OPD classification Project

Mohit kulpe

October 14, 2017

Objective :

classifying patients to one out of two categories, Normal or Abnormal. Use biomechanical features to classify patients according to their Class by using Multiple Logistic Regression Algorithm.

Approach:

1. Check Missing Values 'NA'
2. Correlations between variables
3. Visualize Highly Correlated variables
4. Checking Outliers
5. Handling Outliers
6. Summarizing datasets
7. Model Fitting
8. Overall fraction of correct predictions (Accuracy)
9. Trying to increase fraction of correct predictions (Accuracy)
10. Cross Validation
11. Plotting ROC
12. Testing Model on random data

References:

https://www.kaggle.com/uciml/biomechanical-features-of-orthopedic-patients
Dataset- http://archive.ics.uci.edu/ml/datasets/vertebral+column

Loading Data

```
library(ISLR)
library(MASS)

mydata <- read.csv('OPD_2C.csv')

head(mydata)
```

```
##   pelvic_incidence pelvic_tilt.numeric lumbar_lordosis_angle sacral_slope
## 1         63.02782           22.552586              39.60912     40.47523
## 2         39.05695           10.060991              25.01538     28.99596
## 3         68.83202           22.218482              50.09219     46.61354
## 4         69.29701           24.652878              44.31124     44.64413
## 5         49.71286            9.652075              28.31741     40.06078
## 6         40.25020           13.921907              25.12495     26.32829
##   pelvic_radius degree_spondylolisthesis    class
## 1      98.67292                -0.254400 Abnormal
## 2     114.40543                 4.564259 Abnormal
## 3     105.98514                -3.530317 Abnormal
## 4     101.86850                11.211523 Abnormal
## 5     108.16872                 7.918501 Abnormal
## 6     130.32787                 2.230652 Abnormal
```

```
dim(mydata)
```

```
## [1] 310   7
```

There are 310 observations (Rows) & 7 Columns

1. Checking for Missing Values 'NA'

```
sum(is.na(mydata))
```

```
## [1] 0
```

There are No Missing Values.

2. Correlations between variables

```
cor(mydata[1:6])
```

```
##                           pelvic_incidence pelvic_tilt.numeric
## pelvic_incidence                 1.0000000          0.62919877
## pelvic_tilt.numeric              0.6291988          1.00000000
## lumbar_lordosis_angle            0.7172824          0.43276386
## sacral_slope                     0.8149600          0.06234529
## pelvic_radius                   -0.2474672          0.03266781
## degree_spondylolisthesis         0.6387427          0.39786228
##                           lumbar_lordosis_angle sacral_slope pelvic_radius
## pelvic_incidence                     0.71728236   0.81495999   -0.24746721
## pelvic_tilt.numeric                  0.43276386   0.06234529    0.03266781
## lumbar_lordosis_angle                1.00000000   0.59838689   -0.08034361
## sacral_slope                         0.59838689   1.00000000   -0.34212835
## pelvic_radius                       -0.08034361  -0.34212835    1.00000000
## degree_spondylolisthesis             0.53366701   0.52355746   -0.02606501
##                           degree_spondylolisthesis
## pelvic_incidence                        0.63874275
## pelvic_tilt.numeric                     0.39786228
## lumbar_lordosis_angle                   0.53366701
## sacral_slope                            0.52355746
## pelvic_radius                          -0.02606501
## degree_spondylolisthesis                1.00000000
```

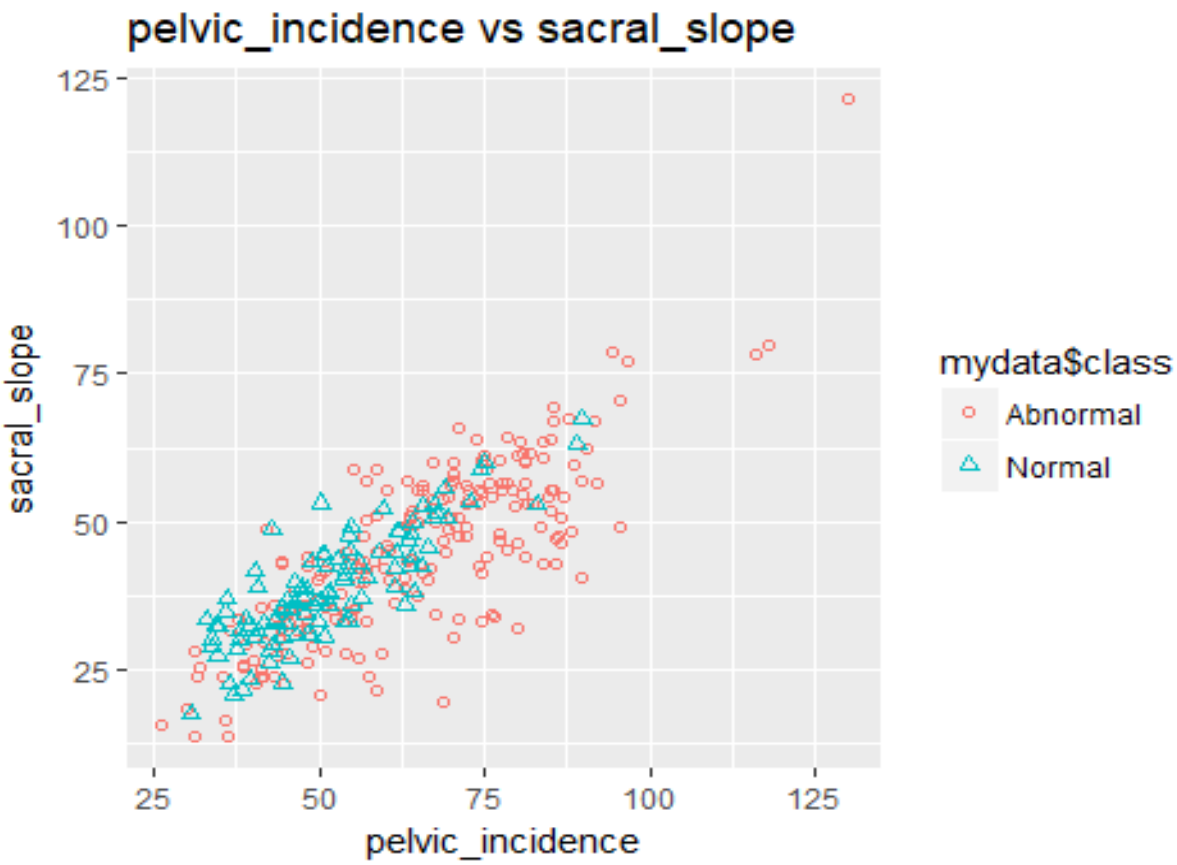
Features pelvic_incidence & sacral_slope are highly correlated with positive correlation value = 0.8149600.

3. Visualize Highly Correlated variables

```
library(ggplot2)
library(gridExtra)
```

```
## Warning: package 'gridExtra' was built under R version 3.4.2
```

```
x <- qplot(x=mydata$pelvic_incidence, y=mydata$sacral_slope,
color=mydata$class, shape=mydata$class, geom='point', main =
'pelvic_incidence vs sacral_slope', ylab = 'sacral_slope', xlab =
'pelvic_incidence')+scale_shape(solid=FALSE)
x
```



The plot shows that there are Outliers present in data, as the three values are far away from rest values.

There might be Outliers in each Feature, So let's Check Outliers in each feature.

4. Checking Outliers

I will mark Outliers as a value which is 3 standard deviations away from the Mean.

```r
# Checking Outliers in pelvic_incidence

urange1 = mean(mydata$pelvic_incidence)+3*sd(mydata$pelvic_incidence)

lrange1 = mean(mydata$pelvic_incidence)-3*sd(mydata$pelvic_incidence)

o1 = which(mydata$pelvic_incidence < lrange1 | mydata$pelvic_incidence >
urange1)


# Checking Outliers in pelvic_tilt.numeric
urange2 = mean(mydata$pelvic_tilt.numeric)+3*sd(mydata$pelvic_tilt.numeric)

lrange2 = mean(mydata$pelvic_tilt.numeric)-3*sd(mydata$pelvic_tilt.numeric)

o2 = which(mydata$pelvic_tilt.numeric < lrange2 | mydata$pelvic_tilt.numeric
> urange2)

outlier <- append(o1,o2)


# Checking Outliers in pelvic_radius
urange3 = mean(mydata$pelvic_radius)+3*sd(mydata$pelvic_radius)

lrange3 = mean(mydata$pelvic_radius)-3*sd(mydata$pelvic_radius)

o3 = which(mydata$pelvic_radius < lrange3 | mydata$pelvic_radius > urange3)

outlier <- append(outlier,o3)


# Checking Outliers in lumbar_lordosis_angle
urange4 =
mean(mydata$lumbar_lordosis_angle)+3*sd(mydata$lumbar_lordosis_angle)

lrange4 = mean(mydata$lumbar_lordosis_angle)-
3*sd(mydata$lumbar_lordosis_angle)

o4 = which(mydata$lumbar_lordosis_angle < lrange4 |
mydata$lumbar_lordosis_angle > urange4)

outlier <- append(outlier,o4)


# Checking Outliers in sacral_slope
urange5 = mean(mydata$sacral_slope)+3*sd(mydata$sacral_slope)

lrange5 = mean(mydata$sacral_slope)-3*sd(mydata$sacral_slope)
```

```
o5 = which(mydata$sacral_slope < lrange5 | mydata$sacral_slope > urange5)

outlier <- append(outlier,o5)


# Checking Outliers in degree_spondylolisthesis
urange6 =
mean(mydata$degree_spondylolisthesis)+3*sd(mydata$degree_spondylolisthesis)

lrange6 = mean(mydata$degree_spondylolisthesis)-
3*sd(mydata$degree_spondylolisthesis)

o6 = which(mydata$degree_spondylolisthesis < lrange6 |
mydata$degree_spondylolisthesis > urange6)

outlier <- append(outlier,o6)



# Outlier present in data with respect to row numbers

uni <- sort(unique(outlier))
uni
```

```
##  [1]  76  86  96 116 123 142 163 164 168 180 198
```

There are total 11 rows in which outliers are present.

5. Handling Outliers

By using two techniques:

1. By deleting entire row in which outlier is present.

2. By replacing outlier with Median.

1.Delete Entire Row

```
newdata <- mydata[-uni,]
dim(newdata)

## [1] 299   7
```

We have deleted 11 rows (310 - 11 = 299)

2. Replace by Median

Here we will use original copy of dataset.

```
mydata1 <- read.csv('OPD_2C.csv')

#replace by Median 1
mydata1$pelvic_incidence[which(mydata1$pelvic_incidence < lrange1 |
mydata1$pelvic_incidence > urange1)] <- median(mydata1$pelvic_incidence)

#replace by Median 2
mydata1$pelvic_tilt.numeric[which(mydata1$pelvic_tilt.numeric < lrange2 |
mydata1$pelvic_tilt.numeric > urange2)] <-
median(mydata1$pelvic_tilt.numeric)

#replace by Median 3
mydata1$pelvic_radius[which(mydata1$pelvic_radius < lrange3 |
mydata1$pelvic_radius > urange3)] <- median(mydata1$pelvic_radius)

#replace by Median 4
mydata1$lumbar_lordosis_angle[which(mydata1$lumbar_lordosis_angle < lrange4 |
mydata1$lumbar_lordosis_angle > urange4)] <-
median(mydata1$lumbar_lordosis_angle)

#replace by Median 5
mydata1$sacral_slope[which(mydata1$sacral_slope < lrange5 |
mydata1$sacral_slope > urange5)] <- median(mydata1$sacral_slope)

#replace by Median 6
mydata1$degree_spondylolisthesis[which(mydata1$degree_spondylolisthesis <
lrange6 | mydata1$degree_spondylolisthesis > urange6)] <-
median(mydata1$degree_spondylolisthesis)
```

Now we have three different datasets:

1. Data with Outliers

2. Data without Outliers (Row Deletion Method)

3. Data without Outliers (Replace with Median Method)

6. Summarizing datasets

```
# Summary of original Dataset with outliers
summary(mydata)
```

```
##  pelvic_incidence pelvic_tilt.numeric lumbar_lordosis_angle
##  Min.   : 26.15   Min.   :-6.555      Min.   : 14.00
##  1st Qu.: 46.43   1st Qu.:10.667      1st Qu.: 37.00
##  Median : 58.69   Median :16.358      Median : 49.56
##  Mean   : 60.50   Mean   :17.543      Mean   : 51.93
##  3rd Qu.: 72.88   3rd Qu.:22.120      3rd Qu.: 63.00
##  Max.   :129.83   Max.   :49.432      Max.   :125.74
##   sacral_slope    pelvic_radius    degree_spondylolisthesis      class
##  Min.   : 13.37   Min.   : 70.08   Min.   :-11.058          Abnormal:210
##  1st Qu.: 33.35   1st Qu.:110.71   1st Qu.:  1.604          Normal  :100
##  Median : 42.40   Median :118.27   Median : 11.768
##  Mean   : 42.95   Mean   :117.92   Mean   : 26.297
##  3rd Qu.: 52.70   3rd Qu.:125.47   3rd Qu.: 41.287
##  Max.   :121.43   Max.   :163.07   Max.   :418.543
```

```
# Summary of Dataset without outliers (Replace with median method)

summary(newdata)
```

```
##  pelvic_incidence pelvic_tilt.numeric lumbar_lordosis_angle
##  Min.   :26.15    Min.   :-6.555      Min.   : 14.00
##  1st Qu.:46.38    1st Qu.:10.600      1st Qu.: 36.66
##  Median :57.30    Median :15.969      Median : 48.50
##  Mean   :59.69    Mean   :16.957      Mean   : 51.44
##  3rd Qu.:72.39    3rd Qu.:21.435      3rd Qu.: 62.78
##  Max.   :96.66    Max.   :46.550      Max.   :100.74
##   sacral_slope   pelvic_radius   degree_spondylolisthesis      class
##  Min.   :13.37   Min.   : 79.0   Min.   :-11.058          Abnormal:199
##  1st Qu.:33.42   1st Qu.:110.9   1st Qu.:  1.496          Normal  :100
##  Median :42.45   Median :118.2   Median : 10.432
```

```
##  Mean   :42.73   Mean   :117.8   Mean   : 23.055
##  3rd Qu.:52.40   3rd Qu.:125.4   3rd Qu.: 39.359
##  Max.   :78.79   Max.   :157.8   Max.   :124.984
```

```
# Summary of Dataset without outliers (Row deletion method)
```

```
summary(mydata1)
```

```
##  pelvic_incidence pelvic_tilt.numeric lumbar_lordosis_angle
##  Min.   :26.15    Min.   :-6.555      Min.   : 14.00
##  1st Qu.:46.43    1st Qu.:10.667      1st Qu.: 37.00
##  Median :58.65    Median :16.328      Median : 49.46
##  Mean   :59.89    Mean   :17.229      Mean   : 51.69
##  3rd Qu.:72.31    3rd Qu.:21.766      3rd Qu.: 62.96
##  Max.   :96.66    Max.   :46.550      Max.   :100.74
##   sacral_slope   pelvic_radius   degree_spondylolisthesis      class
##  Min.   :13.37   Min.   : 79.0   Min.   :-11.058          Abnormal:210
##  1st Qu.:33.35   1st Qu.:110.7   1st Qu.:  1.604          Normal  :100
##  Median :42.39   Median :118.3   Median : 11.616
##  Mean   :42.70   Mean   :117.9   Mean   : 24.112
##  3rd Qu.:52.48   3rd Qu.:125.4   3rd Qu.: 40.234
##  Max.   :79.70   Max.   :157.8   Max.   :124.984
```

By Analyzing Summary we can say that third dataset i.e. dataset in which Outliers are handled by replacing with Median will be the best choice, because Stats of third dataset are close to Original dataset compare to second dataset.

7. Model Fitting

Fitting Model using Logistic Regression & Three datasets

```
# Fit logistic regression with all features & dataset with outliers
glm.fit_mydata <- glm(class ~ . ,data=mydata ,family=binomial)
```

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

```
summary(glm.fit_mydata)
```

```
##
## Call:
## glm(formula = class ~ ., family = binomial, data = mydata)
##
## Deviance Residuals:
```

```
##     Min       1Q   Median       3Q      Max
## -2.2678  -0.3639  -0.0289   0.4081   2.7317
##
## Coefficients:
##                            Estimate Std. Error z value Pr(>|z|)
## (Intercept)              -1.530e+01  3.315e+00  -4.615 3.93e-06 ***
## pelvic_incidence          2.517e+07  4.017e+07   0.627    0.531
## pelvic_tilt.numeric      -2.517e+07  4.017e+07  -0.627    0.531
## lumbar_lordosis_angle     1.794e-02  2.290e-02   0.784    0.433
## sacral_slope             -2.517e+07  4.017e+07  -0.627    0.531
## pelvic_radius             1.077e-01  2.318e-02   4.645 3.39e-06 ***
## degree_spondylolisthesis -1.693e-01  2.335e-02  -7.248 4.23e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 389.86  on 309  degrees of freedom
## Residual deviance: 177.87  on 303  degrees of freedom
## AIC: 191.87
##
## Number of Fisher Scoring iterations: 8

# Fit logistic regression with all features & dataset without outliers (Row
deletion method)
glm.fit_newdata <- glm(class ~ . ,data=newdata ,family=binomial)
summary(glm.fit_newdata)

##
## Call:
## glm(formula = class ~ ., family = binomial, data = newdata)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.26389  -0.38008  -0.03395   0.42245   2.74246
##
## Coefficients:
##                            Estimate Std. Error z value Pr(>|z|)
## (Intercept)              -1.796e+01  3.658e+00  -4.910 9.11e-07 ***
## pelvic_incidence          2.859e+07  4.072e+07   0.702    0.483
## pelvic_tilt.numeric      -2.859e+07  4.072e+07  -0.702    0.483
## lumbar_lordosis_angle     1.802e-02  2.323e-02   0.776    0.438
## sacral_slope             -2.859e+07  4.072e+07  -0.702    0.483
## pelvic_radius             1.279e-01  2.567e-02   4.982 6.29e-07 ***
## degree_spondylolisthesis -1.698e-01  2.353e-02  -7.217 5.31e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
```

```
##     Null deviance: 381.10  on 298  degrees of freedom
## Residual deviance: 172.06  on 292  degrees of freedom
## AIC: 186.06
##
## Number of Fisher Scoring iterations: 8
```

```
# Fit logistic regression with all features & dataset without outliers
(Median replace method)
glm.fit_mydata1 <- glm(class ~ . ,data=mydata1 ,family=binomial)
summary(glm.fit_mydata1)
```

```
##
## Call:
## glm(formula = class ~ ., family = binomial, data = mydata1)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.18907  -0.40135  -0.03885   0.38558   2.75681
##
## Coefficients:
##                            Estimate Std. Error z value Pr(>|z|)
## (Intercept)              -16.30516    3.31276  -4.922 8.57e-07 ***
## pelvic_incidence          -0.12806    0.15346  -0.835    0.404
## pelvic_tilt.numeric        0.04020    0.15250   0.264    0.792
## lumbar_lordosis_angle      0.01486    0.02263   0.657    0.511
## sacral_slope               0.23047    0.15521   1.485    0.138
## pelvic_radius              0.11735    0.02343   5.008 5.51e-07 ***
## degree_spondylolisthesis  -0.16983    0.02347  -7.237 4.58e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 389.86  on 309  degrees of freedom
## Residual deviance: 177.18  on 303  degrees of freedom
## AIC: 191.18
##
## Number of Fisher Scoring iterations: 7
```

8. Overall fraction of correct predictions

1. Overall fraction of correct predictions with all features & dataset with outliers.

```
probability_mydata <- predict(glm.fit_mydata, type = "response")

pred_mydata <- rep("Abnormal", length(probability_mydata))
pred_mydata[probability_mydata > 0.5] <- "Normal"

#Confusion Matrix
table(pred_mydata, mydata$class)

##
## pred_mydata Abnormal Normal
##    Abnormal      186     23
##    Normal         24     77

mean(pred_mydata == mydata$class )

## [1] 0.8483871
```

Overall fraction of correct predictions is 0.8483 (84.83%).

2. Overall fraction of correct predictions all features & dataset without outliers (Row deletion method)

```
probability_newdata <- predict(glm.fit_newdata, type = "response")

pred_newdata <- rep("Abnormal", length(probability_newdata))
pred_newdata[probability_newdata > 0.5] <- "Normal"

#Confusion Matrix
table(pred_newdata, newdata$class)

##
## pred_newdata Abnormal Normal
##     Abnormal      174     22
##     Normal         25     78

mean(pred_newdata == newdata$class)

## [1] 0.8428094
```

Overall fraction of correct predictions is 0.8428 (84.28%).

3. Overall fraction of correct predictions with all features & dataset without outliers (Median replace method)

```
probability_mydata1 <- predict(glm.fit_mydata1, type = "response")

pred_mydata1 <- rep("Abnormal", length(probability_mydata1))
pred_mydata1[probability_mydata1 > 0.5] <- "Normal"

#Confusion Matrix
table(pred_mydata1, mydata1$class)

##
## pred_mydata1 Abnormal Normal
##     Abnormal      187     21
##     Normal         23     79

mean(pred_mydata1 == mydata1$class)

## [1] 0.8580645
```

Overall fraction of correct predictions is 0.8580 (85.80%).

Multiple Logistic Linear Regression Model with all features & dataset without outliers (Median replace method) gives 0.8580 overall fraction of correct predictions.

This implies that replacing Outliers with Median is better method than just deleting observations having Outlier.

This confirms that selection of Third dataset is best choice.

Now We will select Model [glm(class ~ . ,data=mydata1 ,family=binomial)] with third dataset for further use.

9. Trying to increase fraction of correct predictions

Deciding Important Feature to increase fraction of correct predictions.

```
summary(glm.fit_mydata1)

##
## Call:
## glm(formula = class ~ ., family = binomial, data = mydata1)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -2.18907  -0.40135  -0.03885   0.38558   2.75681
##
## Coefficients:
```

```
##                            Estimate Std. Error z value Pr(>|z|)
## (Intercept)              -16.30516    3.31276  -4.922 8.57e-07 ***
## pelvic_incidence          -0.12806    0.15346  -0.835    0.404
## pelvic_tilt.numeric        0.04020    0.15250   0.264    0.792
## lumbar_lordosis_angle      0.01486    0.02263   0.657    0.511
## sacral_slope               0.23047    0.15521   1.485    0.138
## pelvic_radius              0.11735    0.02343   5.008 5.51e-07 ***
## degree_spondylolisthesis  -0.16983    0.02347  -7.237 4.58e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 389.86  on 309  degrees of freedom
## Residual deviance: 177.18  on 303  degrees of freedom
## AIC: 191.18
##
## Number of Fisher Scoring iterations: 7
```

Selecting Important Feature by using Backward Selection Method

Summary shows that pelvic_radius & degree_spondylolisthesis are significant features because their p-value is < 0.05.

So we will use pelvic_radius & degree_spondylolisthesis in this model.

```
# Subset selection 1

glm.fit_mydata1_1 <- glm(class ~ pelvic_radius+degree_spondylolisthesis
,data=mydata1 ,family=binomial)

probability_mydata1_1 <- predict(glm.fit_mydata1_1, type = "response")

pred_mydata1_1 <- rep("Abnormal", length(probability_mydata1_1))
pred_mydata1_1[probability_mydata1_1 > 0.5] <- "Normal"
table(pred_mydata1_1, mydata1$class)

##
## pred_mydata1_1 Abnormal Normal
##       Abnormal      174     27
##       Normal         36     73

mean(pred_mydata1_1 == mydata1$class )

## [1] 0.7967742
```

Overall fraction of correct predictions is 0.7967 (79.67 %).

Now we will try another combination of pelvic_incidence, lumbar_lordosis, pelvic_radius & degree_spondylolisthesis because these 4 features have p-value close to zero.

```
#Subset selection 2

glm.fit_mydata1_2 <- glm(class ~
pelvic_incidence+lumbar_lordosis_angle+pelvic_radius+degree_spondylolisthesis
,data=mydata1 ,family=binomial)

probability_mydata1_2 <- predict(glm.fit_mydata1_2, type = "response")

pred_mydata1_2 <- rep("Abnormal", length(probability_mydata1_2))
pred_mydata1_2[probability_mydata1_2 > 0.5] <- "Normal"
table(pred_mydata1_2, mydata1$class)

##
## pred_mydata1_2 Abnormal Normal
##       Abnormal      178     25
##       Normal         32     75

mean(pred_mydata1_2 == mydata1$class )

## [1] 0.816129
```

Overall fraction of correct predictions is 0.8161 (81.61 %).

Model with all features gives high accuracy of 85.80 %.

Hence we will select Model [glm(class ~ . ,data=mydata1 ,family=binomial)]

But This OFCP is misleading because we trained and tested model on the same set of observations.

In other words, 100 - 85.80 = 14.20 % is the training error rate.
The error rate is often overly optimistic it tends to underestimate the test error rate.

10. Cross Validation

Split data into Training (80%) & Test (20%)

```
set.seed(1)
subset <- sample(nrow(mydata1), nrow(mydata1) * 0.8)
train_mydata1 = mydata1[subset, ]
test_mydata1 = mydata1[-subset, ]
```

Fitting Model with Training Dataset.

```
set.seed(1)
glm.train_mydata1 <- glm(class ~ ., data=train_mydata1,family = binomial)

train_mydata1.probability <- predict(glm.train_mydata1, test_mydata1,
type="response")
train_mydata1_class <- ifelse(train_mydata1.probability > 0.5, 'Normal',
'Abnormal')
table(test_mydata1$class, train_mydata1_class)
```

```
##           train_mydata1_class
##            Abnormal Normal
##   Abnormal       34      6
##   Normal          6     16
```

```
mean(train_mydata1_class == test_mydata1$class)
```

```
## [1] 0.8064516
```

Overall fraction of correct predictions is 0.8064 (80.64 %)

This is the accuracy of our Model i.e 80.64%.

11. Plotting ROC

```
library(ROCR)
```

```
## Warning: package 'ROCR' was built under R version 3.4.2
```

```
## Loading required package: gplots
```

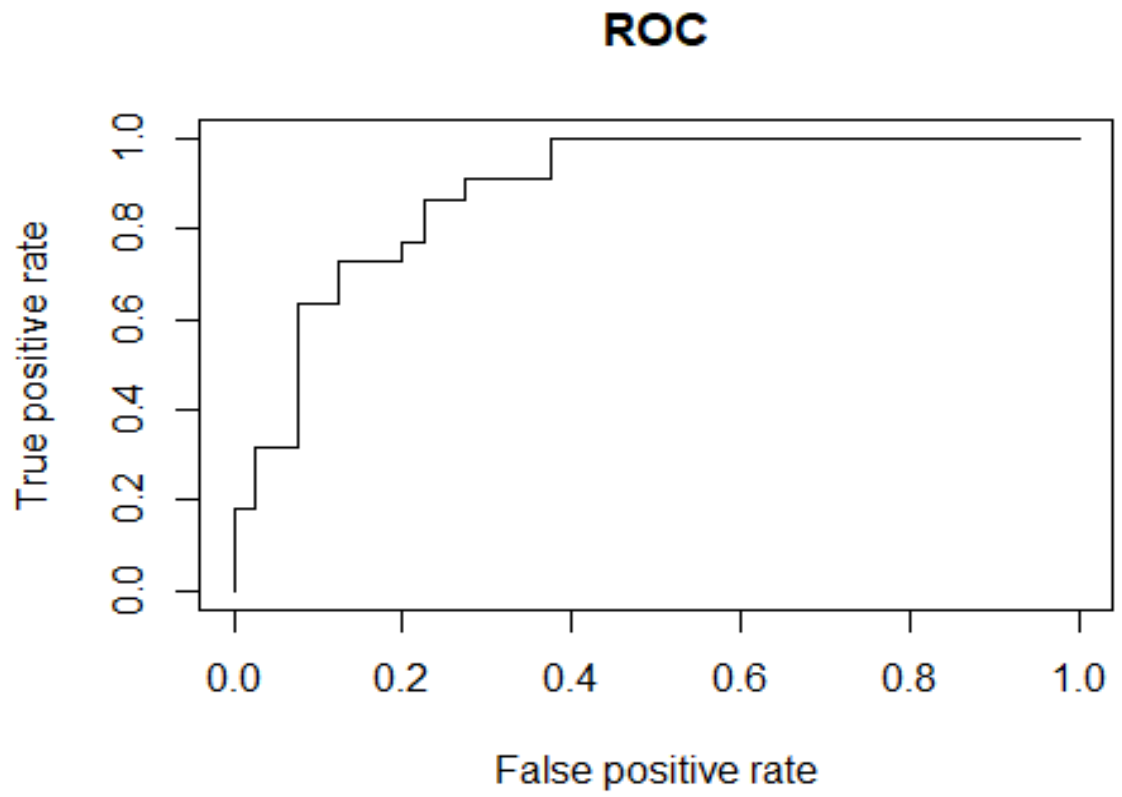
```
## Warning: package 'gplots' was built under R version 3.4.2
```

```
##
## Attaching package: 'gplots'
```

```
## The following object is masked from 'package:stats':
##
##     lowess

library(Metrics)

pr <- prediction(train_mydata1.probability, test_mydata1$class)
perf <- performance(pr,measure = "tpr",x.measure = "fpr")
par(mfrow = c(1,1))
plot(perf, main='ROC')
```



12. Testing Model on random data

Passing predictors:
pelvic_incidence = **40.25020**
pelvic_tilt.numeric = **13.921907**

lumbar_lordosis_angle = **25.12495**
sacral_slope = **26.32829**
pelvic_radius = **130.32787**
degree_spondylolisthesis = **2.230652**

```
testdata = data.frame(pelvic_incidence=40.25020,
pelvic_tilt.numeric=13.921907, lumbar_lordosis_angle=25.12495,
sacral_slope=26.32829, pelvic_radius=130.32787,
degree_spondylolisthesis=2.230652)

glm.fit_mydata1

##
## Call:  glm(formula = class ~ ., family = binomial, data = mydata1)
##
## Coefficients:
##               (Intercept)          pelvic_incidence
##                 -16.30516                   -0.12806
##      pelvic_tilt.numeric     lumbar_lordosis_angle
##                   0.04020                    0.01486
##             sacral_slope             pelvic_radius
##                   0.23047                    0.11735
## degree_spondylolisthesis
##                  -0.16983
##
## Degrees of Freedom: 309 Total (i.e. Null);  303 Residual
## Null Deviance:       389.9
## Residual Deviance: 177.2     AIC: 191.2

result <- predict(glm.fit_mydata1, testdata, type="response")

if (result>=0.0 & result < 0.50) {print('Abnormal')

}else {print('Normal')}

## [1] "Normal"
```

For above predictors our Model classifies patient in Normal Category.

For testing this model pass parameters in *'testdata'* **dataframe and run the model, you will get your patients class i.e.** *Normal or Abnormal.*