## STA 141A Homework5 HaozheGu 999200555

**The codes and results derived by using these codes constitute my own work. I have consulted the following resources regarding this assignment: NONE**

### **Q1**

ii) The least squares regression line is MSE = 0.9624554

iii) The 95% confidence level bootstrap Confidence Interval is:



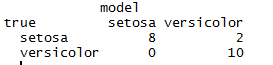
iv) After 10 iterations, the average lengths of two types of Confidence Interval is:



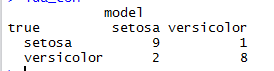
It seems that the average length of Bootstrap confidence interval is narrower than the theoretical confidence interval. But it might because the sample size of the bootstrap is only 400.

Q2

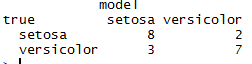
ii) the confusion matrix of logistic regression model on test data is:



iii) The confusion matrix of linear discriminant analysis on test data is:



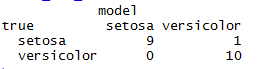
iv) The 3-NN’s Confusion matrix for test data is :



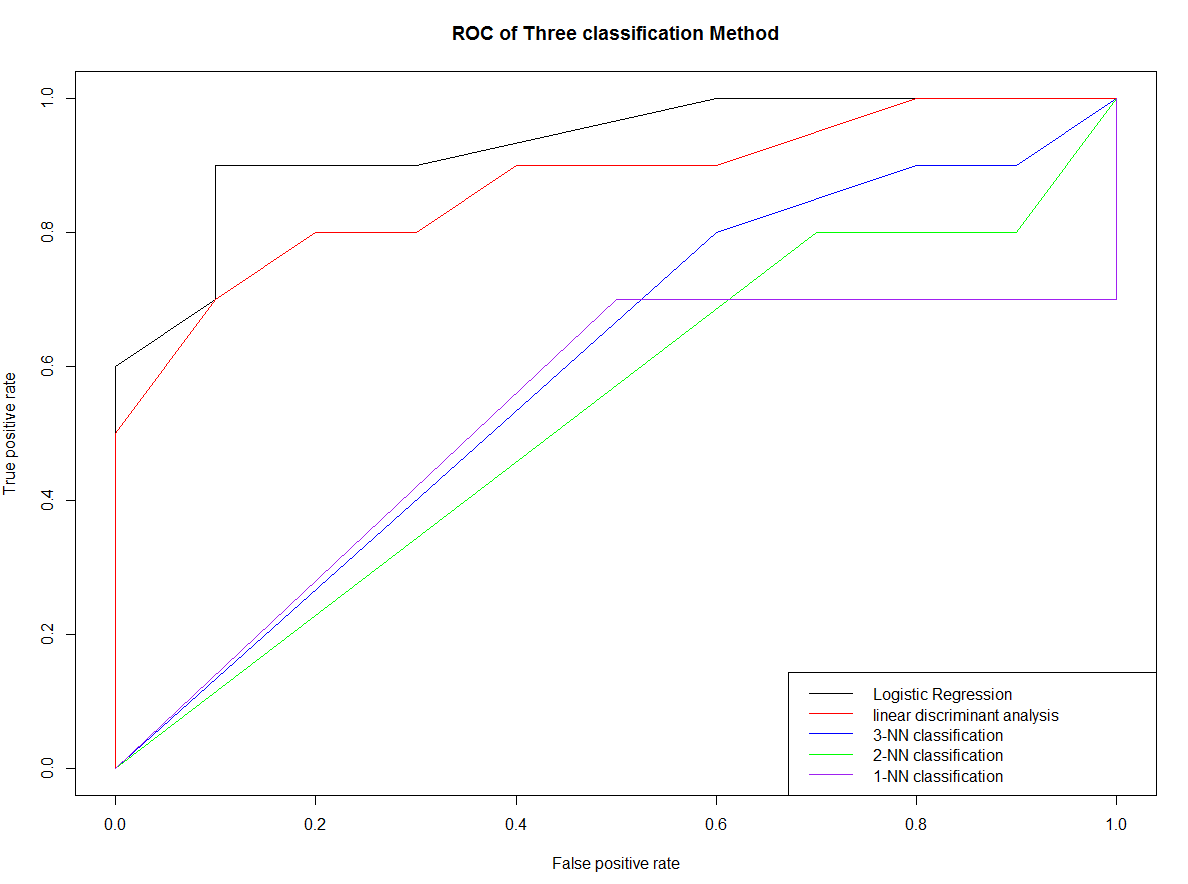
The 2NN’s Confusion matrix for test data is



The 1NN’s Confusion matrix for test data is



v) From the confusion matrix, we can compare the performance of different classification method by the ROC curve, using True positive rate(Sensitivity) against False positive rate.



The larger area under the curve means better performance. From the ROC curve, Logistic Regression seems to be the best classification method, then is Linear Discriminant Analysis, then is KNN.