

CREDIT CARD DEFAULT

PREDICATION

High Level Design



Dnyanesh Yeole

AUGUST 27, 2023

INEURON

# CONTENTS

[**CONTENTS 2**](#_g1pelswuyn44)

[**1. INTRODUCTION 3**](#_x1866x1zw0m8)

[1.1 OVERVIEW 3](#_y3pza317sgwu)

[1.2 OBJECTIVES 3](#_nh00hj2rg6u1)

[1.3 PROBLEM STATEMENT 3](#_o8e8mwbautt4)

[**2. METHODOLOGY 4**](#_sf7fhmggq6v2)

[2.1 DATA COLLECTION 4](#_pke6nf19m9s7)

[2.2 DATASET DESCRIPTION 4](#_3v2k7rs2k7oz)

[**3. DESIGN DETAILS 5**](#_yt3e1hv1ttt0)

[3.1 Tools Used 5](#_dbtxzalb6aq8)

[3.2 Process Flow 5](#_6gsctuyevn5v)

[3.2 Deployment Process 6](#_lbul5b2grwj2)

[**4. CONCLUSION 7**](#_lf8ul5gv4839)

# 

**CHAPTER-1**

# 1. INTRODUCTION

## 1.1 OVERVIEW

In the dynamic landscape of banking and financial institutions, their role in offering crucial financial services is indispensable. Maintaining the integrity of these institutions is paramount, and one key aspect is ensuring judicious investment in customers to mitigate potential financial losses. This overview delves into the significance of credit scoring in this process and how it establishes a crucial link between borrowers' characteristics and default risk

## 1.2 OBJECTIVES

* **Risk Mitigation**: Develop accurate predictive models to identify high-risk borrowers, thereby reducing instances of credit card defaults and minimizing financial losses for the institution.
* **Informed Decision**-Making: Provide financial institutions with actionable insights based on data-driven analysis, enabling well-informed lending decisions that optimize credit management strategies.
* **Customer Trust Building**: Demonstrate responsible lending practices through fair and transparent credit assessments, fostering trust and credibility between financial institutions and customers.
* **Operational Efficiency**: Streamline credit approval processes by automating the evaluation of borrower profiles, improving operational efficiency and resource utilization.
* **Regulatory Compliance**: Ensure that the credit card default prediction process adheres to regulatory standards and industry guidelines, promoting compliance and avoiding legal and reputational risks.

## 1.3 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.

**CHAPTER 2**

# 2. METHODOLOGY

## 2.1 DATA COLLECTION

The UCI Credit dataset is a widely used collection of data designed to facilitate research and analysis related to credit risk assessment and prediction. This dataset is specifically curated to aid in the development and evaluation of predictive models for credit card default.The link is given below:  
<https://www.kaggle.com/datasets/uciml/default-of-credit-card-clients-dataset>

## 2.2 DATASET DESCRIPTION

This dataset consists of 30000 observations, each representing an individual credit card account holder. The dataset encompasses several key attributes that capture various aspects of the account holders' financial characteristics and behaviors.

* 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\_0: 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)
* PAY\_AMT4: Amount of previous payment in June, 2005 (NT dollar)
* PAY\_AMT5: Amount of previous payment in May, 2005 (NT dollar)
* PAY\_AMT6: Amount of previous payment in April, 2005 (NT dollar)
* default.payment.next.month: Default payment (1=yes, 0=no)

**CHAPTER 3**

# 3. DESIGN DETAILS

## 3.1 TOOLS USED

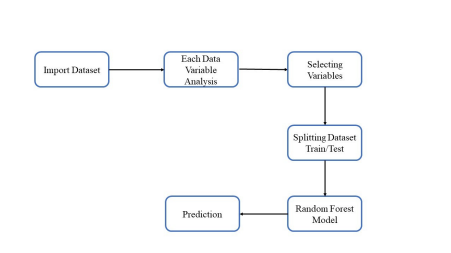
Python programming language and frameworks such as NumPy, Pandas, Scikit-learn, Matplotlib, Seaborn are used to build the whole model.





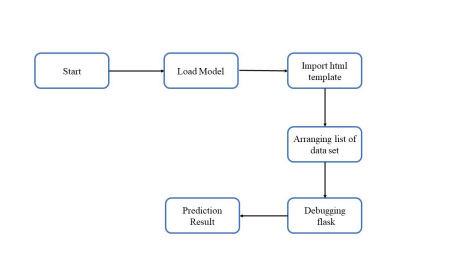


## 3.2 PROCESS FLOW

The following sequence of steps were performed while training the model for the data

## 3.2 DEPLOYMENT PROCESS

The Machine Learning model is deployed using Streamlit. It is a user-friendly Python library that allows you to create interactive and web-based data applications with minimal effort.



**CHAPTER 4**

# 4. CONCLUSION

The project is designed in Streamlit; hence it is accessible to everyone. The above designing process will help banks and loan lenders predict whether customers will default the credit card payment or not, so the bank or respective departments can take necessary action, based on the model's predictions. The UI is made to be user-friendly so that the user will not need much knowledge of any tools but will just need the information for results.