Group Number: 59 Group Members:

● Chentao Jin (Team Lead) - cjin082@uottawa.ca

● Andy Huang

● Amine El Goussi

Dataset:

https://archive.ics.uci.edu/dataset/144/statlog+german+credit+data

Description

The project aims to assess an individual's credit risk using the Statlog German Credit Data. This dataset encompasses 1000 cases, each with 20 distinct features, including categorical and numerical data. It covers aspects like the status of existing checking accounts, credit history, loan purpose, credit amount, savings, and demographic information. Each case is precisely classified into one of two categories: 'good' or 'bad' credit risk. This classification serves as the foundation for our analysis and predictive modeling, focusing on determining the loan amount, rate, duration (numerical attribute contained in fact table) base on other text attributes including the binary creditworthiness, property, age, housing… of individuals based on the provided dataset features.

Motivation

The project is motivated by the significance of credit scoring in the financial world. Credit scoring is a critical tool banks and other financial institutions use to evaluate the risk of lending money to individuals. It impacts consumer access to loans and the terms under which credit is granted. By studying this dataset, we aim to understand how various factors contribute to credit risk and explore the complexities of credit risk assessment.