Case Study: German Credit Analysis

Context:

When a bank receives a loan application, based on the applicant's profile the bank has to decide whether to go ahead with the loan approval or not. Two types of risks are associated with the bank's decision –

If the applicant is a good credit risk, i.e. is likely to repay the loan, then not approving the loan to the person results in a loss of business to the bank
If the applicant is a bad credit risk, i.e. is not likely to repay the loan, then approving the loan to the person results in a financial loss to the bank

To minimize this loss HRE bank wants to automate this process using a predictive model, that will predict if a customer is at risk of making a default or not based on the customer's demographic and socio-economic profiles

You as a Data scientist at HRE bank has been assigned the work of building a predictive model that will predict if a customer is at risk of default or not

Problem:

The data-set aims to answer the following key questions:

What are the key features leading a person to default?
To predict if a customer is going to default or not?

Attribute Information:

The data contains the following characteristics of the people

Age (Numeric: Age in years)
Sex (Categories: male, female)
Job (Categories : 0 - unskilled and non-resident, 1 - unskilled and resident, 2 - skilled, 3 - 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

Pipelines

Hyperparameter tuning

Steps and Tasks:

Import Libraries and Load Dataset

Overview of data

Data Visualization

Data preparation

Choose model

Tune the model

Conclusion