



| | | | | | |
|-----------|---|-----------------------------------------|----------|---|------------------|
| Programme | : | BTech – ECE and ECM | Semester | : | Win 2022 |
| Course | : | Essentials of Data Analytics Lab | Code | : | CSE3506 |
| Faculty | : | Gobinath N | Slot | : | L51 + L52 |

Ex_04/ Logistic Regression glm

Code:

```
#Lab4_AssignedWork
```

```
rm(list=ls())
```

```
setwd("C:\\Users\\Rituraj Anand\\Desktop\\Sem6\\CSE3506\\LAB\\Lab  
4")
```

```
credit=read.csv("CreditWorthiness.csv")
```

```
str(credit)
```

```
credit$credit.status <- as.factor(credit$credit.status)
```

```
credit$education <- as.factor(credit$education)
```

```
credit$m.status <- as.factor(credit$m.status)
```

```
credit$Operties <- as.factor(credit$Operties)
```

```
credit$Duration <- as.factor(credit$Duration)
```

```
credit$inPlans <- as.factor(credit$inPlans)
```

```
credit$JobType <- as.factor(credit$JobType)
```

```
credit$Ndepend <- as.factor(credit$Ndepend)
```

```
credit$telephone <- as.factor(credit$telephone)
```

```
credit$foreign <- as.factor(credit$foreign)
```

```
credit$creditScore <- as.factor(credit$creditScore)
str(credit)
set.seed(325)
credit2 <- sample(2, nrow(credit), replace=T, prob = c(0.71, 0.29))
train <- credit[credit2==1,]
tst <- credit[credit2==2,]
model <-
glm(creditScore~credit.status+Loan.required+education+m.status+Opar
ties+Duration+age+inPlans+JobType+Ndepend+telephone+foreign,
data=train, family='binomial')
summary(model)
model <-
glm(creditScore~credit.status+Loan.required+education+foreign,
data=train, family='binomial')
summary(model)
pre <- predict(model, train, type='response')
head(pre)
head(train)
```

OUTPUT:

```

Coefficients:
(Intercept)                -1.291e+00  7.496e-01  -1.723  0.08491 .
credit.statusall settled till now  1.377e+00  3.882e-01  3.547  0.00039 ***
credit.statusdues not paid earlier  1.886e+00  4.174e-01  4.518  6.24e-06 ***
credit.statusnone taken/all settled -4.044e-01  5.806e-01  -0.697  0.48611
Loan.required                -1.067e-05  3.328e-06  -3.205  0.00135 **
education4 to 7 years          4.822e-01  2.973e-01  1.622  0.10486
educationless than 1 year      -7.298e-01  2.645e-01  -2.759  0.00580 **
educationmore than 7 years     -9.949e-02  2.562e-01  -0.388  0.69777
educationnot employed         -9.254e-02  4.476e-01  -0.207  0.83621
m.statusdivorced or separated or married female  3.510e-03  4.395e-01  0.008  0.99363
m.statusmarried or widowed male  4.263e-01  5.157e-01  0.827  0.40847
m.statussingle male           3.795e-01  4.322e-01  0.878  0.37997
Opartiesyes, co-applicant      -7.032e-01  3.990e-01  -1.763  0.07798 .
Opartiesyes, guarantor         3.954e-01  4.526e-01  0.874  0.38238
Duration2 to 3 years           -9.703e-02  3.027e-01  -0.321  0.74854
Durationless than a year       2.820e-01  3.089e-01  0.913  0.36128
Durationmore than 3 years      5.816e-03  2.279e-01  0.026  0.97964
age                           1.340e-02  9.126e-03  1.469  0.14186
inPlansnone                   4.227e-01  2.569e-01  1.645  0.09988 .
inPlansstores                 7.240e-03  4.538e-01  0.016  0.98727
JobTypeemployee with official position  1.968e-02  2.908e-01  0.068  0.94603
JobTyphenon resident either unemployed or unskilled  1.516e-01  7.585e-01  0.200  0.84155
JobTyperesident unskilled      7.506e-02  3.609e-01  0.208  0.83525
Ndepend2                     1.419e-01  2.667e-01  0.532  0.59456
telephoneyes                  2.484e-01  2.076e-01  1.197  0.23148
foreignyes                    2.231e+00  8.657e-01  2.577  0.00997 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Here the observed influencing factors are 'Credit status', 'Loan required', 'education' and 'foreign'.

Generating the model from these factors:

```

> model <- glm(CreditScore~credit.status+Loan.required+education+foreign, data=train, family='binomial')
> summary(model)

Call:
glm(formula = CreditScore ~ credit.status + Loan.required + education + foreign, family = "binomial", data = train)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.2730  -0.9969   0.6298   0.7959   2.0179

Coefficients:
(Intercept)                -2.300e-01  3.858e-01  -0.596  0.55110
credit.statusall settled till now  1.470e+00  3.674e-01  4.000  6.32e-05 ***
credit.statusdues not paid earlier  2.008e+00  3.962e-01  5.068  4.01e-07 ***
credit.statusnone taken/all settled -2.872e-01  5.664e-01  -0.507  0.61215
Loan.required                -9.699e-06  2.965e-06  -3.272  0.00107 **
education4 to 7 years          5.278e-01  2.826e-01  1.868  0.06180 .
educationless than 1 year      -7.702e-01  2.442e-01  -3.154  0.00161 **
educationmore than 7 years     2.099e-02  2.287e-01  0.092  0.92689
educationnot employed         -8.058e-02  3.675e-01  -0.219  0.82645
foreignyes                    2.113e+00  8.125e-01  2.601  0.00929 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

And using the trained model for prediction:

```
> pre <- predict(model, train, type='response')
> head(pre)
```

| | 1 | 3 | 4 | 5 | 6 | 7 |
|--|-----------|-----------|-----------|-----------|-----------|-----------|
| | 0.7514157 | 0.3352735 | 0.1284502 | 0.5405072 | 0.6810692 | 0.5647218 |

And the credit score of the dataset:

| creditScore |
|-------------|
| good |
| bad |
| bad |
| good |
| good |
| good |
| good |

So, the model predicted the credit scores successfully.

INFERENCE:

Hence, keeping various factors in logistic formula, we test if a user is worthy of giving credit or not .

RESULT:

Hence all the given tasks have been successfully executed using RStudio.