Credit Card Fraud Detection.

Dataset link: <https://drive.google.com/file/d/1CTAlmlREFRaEN3NoHHitewpqAtWS5cVQ/view>

Source Code:

**library(ranger)**

**library(pROC)**

**library(caret)**

**library(data.table)**

**library(caTools)**

**creditcard\_data <- read.csv("creditcard.csv")**

**getwd()**

**setwd("C:/Users/ashwi/Documents/R/Credit Card")**

**head(creditcard\_data)**

**table(creditcard\_data$Class)**

**summary(creditcard\_data$Amount)**

**names(creditcard\_data)**

**var(creditcard\_data$Amount)**

**creditcard\_data$Amount=scale(creditcard\_data$Amount)**

**NewData=creditcard\_data[,-c(1)]**

**head(NewData)**

**set.seed(123)**

**data\_sample = sample.split(NewData$Class,SplitRatio=0.80)**

**train\_data = subset(NewData,data\_sample==TRUE)**

**test\_data = subset(NewData,data\_sample==FALSE)**

**dim(train\_data)**

**dim(test\_data)**

**Logistic\_Model=glm(Class~.,test\_data,family=binomial())**

**summary(Logistic\_Model)**

**plot(Logistic\_Model)**

**library(pROC)**

**lr.predict <- predict(Logistic\_Model,train\_data, probability = TRUE)**

**auc.gbm = roc(test\_data$Class, lr.predict, plot = TRUE, col = "blue")**

**library(rpart)**

**library(rpart.plot)**

**decisionTree\_model <- rpart(Class ~ . , creditcard\_data, method = 'class')**

**predicted\_val <- predict(decisionTree\_model, creditcard\_data, type = 'class')**

**probability <- predict(decisionTree\_model, creditcard\_data, type = 'prob')**

**rpart.plot(decisionTree\_model)**

**library(neuralnet)**

**ANN\_model =neuralnet (Class~.,train\_data,linear.Graph=FALSE)**

**plot(ANN\_model)**

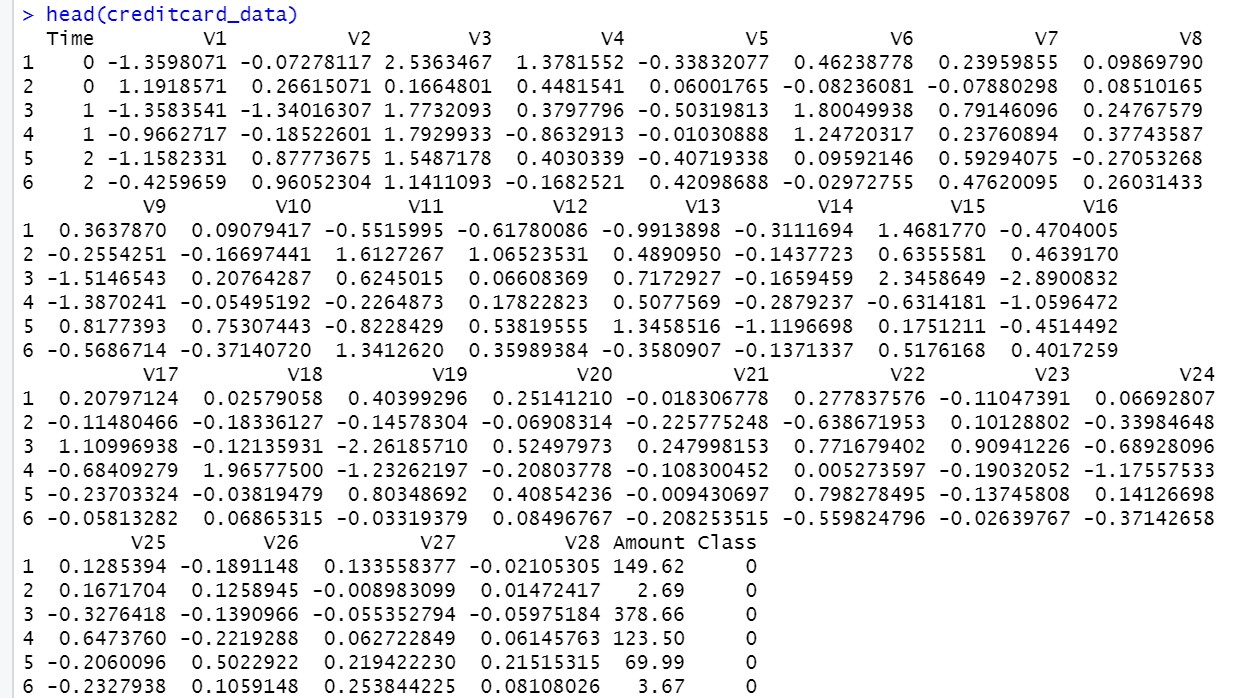
**predANN=compute(ANN\_model,test\_data)**

**resultANN=predANN$net.result**

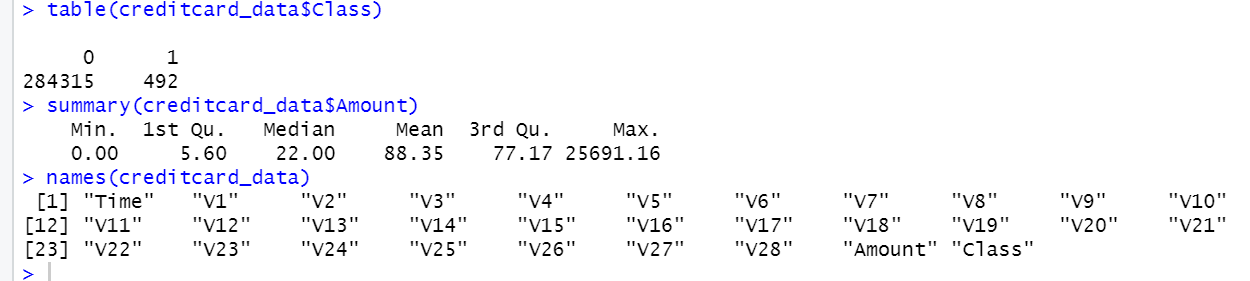
**resultANN=ifelse(resultANN>0.5,1,0)**

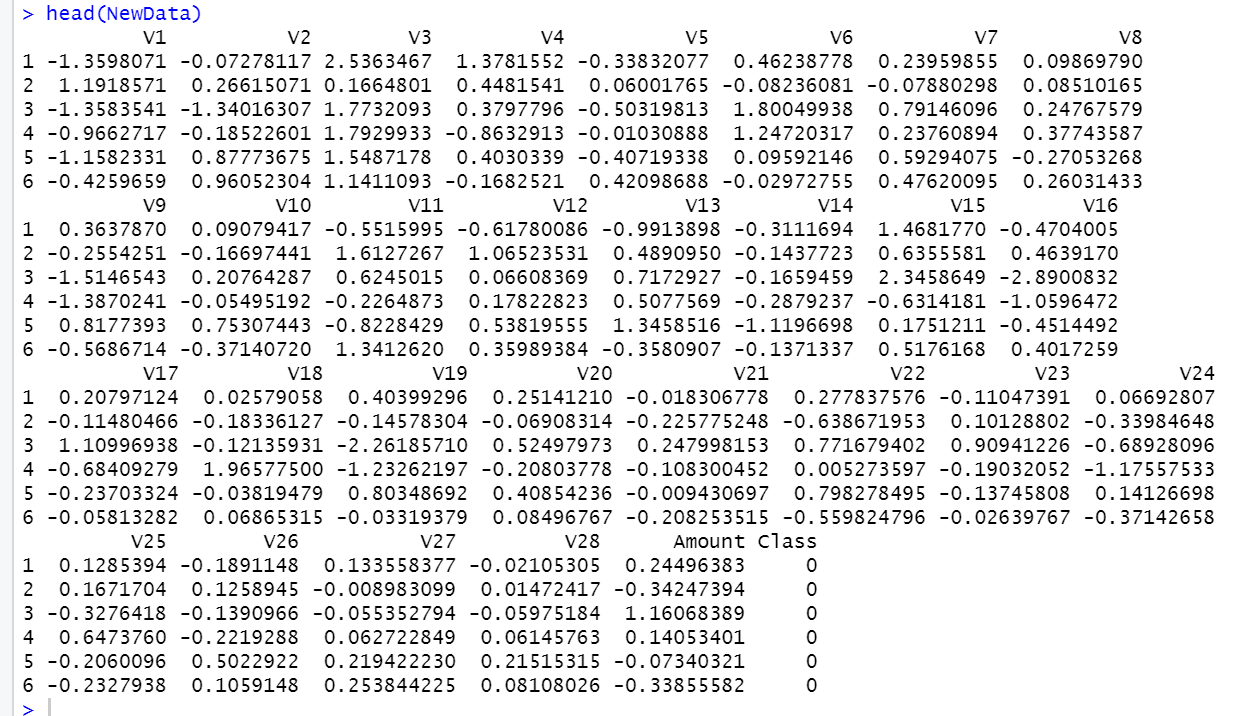
**print(resultANN)**

Output 1: Credit Card info

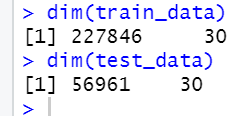


Output 2: Credit Card Data Exploration

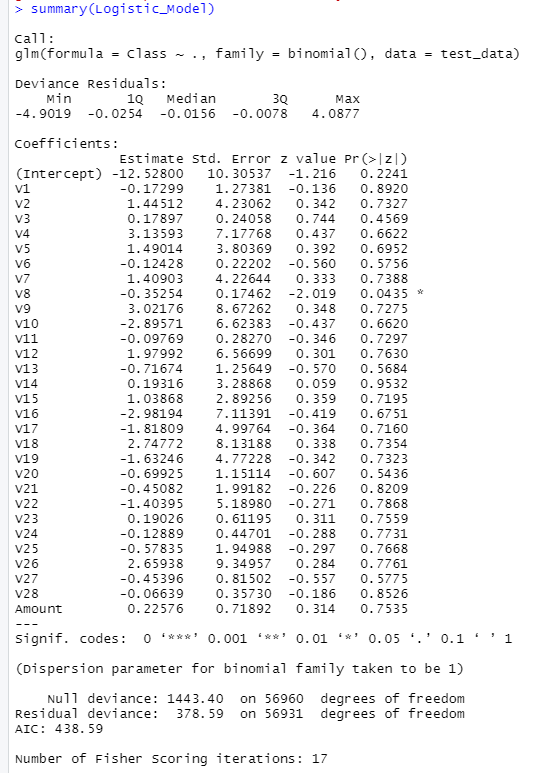


Output 3: Credit Card Data Scaling

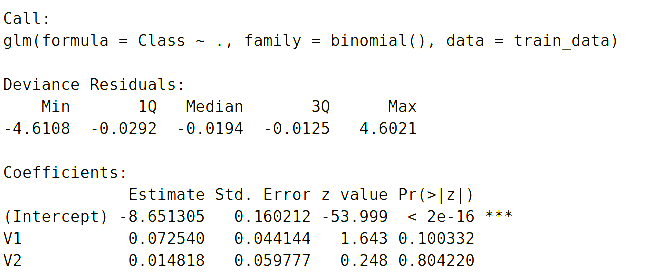
Output 4: Dimensions of Train and Test Data



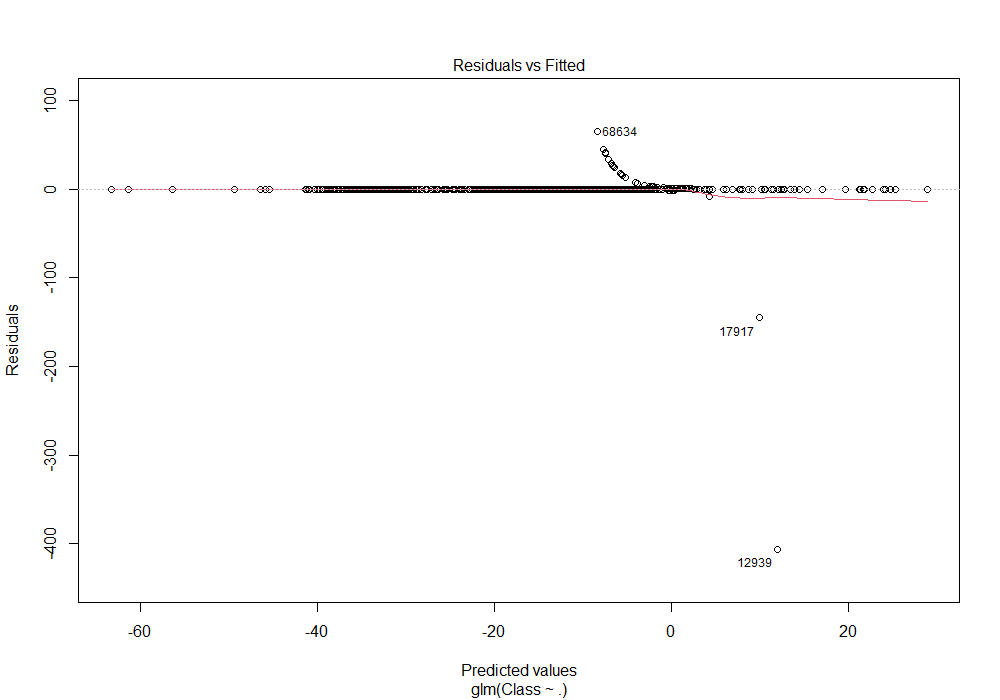
Output 5: Summary of Logistic Model



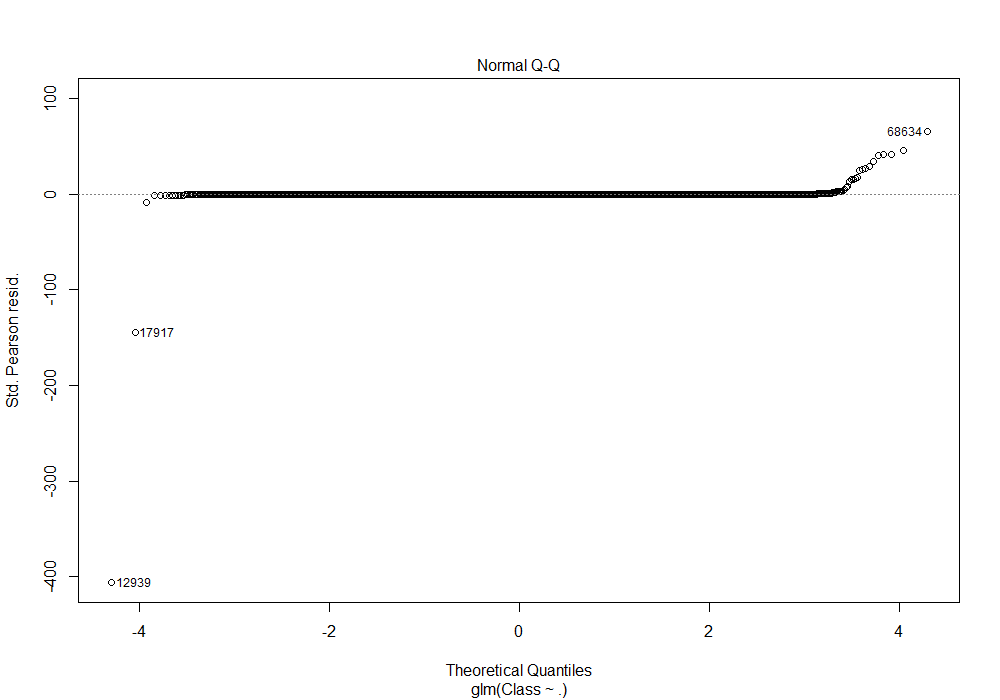
Output 6: Logistic Model ROC values

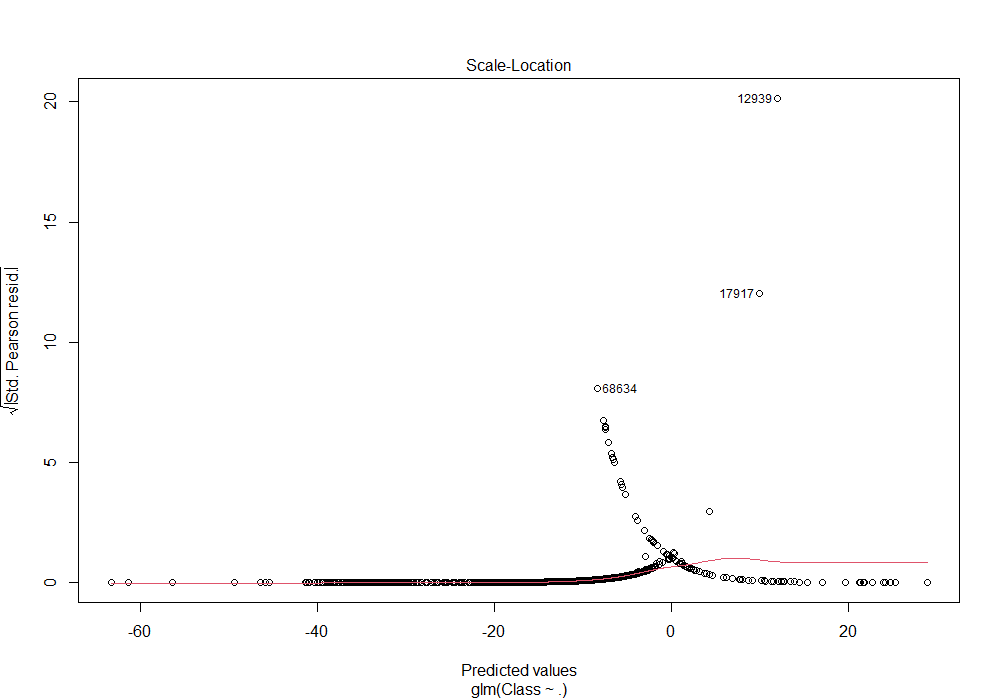


Graph 1: Residuals vs. Fitted Graph using Logistic Regression.

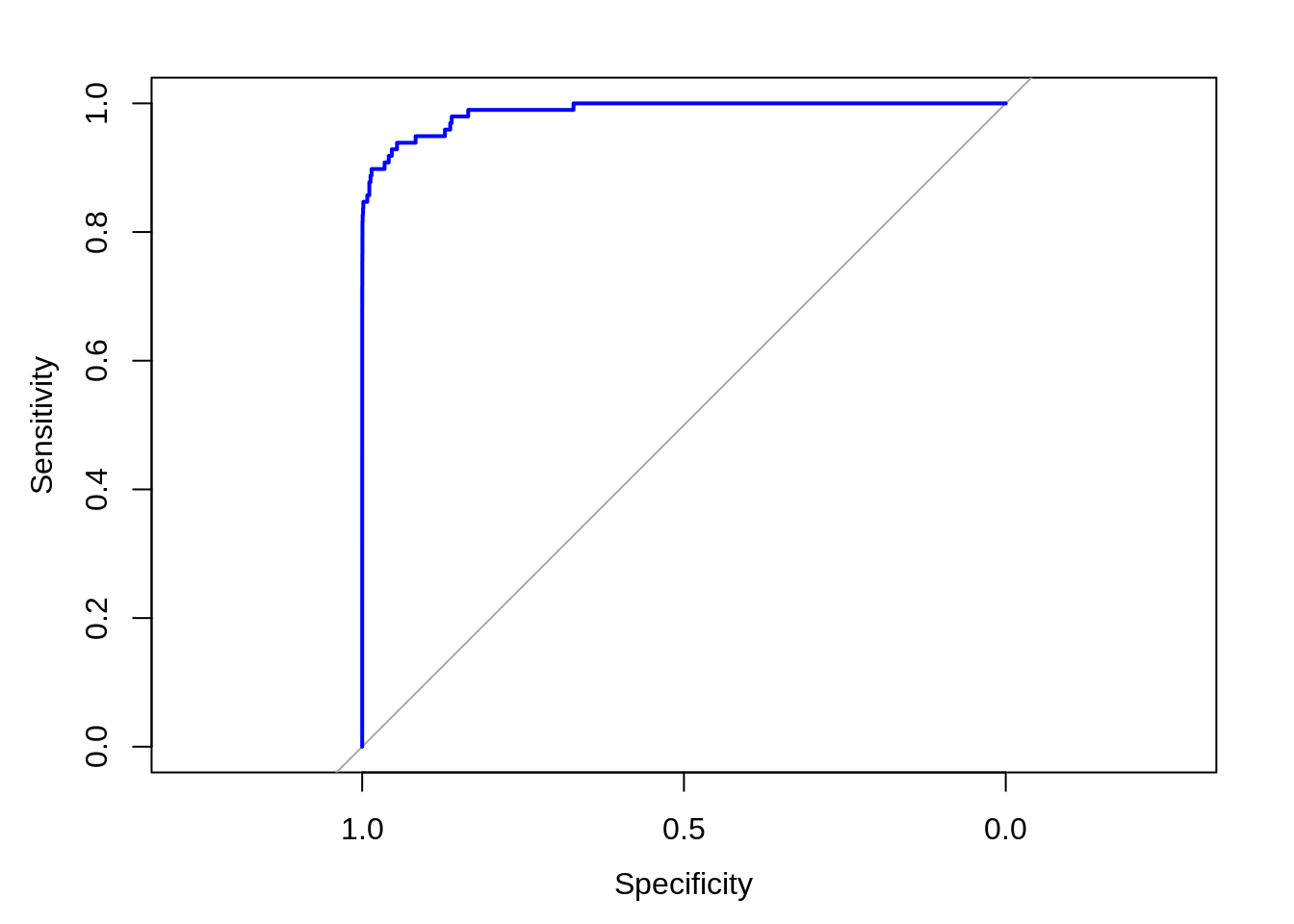


Graph 2: Normal Q-Q Graph along with theoretic quantities using Logistic Regression.

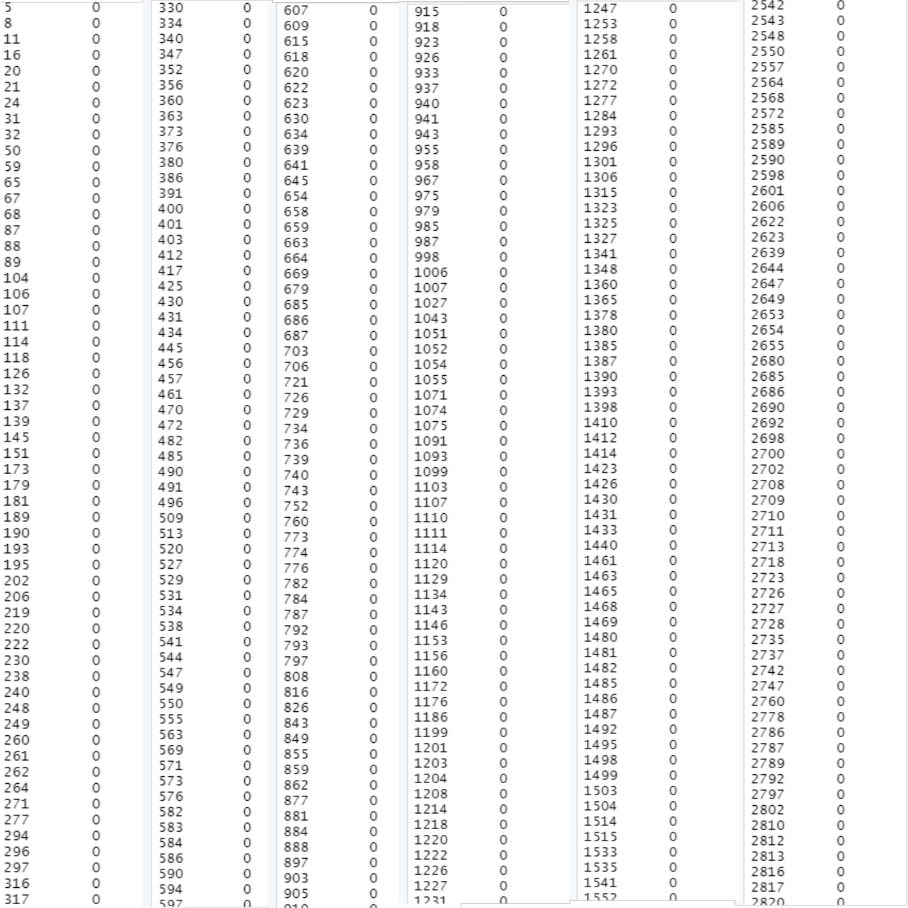
Graph 3: Scale-Location Graph using Logistic Regression



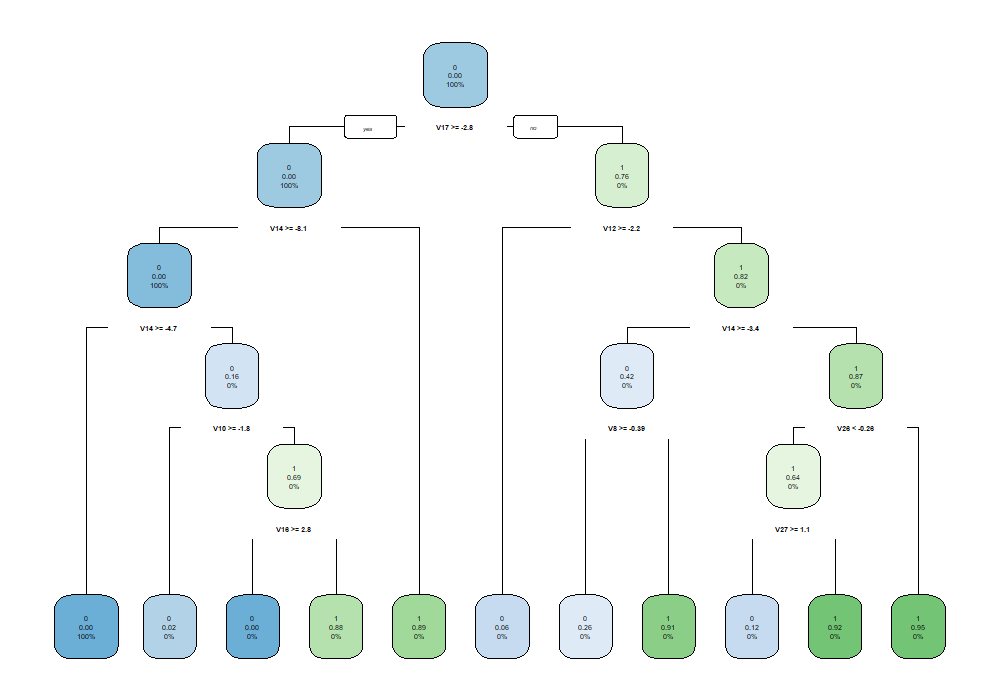
Graph 4: ROC for LR

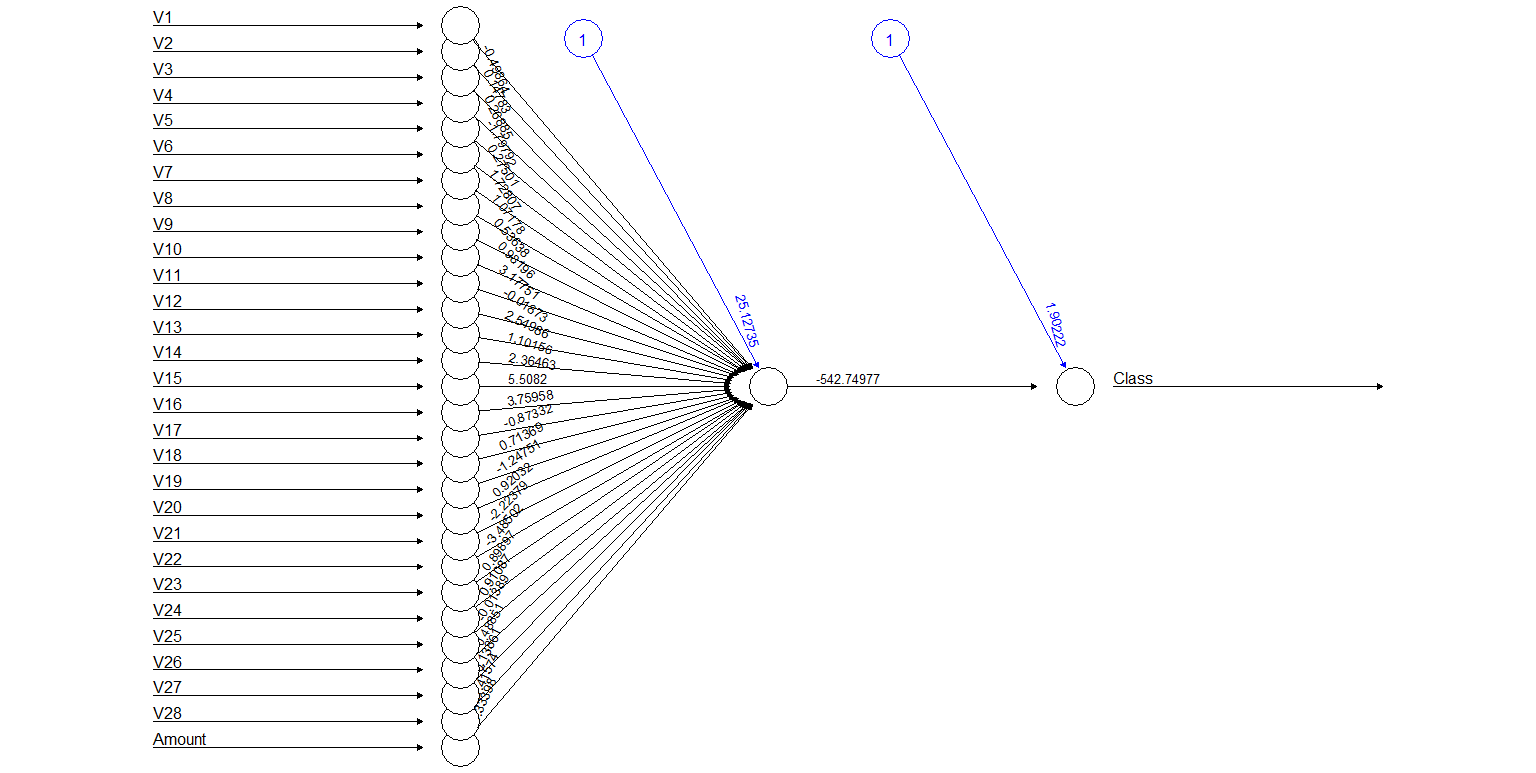


Graph 4: ANN Values



Graph 5: Fraud Detection using Decision Tree.



Graph 6: Fraud Detection using ANN Neurons.

Conclusion:  
We have made a Credit Card Fraud Detection Model using a variety of ML Algorithms along with the plots of performance curves for the respected models.