HIGH LEVEL DESIGN

Credit Card Default Prediction

1. Introduction

a. Why this High-Level Design Document?

The purpose of this High-Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding, and can be used as a reference manual for how the modules interact at a high level.

The HLD will:

- Present all of the design aspects and define them in detail
- Describe the user interface being implemented
- Describe the hardware and software interfaces
- Describe the performance requirements
- Include design features and the architecture of the project
- List and describe the non-functional attributes like:
    - Security
    - Reliability
    - Maintainability
    - Portability
    - Reusability
    - Application Compatibility
    - Resource Utilization
    - Serviceability

1.2 Scope

The HLD documentation presents the structure of the system, such as the database architecture, application architecture (layers), application flow (Navigation), and technology architecture. The HLD uses non-technical to mildly-technical terms which should be understandable to the administrators of the system.

1.3 Definitions

Database - Collection of the information monitored by the system.

AWS - Amazon Web Service

2. Database Information

Our dataset 'Default of Credit Card Clients Dataset' consists of informations about transactions from April 2005 to September 2005 of 30000 clients who were credit holders in a bank in Taiwan. This dataset has binary response variable 'default.payment.next.month' that takes the value 1 if the corresponding client has default payment and 0 otherwise. Out of 30000 clients 6636(22.12%) were with default payment. There are 23 other independent or explanatory variables:

- ID: ID of each client
- LIMIT_BAL: Amount of given credit in NT dollars (includes individual and family/supplementary = credit)
- SEX: Gender (1=male, 2=female)
- EDUCATION: (1=graduate school, 2=university, 3=high school, 4=others, 5=unknown, 6=unknown) MARRIAGE: Marital status (1=married, 2=single, 3=others)
- AGE: Age in years
- PAY_1: Repayment status in September 2005 (-1=pay duly, 1=payment delay for one month, 2=payment delay for two months, … 8=payment delay for eight months,9=payment delay for nine months and above)
- PAY_2: Repayment status in August 2005 (scale same as above)
- PAY_3: Repayment status in July 2005 (scale same as above)
- PAY_4: Repayment status in June 2005 (scale same as above)
- PAY_5: Repayment status in May 2005 (scale same as above)
- PAY_6: Repayment status in April 2005 (scale same as above)
- BILL_AMT1: Amount of bill statement in September 2005 (NT dollar)
- BILL_AMT2: Amount of bill statement in August 2005 (NT dollar)
- BILL_AMT3: Amount of bill statement in July 2005 (NT dollar)
- BILL_AMT4: Amount of bill statement in June 2005 (NT dollar)
- BILL_AMT5: Amount of bill statement in May 2005 (NT dollar)
- BILL_AMT6: Amount of bill statement in April 2005 (NT dollar)
- PAY_AMT1: Amount of previous payment in September 2005 (NT dollar)
- PAY_AMT2: Amount of previous payment in August 2005 (NT dollar)
- PAY_AMT3: Amount of previous payment in July 2005 (NT dollar)

3. Proposed Solution

The solution proposed here is a Credit card default Prediction model can be implemented to perform the above mentioned use case. In this case, we have to enter the last six months bill amounts , paid amounts, payment status, credit limit, personal details. Based on the above details, the model predicts whether customers will default the credit card payment or not.

4. Technical Aspect

This project is divided into three part: