R coding :-

```r
##************* Decision Tree *****************##

rm(list = ls())

setwd("C:/Users/DELL/Desktop/Project 1")

getwd()

train=  read.csv("train.csv", header = T )

colnames(train)

head(train,3)

str(data)

class(train$ID_code)

train$ID_code = as.numeric(train$ID_code)

class(train$ID_code)
```
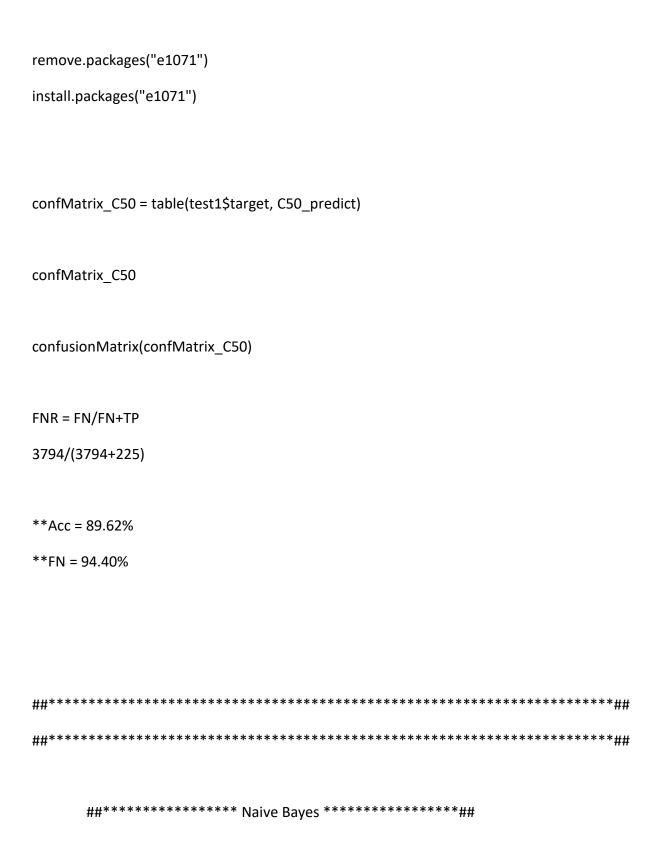
```r
library(DataCombine)


install.packages("numDeriv")


install.packages("caret")
library(caret)


set.seed(1234)
train.index = createDataPartition(train$target, p = .80, list = FALSE)


train1 = train[train.index,]
test1 = train[-train.index,]


colnames(train1)
library(C50)


C50_model = C5.0(target ~.,data = train1, trails = 100, rules = TRUE)


summary(C50_model)


C50_predict = predict(C50_model, test1[,-2], type = 'class')
C50_predict
```

```
remove.packages("e1071")

install.packages("e1071")



confMatrix_C50 = table(test1$target, C50_predict)


confMatrix_C50


confusionMatrix(confMatrix_C50)


FNR = FN/FN+TP

3794/(3794+225)


**Acc = 89.62%

**FN = 94.40%
```

##**##

##**##

##**************** Naive Bayes ****************##

```r
library(e1071)

NB_model = naiveBayes(target ~.,data = train1)

NB_predict = predict(NB_model, test1[,-2], type = 'class')
C50_predict

confMatrix_NB = table(test1$target, NB_predict)

confMatrix_NB

confusionMatrix(confMatrix_NB)
```

acc => 92.2%

FNR = FN/FN+TP

2527/(2527+1492)

=> 62.87%

```
##*****************************************************************************##
##*****************************************************************************##


        ##****************** KNN ******************##
```

```
library(class)

KNN_prediction = knn(train1[, 1:202] , test1[, 1:202], train1$target, k = 1)

conf_matrix = table(KNN_prediction, test1$target )

##*******************************************************************##
##*******************************************************************##


     ##*************** Logistic Regression ***************##




logit_model = glm(target ~.,data = train1, family = "binomial")

summary(logit_model)

logit_prediction = predict(logit_model, newdata = test1, type = "response")

logit_prediction = ifelse(logit_prediction > 0.5, 1, 0)
```

```r
confMatrix = table(test1$target, logit_prediction)

confMatrix


35476+1092


36568/39999

2927/(2927+1092)


##*****************************##


## Now we will predict the target variable in Test data


test_main =  read.csv("test.csv", header = T )


test_main$ID_code = as.numeric(test_main$ID_code)


test_main$Target = with(test_main, ID_code)


logit_prediction_main = predict(logit_model, newdata = test_main, type = "response")
logit_prediction_main = ifelse(logit_prediction_main > 0.5, 1, 0)


logit_prediction_main


table(logit_prediction_main)
```

logit_prediction_main

   0   1

193904 6096