

Trabajo3

Juan Alberto Martinez Lopez / Alberto Armijo Ruiz

27 de mayo de 2017

```
#Librerías utilizadas.  
library("caret")
```

```
## Loading required package: lattice
```

```
## Loading required package: ggplot2
```

```
library("leaps")  
library("e1071")
```

1. default of credit card clients Data Set (Clasificación)

La base de datos se centra en el caso de los pagos por defecto de los clientes en Taiwán y compara la precisión predictiva de la probabilidad de incumplimiento entre seis métodos de minería de datos. Desde la perspectiva de la gestión de riesgos, el resultado de la precisión predictiva de la probabilidad estimada de incumplimiento será más valioso que el resultado binario de la clasificación - clientes creíbles o no creíbles. Con la probabilidad real de default como variable de respuesta (Y), y la probabilidad predictiva de default como variable independiente (X), el resultado de regresión lineal simple ($Y = A + BX$) muestra que el modelo de predicción producido por la red neuronal artificial tiene el mayor coeficiente de determinación. Su intercepción de regresión (A) es cercana a cero, y el coeficiente de regresión (B) a uno. Por lo tanto, entre las seis técnicas de minería de datos, la red neuronal artificial es la única que puede estimar con precisión la probabilidad real de incumplimiento. Dichas características son:

Limit_bat: Cantidad de credito bancario dado(en dolares), incluye el credito individual y de su familia. Sex: genero (0 = mujer, 1 = hombre). Education: Educación recibida en 4 variables: Others: Otros estudios. University: Estudios universitarios. High school: Estudios preparatoria. school: Estudios básicos. Marriage: Estado marital en 3 variables: Married: Casado. Single: soltero. Others: otros. Age: Edad. Pay_1-6: Historia del pago anterior. Seguimos los últimos registros de pagos mensuales (de abril a septiembre de 2005) de la siguiente manera: X6 = el estado de pago en septiembre de 2005; X7 = estado de reembolso en agosto de 2005; . . . X11 = el estado de pago en abril de 2005. La escala de medición para el estado de pago es: -1 = pagar debidamente; 1 = retraso de pago de un mes; 2 = retardo de pago por dos meses; . . . Unesdoc.unesco.org unesdoc.unesco.org 8 = retardo de pago por ocho meses; 9 = retraso de pago por nueve meses y más.

Bill_Amt1-6: Monto del estado de cuenta (dólar NT). X12 = monto del estado de cuenta en septiembre de 2005; X13 = monto del estado de cuenta en agosto de 2005; . . . X17 = monto del estado de cuenta en abril de 2005. Pay_Amt1-6: Monto del pago anterior (dólar NT). X18 = cantidad pagada en septiembre de 2005; X19 = cantidad pagada en agosto de 2005; . . X23 = cantidad pagada en abril de 2005. default payment next month: Variable que aprendemos, basamos si lapersona pagará o no el siguiente mes.

```
credit_card = read.csv("default_of_credict_card_clients.csv"  
                      , sep="," , header = TRUE, row.names =1)  
attach(credit_card)  
summary(credit_card)
```

##	LIMIT_BAL	SEX	EDUCATION	MARRIAGE
##	Min. : 10000	Min. :1.000	Min. :0.000	Min. :0.000
##	1st Qu.: 50000	1st Qu.:1.000	1st Qu.:1.000	1st Qu.:1.000
##	Median : 140000	Median :2.000	Median :2.000	Median :2.000

```

## Mean : 167484 Mean :1.604 Mean :1.853 Mean :1.552
## 3rd Qu.: 240000 3rd Qu.:2.000 3rd Qu.:2.000 3rd Qu.:2.000
## Max. :1000000 Max. :2.000 Max. :6.000 Max. :3.000
## AGE PAY_0 PAY_2 PAY_3
## Min. :21.00 Min. : -2.0000 Min. : -2.0000 Min. : -2.0000
## 1st Qu.:28.00 1st Qu.: -1.0000 1st Qu.: -1.0000 1st Qu.: -1.0000
## Median :34.00 Median : 0.0000 Median : 0.0000 Median : 0.0000
## Mean :35.49 Mean : -0.0167 Mean : -0.1338 Mean : -0.1662
## 3rd Qu.:41.00 3rd Qu.: 0.0000 3rd Qu.: 0.0000 3rd Qu.: 0.0000
## Max. :79.00 Max. : 8.0000 Max. : 8.0000 Max. : 8.0000
## PAY_4 PAY_5 PAY_6 BILL_AMT1
## Min. : -2.0000 Min. : -2.0000 Min. : -2.0000 Min. : -165580
## 1st Qu.: -1.0000 1st Qu.: -1.0000 1st Qu.: -1.0000 1st Qu.: 3559
## Median : 0.0000 Median : 0.0000 Median : 0.0000 Median : 22382
## Mean : -0.2207 Mean : -0.2662 Mean : -0.2911 Mean : 51223
## 3rd Qu.: 0.0000 3rd Qu.: 0.0000 3rd Qu.: 0.0000 3rd Qu.: 67091
## Max. : 8.0000 Max. : 8.0000 Max. : 8.0000 Max. : 964511
## BILL_AMT2 BILL_AMT3 BILL_AMT4 BILL_AMT5
## Min. : -69777 Min. : -157264 Min. : -170000 Min. : -81334
## 1st Qu.: 2985 1st Qu.: 2666 1st Qu.: 2327 1st Qu.: 1763
## Median : 21200 Median : 20089 Median : 19052 Median : 18105
## Mean : 49179 Mean : 47013 Mean : 43263 Mean : 40311
## 3rd Qu.: 64006 3rd Qu.: 60165 3rd Qu.: 54506 3rd Qu.: 50191
## Max. :983931 Max. :1664089 Max. : 891586 Max. :927171
## BILL_AMT6 PAY_AMT1 PAY_AMT2 PAY_AMT3
## Min. : -339603 Min. : 0 Min. : 0 Min. : 0
## 1st Qu.: 1256 1st Qu.: 1000 1st Qu.: 833 1st Qu.: 390
## Median : 17071 Median : 2100 Median : 2009 Median : 1800
## Mean : 38872 Mean : 5664 Mean : 5921 Mean : 5226
## 3rd Qu.: 49198 3rd Qu.: 5006 3rd Qu.: 5000 3rd Qu.: 4505
## Max. : 961664 Max. :873552 Max. :1684259 Max. :896040
## PAY_AMT4 PAY_AMT5 PAY_AMT6
## Min. : 0 Min. : 0.0 Min. : 0.0
## 1st Qu.: 296 1st Qu.: 252.5 1st Qu.: 117.8
## Median : 1500 Median : 1500.0 Median : 1500.0
## Mean : 4826 Mean : 4799.4 Mean : 5215.5
## 3rd Qu.: 4013 3rd Qu.: 4031.5 3rd Qu.: 4000.0
## Max. :621000 Max. :426529.0 Max. :528666.0
## default.payment.next.month
## Min. :0.0000
## 1st Qu.:0.0000
## Median :0.0000
## Mean :0.2212
## 3rd Qu.:0.0000
## Max. :1.0000

```

```

set.seed(1)
train = sample (nrow(credit_card), round(nrow(credit_card)*0.7))
credit_card.train = credit_card[train,]
credit_card.test = credit_card[-train,]

```

2. Preprocesado de los datos.

Lo primero que queremos hacer es comprobar si hay datos perdidos, y si es así; reemplazaremos el valor perdido.

```
anyNA(credit_card.train)
```

```
## [1] FALSE
```

Como no tenemos ningún dato perdido, no tendremos que reemplazar los valores. Si hubieramos tenido valores tenidos, podríamos haber utilizado la función *knnImputation* para reemplazar los valores perdidos por los k vecinos más cercanos (normalmente k=3). También podríamos utilizar la media como sustituto del valor perdido.

Lo siguiente que vamos a hacer es modificar aquellas columnas que separan los datos en variables “clases” como por ejemplo la columna *EDUCATION*, que indica que tipo de estudios tiene cada persona. Por cada tipo en los que los separe, crearemos una nueva columna que indique con 0s y 1s la pertenencia a ese tipo. También tenemos que realizar este proceso con la columna *MARRIAGE*

```
# Modificamos la columna 2, llamada sex, para dividir los datos en 0=mujer, 1=hombre.
```

```
credit_card.train$SEX = ifelse(credit_card.train$SEX == 2, 0, 1)
```

```
summary(credit_card.train)
```

```
##      LIMIT_BAL      SEX      EDUCATION      MARRIAGE
##  Min.   : 10000   Min.   :0.0000   Min.   :0.000   Min.   :0.000
## 1st Qu.: 50000   1st Qu.:0.0000   1st Qu.:1.000   1st Qu.:1.000
## Median :140000   Median :0.0000   Median :2.000   Median :2.000
## Mean   :167948   Mean   :0.3974   Mean   :1.853   Mean   :1.549
## 3rd Qu.:240000   3rd Qu.:1.0000   3rd Qu.:2.000   3rd Qu.:2.000
## Max.   :1000000   Max.   :1.0000   Max.   :6.000   Max.   :3.000
##      AGE      PAY_0      PAY_2      PAY_3
##  Min.   :21.0   Min.   :-2.00000   Min.   :-2.0000   Min.   :-2.0000
## 1st Qu.:28.0   1st Qu.:-1.00000   1st Qu.:-1.0000   1st Qu.:-1.0000
## Median :34.0   Median : 0.00000   Median : 0.0000   Median : 0.0000
## Mean   :35.5   Mean   :-0.01524   Mean   :-0.1321   Mean   :-0.1697
## 3rd Qu.:41.0   3rd Qu.: 0.00000   3rd Qu.: 0.0000   3rd Qu.: 0.0000
## Max.   :79.0   Max.   : 8.00000   Max.   : 8.0000   Max.   : 8.0000
##      PAY_4      PAY_5      PAY_6      BILL_AMT1
##  Min.   :-2.000   Min.   :-2.0000   Min.   :-2.000   Min.   :-154973
## 1st Qu.:-1.000   1st Qu.:-1.0000   1st Qu.:-1.000   1st Qu.: 3508
## Median : 0.000   Median : 0.0000   Median : 0.000   Median : 22143
## Mean   :-0.221   Mean   :-0.2678   Mean   :-0.289   Mean   : 51238
## 3rd Qu.: 0.000   3rd Qu.: 0.0000   3rd Qu.: 0.000   3rd Qu.: 66175
## Max.   : 8.000   Max.   : 8.0000   Max.   : 8.000   Max.   : 964511
##      BILL_AMT2      BILL_AMT3      BILL_AMT4      BILL_AMT5
##  Min.   :-69777   Min.   : -46127   Min.   :-170000   Min.   :-81334
## 1st Qu.: 2956   1st Qu.: 2552   1st Qu.: 2276   1st Qu.: 1730
## Median :20895   Median : 20012   Median : 19044   Median : 18093
## Mean   :49092   Mean   : 47104   Mean   : 43452   Mean   : 40395
## 3rd Qu.:62795   3rd Qu.:59674   3rd Qu.:53536   3rd Qu.:49719
## Max.   :983931   Max.   :1664089   Max.   :891586   Max.   :927171
##      BILL_AMT6      PAY_AMT1      PAY_AMT2      PAY_AMT3
##  Min.   :-339603   Min.   : 0   Min.   : 0.0   Min.   : 0
## 1st Qu.: 1261   1st Qu.: 1000   1st Qu.: 809.8   1st Qu.: 390
## Median :17036   Median : 2100   Median :2000.0   Median :1794
## Mean   :38948   Mean   : 5608   Mean   :5988.9   Mean   : 5218
```

```
## 3rd Qu.: 48740 3rd Qu.: 5006 3rd Qu.: 5000.0 3rd Qu.: 4525
## Max. : 961664 Max. :493358 Max. :1684259.0 Max. :896040
## PAY_AMT4 PAY_AMT5 PAY_AMT6
## Min. : 0 Min. : 0.0 Min. : 0
## 1st Qu.: 300 1st Qu.: 269.8 1st Qu.: 100
## Median : 1500 Median : 1500.0 Median : 1500
## Mean : 4793 Mean : 4838.0 Mean : 5252
## 3rd Qu.: 4010 3rd Qu.: 4037.8 3rd Qu.: 4000
## Max. :528897 Max. :426529.0 Max. :528666
## default.payment.next.month
## Min. :0.0000
## 1st Qu.:0.0000
## Median :0.0000
## Mean :0.2209
## 3rd Qu.:0.0000
## Max. :1.0000
```

```
# También tenemos que modificar la columna EDUCATION, la dividiremos en cuatro columnas diferentes:
# ed.other, ed.university, ed.high_school, ed.school
ed.other = ifelse(credit_card.train$EDUCATION == 4, 1,0)
ed.university = ifelse(credit_card.train$EDUCATION == 2, 1, 0)
ed.high_school = ifelse(credit_card.train$EDUCATION == 3 | credit_card.train$EDUCATION == 2, 1, 0)
ed.school = ifelse(credit_card.train$EDUCATION == 1 | credit_card.train$EDUCATION == 2 | credit_card.train$EDUCATION == 3, 1, 0)

credit_card.train = cbind(credit_card.train,ed.other, ed.high_school, ed.school, ed.university)

# Borramos la columna EDUCATION.
credit_card.train = credit_card.train[,~which(colnames(credit_card.train) == "EDUCATION")]
summary(credit_card.train)
```

```
## LIMIT_BAL SEX MARRIAGE AGE
## Min. : 10000 Min. :0.0000 Min. :0.000 Min. :21.0
## 1st Qu.: 50000 1st Qu.:0.0000 1st Qu.:1.000 1st Qu.:28.0
## Median : 140000 Median :0.0000 Median :2.000 Median :34.0
## Mean : 167948 Mean :0.3974 Mean :1.549 Mean :35.5
## 3rd Qu.: 240000 3rd Qu.:1.0000 3rd Qu.:2.000 3rd Qu.:41.0
## Max. :1000000 Max. :1.0000 Max. :3.000 Max. :79.0
## PAY_0 PAY_2 PAY_3 PAY_4
## Min. : -2.00000 Min. : -2.0000 Min. : -2.0000 Min. : -2.000
## 1st Qu.: -1.00000 1st Qu.: -1.0000 1st Qu.: -1.0000 1st Qu.: -1.000
## Median : 0.00000 Median : 0.0000 Median : 0.0000 Median : 0.000
## Mean : -0.01524 Mean : -0.1321 Mean : -0.1697 Mean : -0.221
## 3rd Qu.: 0.00000 3rd Qu.: 0.0000 3rd Qu.: 0.0000 3rd Qu.: 0.000
## Max. : 8.00000 Max. : 8.0000 Max. : 8.0000 Max. : 8.000
## PAY_5 PAY_6 BILL_AMT1 BILL_AMT2
## Min. : -2.0000 Min. : -2.000 Min. : -154973 Min. : -69777
## 1st Qu.: -1.0000 1st Qu.: -1.000 1st Qu.: 3508 1st Qu.: 2956
## Median : 0.0000 Median : 0.000 Median : 22143 Median : 20895
## Mean : -0.2678 Mean : -0.289 Mean : 51238 Mean : 49092
## 3rd Qu.: 0.0000 3rd Qu.: 0.000 3rd Qu.: 66175 3rd Qu.: 62795
## Max. : 8.0000 Max. : 8.000 Max. : 964511 Max. : 983931
## BILL_AMT3 BILL_AMT4 BILL_AMT5 BILL_AMT6
## Min. : -46127 Min. : -170000 Min. : -81334 Min. : -339603
## 1st Qu.: 2552 1st Qu.: 2276 1st Qu.: 1730 1st Qu.: 1261
## Median : 20012 Median : 19044 Median : 18093 Median : 17036
```

```
## Mean : 47104 Mean : 43452 Mean : 40395 Mean : 38948
## 3rd Qu.: 59674 3rd Qu.: 53536 3rd Qu.: 49719 3rd Qu.: 48740
## Max. :1664089 Max. : 891586 Max. :927171 Max. : 961664
## PAY_AMT1 PAY_AMT2 PAY_AMT3 PAY_AMT4
## Min. : 0 Min. : 0.0 Min. : 0 Min. : 0
## 1st Qu.: 1000 1st Qu.: 809.8 1st Qu.: 390 1st Qu.: 300
## Median : 2100 Median : 2000.0 Median : 1794 Median : 1500
## Mean : 5608 Mean : 5988.9 Mean : 5218 Mean : 4793
## 3rd Qu.: 5006 3rd Qu.: 5000.0 3rd Qu.: 4525 3rd Qu.: 4010
## Max. :493358 Max. :1684259.0 Max. :896040 Max. :528897
## PAY_AMT5 PAY_AMT6 default.payment.next.month
## Min. : 0.0 Min. : 0 Min. :0.0000
## 1st Qu.: 269.8 1st Qu.: 100 1st Qu.:0.0000
## Median : 1500.0 Median : 1500 Median :0.0000
## Mean : 4838.0 Mean : 5252 Mean :0.2209
## 3rd Qu.: 4037.8 3rd Qu.: 4000 3rd Qu.:0.0000
## Max. :426529.0 Max. :528666 Max. :1.0000
## ed.other ed.high_school ed.school ed.university
## Min. :0.000000 Min. :0.0000 Min. :0.0000 Min. :0.0000
## 1st Qu.:0.000000 1st Qu.:0.0000 1st Qu.:1.0000 1st Qu.:0.0000
## Median :0.000000 Median :1.0000 Median :1.0000 Median :0.0000
## Mean :0.003857 Mean :0.6332 Mean :0.9844 Mean :0.4708
## 3rd Qu.:0.000000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.0000
## Max. :1.000000 Max. :1.0000 Max. :1.0000 Max. :1.0000
```

También tenemos que modificar la columna marriage. Introduciremos tres nueva columnas: marriage.married, marriage.single, marriage.others

```
marriage.married = ifelse(credit_card.train$MARRIAGE == 1, 1,0)
marriage.single = ifelse(credit_card.train$MARRIAGE == 2, 1,0)
marriage.others = ifelse(credit_card.train$MARRIAGE == 3, 1,0)
```

Introducimos los datos.

```
credit_card.train = cbind(credit_card.train, marriage.married, marriage.single, marriage.others)
```

Borramos la variable MARRIAGE.

```
credit_card.train = credit_card.train[, -which(colnames(credit_card.train) == "MARRIAGE")]
summary(credit_card.train)
```

```
## LIMIT_BAL SEX AGE PAY_0
## Min. : 10000 Min. :0.0000 Min. :21.0 Min. : -2.00000
## 1st Qu.: 50000 1st Qu.:0.0000 1st Qu.:28.0 1st Qu.: -1.00000
## Median : 140000 Median :0.0000 Median :34.0 Median : 0.00000
## Mean : 167948 Mean :0.3974 Mean :35.5 Mean : -0.01524
## 3rd Qu.: 240000 3rd Qu.:1.0000 3rd Qu.:41.0 3rd Qu.: 0.00000
## Max. :1000000 Max. :1.0000 Max. :79.0 Max. : 8.00000
## PAY_2 PAY_3 PAY_4 PAY_5
## Min. : -2.0000 Min. : -2.0000 Min. : -2.000 Min. : -2.0000
## 1st Qu.: -1.0000 1st Qu.: -1.0000 1st Qu.: -1.000 1st Qu.: -1.0000
## Median : 0.0000 Median : 0.0000 Median : 0.000 Median : 0.0000
## Mean : -0.1321 Mean : -0.1697 Mean : -0.221 Mean : -0.2678
## 3rd Qu.: 0.0000 3rd Qu.: 0.0000 3rd Qu.: 0.000 3rd Qu.: 0.0000
## Max. : 8.0000 Max. : 8.0000 Max. : 8.000 Max. : 8.0000
## PAY_6 BILL_AMT1 BILL_AMT2 BILL_AMT3
## Min. : -2.000 Min. : -154973 Min. : -69777 Min. : -46127
## 1st Qu.: -1.000 1st Qu.: 3508 1st Qu.: 2956 1st Qu.: 2552
## Median : 0.000 Median : 22143 Median : 20895 Median : 20012
```

```
## Mean      :-0.289      Mean      : 51238      Mean      : 49092      Mean      : 47104
## 3rd Qu.: 0.000      3rd Qu.: 66175      3rd Qu.: 62795      3rd Qu.: 59674
## Max.      : 8.000      Max.      : 964511      Max.      :983931      Max.      :1664089
## BILL_AMT4      BILL_AMT5      BILL_AMT6      PAY_AMT1
## Min.      :-170000      Min.      :-81334      Min.      :-339603      Min.      : 0
## 1st Qu.: 2276      1st Qu.: 1730      1st Qu.: 1261      1st Qu.: 1000
## Median : 19044      Median : 18093      Median : 17036      Median : 2100
## Mean      : 43452      Mean      : 40395      Mean      : 38948      Mean      : 5608
## 3rd Qu.: 53536      3rd Qu.: 49719      3rd Qu.: 48740      3rd Qu.: 5006
## Max.      : 891586      Max.      :927171      Max.      : 961664      Max.      :493358
## PAY_AMT2      PAY_AMT3      PAY_AMT4      PAY_AMT5
## Min.      : 0.0      Min.      : 0      Min.      : 0      Min.      : 0.0
## 1st Qu.: 809.8      1st Qu.: 390      1st Qu.: 300      1st Qu.: 269.8
## Median : 2000.0      Median : 1794      Median : 1500      Median : 1500.0
## Mean      : 5988.9      Mean      : 5218      Mean      : 4793      Mean      : 4838.0
## 3rd Qu.: 5000.0      3rd Qu.: 4525      3rd Qu.: 4010      3rd Qu.: 4037.8
## Max.      :1684259.0      Max.      :896040      Max.      :528897      Max.      :426529.0
## PAY_AMT6      default.payment.next.month      ed.other
## Min.      : 0      Min.      :0.0000      Min.      :0.000000
## 1st Qu.: 100      1st Qu.:0.0000      1st Qu.:0.000000
## Median : 1500      Median :0.0000      Median :0.000000
## Mean      : 5252      Mean      :0.2209      Mean      :0.003857
## 3rd Qu.: 4000      3rd Qu.:0.0000      3rd Qu.:0.000000
## Max.      :528666      Max.      :1.0000      Max.      :1.000000
## ed.high_school      ed.school      ed.university      marriage.married
## Min.      :0.0000      Min.      :0.0000      Min.      :0.0000      Min.      :0.000
## 1st Qu.:0.0000      1st Qu.:1.0000      1st Qu.:0.0000      1st Qu.:0.000
## Median :1.0000      Median :1.0000      Median :0.0000      Median :0.000
## Mean      :0.6332      Mean      :0.9844      Mean      :0.4708      Mean      :0.458
## 3rd Qu.:1.0000      3rd Qu.:1.0000      3rd Qu.:1.0000      3rd Qu.:1.000
## Max.      :1.0000      Max.      :1.0000      Max.      :1.0000      Max.      :1.000
## marriage.single      marriage.others
## Min.      :0.0000      Min.      :0.000000
## 1st Qu.:0.0000      1st Qu.:0.000000
## Median :1.0000      Median :0.000000
## Mean      :0.5296      Mean      :0.01076
## 3rd Qu.:1.0000      3rd Qu.:0.000000
## Max.      :1.0000      Max.      :1.000000
```

Tambi n vamos a cambiar el nombre de la columna *PAY0* por *PAY1*.

```
colnames(credit_card.train)[which(colnames(credit_card.train)=="PAY_0")]="PAY_1"
summary(credit_card.train)
```

```
## LIMIT_BAL      SEX      AGE      PAY_1
## Min.      : 10000      Min.      :0.0000      Min.      :21.0      Min.      : -2.00000
## 1st Qu.: 50000      1st Qu.:0.0000      1st Qu.:28.0      1st Qu.: -1.00000
## Median : 140000      Median :0.0000      Median :34.0      Median : 0.00000
## Mean      : 167948      Mean      :0.3974      Mean      :35.5      Mean      : -0.01524
## 3rd Qu.: 240000      3rd Qu.:1.0000      3rd Qu.:41.0      3rd Qu.: 0.00000
## Max.      :1000000      Max.      :1.0000      Max.      :79.0      Max.      : 8.00000
## PAY_2      PAY_3      PAY_4      PAY_5
## Min.      : -2.0000      Min.      : -2.0000      Min.      : -2.000      Min.      : -2.0000
## 1st Qu.: -1.0000      1st Qu.: -1.0000      1st Qu.: -1.000      1st Qu.: -1.0000
## Median : 0.0000      Median : 0.0000      Median : 0.000      Median : 0.0000
```

```

## Mean      :-0.1321    Mean      :-0.1697    Mean      :-0.221    Mean      :-0.2678
## 3rd Qu.: 0.0000    3rd Qu.: 0.0000    3rd Qu.: 0.000    3rd Qu.: 0.0000
## Max.      : 8.0000    Max.      : 8.0000    Max.      : 8.000    Max.      : 8.0000
## PAY_6      BILL_AMT1      BILL_AMT2      BILL_AMT3
## Min.      :-2.000    Min.      :-154973    Min.      :-69777    Min.      : -46127
## 1st Qu.: -1.000    1st Qu.: 3508    1st Qu.: 2956    1st Qu.: 2552
## Median : 0.000    Median : 22143    Median : 20895    Median : 20012
## Mean      :-0.289    Mean      : 51238    Mean      : 49092    Mean      : 47104
## 3rd Qu.: 0.000    3rd Qu.: 66175    3rd Qu.: 62795    3rd Qu.: 59674
## Max.      : 8.000    Max.      : 964511    Max.      :983931    Max.      :1664089
## BILL_AMT4      BILL_AMT5      BILL_AMT6      PAY_AMT1
## Min.      :-170000    Min.      :-81334    Min.      :-339603    Min.      : 0
## 1st Qu.: 2276    1st Qu.: 1730    1st Qu.: 1261    1st Qu.: 1000
## Median : 19044    Median : 18093    Median : 17036    Median : 2100
## Mean      : 43452    Mean      : 40395    Mean      : 38948    Mean      : 5608
## 3rd Qu.: 53536    3rd Qu.: 49719    3rd Qu.: 48740    3rd Qu.: 5006
## Max.      : 891586    Max.      :927171    Max.      : 961664    Max.      :493358
## PAY_AMT2      PAY_AMT3      PAY_AMT4      PAY_AMT5
## Min.      : 0.0    Min.      : 0    Min.      : 0    Min.      : 0.0
## 1st Qu.: 809.8    1st Qu.: 390    1st Qu.: 300    1st Qu.: 269.8
## Median : 2000.0    Median : 1794    Median : 1500    Median : 1500.0
## Mean      : 5988.9    Mean      : 5218    Mean      : 4793    Mean      : 4838.0
## 3rd Qu.: 5000.0    3rd Qu.: 4525    3rd Qu.: 4010    3rd Qu.: 4037.8
## Max.      :1684259.0    Max.      :896040    Max.      :528897    Max.      :426529.0
## PAY_AMT6      default.payment.next.month      ed.other
## Min.      : 0    Min.      :0.0000    Min.      :0.000000
## 1st Qu.: 100    1st Qu.:0.0000    1st Qu.:0.000000
## Median : 1500    Median :0.0000    Median :0.000000
## Mean      : 5252    Mean      :0.2209    Mean      :0.003857
## 3rd Qu.: 4000    3rd Qu.:0.0000    3rd Qu.:0.000000
## Max.      :528666    Max.      :1.0000    Max.      :1.000000
## ed.high_school      ed.school      ed.university      marriage.married
## Min.      :0.0000    Min.      :0.0000    Min.      :0.0000    Min.      :0.000
## 1st Qu.:0.0000    1st Qu.:1.0000    1st Qu.:0.0000    1st Qu.:0.000
## Median :1.0000    Median :1.0000    Median :0.0000    Median :0.000
## Mean      :0.6332    Mean      :0.9844    Mean      :0.4708    Mean      :0.458
## 3rd Qu.:1.0000    3rd Qu.:1.0000    3rd Qu.:1.0000    3rd Qu.:1.000
## Max.      :1.0000    Max.      :1.0000    Max.      :1.0000    Max.      :1.000
## marriage.single      marriage.others
## Min.      :0.0000    Min.      :0.000000
## 1st Qu.:0.0000    1st Qu.:0.000000
## Median :1.0000    Median :0.000000
## Mean      :0.5296    Mean      :0.01076
## 3rd Qu.:1.0000    3rd Qu.:0.000000
## Max.      :1.0000    Max.      :1.000000

```

Por Ãltimo.

```

# Por Ãltimo, utilizamos la funciÃ³n preprocess.
trans = preProcess(credit_card.train, c("BoxCox") )
trainTransformado = predict(trans, credit_card.train)
summary(trainTransformado)

```

```

## LIMIT_BAL      SEX      AGE      PAY_1
## Min.      : 49.50    Min.      :0.0000    Min.      :1.564    Min.      : -2.00000

```

##	1st Qu.: 82.29	1st Qu.:0.0000	1st Qu.:1.622	1st Qu.: -1.00000
##	Median :113.27	Median :0.0000	Median :1.657	Median : 0.00000
##	Mean :111.23	Mean :0.3974	Mean :1.656	Mean : -0.01524
##	3rd Qu.:133.74	3rd Qu.:1.0000	3rd Qu.:1.688	3rd Qu.: 0.00000
##	Max. :206.99	Max. :1.0000	Max. :1.775	Max. : 8.00000
##	PAY_2	PAY_3	PAY_4	PAY_5
##	Min. : -2.0000	Min. : -2.0000	Min. : -2.000	Min. : -2.0000
##	1st Qu.: -1.0000	1st Qu.: -1.0000	1st Qu.: -1.000	1st Qu.: -1.0000
##	Median : 0.0000	Median : 0.0000	Median : 0.000	Median : 0.0000
##	Mean : -0.1321	Mean : -0.1697	Mean : -0.221	Mean : -0.2678
##	3rd Qu.: 0.0000	3rd Qu.: 0.0000	3rd Qu.: 0.000	3rd Qu.: 0.0000
##	Max. : 8.0000	Max. : 8.0000	Max. : 8.000	Max. : 8.0000
##	PAY_6	BILL_AMT1	BILL_AMT2	BILL_AMT3
##	Min. : -2.000	Min. : -154973	Min. : -69777	Min. : -46127
##	1st Qu.: -1.000	1st Qu.: 3508	1st Qu.: 2956	1st Qu.: 2552
##	Median : 0.000	Median : 22143	Median : 20895	Median : 20012
##	Mean : -0.289	Mean : 51238	Mean : 49092	Mean : 47104
##	3rd Qu.: 0.000	3rd Qu.: 66175	3rd Qu.: 62795	3rd Qu.: 59674
##	Max. : 8.000	Max. : 964511	Max. : 983931	Max. : 1664089
##	BILL_AMT4	BILL_AMT5	BILL_AMT6	PAY_AMT1
##	Min. : -170000	Min. : -81334	Min. : -339603	Min. : 0
##	1st Qu.: 2276	1st Qu.: 1730	1st Qu.: 1261	1st Qu.: 1000
##	Median : 19044	Median : 18093	Median : 17036	Median : 2100
##	Mean : 43452	Mean : 40395	Mean : 38948	Mean : 5608
##	3rd Qu.: 53536	3rd Qu.: 49719	3rd Qu.: 48740	3rd Qu.: 5006
##	Max. : 891586	Max. : 927171	Max. : 961664	Max. : 493358
##	PAY_AMT2	PAY_AMT3	PAY_AMT4	PAY_AMT5
##	Min. : 0.0	Min. : 0	Min. : 0	Min. : 0.0
##	1st Qu.: 809.8	1st Qu.: 390	1st Qu.: 300	1st Qu.: 269.8
##	Median : 2000.0	Median : 1794	Median : 1500	Median : 1500.0
##	Mean : 5988.9	Mean : 5218	Mean : 4793	Mean : 4838.0
##	3rd Qu.: 5000.0	3rd Qu.: 4525	3rd Qu.: 4010	3rd Qu.: 4037.8
##	Max. : 1684259.0	Max. : 896040	Max. : 528897	Max. : 426529.0
##	PAY_AMT6	default.payment.next.month	ed.other	
##	Min. : 0	Min. : 0.0000	Min. : 0.000000	
##	1st Qu.: 100	1st Qu.: 0.0000	1st Qu.: 0.000000	
##	Median : 1500	Median : 0.0000	Median : 0.000000	
##	Mean : 5252	Mean : 0.2209	Mean : 0.003857	
##	3rd Qu.: 4000	3rd Qu.: 0.0000	3rd Qu.: 0.000000	
##	Max. : 528666	Max. : 1.0000	Max. : 1.000000	
##	ed.high_school	ed.school	ed.university	marriage.married
##	Min. : 0.0000	Min. : 0.0000	Min. : 0.0000	Min. : 0.000
##	1st Qu.: 0.0000	1st Qu.: 1.0000	1st Qu.: 0.0000	1st Qu.: 0.000
##	Median : 1.0000	Median : 1.0000	Median : 0.0000	Median : 0.000
##	Mean : 0.6332	Mean : 0.9844	Mean : 0.4708	Mean : 0.458
##	3rd Qu.: 1.0000	3rd Qu.: 1.0000	3rd Qu.: 1.0000	3rd Qu.: 1.000
##	Max. : 1.0000	Max. : 1.0000	Max. : 1.0000	Max. : 1.000
##	marriage.single	marriage.others		
##	Min. : 0.0000	Min. : 0.00000		
##	1st Qu.: 0.0000	1st Qu.: 0.00000		
##	Median : 1.0000	Median : 0.00000		
##	Mean : 0.5296	Mean : 0.01076		
##	3rd Qu.: 1.0000	3rd Qu.: 0.00000		
##	Max. : 1.0000	Max. : 1.00000		


```

# Para hacer más sencillo hacer las transformaciones al conjunto de test, se crearán funciones para r

# Función para comprobar si hay datos perdidos y reemplazarlos.
reemplazarCol = function(x){
  # Columna SEX
  x$SEX = ifelse(x$SEX == 2, 0, 1)

  # Columna EDUCATION.
  x.other = ifelse(x$EDUCATION == 4, 1, 0)
  x.university = ifelse(x$EDUCATION == 2, 1, 0)
  x.high_school = ifelse(x$EDUCATION == 3 | x$EDUCATION == 2, 1, 0)
  x.school = ifelse(x$EDUCATION == 1 | x$EDUCATION == 2 | x$EDUCATION == 3, 1, 0)
  x = cbind(x, x.other, x.high_school, x.school, x.university)

  # Borramos la columna EDUCATION.
  x = x[, -which(colnames(x) == "EDUCATION")]

  # Columna MARRIAGE.
  x.married = ifelse(x$MARRIAGE == 1, 1, 0)
  x.single = ifelse(x$MARRIAGE == 2, 1, 0)
  x.others = ifelse(x$MARRIAGE == 3, 1, 0)

  # Introducimos los datos.
  x = cbind(x.train, x.married, x.single, x.others)

  # Borramos la variable MARRIAGE.
  x = x[, -which(colnames(x) == "MARRIAGE")]

  # Cambiamos el nombre de la variables PAY_0.
  colnames(x)[which(colnames(x) == "PAY_0")] = "PAY_1"
  x
}

preprocesar = function(x, pred=trans){
  transTest = predict(pred, x)
  transTest
}

```

Por último, vamos a utilizar el filtro PCA para comprobar si todos los datos son relevantes. Si encontramos algún dato redundante, lo eliminaremos de nuestro conjunto de datos. Para saber si un dato es redundante, debemos comprobar si todos los atributos PC (PC1, PC2, ..., PCx) son 0 o muy cercanos a 0; si encontramos algún atributo PC que se aleje de 0, no podemos asegurar que el atributo sea redundante, y por lo tanto no lo podremos quitar.

```

pcaTransformation = prcomp(trainTransformado[, -default.payment.next.month], center=F, scale=F)
pcaTransformation$rotation

```

	PC1	PC2	PC3
## SEX	-1.246341e-06	-4.904872e-07	6.438891e-07
## AGE	-4.969286e-06	-4.554302e-07	2.875395e-06
## PAY_1	-8.402099e-07	-1.174667e-06	-3.702498e-06
## PAY_2	-7.831655e-07	-1.423524e-06	-4.412710e-06
## PAY_3	-6.673071e-07	-4.589883e-07	-3.545257e-06
## PAY_4	-5.207917e-07	1.417210e-06	-1.570775e-06
## PAY_5	-4.037429e-07	2.881611e-06	-2.724843e-06

## PAY_6	-3.766899e-07	3.805578e-06	-4.208659e-06
## BILL_AMT1	-4.504475e-01	-5.518948e-01	-1.939668e-01
## BILL_AMT2	-4.392710e-01	-4.001284e-01	-5.929929e-02
## BILL_AMT3	-4.267653e-01	-3.653856e-02	5.589409e-01
## BILL_AMT4	-3.965494e-01	2.495663e-01	7.254672e-02
## BILL_AMT5	-3.674088e-01	4.259793e-01	-2.116082e-01
## BILL_AMT6	-3.519808e-01	4.823144e-01	-3.119262e-01
## PAY_AMT1	-3.243380e-02	2.060446e-02	1.527775e-01
## PAY_AMT2	-3.775548e-02	1.711008e-01	6.730651e-01
## PAY_AMT3	-3.242933e-02	1.224791e-01	-3.265661e-03
## PAY_AMT4	-2.677187e-02	9.158369e-02	-5.831541e-02
## PAY_AMT5	-2.730621e-02	6.424338e-02	3.452144e-02
## PAY_AMT6	-2.977955e-02	-1.973644e-02	1.347066e-01
## default.payment.next.month	-6.351776e-07	-4.214420e-08	-3.104142e-07
## ed.other	-9.624187e-09	-2.745489e-08	2.474927e-08
## ed.high_school	-1.896296e-06	-6.830605e-07	4.915918e-07
## ed.school	-2.934032e-06	-1.293661e-07	1.589245e-06
## ed.university	-1.456700e-06	-4.301804e-07	2.616470e-07
## marriage.married	-1.425640e-06	-7.724235e-08	9.020876e-07
## marriage.single	-1.540599e-06	-1.426100e-07	8.016949e-07
## marriage.others	-2.413312e-08	-4.908935e-08	2.310480e-08
##	PC4	PC5	PC6
## SEX	-3.566762e-06	-1.520925e-06	5.286491e-07
## AGE	-1.461946e-05	-5.976859e-06	3.173099e-06
## PAY_1	8.990123e-06	5.326504e-06	-2.426989e-06
## PAY_2	1.002619e-05	6.261145e-06	-3.382277e-06
## PAY_3	1.159672e-05	5.567521e-06	2.979077e-06
## PAY_4	1.226031e-05	6.646866e-06	4.967272e-07
## PAY_5	1.250268e-05	3.280485e-06	-1.587765e-06
## PAY_6	1.088009e-05	5.014483e-06	-8.464995e-07
## BILL_AMT1	-2.931502e-01	-3.297459e-03	-4.848133e-01
## BILL_AMT2	3.499226e-02	-2.453157e-02	5.131682e-01
## BILL_AMT3	2.082730e-01	2.751252e-01	5.838854e-02
## BILL_AMT4	3.800910e-01	-4.508734e-01	-1.622145e-01
## BILL_AMT5	6.119784e-02	-9.923400e-02	-6.073123e-03
## BILL_AMT6	-2.494273e-01	3.956564e-01	6.400747e-02
## PAY_AMT1	-1.580537e-01	-1.890328e-01	5.668722e-01
## PAY_AMT2	-2.936760e-01	8.716394e-02	-2.619771e-01
## PAY_AMT3	-2.448393e-01	-5.475542e-01	-1.790583e-01
## PAY_AMT4	-3.731955e-01	-3.350907e-03	7.561300e-02
## PAY_AMT5	-4.698687e-01	1.592771e-01	6.227597e-02
## PAY_AMT6	-3.575028e-01	-4.306835e-01	1.900638e-01
## default.payment.next.month	-2.441713e-07	8.554837e-08	1.441512e-07
## ed.other	-5.952895e-08	-6.177029e-08	-4.029573e-08
## ed.high_school	-4.060606e-06	-9.485931e-07	7.318980e-07
## ed.school	-8.651648e-06	-3.320632e-06	2.003742e-06
## ed.university	-3.086885e-06	-7.760970e-07	5.044920e-07
## marriage.married	-3.977989e-06	-1.701712e-06	6.245093e-07
## marriage.single	-4.704388e-06	-1.882783e-06	1.280249e-06
## marriage.others	-1.013772e-07	7.803180e-09	-1.311795e-08
##	PC7	PC8	PC9
## SEX	-1.174610e-07	1.187276e-06	-7.100879e-07
## AGE	1.514887e-07	3.516323e-06	-2.187484e-06
## PAY_1	-7.950970e-07	-3.288392e-06	2.428264e-06

## PAY_2	-8.862946e-07	-4.894956e-06	1.669249e-06
## PAY_3	1.712987e-07	-3.554414e-06	1.068221e-06
## PAY_4	-1.059430e-06	-5.071150e-06	2.949098e-06
## PAY_5	9.871106e-07	-3.844836e-06	5.277723e-06
## PAY_6	-3.233630e-06	-2.176669e-06	4.700824e-06
## BILL_AMT1	-5.784232e-02	2.612453e-02	1.639371e-01
## BILL_AMT2	7.860051e-02	1.918421e-01	3.484413e-02
## BILL_AMT3	-6.908146e-02	-1.534870e-01	-3.474197e-01
## BILL_AMT4	2.872122e-01	-1.086755e-01	-1.511017e-01
## BILL_AMT5	-4.807444e-01	1.034837e-01	1.026033e-01
## BILL_AMT6	2.294940e-01	-1.042140e-01	2.420389e-01
## PAY_AMT1	9.399958e-02	3.042951e-01	2.706315e-01
## PAY_AMT2	-6.229616e-02	2.569055e-01	3.745537e-01
## PAY_AMT3	2.838555e-01	3.527173e-01	-1.530299e-01
## PAY_AMT4	-4.703760e-01	2.596694e-01	-6.002146e-01
## PAY_AMT5	4.883207e-01	-1.949813e-01	-3.736668e-01
## PAY_AMT6	-2.577384e-01	-7.191139e-01	1.497786e-01
## default.payment.next.month	-1.028962e-07	-2.499193e-07	-1.679699e-07
## ed.other	4.080268e-08	7.368398e-08	-9.062423e-08
## ed.high_school	3.044095e-08	1.277443e-06	-1.851432e-07
## ed.school	8.033230e-08	1.986790e-06	-1.117627e-06
## ed.university	2.021123e-07	5.957239e-07	-7.079063e-08
## marriage.married	-4.819974e-07	1.138148e-06	-7.005156e-07
## marriage.single	5.108585e-07	9.349753e-07	-5.659778e-07
## marriage.others	6.052524e-08	2.104871e-08	-6.005771e-08
##	PC10	PC11	PC12
## SEX	2.036624e-06	-7.436572e-07	1.129676e-06
## AGE	8.792866e-06	-3.923402e-06	2.457274e-06
## PAY_1	-4.146034e-06	2.369751e-06	-2.616384e-06
## PAY_2	-6.314026e-06	4.068402e-06	-1.224670e-06
## PAY_3	-5.350430e-06	1.189033e-06	-1.563536e-06
## PAY_4	-2.355507e-06	2.953267e-06	2.803969e-08
## PAY_5	-4.564415e-06	1.561645e-06	-2.916779e-06
## PAY_6	-6.017116e-06	4.932324e-07	8.077485e-07
## BILL_AMT1	2.478302e-01	-2.037808e-01	6.647957e-03
## BILL_AMT2	-3.210795e-01	4.804633e-01	1.101541e-02
## BILL_AMT3	-1.827823e-01	-4.374404e-01	-8.857876e-02
## BILL_AMT4	3.769429e-01	2.292044e-01	2.989633e-01
## BILL_AMT5	6.006808e-02	7.354659e-02	-6.031386e-01
## BILL_AMT6	-1.972868e-01	-1.425941e-01	3.767805e-01
## PAY_AMT1	4.808991e-01	-4.325673e-01	8.535603e-03
## PAY_AMT2	-4.298254e-03	3.726952e-01	6.866880e-02
## PAY_AMT3	-5.275991e-01	-2.733176e-01	-1.031887e-01
## PAY_AMT4	1.413774e-01	9.280919e-02	4.061418e-01
## PAY_AMT5	2.639755e-01	2.206015e-01	-4.576697e-01
## PAY_AMT6	-1.347172e-01	4.938737e-04	8.378889e-02
## default.payment.next.month	1.118034e-07	1.525853e-07	-1.476304e-07
## ed.other	-5.982512e-08	-1.087642e-07	-1.822524e-08
## ed.high_school	2.824303e-06	-1.328202e-06	2.085331e-06
## ed.school	5.245684e-06	-2.347095e-06	1.718855e-06
## ed.university	2.245202e-06	-4.064207e-07	1.662469e-06
## marriage.married	2.420860e-06	-1.849774e-06	2.169335e-07
## marriage.single	2.849312e-06	-3.909275e-07	1.299498e-06
## marriage.others	2.200491e-08	-1.101321e-07	-1.341060e-08

	PC13	PC14	PC15
## SEX	-4.994388e-02	1.758165e-01	2.101293e-02
## AGE	-2.596834e-01	6.524050e-01	1.111274e-01
## PAY_1	2.845449e-01	2.077701e-01	-6.139335e-01
## PAY_2	3.829019e-01	1.854558e-01	-4.240037e-01
## PAY_3	4.076562e-01	1.668681e-01	-1.036243e-01
## PAY_4	4.146728e-01	1.292552e-01	2.389101e-01
## PAY_5	4.050352e-01	8.311893e-02	3.917349e-01
## PAY_6	3.947186e-01	5.291373e-02	4.403649e-01
## BILL_AMT1	1.198614e-05	-6.690806e-06	1.697121e-06
## BILL_AMT2	-4.397442e-06	-1.336072e-06	2.072544e-06
## BILL_AMT3	-4.387547e-06	-5.561526e-07	2.887197e-07
## BILL_AMT4	3.533276e-07	-2.319828e-06	-2.162308e-06
## BILL_AMT5	-5.729437e-06	-6.515125e-07	-1.071000e-06
## BILL_AMT6	-4.331925e-06	-3.200078e-06	-2.163376e-06
## PAY_AMT1	2.204872e-05	-8.197985e-06	-5.764811e-06
## PAY_AMT2	1.028501e-05	-1.219756e-06	-1.911368e-06
## PAY_AMT3	1.398044e-05	4.959350e-08	-1.034443e-06
## PAY_AMT4	2.118311e-05	-5.844697e-06	-1.279737e-06
## PAY_AMT5	1.554901e-05	-3.765884e-06	1.462420e-06
## PAY_AMT6	8.854075e-06	-3.437856e-06	-1.349495e-06
## default.payment.next.month	3.565418e-03	1.389473e-01	-5.492867e-02
## ed.other	-1.083454e-03	1.140839e-04	2.688836e-04
## ed.high_school	-7.551419e-02	3.187417e-01	4.941901e-02
## ed.school	-1.531398e-01	3.949888e-01	6.825188e-02
## ed.university	-5.780995e-02	2.414498e-01	4.875016e-02
## marriage.married	-7.853307e-02	1.762228e-01	1.671912e-02
## marriage.single	-7.627864e-02	2.129252e-01	5.047921e-02
## marriage.others	-1.034405e-03	5.002343e-03	3.949675e-04
	PC16	PC17	PC18
## SEX	1.627752e-02	8.750777e-02	-6.665852e-02
## AGE	-2.995491e-03	5.744671e-02	-6.626594e-02
## PAY_1	-6.173908e-01	6.414465e-02	-2.587711e-01
## PAY_2	2.871176e-01	-2.202249e-02	5.427556e-01
## PAY_3	5.637225e-01	-3.201344e-02	-1.454100e-01
## PAY_4	1.583899e-01	-1.742211e-02	-5.598575e-01
## PAY_5	-1.962281e-01	1.550347e-02	-5.120993e-02
## PAY_6	-3.779363e-01	1.789307e-02	4.471568e-01
## BILL_AMT1	9.071210e-07	1.563716e-07	-1.181259e-06
## BILL_AMT2	-2.002981e-06	-4.971292e-07	1.673945e-07
## BILL_AMT3	-1.182029e-06	7.465039e-07	1.339919e-06
## BILL_AMT4	-1.326285e-07	-4.726019e-07	1.113125e-06
## BILL_AMT5	2.382202e-06	2.480886e-07	-2.326239e-07
## BILL_AMT6	9.022128e-07	-1.679566e-07	-1.558364e-06
## PAY_AMT1	-1.194680e-06	-2.048297e-07	2.436347e-06
## PAY_AMT2	1.247216e-06	-4.297826e-07	-2.029782e-07
## PAY_AMT3	5.342328e-07	-1.082609e-07	-1.891098e-06
## PAY_AMT4	-1.664132e-06	-1.993218e-07	-8.711472e-07
## PAY_AMT5	-2.171168e-06	-7.178214e-07	1.838785e-06
## PAY_AMT6	-6.307774e-07	-4.082202e-07	-3.308277e-08
## default.payment.next.month	-7.173902e-02	-8.771720e-04	-5.920080e-02
## ed.other	-7.705337e-04	1.797029e-03	-2.304360e-03
## ed.high_school	4.699754e-03	-2.166173e-01	1.747435e-01
## ed.school	9.975398e-04	4.730832e-02	-2.564323e-02

## ed.university	8.158430e-03	-1.705680e-01	2.176285e-01
## marriage.married	-5.806579e-02	-6.455242e-01	-8.811367e-02
## marriage.single	5.742924e-02	6.982767e-01	5.131589e-02
## marriage.others	-2.065799e-04	-1.827002e-03	-8.426785e-04
##	PC19	PC20	PC21
## SEX	-3.011371e-01	-5.439066e-02	-9.269997e-01
## AGE	-2.605972e-01	2.322482e-03	2.268556e-01
## PAY_1	1.030241e-01	1.444884e-01	-1.842676e-02
## PAY_2	-1.848766e-01	-4.601682e-01	5.295734e-02
## PAY_3	-9.123618e-03	6.248619e-01	-3.065796e-03
## PAY_4	1.689886e-01	-3.203558e-01	-1.227148e-04
## PAY_5	4.500454e-03	-3.466360e-01	4.049519e-02
## PAY_6	-1.704839e-01	3.882150e-01	-1.424139e-02
## BILL_AMT1	5.299544e-07	1.303697e-06	2.534744e-07
## BILL_AMT2	-3.737755e-07	-1.141532e-06	2.102662e-07
## BILL_AMT3	1.286125e-07	-7.714272e-07	-2.177252e-07
## BILL_AMT4	-5.381460e-07	1.608945e-06	2.150151e-08
## BILL_AMT5	1.224179e-06	-2.527579e-07	-6.550249e-07
## BILL_AMT6	-9.827655e-08	-8.248516e-07	-6.730994e-08
## PAY_AMT1	7.972164e-07	-2.254360e-06	-2.572910e-07
## PAY_AMT2	6.795517e-08	1.892255e-06	1.468281e-07
## PAY_AMT3	1.843346e-06	-9.428161e-07	-4.989937e-08
## PAY_AMT4	1.167263e-06	-1.423284e-06	5.328227e-07
## PAY_AMT5	4.860265e-07	1.235560e-06	-5.761471e-07
## PAY_AMT6	8.858846e-07	4.280460e-07	-3.692929e-07
## default.payment.next.month	-8.387689e-02	1.283156e-02	6.368024e-02
## ed.other	-5.287202e-03	8.967046e-04	-5.917765e-05
## ed.high_school	5.012724e-01	7.962890e-03	-1.206010e-01
## ed.school	-1.267566e-01	3.491108e-03	1.384555e-01
## ed.university	6.150463e-01	1.401235e-02	-1.904139e-01
## marriage.married	-2.690614e-01	-2.538711e-03	7.052401e-02
## marriage.single	1.174594e-01	5.750873e-04	7.215873e-02
## marriage.others	3.094603e-03	3.576971e-03	-2.084945e-03
##	PC22	PC23	PC24
## SEX	-2.272875e-02	-7.700855e-03	-1.221135e-02
## AGE	1.563736e-02	1.275858e-01	6.968118e-02
## PAY_1	-2.515380e-02	1.190260e-01	8.279088e-03
## PAY_2	1.133896e-01	5.308730e-03	8.391557e-03
## PAY_3	-2.535791e-01	1.969639e-02	6.448905e-03
## PAY_4	5.328498e-01	-1.617403e-02	7.003669e-03
## PAY_5	-7.145802e-01	4.150947e-02	4.710116e-03
## PAY_6	3.523879e-01	-1.360022e-02	-3.681348e-03
## BILL_AMT1	-2.313214e-07	-6.527956e-07	3.440264e-08
## BILL_AMT2	1.999122e-07	1.954971e-07	-5.419416e-08
## BILL_AMT3	-3.088154e-07	-1.540538e-08	1.553855e-07
## BILL_AMT4	5.348892e-07	-9.749554e-08	-2.426431e-07
## BILL_AMT5	6.851993e-07	-1.891170e-08	-2.367308e-07
## BILL_AMT6	-1.038164e-06	-3.287585e-09	7.421536e-08
## PAY_AMT1	6.217775e-07	-1.035258e-06	-3.598166e-08
## PAY_AMT2	-7.740161e-07	-2.707333e-07	-7.879775e-08
## PAY_AMT3	2.481757e-06	-2.825527e-07	7.292261e-09
## PAY_AMT4	-3.089837e-06	-2.474531e-07	-1.977012e-08
## PAY_AMT5	1.119133e-06	-2.857288e-07	-4.870618e-07
## PAY_AMT6	-2.829919e-07	-1.951784e-07	-3.777309e-07

## default.payment.next.month	-4.800197e-02	-9.774235e-01	-7.388131e-03
## ed.other	1.287133e-04	2.006302e-03	7.356346e-03
## ed.high_school	-1.058611e-02	-1.286482e-02	-7.377545e-01
## ed.school	9.579186e-03	6.751613e-02	2.515492e-02
## ed.university	-1.166051e-02	-5.175566e-02	6.642126e-01
## marriage.married	-1.885696e-03	6.257076e-02	8.768240e-02
## marriage.single	7.377785e-03	1.502572e-02	8.398431e-03
## marriage.others	3.679473e-03	-2.343316e-03	-2.774480e-02
##	PC25	PC26	PC27
## SEX	-1.098827e-03	-9.080783e-03	-1.157969e-03
## AGE	6.552254e-02	4.633396e-01	1.661856e-01
## PAY_1	-2.668582e-04	-1.975495e-03	-3.450206e-06
## PAY_2	1.762340e-03	3.197400e-03	-4.811322e-04
## PAY_3	-2.006863e-03	5.812044e-04	4.312657e-04
## PAY_4	-1.713267e-03	-1.433924e-03	-7.021459e-04
## PAY_5	4.996782e-03	1.125720e-03	-3.671973e-04
## PAY_6	-3.357714e-03	1.973568e-03	1.212946e-03
## BILL_AMT1	-1.616656e-08	-1.106234e-07	-5.425752e-09
## BILL_AMT2	5.079917e-08	-2.879529e-08	-3.599314e-08
## BILL_AMT3	-3.631167e-08	-8.000041e-08	4.779923e-08
## BILL_AMT4	5.283877e-08	-1.715224e-08	-3.930950e-08
## BILL_AMT5	5.696305e-08	-2.114595e-08	-5.038830e-08
## BILL_AMT6	-6.921757e-08	2.245749e-07	5.992019e-08
## PAY_AMT1	-2.370285e-08	4.074902e-08	-3.536513e-09
## PAY_AMT2	3.617539e-08	5.421583e-09	-4.055122e-08
## PAY_AMT3	-1.590220e-08	-2.214886e-07	7.734258e-08
## PAY_AMT4	4.652286e-08	-9.243737e-09	8.456886e-09
## PAY_AMT5	-6.327530e-08	-1.044964e-07	-3.507302e-08
## PAY_AMT6	7.958019e-09	3.563298e-08	-3.530966e-08
## default.payment.next.month	-4.693718e-03	4.323984e-03	4.404874e-04
## ed.other	-7.698843e-02	3.299730e-01	-9.406210e-01
## ed.high_school	-4.431952e-02	1.706590e-02	1.336661e-04
## ed.school	2.933921e-01	-7.744452e-01	-2.940008e-01
## ed.university	2.240248e-02	1.100169e-02	3.112533e-03
## marriage.married	-3.919074e-01	-2.735200e-02	1.360402e-02
## marriage.single	-3.935797e-01	-1.743775e-03	2.221317e-02
## marriage.others	7.698497e-01	2.745588e-01	2.204228e-02
##	PC28		
## SEX	-2.526586e-03		
## AGE	3.154843e-01		
## PAY_1	4.450828e-04		
## PAY_2	-7.801586e-05		
## PAY_3	2.016428e-04		
## PAY_4	4.108864e-04		
## PAY_5	3.252575e-04		
## PAY_6	6.809812e-06		
## BILL_AMT1	4.225614e-09		
## BILL_AMT2	-1.503755e-08		
## BILL_AMT3	1.305471e-08		
## BILL_AMT4	-4.364877e-09		
## BILL_AMT5	-6.587599e-09		
## BILL_AMT6	-2.826747e-09		
## PAY_AMT1	-1.678314e-08		
## PAY_AMT2	-4.651063e-09		

```
## PAY_AMT3                -1.739533e-09
## PAY_AMT4                 4.393709e-09
## PAY_AMT5                -8.656330e-09
## PAY_AMT6                -8.000439e-09
## default.payment.next.month 5.074205e-04
## ed.other                 1.804076e-02
## ed.high_school          -9.043577e-03
## ed.school               1.280622e-02
## ed.university           9.999352e-03
## marriage.married        -5.395885e-01
## marriage.single         -5.272744e-01
## marriage.others         -5.749932e-01
```

Según los resultados del filtro PCA, no hay ningún dato que sea redundante. Por lo tanto, no eliminaremos ninguno de los atributos del conjunto de datos.

Estudio de los parámetros e hiperparámetros.

Vamos a realizar un estudio sobre los parámetros para saber cuáles de ellos son los más importantes. Con los más importantes crearemos los modelos lineales que vamos a ajustar.

```
modelo_step = glm(default.payment.next.month ~ .,family = gaussian, data=trainTransformado)
modelo_principal = step(modelo_step)
```

```
## Start: AIC=19817.48
## default.payment.next.month ~ LIMIT_BAL + SEX + AGE + PAY_1 +
## PAY_2 + PAY_3 + PAY_4 + PAY_5 + PAY_6 + BILL_AMT1 + BILL_AMT2 +
## BILL_AMT3 + BILL_AMT4 + BILL_AMT5 + BILL_AMT6 + PAY_AMT1 +
## PAY_AMT2 + PAY_AMT3 + PAY_AMT4 + PAY_AMT5 + PAY_AMT6 + ed.other +
## ed.high_school + ed.school + ed.university + marriage.married +
## marriage.single + marriage.others
##
##           Df Deviance   AIC
## - ed.university    1   3150.2 19816
## - BILL_AMT3         1   3150.2 19816
## - BILL_AMT6         1   3150.3 19816
## - BILL_AMT4         1   3150.3 19816
## - PAY_AMT5          1   3150.3 19816
## - ed.other          1   3150.3 19816
## - PAY_6             1   3150.3 19816
## - PAY_AMT3          1   3150.3 19816
## - BILL_AMT5         1   3150.3 19816
## - PAY_AMT6          1   3150.3 19816
## - PAY_4             1   3150.4 19816
## - BILL_AMT2         1   3150.4 19816
## <none>              3150.2 19818
## - PAY_AMT4          1   3150.5 19818
## - ed.high_school    1   3150.6 19818
## - PAY_AMT2          1   3150.6 19818
## - PAY_3             1   3150.9 19820
## - PAY_5             1   3150.9 19820
## - marriage.others   1   3151.0 19821
## - marriage.single   1   3151.1 19821
## - SEX               1   3151.3 19823
```

```

## - marriage.married 1 3151.4 19823
## - AGE 1 3152.4 19830
## - PAY_AMT1 1 3152.6 19831
## - PAY_2 1 3152.8 19833
## - BILL_AMT1 1 3152.9 19833
## - LIMIT_BAL 1 3153.8 19839
## - ed.school 1 3154.5 19844
## - PAY_1 1 3276.2 20639
##
## Step: AIC=19815.48
## default.payment.next.month ~ LIMIT_BAL + SEX + AGE + PAY_1 +
## PAY_2 + PAY_3 + PAY_4 + PAY_5 + PAY_6 + BILL_AMT1 + BILL_AMT2 +
## BILL_AMT3 + BILL_AMT4 + BILL_AMT5 + BILL_AMT6 + PAY_AMT1 +
## PAY_AMT2 + PAY_AMT3 + PAY_AMT4 + PAY_AMT5 + PAY_AMT6 + ed.other +
## ed.high_school + ed.school + marriage.married + marriage.single +
## marriage.others
##
## Df Deviance AIC
## - BILL_AMT3 1 3150.2 19814
## - BILL_AMT6 1 3150.3 19814
## - BILL_AMT4 1 3150.3 19814
## - PAY_AMT5 1 3150.3 19814
## - ed.other 1 3150.3 19814
## - PAY_6 1 3150.3 19814
## - PAY_AMT3 1 3150.3 19814
## - BILL_AMT5 1 3150.3 19814
## - PAY_AMT6 1 3150.3 19814
## - PAY_4 1 3150.4 19814
## - BILL_AMT2 1 3150.4 19814
## <none> 3150.2 19816
## - PAY_AMT4 1 3150.5 19816
## - PAY_AMT2 1 3150.6 19816
## - PAY_3 1 3150.9 19818
## - PAY_5 1 3150.9 19818
## - ed.high_school 1 3151.0 19818
## - marriage.others 1 3151.0 19819
## - marriage.single 1 3151.1 19819
## - SEX 1 3151.3 19821
## - marriage.married 1 3151.4 19821
## - AGE 1 3152.5 19829
## - PAY_AMT1 1 3152.6 19829
## - PAY_2 1 3152.8 19831
## - BILL_AMT1 1 3152.9 19831
## - LIMIT_BAL 1 3153.8 19837
## - ed.school 1 3154.5 19842
## - PAY_1 1 3276.2 20637
##
## Step: AIC=19813.49
## default.payment.next.month ~ LIMIT_BAL + SEX + AGE + PAY_1 +
## PAY_2 + PAY_3 + PAY_4 + PAY_5 + PAY_6 + BILL_AMT1 + BILL_AMT2 +
## BILL_AMT4 + BILL_AMT5 + BILL_AMT6 + PAY_AMT1 + PAY_AMT2 +
## PAY_AMT3 + PAY_AMT4 + PAY_AMT5 + PAY_AMT6 + ed.other + ed.high_school +
## ed.school + marriage.married + marriage.single + marriage.others
##

```



```

##              Df Deviance   AIC
## - BILL_AMT6      1   3150.3 19812
## - BILL_AMT4      1   3150.3 19812
## - ed.other        1   3150.3 19812
## - PAY_AMT5        1   3150.3 19812
## - PAY_6           1   3150.3 19812
## - BILL_AMT5        1   3150.3 19812
## - PAY_AMT3         1   3150.3 19812
## - PAY_AMT6         1   3150.3 19812
## - PAY_4           1   3150.4 19812
## - BILL_AMT2        1   3150.4 19813
## <none>              3150.2 19814
## - PAY_AMT4         1   3150.5 19814
## - PAY_3           1   3150.9 19816
## - PAY_5           1   3150.9 19816
## - PAY_AMT2         1   3151.0 19816
## - ed.high_school   1   3151.0 19816
## - marriage.others   1   3151.0 19817
## - marriage.single   1   3151.1 19817
## - SEX              1   3151.3 19819
## - marriage.married  1   3151.4 19819
## - AGE              1   3152.5 19827
## - PAY_AMT1         1   3152.7 19828
## - PAY_2            1   3152.8 19829
## - BILL_AMT1         1   3152.9 19829
## - LIMIT_BAL         1   3153.8 19835
## - ed.school         1   3154.5 19840
## - PAY_1            1   3276.2 20635
##
## Step:  AIC=19811.59
## default.payment.next.month ~ LIMIT_BAL + SEX + AGE + PAY_1 +
##   PAY_2 + PAY_3 + PAY_4 + PAY_5 + PAY_6 + BILL_AMT1 + BILL_AMT2 +
##   BILL_AMT4 + BILL_AMT5 + PAY_AMT1 + PAY_AMT2 + PAY_AMT3 +
##   PAY_AMT4 + PAY_AMT5 + PAY_AMT6 + ed.other + ed.high_school +
##   ed.school + marriage.married + marriage.single + marriage.others
##
##              Df Deviance   AIC
## - BILL_AMT4      1   3150.3 19810
## - ed.other        1   3150.3 19810
## - PAY_6           1   3150.3 19810
## - BILL_AMT5        1   3150.3 19810
## - PAY_AMT6         1   3150.3 19810
## - PAY_AMT3         1   3150.3 19810
## - PAY_AMT5         1   3150.4 19810
## - PAY_4           1   3150.4 19811
## - BILL_AMT2        1   3150.4 19811
## - PAY_AMT4         1   3150.5 19812
## <none>              3150.3 19812
## - PAY_3           1   3151.0 19814
## - PAY_5           1   3151.0 19814
## - PAY_AMT2         1   3151.0 19814
## - ed.high_school   1   3151.0 19815
## - marriage.others   1   3151.0 19815
## - marriage.single   1   3151.1 19815

```

```

## - SEX                1    3151.3 19817
## - marriage.married   1    3151.4 19817
## - AGE                1    3152.5 19825
## - PAY_AMT1           1    3152.7 19826
## - PAY_2              1    3152.9 19827
## - BILL_AMT1          1    3152.9 19827
## - LIMIT_BAL          1    3153.8 19834
## - ed.school          1    3154.5 19838
## - PAY_1              1    3276.3 20633
##
## Step:  AIC=19809.76
## default.payment.next.month ~ LIMIT_BAL + SEX + AGE + PAY_1 +
##     PAY_2 + PAY_3 + PAY_4 + PAY_5 + PAY_6 + BILL_AMT1 + BILL_AMT2 +
##     BILL_AMT5 + PAY_AMT1 + PAY_AMT2 + PAY_AMT3 + PAY_AMT4 + PAY_AMT5 +
##     PAY_AMT6 + ed.other + ed.high_school + ed.school + marriage.married +
##     marriage.single + marriage.others
##
##               Df Deviance   AIC
## - BILL_AMT5      1    3150.3 19808
## - ed.other        1    3150.3 19808
## - PAY_6           1    3150.3 19808
## - PAY_AMT6        1    3150.3 19808
## - PAY_AMT3        1    3150.4 19808
## - PAY_AMT5        1    3150.4 19809
## - PAY_4           1    3150.4 19809
## - BILL_AMT2       1    3150.4 19809
## - PAY_AMT4        1    3150.6 19810
## <none>            3150.3 19810
## - PAY_5           1    3151.0 19812
## - PAY_3           1    3151.0 19813
## - ed.high_school  1    3151.0 19813
## - marriage.others 1    3151.0 19813
## - PAY_AMT2        1    3151.1 19813
## - marriage.single 1    3151.1 19813
## - SEX             1    3151.3 19815
## - marriage.married 1    3151.4 19815
## - AGE             1    3152.6 19823
## - PAY_AMT1        1    3152.7 19824
## - PAY_2           1    3152.9 19825
## - BILL_AMT1       1    3153.0 19826
## - LIMIT_BAL       1    3153.9 19832
## - ed.school       1    3154.6 19836
## - PAY_1           1    3276.4 20632
##
## Step:  AIC=19807.97
## default.payment.next.month ~ LIMIT_BAL + SEX + AGE + PAY_1 +
##     PAY_2 + PAY_3 + PAY_4 + PAY_5 + PAY_6 + BILL_AMT1 + BILL_AMT2 +
##     PAY_AMT1 + PAY_AMT2 + PAY_AMT3 + PAY_AMT4 + PAY_AMT5 + PAY_AMT6 +
##     ed.other + ed.high_school + ed.school + marriage.married +
##     marriage.single + marriage.others
##
##               Df Deviance   AIC
## - PAY_6           1    3150.4 19806
## - ed.other        1    3150.4 19806

```

```

## - PAY_AMT6      1  3150.4 19806
## - PAY_AMT3      1  3150.4 19807
## - PAY_4         1  3150.4 19807
## - PAY_AMT5      1  3150.4 19807
## - PAY_AMT4      1  3150.6 19808
## - BILL_AMT2     1  3150.6 19808
## <none>          3150.3 19808
## - PAY_3         1  3151.0 19811
## - PAY_5         1  3151.0 19811
## - ed.high_school 1  3151.1 19811
## - marriage.others 1  3151.1 19811
## - PAY_AMT2      1  3151.1 19811
## - marriage.single 1  3151.1 19812
## - SEX           1  3151.4 19813
## - marriage.married 1  3151.4 19814
## - AGE           1  3152.6 19821
## - PAY_AMT1      1  3152.8 19823
## - PAY_2         1  3152.9 19823
## - BILL_AMT1     1  3153.0 19824
## - LIMIT_BAL     1  3153.9 19830
## - ed.school     1  3154.6 19835
## - PAY_1         1  3276.5 20631
##
## Step:  AIC=19806.24
## default.payment.next.month ~ LIMIT_BAL + SEX + AGE + PAY_1 +
##   PAY_2 + PAY_3 + PAY_4 + PAY_5 + BILL_AMT1 + BILL_AMT2 + PAY_AMT1 +
##   PAY_AMT2 + PAY_AMT3 + PAY_AMT4 + PAY_AMT5 + PAY_AMT6 + ed.other +
##   ed.high_school + ed.school + marriage.married + marriage.single +
##   marriage.others
##
##           Df Deviance  AIC
## - ed.other      1  3150.4 19805
## - PAY_AMT6      1  3150.4 19805
## - PAY_AMT3      1  3150.4 19805
## - PAY_4         1  3150.5 19805
## - PAY_AMT5      1  3150.5 19805
## - BILL_AMT2     1  3150.6 19806
## - PAY_AMT4      1  3150.6 19806
## <none>          3150.4 19806
## - PAY_3         1  3151.0 19809
## - ed.high_school 1  3151.1 19809
## - marriage.others 1  3151.1 19809
## - PAY_AMT2      1  3151.1 19809
## - PAY_5         1  3151.1 19810
## - marriage.single 1  3151.2 19810
## - SEX           1  3151.4 19811
## - marriage.married 1  3151.5 19812
## - AGE           1  3152.6 19819
## - PAY_AMT1      1  3152.9 19821
## - PAY_2         1  3152.9 19822
## - BILL_AMT1     1  3153.0 19822
## - LIMIT_BAL     1  3153.9 19828
## - ed.school     1  3154.6 19833
## - PAY_1         1  3276.6 20629

```

```

##
## Step: AIC=19804.5
## default.payment.next.month ~ LIMIT_BAL + SEX + AGE + PAY_1 +
## PAY_2 + PAY_3 + PAY_4 + PAY_5 + BILL_AMT1 + BILL_AMT2 + PAY_AMT1 +
## PAY_AMT2 + PAY_AMT3 + PAY_AMT4 + PAY_AMT5 + PAY_AMT6 + ed.high_school +
## ed.school + marriage.married + marriage.single + marriage.others
##
##          Df Deviance  AIC
## - PAY_AMT6      1   3150.5 19803
## - PAY_AMT3      1   3150.5 19803
## - PAY_4          1   3150.5 19803
## - PAY_AMT5      1   3150.5 19803
## - BILL_AMT2     1   3150.6 19804
## - PAY_AMT4      1   3150.7 19804
## <none>          3150.4 19805
## - PAY_3          1   3151.1 19807
## - ed.high_school 1   3151.1 19808
## - marriage.others 1   3151.2 19808
## - PAY_AMT2      1   3151.2 19808
## - PAY_5          1   3151.2 19808
## - marriage.single 1   3151.2 19808
## - SEX            1   3151.5 19810
## - marriage.married 1   3151.5 19810
## - AGE            1   3152.7 19818
## - PAY_AMT1      1   3152.9 19819
## - PAY_2          1   3153.0 19820
## - BILL_AMT1     1   3153.1 19821
## - LIMIT_BAL     1   3153.9 19826
## - ed.school      1   3155.5 19837
## - PAY_1          1   3276.6 20627
##
## Step: AIC=19802.95
## default.payment.next.month ~ LIMIT_BAL + SEX + AGE + PAY_1 +
## PAY_2 + PAY_3 + PAY_4 + PAY_5 + BILL_AMT1 + BILL_AMT2 + PAY_AMT1 +
## PAY_AMT2 + PAY_AMT3 + PAY_AMT4 + PAY_AMT5 + ed.high_school +
## ed.school + marriage.married + marriage.single + marriage.others
##
##          Df Deviance  AIC
## - PAY_AMT3      1   3150.6 19802
## - PAY_4          1   3150.6 19802
## - PAY_AMT5      1   3150.6 19802
## - BILL_AMT2     1   3150.7 19803
## - PAY_AMT4      1   3150.7 19803
## <none>          3150.5 19803
## - PAY_3          1   3151.1 19806
## - ed.high_school 1   3151.2 19806
## - marriage.others 1   3151.2 19806
## - PAY_5          1   3151.3 19806
## - PAY_AMT2      1   3151.3 19806
## - marriage.single 1   3151.3 19806
## - SEX            1   3151.5 19808
## - marriage.married 1   3151.6 19809
## - AGE            1   3152.7 19816
## - PAY_2          1   3153.0 19818

```

```

## - PAY_AMT1          1    3153.1 19819
## - BILL_AMT1         1    3153.2 19819
## - LIMIT_BAL         1    3154.1 19825
## - ed.school         1    3155.6 19835
## - PAY_1             1    3276.9 20627
##
## Step:  AIC=19801.57
## default.payment.next.month ~ LIMIT_BAL + SEX + AGE + PAY_1 +
##     PAY_2 + PAY_3 + PAY_4 + PAY_5 + BILL_AMT1 + BILL_AMT2 + PAY_AMT1 +
##     PAY_AMT2 + PAY_AMT4 + PAY_AMT5 + ed.high_school + ed.school +
##     marriage.married + marriage.single + marriage.others
##
##              Df Deviance   AIC
## - PAY_4          1    3150.7 19801
## - PAY_AMT5        1    3150.7 19801
## - BILL_AMT2       1    3150.8 19802
## <none>            3150.6 19802
## - PAY_AMT4        1    3150.9 19802
## - PAY_3           1    3151.2 19804
## - ed.high_school  1    3151.3 19805
## - PAY_5           1    3151.3 19805
## - marriage.others 1    3151.3 19805
## - marriage.single 1    3151.4 19805
## - PAY_AMT2        1    3151.4 19805
## - SEX             1    3151.6 19807
## - marriage.married 1    3151.7 19807
## - AGE             1    3152.8 19815
## - PAY_2           1    3153.1 19817
## - PAY_AMT1        1    3153.4 19818
## - BILL_AMT1       1    3153.4 19819
## - LIMIT_BAL       1    3154.3 19824
## - ed.school       1    3155.7 19834
## - PAY_1           1    3277.2 20627
##
## Step:  AIC=19800.58
## default.payment.next.month ~ LIMIT_BAL + SEX + AGE + PAY_1 +
##     PAY_2 + PAY_3 + PAY_5 + BILL_AMT1 + BILL_AMT2 + PAY_AMT1 +
##     PAY_AMT2 + PAY_AMT4 + PAY_AMT5 + ed.high_school + ed.school +
##     marriage.married + marriage.single + marriage.others
##
##              Df Deviance   AIC
## - PAY_AMT5        1    3150.9 19800
## - BILL_AMT2       1    3151.0 19801
## <none>            3150.7 19801
## - PAY_AMT4        1    3151.0 19801
## - ed.high_school  1    3151.4 19803
## - marriage.others 1    3151.5 19804
## - PAY_AMT2        1    3151.5 19804
## - marriage.single 1    3151.5 19804
## - SEX             1    3151.8 19806
## - marriage.married 1    3151.8 19806
## - PAY_3           1    3151.9 19806
## - PAY_5           1    3152.6 19811
## - AGE             1    3153.0 19814

```

```

## - PAY_2          1    3153.3 19816
## - PAY_AMT1       1    3153.6 19818
## - BILL_AMT1      1    3153.6 19818
## - LIMIT_BAL      1    3154.5 19824
## - ed.school      1    3155.9 19833
## - PAY_1          1    3278.4 20633
##
## Step:  AIC=19799.67
## default.payment.next.month ~ LIMIT_BAL + SEX + AGE + PAY_1 +
##     PAY_2 + PAY_3 + PAY_5 + BILL_AMT1 + BILL_AMT2 + PAY_AMT1 +
##     PAY_AMT2 + PAY_AMT4 + ed.high_school + ed.school + marriage.married +
##     marriage.single + marriage.others
##
##              Df Deviance   AIC
## <none>              3150.9 19800
## - BILL_AMT2        1    3151.2 19800
## - PAY_AMT4         1    3151.2 19800
## - ed.high_school   1    3151.6 19803
## - marriage.others  1    3151.6 19803
## - marriage.single  1    3151.7 19803
## - PAY_AMT2         1    3151.8 19804
## - SEX              1    3151.9 19805
## - marriage.married 1    3152.0 19805
## - PAY_3            1    3152.0 19805
## - PAY_5            1    3152.8 19810
## - AGE              1    3153.1 19813
## - PAY_2            1    3153.4 19815
## - PAY_AMT1         1    3153.9 19818
## - BILL_AMT1        1    3153.9 19818
## - LIMIT_BAL        1    3154.8 19824
## - ed.school        1    3156.0 19832
## - PAY_1            1    3278.8 20634

```