**R MARKDOWN**

**ANGELA MARIA BLANCO MORENO**

---

title: "creditcard\_ambm"

author: "Angela"

date: "5/9/2021"

output:

word\_document: default

html\_document: default

pdf\_document: default

---

```{r setup, include=FALSE}

knitr::opts\_chunk$set(echo = TRUE)

```

## business problem/goal

In this R project, I will present how to perform detention of credit card frauds, through the verious algortihms like desicion tree model, and logistics regrresion model.

For carrying out the credit card fraud detection, I will make use of the Card Transactions dataset that contains a mix of fraud as well as non-fraudulent transactions.

The file has 56 observation of 8 variables, it was imported from a CSV file

## getting the data

```{r creditcard}

summary(creditcard)

```

Table

Description automatically generated

##see first 6 rows of my data

```{r}

head(creditcard, 6)

```

Table

Description automatically generated

##view and rename a column

##now rename V1

```{r creditcard}

colnames(creditcard)[colnames(creditcard)=='V1']<-'1V'

colnames(creditcard)

```

Table

Description automatically generated

##characteristics of my data

The creditcard data has `r ncol(creditcard)` rows and `r nrow(creditcard)`rows. Bellow is a summary of the types of variables in the dataset:

```{r creditcard}

str(creditcard)

```

Text

Description automatically generated

##the characteristics tables

```{r}

text\_tbl<-data.frame(names = c("Amount", "Class"), Description = c("this is the amount expend per transaccion", "this is the class identified per transaccion"))

#to print the table in my output:

print(text\_tbl)

```

Application

Description automatically generated

## ## Data exploration

```{r}

dim(creditcard)

tail(creditcard)

table(creditcard$Class)

summary(creditcard$Amount)

names(creditcard)

var(creditcard$Amount)

sd(creditcard$Amount)

```

Table

Description automatically generated

## Data Manipulation

```{r}

head(creditcard)

creditcard$Amount=scale(creditcard$Amount)

newcreditcard=creditcard[,-c(1)]

head(newcreditcard)

```

Table

Description automatically generated

##Data Modeling

```{r}

library(caTools)

set.seed(123)

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

train\_data= subset(newcreditcard, data\_sample==TRUE)

test\_data= subset(newcreditcard, data\_sample==FALSE)

dim(train\_data)

dim(test\_data)

```



## decision tree model

```{r}

library(rpart)

install.packages("rpart.plot")

library(rpart.plot)

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

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

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

rpart.plot(decisionTree\_model)

```

## Logistics regression model

```{r}

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

summary(Logistic\_Model)

plot(Logistic\_Model)

```

A picture containing diagram

Description automatically generated