Problem Statement - German Credit Risk Analysis
Case Study: German Credit Analysis

Context:

To minimize loss from the bank's perspective, the bank needs a decision rule regarding whom to approve the loan and whom not to. An applicant's demographic and socio-economic profiles are considered by loan managers before a decision is taken regarding his/her loan application.

In this dataset, each entry represents a person who takes credit from a bank. Each person is classified as a good or bad credit risk according to the set of attributes.

Objective:

The objective is to build a predictive model on this data to help the bank take a decision on whether to approve a loan to a prospective applicant.

Attribute Information:

- The data contains characteristics of the people

  ◦ Age (Numeric: Age in years)
  ◦ Sex (Categories: male, female)
  ◦ Job (Categories: unskilled and non-resident, unskilled and resident, skilled, highly skilled)
  ◦ Housing (Categories: own, rent, or free)
  ◦ Saving accounts (Categories: little, moderate, quite rich, rich)
  ◦ Checking account (Categories: little, moderate, rich)
  ◦ Credit amount (Numeric: Amount of credit in DM - Deutsche Mark)
  ◦ Duration (Numeric: Duration for which the credit is given in months)

- Purpose (Categories: car, furniture/equipment, radio/TV, domestic appliances, repairs, education, business, vacation/others)
- Risk (0 - Person is not at risk, 1 - Person is at risk(defaulter))

Learning Outcomes:

- Exploratory Data Analysis
- Preparing the data to train a model
- Training and understanding of data using a logistic regression model
- Model evaluation

Steps and Tasks:

- Import Libraries and Load Dataset
- Overview of data
- Data Visualization
- Data preparation
- Choose Model, Train, and Evaluate
- Conclusion