Prediction Assignment Writeup

Abhijeet Mohite

25/06/2020

##Enviornment Setup

```
## Warning: package 'caret' was built under R version 3.6.3
```

```
## Loading required package: lattice
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 3.6.3
```

```
## Warning: package 'randomForest' was built under R version 3.6.3
```

```
## randomForest 4.6-14
```

```
## Type rfNews() to see new features/changes/bug fixes.
```

```
##
## Attaching package: 'randomForest'
```

```
## The following object is masked from 'package:ggplot2':
##
##     margin
```

```
## Warning: package 'e1071' was built under R version 3.6.3
```

```
## Warning: package 'gbm' was built under R version 3.6.3
```

```
## Loaded gbm 2.1.5
```

```
## [1] 19622   160
```
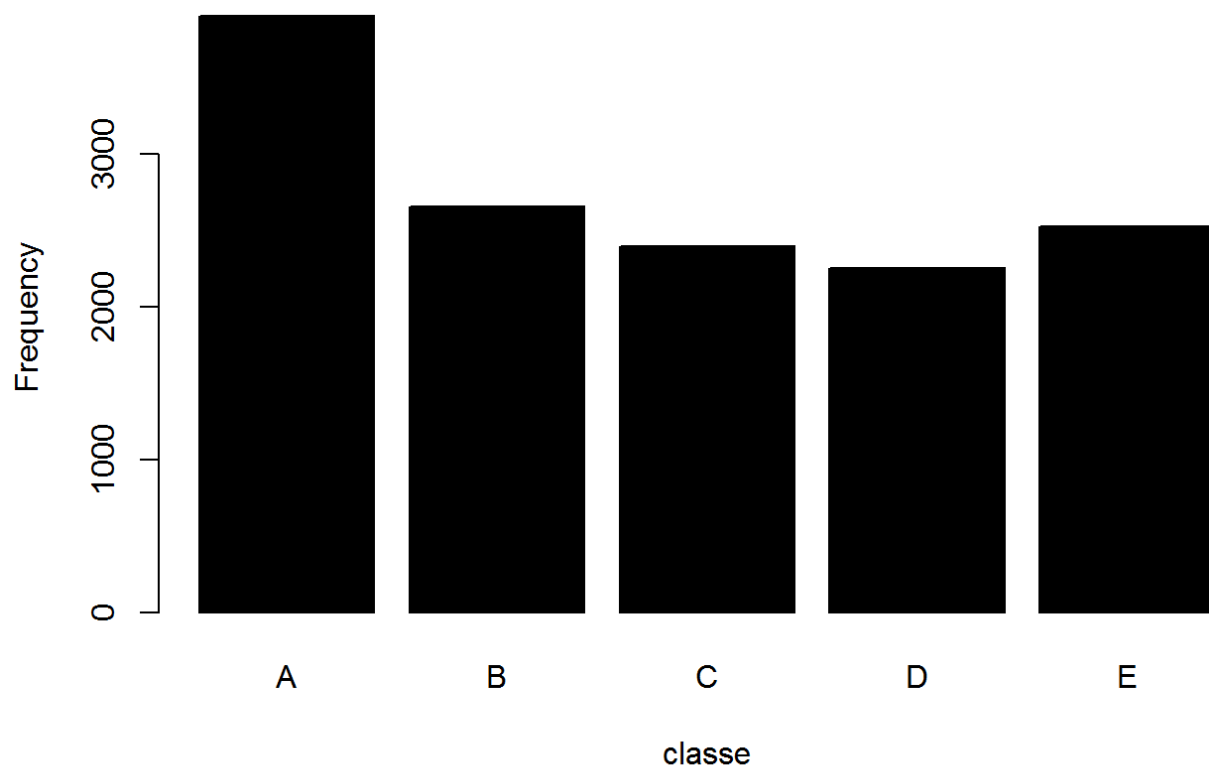
```
## [1]  20 160
```

```
## [1] 19622    53
```

```
## [1] 20 53
```

```
## [1] 13737      53
```

```
## [1] 5885     53
```

Frequency of different levels

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 1674    3    0    0    0
##          B    0 1132   14    0    0
##          C    0    4 1012    7    0
##          D    0    0    0  956    0
##          E    0    0    0    1 1082
##
## Overall Statistics
##
##                Accuracy : 0.9951
##                  95% CI : (0.9929, 0.9967)
##     No Information Rate : 0.2845
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.9938
##
##  Mcnemar's Test P-Value : NA
##
## Statistics by Class:
##
##                      Class: A Class: B Class: C Class: D Class: E
## Sensitivity            1.0000   0.9939   0.9864   0.9917   1.0000
## Specificity            0.9993   0.9971   0.9977   1.0000   0.9998
## Pos Pred Value         0.9982   0.9878   0.9892   1.0000   0.9991
## Neg Pred Value         1.0000   0.9985   0.9971   0.9984   1.0000
## Prevalence             0.2845   0.1935   0.1743   0.1638   0.1839
## Detection Rate         0.2845   0.1924   0.1720   0.1624   0.1839
## Detection Prevalence   0.2850   0.1947   0.1738   0.1624   0.1840
## Balanced Accuracy      0.9996   0.9955   0.9920   0.9959   0.9999
```

```
## Confusion Matrix and Statistics
##
##           Reference
## Prediction    A    B    C    D    E
##          A 1651   30    0    4    1
##          B   14 1080   35    4    6
##          C    5   29  979   20   17
##          D    4    0   11  929    7
##          E    0    0    1    7 1051
##
## Overall Statistics
##
##                Accuracy : 0.9669
##                  95% CI : (0.962, 0.9713)
##     No Information Rate : 0.2845
##     P-Value [Acc > NIR] : < 2.2e-16
##
##                   Kappa : 0.9581
##
##  Mcnemar's Test P-Value : 2.327e-05
##
## Statistics by Class:
##
##                     Class: A Class: B Class: C Class: D Class: E
## Sensitivity           0.9863   0.9482   0.9542   0.9637   0.9713
## Specificity           0.9917   0.9876   0.9854   0.9955   0.9983
## Pos Pred Value        0.9792   0.9482   0.9324   0.9769   0.9924
## Neg Pred Value        0.9945   0.9876   0.9903   0.9929   0.9936
## Prevalence            0.2845   0.1935   0.1743   0.1638   0.1839
## Detection Rate        0.2805   0.1835   0.1664   0.1579   0.1786
## Detection Prevalence  0.2865   0.1935   0.1784   0.1616   0.1799
## Balanced Accuracy     0.9890   0.9679   0.9698   0.9796   0.9848
```

```
##  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
##  B  A  B  A  A  E  D  B  A  A  B  C  B  A  E  E  A  B  B  B
## Levels: A B C D E
```