**High Level Design(HLD)**

Credit Card Default Prediction



Revision Number :

Last Date of Revision :

|  |  |  |  |
| --- | --- | --- | --- |
| Date Issued | Version | Description | Author |
|  | 1 | Abstract,  Introduction,  General Description | D.Jagannath |
|  | 2 | Design Detail ,KPI,  Deployment | D.Jagannath |
|  | 3 | Final Revision | D.Jagannath |

Contents

Document Version Control…………………………………………………………

Abstract…………………………………………………………………………………………………-4

1.Introduction ………………………………………………………………………..-5

1.1- Why this High Level Design Document ?.............................-5

1.2- Scope……………………………………………………………………………….-5

1.3- Definations……………………………………………………………-6

2.General Description………………………………………………………………..-6

2.1-Product perspective……………………………………………………………………..-6

2.2-Problem Statement……………………………………………………………………….-6

2.3-Proposed Solution…………………………………………………………………………-6

2.4-Tools Used……………………………………………………………………………………-7

3.Design Details…………………………………………………………………………..-8

3.1-Process Flow ……………………………………………………………………………………-8

3.1.1-Model Training and Evaluation………………………………………………………..-9

3.2- Event Log……………………………………………………………………………………..-10

3.3-Error Handling……………………………………………………………………………….-10

4.Performance………………………………………………………………………….-10

4.1-Reusabilty………………………………………………………………………………………-10

4.2- Application Compatibility…………………………………………………………….-10

4.3- Deployment………………………………………………………………………………..-11

5.KPIs(Key Performance Indicators)…………………………………………-12

6.Conclusion…………………………………………………………………-12

Abstract

Credit risk plays a major role in the banking industry business. Banks' main activities involve granting loan, credit card, investment, mortgage, and others. Credit card has been one of the most booming financial services by banks over the past years. However, with the growing number of credit card users, banks have been facing an escalating credit card default rate. As such data analytics can provide solutions to tackle the current phenomenon and management credit risks. This paper provides a performance evaluation of credit card default prediction. Thus, logistic regression, rpart decision tree, and random forest are used to test the variable in predicting credit default and random forest proved to have the higher accuracy and area under the curve. This result shows that random forest best describe which factors should be considered with an accuracy of 82 % and an Area under Curve of 77 % when assessing the credit risk of credit card customers.

Introduction

* 1. 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 o Maintainability

* + - Portability
    - Reusability

Application compatibility o 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. Definitions :

|  |  |
| --- | --- |
| Database | Collection of all the information monitored by this system |
| IDE | Integrated Development Environment |
| AWS | Amazon Web Services |

2.General Description:

2.1 Product Perspective:

Credit card default payment prediction solution is a Machine Learning based model which will help us in prediction of further default payments and to take decision about new clients to give approval for new clients to sanction the loans by use of Machine learning model .

2.2 Problem Statement :

Financial threats are displaying a trend about the credit risk of commercial banks as the incredible improvement in the financial industry has arisen. In this way, one of the biggest threats faces by commercial banks is the risk prediction of credit clients. The goal is to predict the probability of credit default based on credit card owner's characteristics and payment history.

2.3 PROPOSED SOLUTION:

The Proposed Solution is Credit Card Default Payment based on default payments prediction can be implemented to perform the operations of finding solutions of remaining default payment predictions and these transactions will be recorded and those scores will be helpful in further loan approvals. The model verify the details of the clients based on age, sex, martial status, income, previous loan statements, bank statements, monthly transactions will be considered these details and verify the clients status is eligible for the loan approvals to avoid the financial risk involving in the payments.

2.4. Tools Used

Python programming language and frameworks such as NumPy, Pandas, Scikit-learn, Matplotlib and Seaborn are used to build the whole model.

* + Jupyter Notebook is used as IDE.
  + For visualization of the plots, Matplotlib, Seaborn and Plotly are used.
  + AWS is used for deployment of the model.

Cassandra DB is used to retrieve, insert, delete, and update the database. Front end development is done using HTML/CSS . Python Django is used for backend development. GitHub is used as version control system.

3.Design Details

3.1-Process Flow : For Finding Next EMI’s of Credit Card Default payments, we will use a Machine Learning based model as shown in the below figure.

Proposed Methodology :

Dataset from UCI Library /Data Bases ets

Exploratory Data Analysis

Splitting Dataset into

Train and Test Datasets

Selecting Best Suitable Algorithm

Prediction

3.1.1- Modal Training and Evaluation :

Data Splitting Process

Data Pre-processing

Dataset

Evaluation Process

Deployment Process

Prediction

3.2- Event Log :

The system should log every event so that the user will know what process is running internally

Initial Step-By-Step Description:

1. The System identifies at what step logging required
2. The System should be able to each and every system flow
3. Developer can choose logging method. You choose database logging / File logging as well
4. System should not hang even after using so many logging. Logging just because we can easily debug issues so logging is mandatory to do

3.3-Error Handling :

Should errors be encountered, an explanation as to what went wrong . An error as anything that falls outside normal and intended usage.

4.Performance :

The Machine Learning Credit Card Default Prediction solution will be used for to know whether a person able to pay his payment or not by using his past 6 months data . So that Finance company will take necessary action.

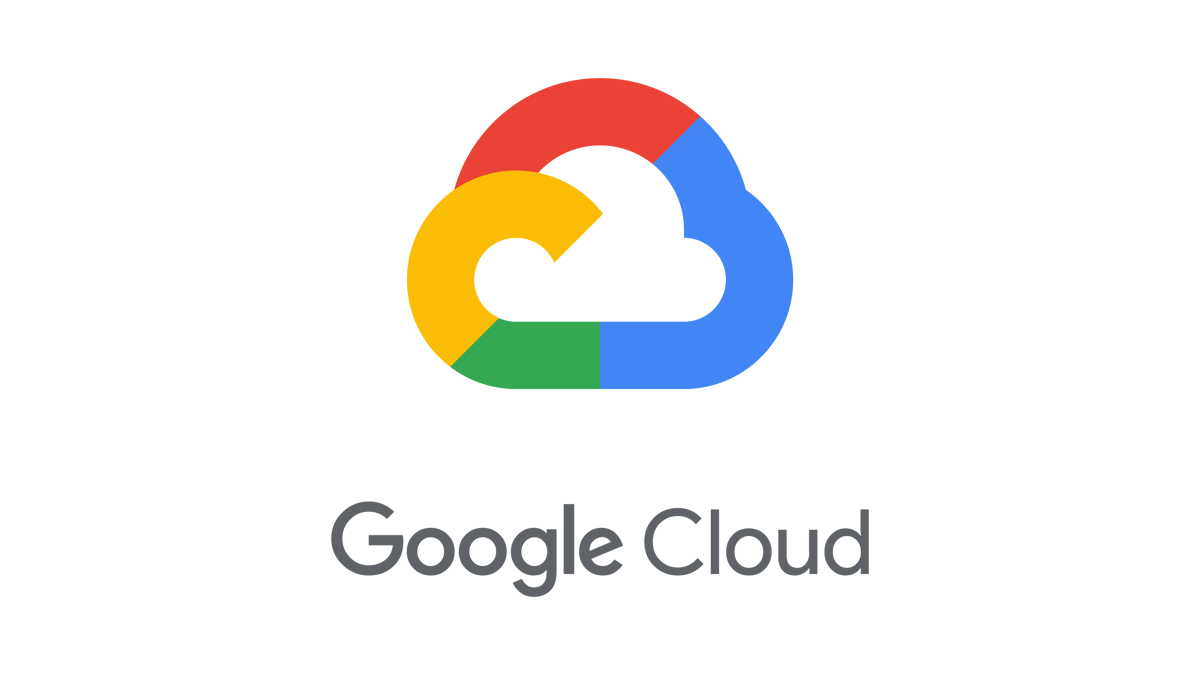
4.1-Reusabilty :

The code written and the components used should have the ability to be reused with no problems

4.2- Application Compatibility :

The different components for this project will be using python as an interface between them. Each component will have its own task to perform, and it is the job of the Python to ensure proper transfer of information.

4.3- Deployment :

5.KPIs(Key Performance Indicators) :

1. Key Indicators displaying a summary of Credit Card Default Prediction.
2. Time and Workload reduction using Credit Card Default Prediction.
3. To detect persons who are not able to pay their payment by using their past 6 months data and help to take necessary actions.
4. It gives persons previous transaction details and credit card score.

6.Conclusion :

Credit Card Default Prediction is takes credit score, via certain score mapping . The model is effective in predicting users willingness to repay credit card dept by using their past 6 months data . The users (his/her) credit card rating can be directly expressed via the credit score . The higher the score, the better credit rating, the lower the score, the lower credit rating. With help of model companies will find out about peoples willingness of repay their debt.