**VAIBHAVI PATIL ROLL.NO23**

**LDA QDA**

library(MASS)

library(caret)

library(ROCR)

**## Loading Data & Preprocessing**

df = read.csv("breast-cancer.csv")

dim(df)

df = df[,2:32]

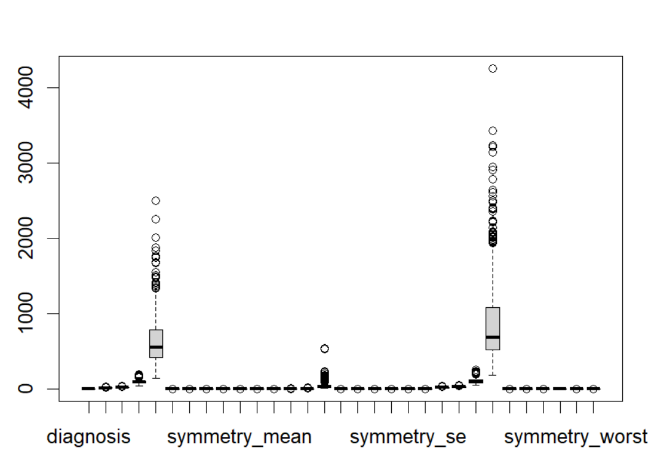
str(df)

colSums(is.na(df))

df$diagnosis = as.factor(df$diagnosis)

str(df)

**## Checking for Outlier**

b=boxplot(df)

**## Making LDA Model**

actual\_lda = lda(diagnosis ~ . , data=df)

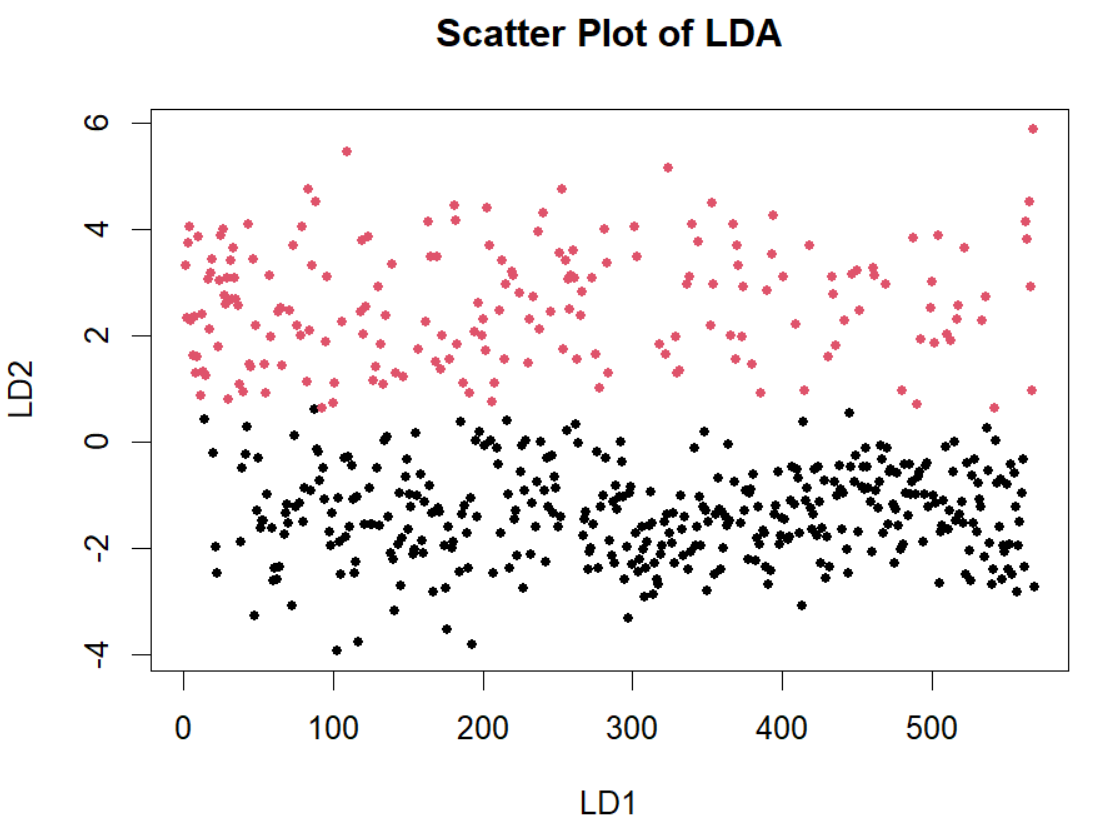
**## Predicting**

predicted\_lda = predict(actual\_lda)

**## Plotting Prediction**

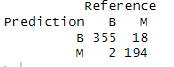
plot(predicted\_lda$x, col = as.numeric(predicted\_lda$class), pch = 19,

xlab = "LD1", ylab = "LD2", main = "Scatter Plot of LDA")



**## Confusion Matrix**

confusionMatrix(predicted\_lda$class,df$diagnosis)$table



**## Accuracy**

mean(predicted\_lda$class==df$diagnosis)



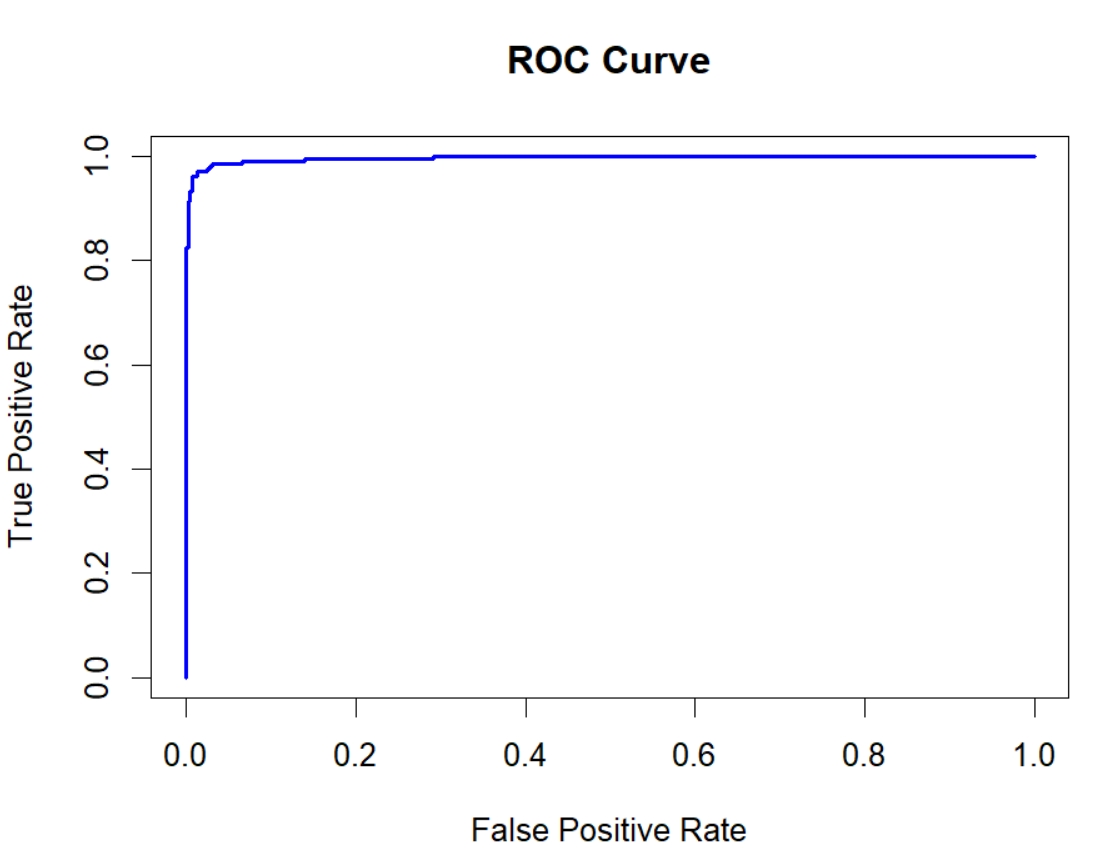
**## Finding Prediction True Positive vs False Positive**

pred\_lda <- prediction(predicted\_lda$posterior[, "M"], df$diagnosis == "M")

perf\_lda <- performance(pred\_lda, "tpr", "fpr")

**## Ploting ROC Curve for LDA**

plot(perf\_lda, col = "blue", lwd = 2, main = "ROC Curve", xlim = c(0, 1), ylim = c(0, 1), xlab = "False Positive Rate", ylab = "True Positive Rate")



**## AUC Score**

auc\_lda <- performance(pred\_lda, "auc")

print(paste("AUC for LDA:", [auc\_lda@y.values[[1]])](mailto:auc_lda@y.values[[1]])))



**## QDA Model**

actual\_qda = qda(diagnosis ~ . , data=df)

**## Prediction using Model**

predicted\_qda = predict(actual\_qda)

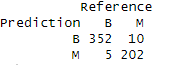
**## Accuracy**

mean(predicted\_qda$class==df$diagnosis)



**## Confution Matrix**

confusionMatrix(predicted\_qda$class,df$diagnosis)$table



**## Finding Prediction True Positive vs False Positive**

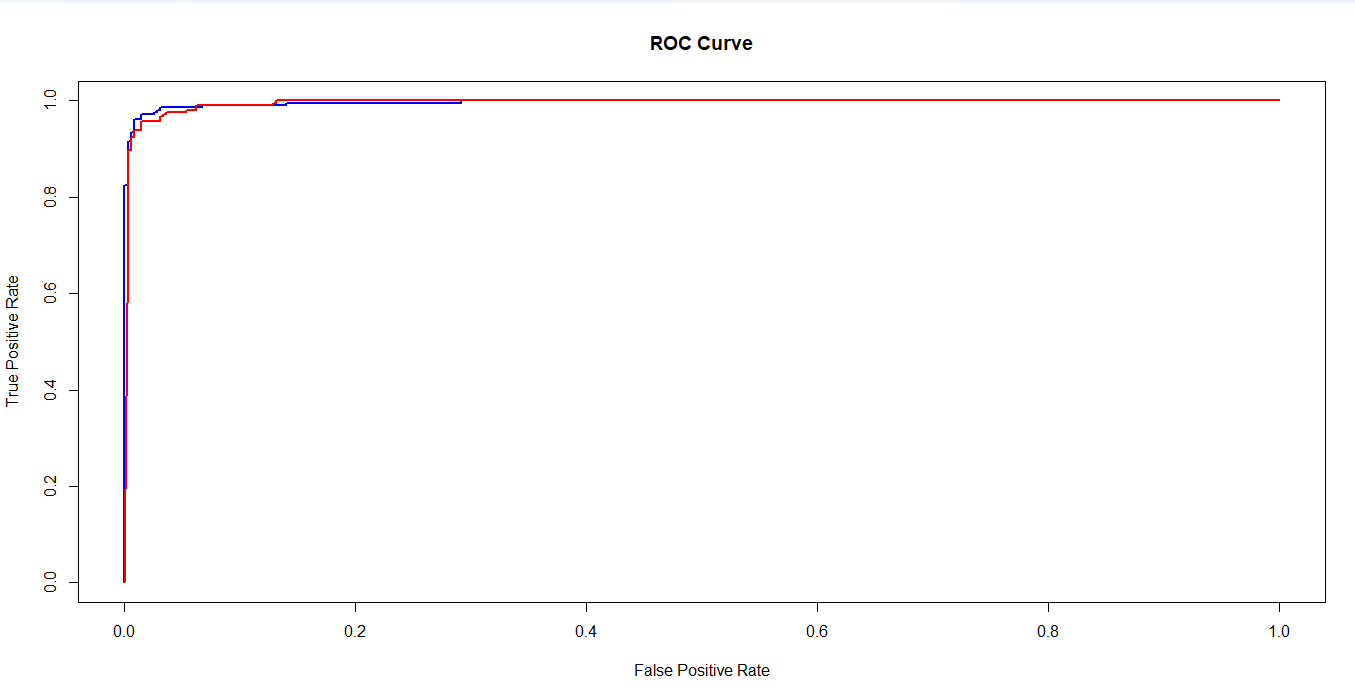
pred\_qda <- prediction(predicted\_qda$posterior[, "M"], df$diagnosis == "M")

**# Create performance object for ROC curve**

perf\_qda <- performance(pred\_qda, "tpr", "fpr")

**# Plot ROC curve**

plot(perf\_qda, col = "red", lwd = 2, add = TRUE)



**# Calculate AUC**

auc\_qda <- performance(pred\_qda, "auc")

print(paste("AUC for QDA:", [auc\_qda@y.values[[1]])](mailto:auc_qda@y.values[[1]])))



**Interpretation:**

1. **Based on Confusion Matrix LDA accuracy was slightly lower than QDA.**
2. **Based on ROC Curve & AUC Score, LDA was 0.996 and QDA was 0.995**
3. **Finally, we can conclude that LDA (20 misclassified) has predicted more misclassified points compared to QDA (15 misclassified)**