Architecture

Credit Card Default Prediction

Abstract

The facilities provided by banks are numerous and help us in different areas of financial needs. It involves granting of loans, mortgage, credit card etc. Credit card is one of the most common utility everyone has today and has been booming for many years. There are also many risks involved with the same. As the number of consumers have been increasing exponentially, the danger of credit card defaulting also has been on the rise. Understanding the characteristics of consumers and their banking details can shed some light on the possibility of them defaulting. This project provides a method to predict credit card defaulters using machine learning models like Logistic regression, Random Forest, Decision tree, SVM and XGB.

Table of contents

1. Introduction	1
1.1 Scope	
2. Technical specifications	
2.1 Dataset	
2.2 Prediction	
2.3 Logging	
2.4 Deployment	
3. Proposed Solution	
4. Model Training workflow	

1. Introduction

The purpose of this document is to present a detailed description of the credit card default prediction system. It covers the purposes and features of the system and how the system performs the actions

The objective of the project is to predict if the person is a credit defaulter or not based on the person's characteristics and payment history

Predicting if a person will default in their credit card payments depends upon a variety of factors like the balance in their account, age, marital status, gender, education, their repayment status for each month for 6 months, their total bill amount and the amount they have paid back.

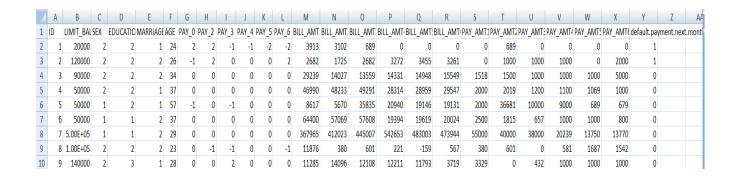
1.1 Scope

The credit card default prediction system is a machine learning based model that will help us to detect credit card defaulters. Early detection of defaulters will be key to proper and safe bank management. It is able to make predictions given a number of independent variables like age, education, marital status, repayment status etc

2. Technical specifications

2.1 Dataset

The dataset used for the study consists of 3000 rows of data and 25 columns which relate to data related to each customer in the bank



2.2 Prediction

The system prompts the user to input the data related to the customer. After the information has been loaded the system will output if the person is a defaulter or not based on the input provided

2.3 Logging

The system logs all the steps so that the user will be aware of the processes working internally. The type of logging chosen for each depends upon the type of operation that is carried out like info, critical, debug or warning logging. Logging helps to debug issues, so it is a mandatory step

2.4 Deployment

The machine learning model will be deployed as an API using FastAPI

3. Proposed Solution

The proposed solution is to build various machine learning models like Logistic regression, Random Forest, Decision tree, SVM and XGB and find which model is able to best predict the credit card defaulter given a set of training data.

4. Model Training workflow

The different models used in the project will be used to detect if a consumer is a credit card defaulter or not based on the data used in the training stage. The goal is to identify the model which is able to make predictions with the highest accuracy so that defaulters can be indentified easily.

