



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 Logistic Regression (using Statsmodel Library) 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.

It is worse to class a customer as good when they are bad, than it is to class a customer as bad when they are good.

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 (20%)
- Feature Selection & Engineering (20%)
- Model Comparison (35%)
- Model Selection (10%)
- Presentation (15%)

Data Preparation.

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

Feature Selection.

Select the right features based on importance and significance.

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 Pseudo R-squared, ROC AUC, Accuracy etc.

Model Selection.

Select the best model and specify the reason for your selection



Attribute Information:

Attribute 1: (qualitative)

Status of existing checking account

A11 : ... < 0 DM

A12 : 0 <= ... < 200 DM

A13 : ... >= 200 DM / salary assignments for at least 1 year

A14 : no checking account

Attribute 2: (numerical)

Duration in month

Attribute 3: (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)

Attribute 4: (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):

Attribute 5: (numerical)
Credit amount

Attribute 6: (qualitative)
Savings account/bonds
A61 : ... < 100 DM
A62 : 100 <= ... < 500 DM
A63 : 500 <= ... < 1000 DM
A64 : .. >= 1000 DM
A65 : unknown/ no savings account

Attribute 7: (qualitative)
Present employment since
A71 : unemployed
A72 : ... < 1 year
A73 : 1 <= ... < 4 years
A74 : 4 <= ... < 7 years
A75 : .. >= 7 years

Attribute 8: (numerical)
Installment rate in percentage of disposable income

Attribute 9: (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):

Attribute 10: (qualitative)
Other debtors / guarantors

A101 : none

A102 : co-applicant

A103 : guarantor

Attribute 11: (numerical)
Present residence since

Attribute 12: (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

Attribute 13: (numerical)
Age in years

Attribute 14: (qualitative)
Other installment plans

A141 : bank

A142 : stores

A143 : none

Attribute 15: (qualitative)
Housing

A151 : rent

A152 : own

A153 : for free



Attribute Information (Continued):

Attribute 16: (numerical)

Number of existing credits at this bank

Attribute 17: (qualitative)

Job

A171 : unemployed/ unskilled - non-resident

A172 : unskilled - resident

A173 : skilled employee / official

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

Attribute 18: (numerical)

Number of people being liable to provide maintenance for

Attribute 19: (qualitative)

Telephone

A191 : none

A192 : yes, registered under the customers name

Attribute 20: (qualitative)

foreign worker

A201 : yes

A202 : no