Lab-2

Task-1: Loan repay prediction using bank data.

1. Download bank dataset-

<https://drive.google.com/open?id=1vHVyDJz6AJFg2lCLjNaijdPGbiSh5DWu>

1. Develop a binomial logistic regression using single attribute “age”

Observe all parameters (Null Deviance, Residual deviance, P value, Fisher score, AIC , Degree of freedom)

1. Develop a binomial logistic regression using two attribute “age + salary”

Observe changes in all parameters (Null Deviance, Residual deviance, P value, Fisher score, AIC , Degree of freedom)

Task2-

Loan repay prediction using detailed bank loan data.

1. Download bank dataset-

<https://drive.google.com/open?id=1_XR6HytlYi0p2rUuplO6QuA8L2SByI8d>

1. Develop a final (as per your criteria and knowledge) binomial logistic regression model.

Hint- You may use wither forward or backward selection process to get acceptable number of features. Meanwhile, see AIC/deviance/pvalue to check the significance of features.

1. Think- Normalization of variable will improve the performance (if yes, deduce reasoning )

Task3-

1. Download bank dataset-

<https://drive.google.com/open?id=1N7raOWmOSyJPhZtzulVxNx8aYYqOjPAX>

1. Develop a binomial logistic regression using attributes BP + weight

Observe all parameters (Null Deviance, Residual deviance, P value, Fisher score, AIC , Degree of freedom)

1. Develop a binomial logistic regression using single attribute “BP”

Observe changes in all parameters (Null Deviance, Residual deviance, P value, Fisher score, AIC , Degree of freedom)

1. Deduce the conclusion from point 2 and 3
2. Think to improve the model…

Hint- Polynomial logistic regression

Hint- to confirm it – see the data distribution (plot(weight,BP,col=Suffered.Heart.storke+1)).

1. One of many solutions can be-

BPS = BP\*BP

weights = weight\*weight

BPweight = BP\*weight

m1=glm(Suffered.Heart.storke~BP+weight+weights+BPS+BPweight,family=binomial)

m1=glm(Suffered.Heart.storke~BP+weight+weights+BPS,family=binomial)

Task4-

1. Download bank dataset-

<https://drive.google.com/open?id=16n-ok41MPUDNehxvwLajYQtpHCqN60e0>

1. Develop a final (as per your criteria and knowledge) binomial logistic regression model.
2. Calculate log likelihood of your model
3. Calculate p-value of model (using pchisq)
4. Train to nested models and measure lr test.
5. Calculate the p value using lr test and it should be equal to as point 4.
6. Separate the in train and test (70:30)
7. Train your model in train dataset.
8. Display the probability of each observations of train dataset.
9. Display confusion table and calculate all dependent measures (accuracy, recall, ……
10. Test your model on test dataset