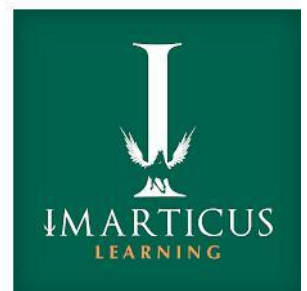


Identifying Good & Bad Customers for Granting Credit

This dataset classifies people described by a set of attributes as good or bad credit risks

You will need to apply Decision Tree to build model to predict which customer is likely to be classified as bad or good. You will use “Creditability “as your response variable while rest of the variables as your Predictor variables.

So, while you build a classification model, you should not only focus on accuracy but also keep a tab on False Positive & False Negative as well.



Evaluation

Evaluation will be based on:

- Data Exploration & Preparation (25%)
- Feature Engineering (20%)
- Model Comparison (20%)
- Model Selection (25%)
- Feature Importance (10%)



Data Preparation

Perform Missing Value Analysis & Treatment. Treat Outliers accordingly. Treat Multicollinearity as applicable.

Feature Engineering

Apply feature engineering techniques to see how new features can be created to improve the model.

Model Comparison

Build multiple Logistic Regression models and compare model evaluation metrics like ROC AUC, Accuracy, Precision Score, Recall Score and Kappa Score

Model Selection

Select the best model and specify the reason for your selection



Attribute Information:

1. Attribute 1: Status_of_existing_account (qualitative)
Status of existing checking account
A11 : ... < 0 DM
A12 : $0 \leq \dots < 200$ DM
A13 : ... ≥ 200 DM / salary assignments for at least 1 year
A14 : no checking account
2. Attribute 2: Duration_of_Credit_month (numerical)
Duration in month
3. Attribute 3: Payment_Status_of_Previous_Credit (Credit_history) (qualitative)
Credit history
A30 : no credits taken/all credits paid back duly
A31 : all credits at this bank paid back duly
A32 : existing credits paid back duly till now
A33 : delay in paying off in the past
A34 : critical account/ other credits existing (not at this bank)
4. Attribute 4: Purpose_of_loan (qualitative)
Purpose
A40 : car (new)
A41 : car (used)
A42 : furniture/equipment
A43 : radio/television
A44 : domestic appliances
A45 : repairs
A46 : education
A47 : (vacation - does not exist?)
A48 : retraining
A49 : business
A410: others

Attribute Information (Continued):

5. Attribute 5: Credit_Amount numerical)

Credit amount

6. Attribute 6: Value_of_Savings_account&bonds (qualitative)

Savings account/bonds

A61 : ... < 100 DM

A62 : 100 <= ... < 500 DM

A63 : 500 <= ... < 1000 DM

A64 : .. >= 1000 DM

A65 : unknown/ no savings account

7. Attribute 7: Years_of_Present_Employment (qualitative)

Present employment since

A71 : unemployed

A72 : ... < 1 year

A73 : 1 <= ... < 4 years

A74 : 4 <= ... < 7 years

A75 : .. >= 7 years

8. Attribute 8: Percentage_of_disposable_income (numerical)

Installment rate in percentage of disposable income

9. Attribute 9: Sex_&_Marital_Status (qualitative)

Personal status and sex

A91 : male : divorced/separated

A92 : female : divorced/separated/married

A93 : male : single

A94 : male : married/widowed

A95 : female : single

Attribute Information (Continued):

10. Attribute 10: Guarantors/Debtors (qualitative)

Other debtors / guarantors

A101 : none

A102 : co-applicant

A103 : guarantor

11. Attribute 11: Duration_in_Present_Residence (numerical)

Present residence since

12. Attribute 12: Property (qualitative)

Property

A121 : real estate

A122 : if not A121 : building society savings agreement/life insurance

A123 : if not A121/A122 : car or other, not in attribute 6

A124 : unknown / no property

13. Attribute 13: Age_in_years (numerical)

Age in years

14. Attribute 14: Concurrent_Credits (qualitative)

Other installment plans

A141 : bank

A142 : stores

A143 : none

15. Attribute 15: Housing (qualitative)

Housing

A151 : rent

A152 : own

A153 : for free

Attribute Information (Continued):

16. Attribute 16: No_of_Credits_at_this__Bank (numerical)

Number of existing credits at this bank

17. Attribute 17: Occupation (qualitative)

Job

A171 : unemployed/ unskilled - non-resident

A172 : unskilled - resident

A173 : skilled employee / official

A174 : management/ self-employed/ highly qualified employee/
officer

18. Attribute 18: No_of_dependents (numerical)

Number of people being liable to provide maintenance

19. Attribute 19: Telephone (qualitative)

Telephone

A191 : none

A192 : yes, registered under the customer's name

20. Attribute 20: Foreign_Worker (qualitative)

foreign worker

A201 : yes

A202 : no