       0.9231         0.40000       0.9474
## Prevalence                  0.1739       0.2609         0.13043       0.4348
## Detection Rate              0.1739       0.2609         0.04348       0.3913
## Detection Prevalence        0.2174       0.3043         0.08696       0.3913
## Balanced Accuracy           0.9737       0.9706         0.64167       0.9500
```

SIMCA

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece        7     0        0     0
##   Italy         0    21        0     1
##   Portugal      9     7       13     8
##   Spain         0     0        0    31
##
## Overall Statistics
##
##                Accuracy : 0.7423
##                  95% CI : (0.6435, 0.8258)
##     No Information Rate : 0.4124
##     P-Value [Acc > NIR] : 4.477e-11
##
##                   Kappa : 0.6516
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                0.43750       0.7500          1.0000       0.7750
## Specificity                1.00000       0.9855          0.7143       1.0000
## Pos Pred Value             1.00000       0.9545          0.3514       1.0000
## Neg Pred Value             0.90000       0.9067          1.0000       0.8636
## Precision                  1.00000       0.9545          0.3514       1.0000
## Recall                     0.43750       0.7500          1.0000       0.7750
## F1                         0.60870       0.8400          0.5200       0.8732
## Prevalence                 0.16495       0.2887          0.1340       0.4124
## Detection Rate             0.07216       0.2165          0.1340       0.3196
## Detection Prevalence       0.07216       0.2268          0.3814       0.3196
## Balanced Accuracy          0.71875       0.8678          0.8571       0.8875
```

```
## Confusion Matrix and Statistics
##
##           Reference
```

```
## Prediction Greece Italy Portugal Spain
##   Greece        0     0        0     0
##   Italy         0     5        0     0
##   Portugal      4     1        3     3
##   Spain         0     0        0     7
##
## Overall Statistics
##
##                Accuracy : 0.6522
##                  95% CI : (0.4273, 0.8362)
##     No Information Rate : 0.4348
##     P-Value [Acc > NIR] : 0.0296
##
##                   Kappa : 0.5354
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 0.0000       0.8333          1.0000       0.7000
## Specificity                 1.0000       1.0000          0.6000       1.0000
## Pos Pred Value                 NaN       1.0000          0.2727       1.0000
## Neg Pred Value              0.8261       0.9444          1.0000       0.8125
## Precision                      NA       1.0000          0.2727       1.0000
## Recall                      0.0000       0.8333          1.0000       0.7000
## F1                             NA       0.9091          0.4286       0.8235
## Prevalence                  0.1739       0.2609          0.1304       0.4348
## Detection Rate              0.0000       0.2174          0.1304       0.3043
## Detection Prevalence        0.0000       0.2174          0.4783       0.3043
## Balanced Accuracy           0.5000       0.9167          0.8000       0.8500
```

Random Forest

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece       16     0        0     0
##   Italy         0    28        0     0
##   Portugal      0     0       13     0
##   Spain         0     0        0    40
##
## Overall Statistics
##
##                Accuracy : 1
##                  95% CI : (0.9627, 1)
##     No Information Rate : 0.4124
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 1
##
##  Mcnemar's Test P-Value : NA
##
```

```
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 1.0000       1.0000           1.000       1.0000
## Specificity                 1.0000       1.0000           1.000       1.0000
## Pos Pred Value              1.0000       1.0000           1.000       1.0000
## Neg Pred Value              1.0000       1.0000           1.000       1.0000
## Precision                   1.0000       1.0000           1.000       1.0000
## Recall                      1.0000       1.0000           1.000       1.0000
## F1                          1.0000       1.0000           1.000       1.0000
## Prevalence                  0.1649       0.2887           0.134       0.4124
## Detection Rate              0.1649       0.2887           0.134       0.4124
## Detection Prevalence        0.1649       0.2887           0.134       0.4124
## Balanced Accuracy           1.0000       1.0000           1.000       1.0000
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction Greece Italy Portugal Spain
##   Greece        4     0        0     0
##   Italy         0     6        0     0
##   Portugal      0     0        3     1
##   Spain         0     0        0     9
##
## Overall Statistics
##
##                Accuracy : 0.9565
##                  95% CI : (0.7805, 0.9989)
##     No Information Rate : 0.4348
##     P-Value [Acc > NIR] : 1.48e-07
##
##                   Kappa : 0.9387
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: Greece Class: Italy Class: Portugal Class: Spain
## Sensitivity                 1.0000       1.0000          1.0000       0.9000
## Specificity                 1.0000       1.0000          0.9500       1.0000
## Pos Pred Value              1.0000       1.0000          0.7500       1.0000
## Neg Pred Value              1.0000       1.0000          1.0000       0.9286
## Precision                   1.0000       1.0000          0.7500       1.0000
## Recall                      1.0000       1.0000          1.0000       0.9000
## F1                          1.0000       1.0000          0.8571       0.9474
## Prevalence                  0.1739       0.2609          0.1304       0.4348
## Detection Rate              0.1739       0.2609          0.1304       0.3913
## Detection Prevalence        0.1739       0.2609          0.1739       0.3913
## Balanced Accuracy           1.0000       1.0000          0.9750       0.9500
```