R_Implementation

Rachael Joan Dias

12/10/2018

```
library(readr)
#Load the csv file and save as a data frame
cardio<-read.csv("Cardiotocographic.csv")
str(cardio)
```

```
## 'data.frame':    2126 obs. of  22 variables:
##  $ LB      : int  120 132 133 134 132 134 134 122 122 122 ...
##  $ AC      : num  0 0.00638 0.00332 0.00256 0.00651 ...
##  $ FM      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ UC      : num  0 0.00638 0.00831 0.00768 0.00814 ...
##  $ DL      : num  0 0.00319 0.00332 0.00256 0 ...
##  $ DS      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ DP      : num  0 0 0 0 0 ...
##  $ ASTV    : int  73 17 16 16 16 26 29 83 84 86 ...
##  $ MSTV    : num  0.5 2.1 2.1 2.4 2.4 5.9 6.3 0.5 0.5 0.3 ...
##  $ ALTV    : int  43 0 0 0 0 0 0 6 5 6 ...
##  $ MLTV    : num  2.4 10.4 13.4 23 19.9 0 0 15.6 13.6 10.6 ...
##  $ Width   : int  64 130 130 117 117 150 150 68 68 68 ...
##  $ Min     : int  62 68 68 53 53 50 50 62 62 62 ...
##  $ Max     : int  126 198 198 170 170 200 200 130 130 130 ...
##  $ Nmax    : int  2 6 5 11 9 5 6 0 0 1 ...
##  $ Nzeros  : int  0 1 1 0 0 3 3 0 0 0 ...
##  $ Mode    : int  120 141 141 137 137 76 71 122 122 122 ...
##  $ Mean    : int  137 136 135 134 136 107 107 122 122 122 ...
##  $ Median  : int  121 140 138 137 138 107 106 123 123 123 ...
##  $ Variance: int  73 12 13 13 11 170 215 3 3 1 ...
##  $ Tendency: int  1 0 0 1 1 0 0 1 1 1 ...
##  $ NSP     : int  2 1 1 1 1 3 3 3 3 3 ...
```

```
#Convert the outcome variable NSP to factors
cardio$NSPF<-as.factor(cardio$NSP)
str(cardio)
```

```
## 'data.frame':    2126 obs. of  23 variables:
##  $ LB      : int  120 132 133 134 132 134 134 122 122 122 ...
##  $ AC      : num  0 0.00638 0.00332 0.00256 0.00651 ...
##  $ FM      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ UC      : num  0 0.00638 0.00831 0.00768 0.00814 ...
##  $ DL      : num  0 0.00319 0.00332 0.00256 0 ...
##  $ DS      : num  0 0 0 0 0 0 0 0 0 0 ...
##  $ DP      : num  0 0 0 0 0 ...
##  $ ASTV    : int  73 17 16 16 16 26 29 83 84 86 ...
##  $ MSTV    : num  0.5 2.1 2.1 2.4 2.4 5.9 6.3 0.5 0.5 0.3 ...
##  $ ALTV    : int  43 0 0 0 0 0 0 6 5 6 ...
##  $ MLTV    : num  2.4 10.4 13.4 23 19.9 0 0 15.6 13.6 10.6 ...
##  $ Width   : int  64 130 130 117 117 150 150 68 68 68 ...
##  $ Min     : int  62 68 68 53 53 50 50 62 62 62 ...
##  $ Max     : int  126 198 198 170 170 200 200 130 130 130 ...
```

```
##  $ Nmax    : int  2 6 5 11 9 5 6 0 0 1 ...
##  $ Nzeros  : int  0 1 1 0 0 3 3 0 0 0 ...
##  $ Mode    : int  120 141 141 137 137 76 71 122 122 122 ...
##  $ Mean    : int  137 136 135 134 136 107 107 122 122 122 ...
##  $ Median  : int  121 140 138 137 138 107 106 123 123 123 ...
##  $ Variance: int  73 12 13 13 11 170 215 3 3 1 ...
##  $ Tendency: int  1 0 0 1 1 0 0 1 1 1 ...
##  $ NSP     : int  2 1 1 1 1 3 3 3 3 3 ...
##  $ NSPF    : Factor w/ 3 levels "1","2","3": 2 1 1 1 1 3 3 3 3 3 ...
```

```
#Multinomial Logistic regression
library(nnet)

#Considering level 1 which represents a normal patient as the reference level
cardio$out<-relevel(cardio$NSPF, ref="1")

#Fit a multinomial logistic regression model
model<-multinom(out~LB+AC+FM, data=cardio)
```

```
## # weights:  15 (8 variable)
## initial  value 2335.649726
## iter  10 value 1289.818570
## iter  20 value 1041.240748
## iter  30 value 1036.259309
## final  value 1019.987655
## converged
```

```
#Print summary of the model
summary(model)
```

```
## Call:
## multinom(formula = out ~ LB + AC + FM, data = cardio)
##
## Coefficients:
##   (Intercept)           LB        AC       FM
## 2 -16.2182977  0.112918884 -829.1624 6.137294
## 3  -0.4208594 -0.006730701 -789.8814 8.231494
##
## Std. Errors:
##   (Intercept)          LB          AC       FM
## 2    1.261066 0.009050217 0.005518354 1.746013
## 3    1.212057 0.009170115 0.009872315 1.372880
##
## Residual Deviance: 2039.975
## AIC: 2055.975
```

```
#2-tailed z test
z <- summary(model)$coefficients/summary(model)$standard.errors
p <- (1 - pnorm(abs(z), 0, 1)) * 2
p
```

```
##   (Intercept)        LB AC           FM
## 2   0.0000000 0.0000000  0 4.396984e-04
## 3   0.7284205 0.4629596  0 2.025026e-09
```

Predictions

```r
#Printing prediction of our model
head(predict(model,cardio, type="prob"))
```

```
##           1           2           3
## 1 0.7341566 0.050942148 0.214901289
## 2 0.9969034 0.001352476 0.001744078
## 3 0.9628285 0.018450599 0.018720927
## 4 0.9297324 0.037502013 0.032765629
## 5 0.9972224 0.001209563 0.001568084
## 6 0.7951989 0.112322025 0.092479054
```

Misclassification Error

```r
#Print confusion matrix of the model
x<-table(predict(model),cardio$NSPF)
print(x)
```

```
##
##        1    2    3
##   1 1592  165  137
##   2   61  128   27
##   3    2    2   12
```

```r
#Error of the model
1-sum(diag(x))/sum(x)
```

```
## [1] 0.1853246
```