#Classify whether application accepted or not using Logistic regression

#Problem Statement for Credit Card Data

**Credit Code Output:-**

> sum(is.na(Credit)) # To get the count of NA Values

[1] 0

> str(Credit)

'data.frame': 1319 obs. of 13 variables:

$ Sr.No. : int 1 2 3 4 5 6 7 8 9 10 ...

$ card : chr "yes" "yes" "yes" "yes" ...

$ reports : int 0 0 0 0 0 0 0 0 0 0 ...

$ age : num 37.7 33.2 33.7 30.5 32.2 ...

$ income : num 4.52 2.42 4.5 2.54 9.79 ...

$ share : num 0.03327 0.00522 0.00416 0.06521 0.06705 ...

$ expenditure: num 124.98 9.85 15 137.87 546.5 ...

$ owner : chr "yes" "no" "yes" "no" ...

$ selfemp : chr "no" "no" "no" "no" ...

$ dependents : int 3 3 4 0 2 0 2 0 0 0 ...

$ months : int 54 34 58 25 64 54 7 77 97 65 ...

$ majorcards : int 1 1 1 1 1 1 1 1 1 1 ...

$ active : int 12 13 5 7 5 1 5 3 6 18 ...

> summary(logit)

Call:

glm(formula = factor(card) ~ ., family = binomial, data = Credit)

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) 1.938e+14 8.850e+06 21898781 <2e-16 \*\*\*

Sr.No. 8.203e+10 4.908e+03 16712515 <2e-16 \*\*\*

reports -5.265e+14 1.437e+06 -366540548 <2e-16 \*\*\*

age 3.978e+12 2.211e+05 17992987 <2e-16 \*\*\*

income -4.678e+12 1.472e+06 -3178985 <2e-16 \*\*\*

share 1.764e+16 4.362e+07 404458593 <2e-16 \*\*\*

expenditure -6.835e+11 1.568e+04 -43584653 <2e-16 \*\*\*

owneryes -8.506e+13 4.361e+06 -19501393 <2e-16 \*\*\*

selfempyes 1.665e+14 7.375e+06 22571127 <2e-16 \*\*\*

dependents -1.451e+13 1.621e+06 -8948607 <2e-16 \*\*\*

months -4.196e+11 3.145e+04 -13342062 <2e-16 \*\*\*

majorcards 8.479e+13 4.858e+06 17453988 <2e-16 \*\*\*

active 1.069e+13 3.175e+05 33660613 <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: 11606.1 on 1306 degrees of freedom

AIC: 11632

Number of Fisher Scoring iterations: 25

|  |
| --- |
| > exp(coef(logit))  (Intercept) Sr.No. reports age income share expenditure owneryes selfempyes dependents  Inf Inf 0 Inf 0 Inf 0 0 Inf 0  months majorcards active  0 Inf Inf  > table(Credit$card)  no yes  296 1023 |
|  |
| |  | | --- | | > table(probo)  probo  FALSE TRUE  147 1172  > confusion    no yes  FALSE 141 6  TRUE 155 1017  > Accuracy  [1] 0.8779378  > Error  [1] 0.1220622 | |

ROC Curve-

