Classify whether application accepted or not using Logistic regression

**Ans :**

**R Code :**

## Logistic Regression

########## Creditcard Data Set #########

creditcard <- read.csv('D:\\Data Science\\Excelr\\Assignments\\Assignment\\Logistic Regression\\creditcard.csv')

creditcard <- creditcard[,-1] # Removing the first column which is is an Index

# GLM function use sigmoid curve to produce desirable results

# The output of sigmoid function lies in between 0-1

model <- glm(card~.,data=creditcard,family = "binomial")

summary(model) # Confusion matrix table

prob <- predict(model,creditcard,type="response")

prob

# Confusion matrix and considering the threshold value as 0.5

confusion<-table(prob>0.5,creditcard$card)

confusion

# Model Accuracy

Accuracy<-sum(diag(confusion)/sum(confusion))

Accuracy

##ROC

library(ROCR)

rocrpred<-prediction(prob,creditcard$card)

rocrperf<-performance(rocrpred,'tpr','fpr')

plot(rocrperf,colorize=T,text.adj=c(-0.2,1.7))

**Results :**

> summary(model) # Confusion matrix table

Call:

glm(formula = card ~ ., family = "binomial", data = creditcard)

Deviance Residuals:

Min 1Q Median 3Q Max

-8.49 0.00 0.00 0.00 8.49

Coefficients:

Estimate Std. Error z value Pr(>|z|)

(Intercept) 6.610e+14 8.418e+06 78529581 <2e-16 \*\*\*

reports -5.411e+14 1.436e+06 -376913075 <2e-16 \*\*\*

age 1.015e+12 2.208e+05 4595721 <2e-16 \*\*\*

income -4.840e+13 1.470e+06 -32917418 <2e-16 \*\*\*

share 1.732e+16 4.353e+07 397938311 <2e-16 \*\*\*

expenditure -7.184e+11 1.568e+04 -45826822 <2e-16 \*\*\*

owneryes 1.262e+11 4.361e+06 28936 <2e-16 \*\*\*

selfempyes 2.868e+14 7.375e+06 38894429 <2e-16 \*\*\*

dependents 2.155e+13 1.619e+06 13313846 <2e-16 \*\*\*

months -1.194e+12 3.140e+04 -38025935 <2e-16 \*\*\*

majorcards 4.265e+13 4.858e+06 8780075 <2e-16 \*\*\*

active 8.221e+12 3.175e+05 25897563 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 1404.6 on 1318 degrees of freedom

Residual deviance: 13336.2 on 1307 degrees of freedom

AIC: 13360

Number of Fisher Scoring iterations: 25

> # Confusion matrix and considering the threshold value as 0.5

> confusion<-table(prob>0.5,creditcard$card)

> confusion

no yes

FALSE 116 5

TRUE 180 1018

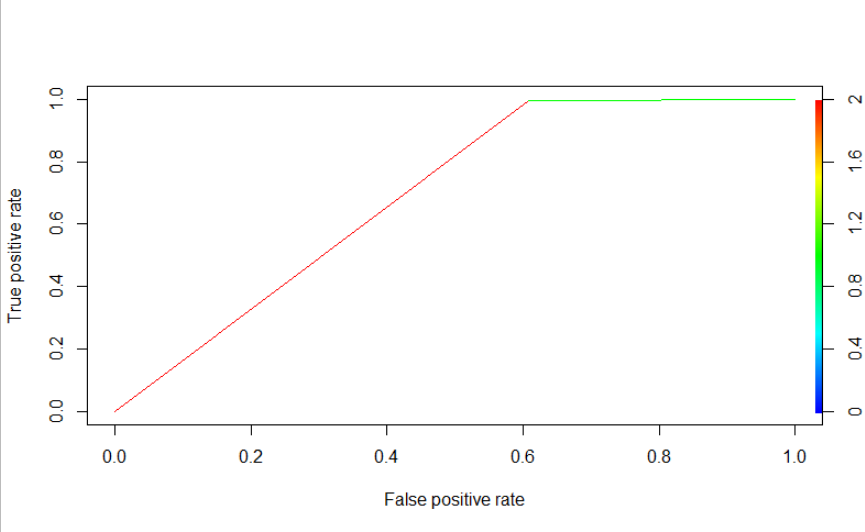
> # Model Accuracy

> Accuracy<-sum(diag(confusion)/sum(confusion))

> Accuracy

[1] 0.8597422

**Plots :**



**Inference :**

Getting good Accuracy in Model.