A Mini Project Report on

LOAN PREDICTION SYSTEM

T.E. - I.T Engineering

Submitted By

Om Chavan **20104056**

Sampada Mahadik **20104092**

Parthavi Khatu **20104108**

Under The Guidance Of

Prof. Geetanjali Kalme

DEPARTMENT OF INFORMATION TECHNOLOGY

A.P.SHAH INSTITUTE OF TECHNOLOGY

G.B. Road, Kasarvadavali, Thane (W), Mumbai-400615

UNIVERSITY OF MUMBAI

Academic year: 2022-23

CERTIFICATE

This to certify that the Mini Project report on Loan Prediction System has been submitted by **Om Chavan (20104056)**, **Sampada Mahadik (20104092)** and **Parthavi Khatu (20104108)** who are a Bonafede students of A. P. Shah Institute of Technology, Thane, Mumbai, as a partial fulfilment of the requirement for the degree in **Information Technology**, during the academic year **2022-2023** in the satisfactory manner as per the curriculum laid down by University of Mumbai.

Prof. Geetanjali Kalme
Guide

Dr. Kiran Deshpande
HOD , Information Technology

Dr. Uttam D. Kolekar
Principal

External Examiner(s)
1.
2.

Place: A.P. Shah Institute of Technology, Thane
Date:

ACKNOWLEDGEMENT

This project would not have come to fruition without the invaluable help of our guide **Prof. Geetanjali Kalme**. Expressing gratitude towards our HOD, **Dr. Kiran Deshpande**, and the Department of Information Technology for providing us with the opportunity as well as the support required to pursue this project. We would also like to thank our teacher **Ms. Charul Singh** who gave us her valuable suggestions and ideas when we were in need of them. We would also like to thank our peers for their helpful suggestions.

ABSTRACT

In today's world, taking loans from financial institutions has become a very common phenomenon. Everyday a large number of people make application for loans, for a variety of purposes. But all these applicants are not reliable and everyone cannot be approved. Every year, we read about a number of cases where people do not repay bulk of the loan amount to the banks due to which they suffer huge losses .The risk associated with making a decision on loan approval is immense. Loan approval is a very important process for banking organizations. The system predicts loan eligibility for the loan applications. Recovery of loans is a major contributing parameter in the financial statements of a bank. It is very difficult to predict the possibility of payment of loan by the customer. In recent years many researchers worked on loan approval prediction systems. Machine Learning (ML) techniques are very useful in predicting outcomes for large amount of data. Nowadays, due to the continuous increase in demands and needs of People which includes financial, educational growth, etc. The growth of a person requires Money which leads to rendering Loans from Financial Institutions. Loan approval is a quite hectic technique that requires lots of human effort it may also give rise to human errors. In this paper, a technology-based approach is used in terms of Machine Learning Algorithms. Loan Prediction is done using Machine learning algorithms like Support Vector Machine (SVM), Gaussian Naïve Bayes, K-Nearest Neighbor(KNN), and Decision Tree(DT). The primary goal to implement this eligibility loan prediction project is to reduce human errors and to develop system which gives significant good accuracy rate to smoothen the Loan processing stage. . So, the idea of this project is to gather loan data from multiple data sources and use various machine learning algorithms on this data to extract important information. This model can be used by the organizations in making the right decision to approve or reject the loan request of the customers. Model is marginally better because it includes variables (personal attributes of customer like age, purpose, credit history, credit amount, credit duration, etc.) other than checking account information (which shows wealth of a customer) that should be taken into account to calculate the probability of default on loan correctly.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Due to Digitization, every individual is reliable on to the internet for speed and quick process. In today's era, banks are boosting their financial status by providing different kinds of loans to people who are in need. Due to necessity of needs increasing day by day people tends to get Loan from these banking institution to satisfy their requirements. Logically and Practically, banks cannot provide loans to everyone as there are only limited reserves associated with each of them. So, banks must follow some standard verified process to approve the eligibility of loan, because if the one who got his/her loan approved failed to pay back his loan it may have a direct impact on the financial status of the bank. Therefore, banks provide loans only for a limited set of people who are capable of repaying their loans. But finding out who is eligible for the loan is a much more hectic and time consuming process.

In this project, we will develop a model to predict who is eligible for a loan in order to reduce the risk associated with the decision process and to modify the typical loan approval process into a much easier one. Moreover, we will make use of previous data of loan decisions made by the company and with the help of various data mining techniques, we will develop a loan approval decision predicting model which can draw decisions for each individual based on the information provided by them. We will use a machine-learning-based KNN, SVM, Naïve Bayes(Gaussian|) and Decision Tree algorithms to train the model. This project primary goal is to develop a loan prediction model with a better accuracy rate.

A loan is the core business part of banks. The main portion the bank's profit is directly come from the profit earned from the loans. Though bank approves loan after a regress process of verification and testimonial but still there's no surety whether the chosen hopeful is the right hopeful or not. This process takes fresh time while doing it manually. We can prophesy whether that particular hopeful is safe or not and the whole process of testimonial is automated by machine literacy style. Loan Prognostic is really helpful for retainer of banks as well as for the hopeful also.

- Problem Identified :
- The major problem faced by individual was they generally used to go to bank or any type of finance company to get the eligibility or knowledge about the loans.
- Loan prediction is a very common real-life problem that every finance company faces in their lending operations.

- Solution Proposed :

- By using our Loan Prediction System user will get and proper idea about his/her loan eligibility status and information about the loans within few clicks.

- If the loan approval process is automated, it can save a lot of man hours and improve the accuracy and speed of service to the customers.

1.1 Purpose :

Loan Prediction system is a system which provides you a interface for loan approval to the applicants application of loan**.** Applicants provides the system about their personal information and according to their information system analyses the data and specifies his status of availability for loan. Applicant can be able to calculate their Simple interest rate of any amount at the feature present in the system named as simple interest rate calculator. User can get a proper idea and knowledge regarding the different types of loans as we have provided a loan information section in the system which will help the applicants to get knowledge about the current loan percentages. This all the attributes will help the applicant to know his/her loan eligibility status within few clicks by just entering few details. Previously applicants used to go to Banks and other financial institute to check all this entities but by using our loan eligibility prediction system user will save his/her time and resources simultaneously.

1.2 Problem Statement :

A loan prediction system is an essential tool for financial institutions to determine whether or not to grant a loan to a borrower. The goal of this project is to develop a machine learning model that can accurately predict whether a loan applicant is likely to repay the loan or default on it based on their financial and personal information. The dataset for this project contains information about loan applicants, such as their income, credit score, employment status, loan amount, loan term, and other relevant factors that can impact their ability to repay the loan. The data also includes whether the loan was approved or not. The loan prediction system should be scalable, easy to use, and have a high level of accuracy. The system should also be able to handle new loan applications and adjust its predictions accordingly based on new data. Overall, the goal of this project is to develop a loan prediction system that can help financial institutions make informed lending decisions, minimize the risk of default, and ensure the profitability of their lending operations.

1.3 Objectives:

- To develop user friendly interface for the customer.
- To construct a loan eligibility prediction system with less errors and high accuracy percentage.
- To develop a system which will give a basic information related to different types of financial loans.
- To construct a loan prediction system which can save a lot of man hours and improve the accuracy and speed of service to the customers.
- To develop a system which will calculate the simple interest according to the amount entered.

1.4 Scope

- Can be helpful for the customer to understand the loan eligibility status.
- Can be useful for the customer to get the accurate prediction of the loan.
- Can be useful for the customer to get the basic knowledge regarding the types of financial loans.
- Can be useful for the user to calculate the simple interest rate within few clicks.
- Can be useful for any individual as it will save a lot of time of the user to predict his/her loan approval.

CHAPTER 2

Literature Review

Sr.no	Title	Author(s)	Year	Algorithms	Limitations	Result
1	Loan prediction using Logistic Regression in Machine Learning.	S. Sreesouthry, A. Ayubkhan, M. Mohamed, Rizwan D, Prithivi Raj, Lokesh K.	2021	Min-Max normalization, Linear Regression	Imbalanced dataset was used.	Loan prediction using Logistic Regression in Machine Learning was successfully implemented.
2	Loan approval prediction.	Veeraballi Nagajyothi.	2020	KNN, Decision tree, Naïve Bayes.	Large number of attributes were prime for consideration.	Loan approval prediction was successfully implemented.
3	Predicting loan status using Logistic Regression and Binary Tree.	T Sunitha, M. Chandravallika, M Ranganayak.	2020	Logistic Regression, Binary tree.	No detailed information about dataset is mentioned in the paper i.e. no of instances used in it.	Predicting loan status using Logistic Regression and Binary Tree was successfully implemented.

CHAPTER 3

PROPOSED SYSTEM

3.1 Features and Functionality

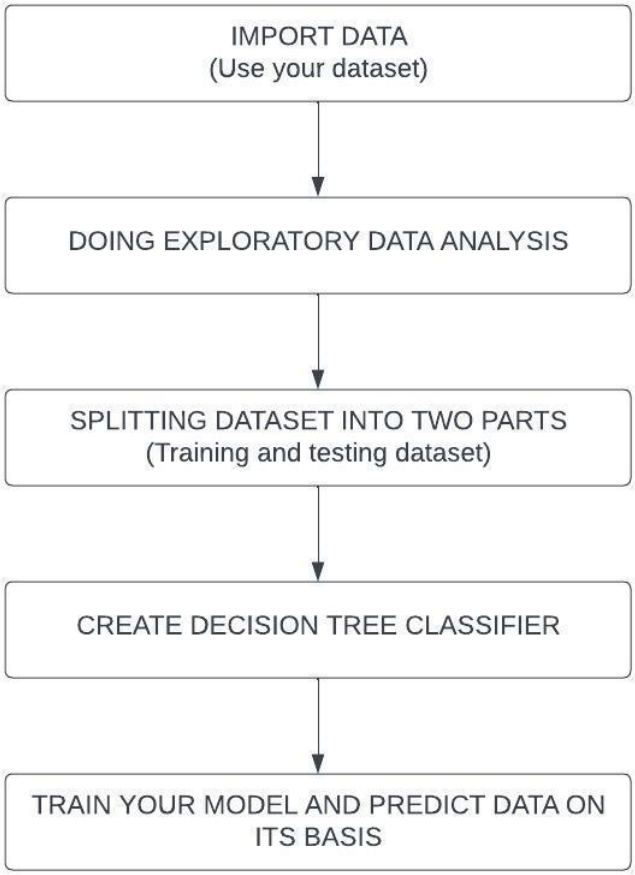
1. User-friendly and easy to Use:
   - Customer panel allows user to enter the personal details and predict the eligibility of the loan.
   - Our loan prediction system is language independent.
2. Loans information section:
   - Customer can get an basic information regarding different types of financial loans.
3. Simple interest rate calculator:
   - Customer can calculate his/her simple interest rate by entering the values within few clicks.
4. Eligibility status:
   - After entering all the detail information precisely system will analyse the data and reflect the eligibility status of the applicant.
5. Review section:
   - Customer can share their experience/opinion of prediction in the review section.

3.2 Algorithm Working with flowchart

The Decision Tree algorithm is a popular machine learning algorithm for classification and regression tasks. It is a supervised learning algorithm that can be used for both categorical and continuous input and output variables.

The algorithm works by recursively partitioning the input space into smaller and smaller regions, based on the values of input features, in a tree-like structure. At each node of the tree, a decision is made based on the value of one of the input features, and the data is split into two or more subsets, each corresponding to a different branch of the tree. The process continues recursively until a stopping criterion is met, such as a maximum tree depth or a minimum number of samples in a leaf node. To build the decision tree, the algorithm uses a greedy strategy to select the best feature at each node to split the data. The best feature is chosen based on a criterion that measures the degree of impurity or uncertainty of the data at the current node. Common impurity criteria include entropy, Gini index, and misclassification error.Once the decision tree is built, it can be used to predict the output variable for new data points. The data point is traversed down the tree from the root node to a leaf node based on the values of the input features, and the output variable is assigned the value associated with the leaf node.

Decision trees have several advantages, including being easy to interpret, handling both categorical and continuous input features, and requiring little data preparation. However, they can be prone to overfitting, especially if the tree is deep or the data is noisy. To address this issue, several techniques have been developed, such as pruning, ensemble methods (e.g. random forest), and boosting.



CHAPTER 4

REQUIREMENT ANALYSIS

Requirements analysis is an important process in software development that involves gathering and documenting the needs and constraints of stakeholders. Here are some potential requirements for a loan prediction system project:

- Data Collection: The system must be able to collect and store data related to loan applications, including personal and financial information about the applicants.
- Data Preprocessing: The system must be able to preprocess and clean the collected data to ensure that it is accurate and consistent.
- Feature Engineering: The system must be able to perform feature engineering on the preprocessed data to extract relevant features that can be used to predict loan approval or rejection.
- Model Selection: The system must be able to select an appropriate machine learning model for loan prediction based on the nature of the problem and the available data.
- Model Training and Evaluation: The system must be able to train the selected model on the preprocessed data and evaluate its performance using appropriate metrics such as accuracy, precision, recall, and F1 score.
- User Interface: The system must provide a user-friendly interface for stakeholders to interact with the system, input data, and view the prediction results.
- Security: The system must be designed with appropriate security measures to protect the sensitive personal and financial information of the loan applicants.
- Scalability: The system must be scalable to handle a large volume of loan applications and provide predictions in a timely manner.
- Maintenance and Support: The system must be maintained and supported to ensure that it remains accurate, up-to-date, and meets the changing needs of stakeholders.
- Compliance: The system must comply with applicable laws, regulations, and industry standards related to loan applications and data privacy.

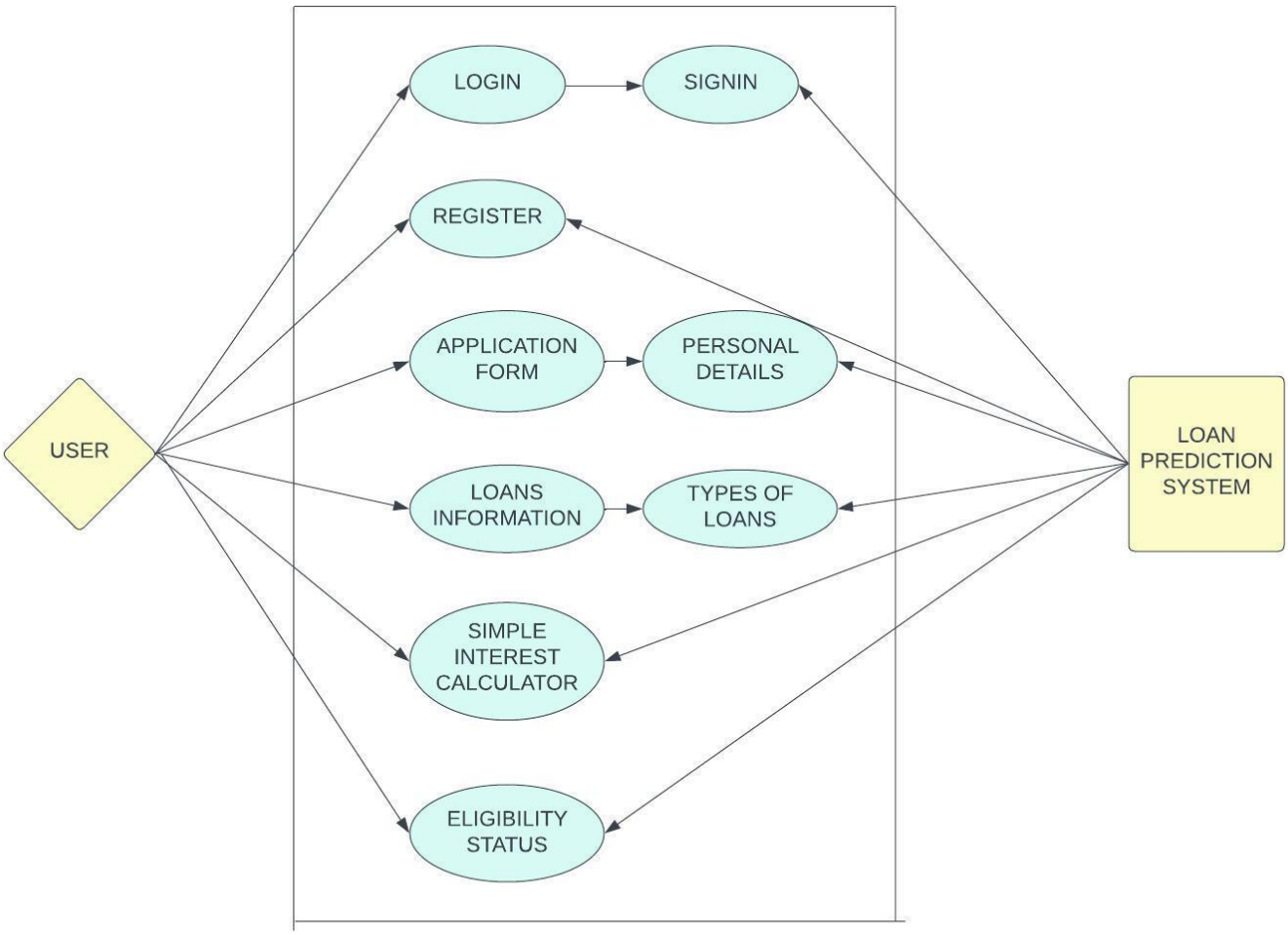
Main requirements:

The Main Requirements include Microsoft Visual Studio code, Jupyter Notebook, html/CSS and Python (Flask) to develop the web application. Dataset for training and testing the data for loan prediction analysis.

CHAPTER 5

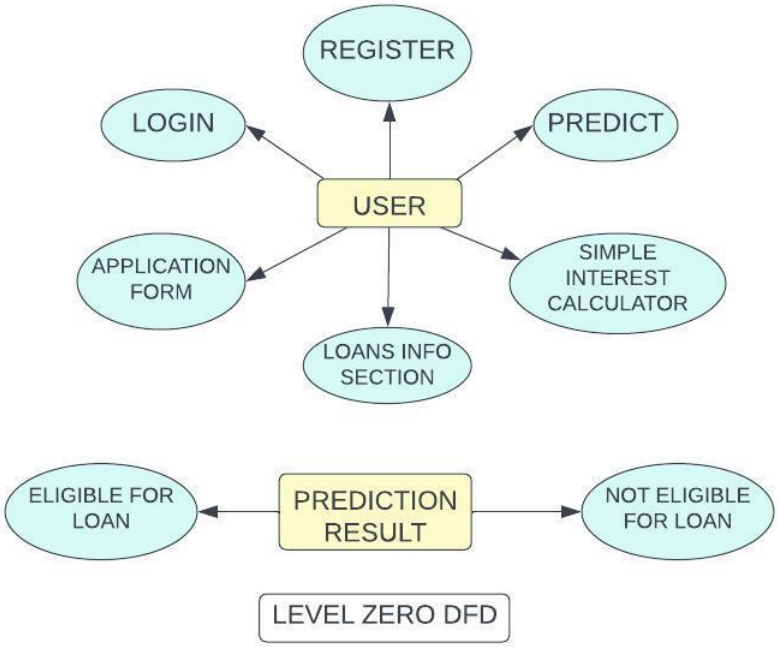
PROJECT DESIGN

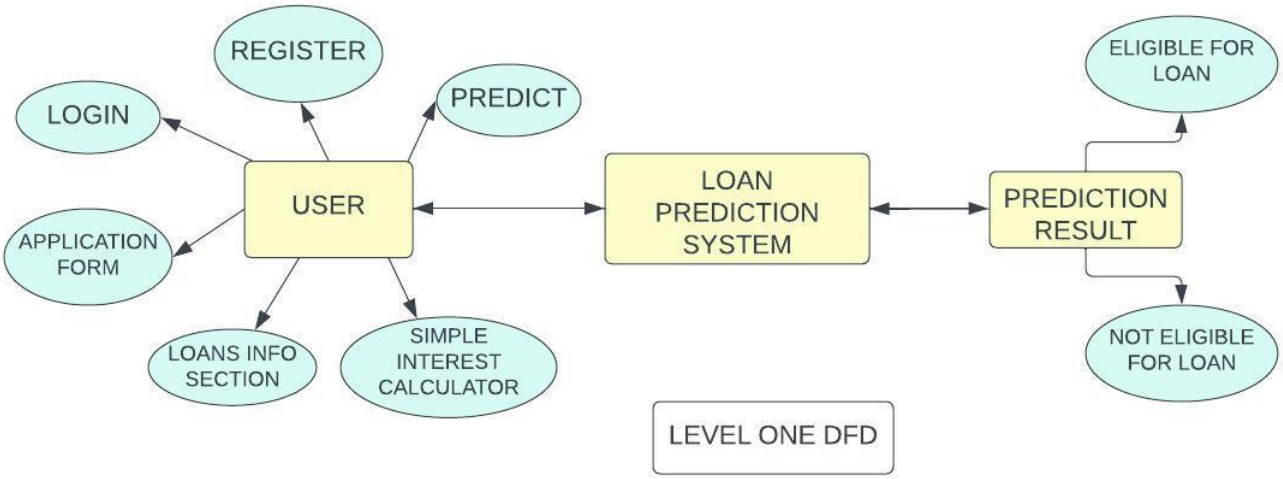
- **Use case diagram:-**



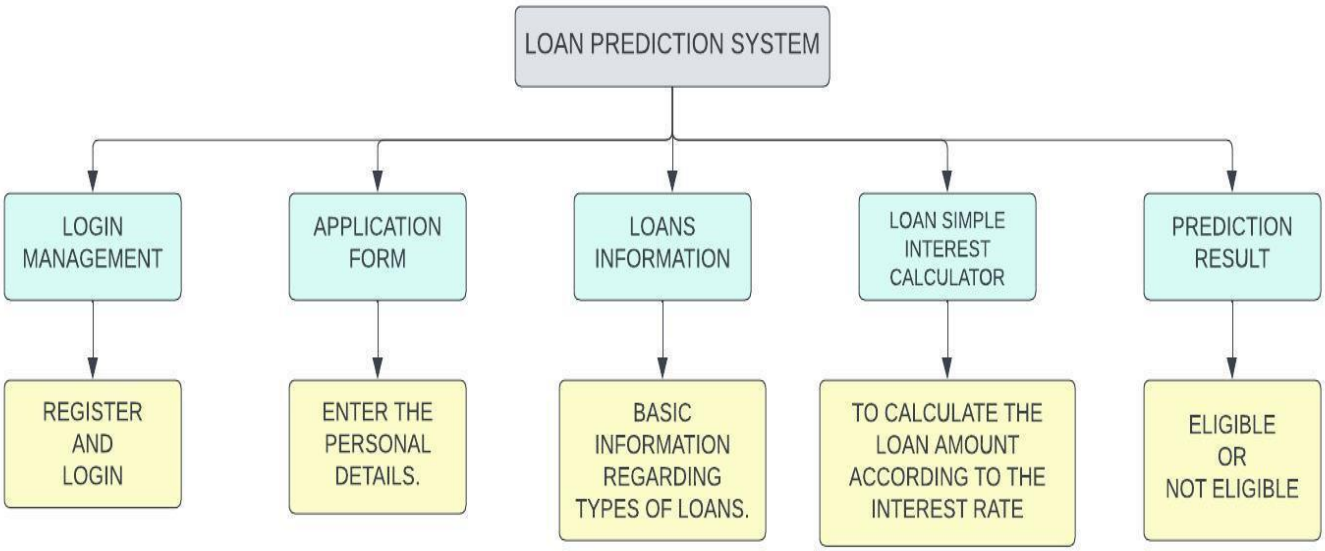
USE CASE DIAGRAM

- **DFD (Data Flow Diagram)**





- **Flowchart**


LOAN PREDICTION SYSTEM
LOGIN MANAGEMENT
APPLICATION FORM
LOANS INFORMATION
LOAN SIMPLE INTEREST CALCULATOR
PREDICTION RESULT
REGISTER AND LOGIN
ENTER THE PERSONAL DETAILS.
BASIC INFORMATION REGARDING TYPES OF LOANS.
TO CALCULATE THE LOAN AMOUNT ACCORDING TO THE INTEREST RATE
ELIGIBLE OR NOT ELIGIBLE

CHAPTER 6

TECHNICAL SPECIFICATION

Development:

Visual Studio Code.

VS Code also known as Visual Studio Code is a source code editor made by Microsoft for Windows, Linux, MacOS. It has various features such as Debugging, Syntax highlighting, extension, intelligent code completion.

Jupyter Notebook.

Jupyter Notebook is an open-source web-based interactive development environment (IDE) that allows users to create and share documents that contain live code, equations, visualizations, and narrative text. It supports a wide range of programming languages, including Python, and others.

Frontend: Html, CSS, JavaScript, Python, Flask.

HTML (Hypertext Markup Language) and CSS (Cascading Style Sheets) are two fundamental languages used to create web pages and web applications.HTML is used to structure content on the web page, defining the layout and organization of text, images, and other elements. It uses tags to mark up content, which are enclosed in angle brackets (< >). HTML also supports hyperlinks, which allow users to navigate between different web pages.

CSS, on the other hand, is used to define the visual appearance of the web page. It controls the color, font, layout, and other visual aspects of the content. CSS works by defining style rules that apply to specific HTML elements. Style rules consist of a selector (which identifies the HTML element) and a set of properties (which define the visual attributes of the element).

Together, HTML and CSS allow web developers to create visually appealing and interactive web pages and web applications. By combining HTML markup and CSS styling, developers can create dynamic and responsive designs that work seamlessly across different devices and screen sizes.

JavaScript is a high-level programming language used primarily to create dynamic and interactive web pages and web applications. It is an interpreted language, meaning that it does not need to be compiled before execution.

Python is a high-level, interpreted programming language that is widely used in data science, machine learning, web development, scientific computing, and many other fields. Python is known for its simplicity and ease of use, making it an ideal language for beginners to learn programming. It uses a clean and readable syntax, which emphasizes code readability and maintainability. Python supports a wide range of programming paradigms, including object-oriented, functional, and procedural programming. It also has a large standard library that provides a rich set of modules and tools for many common tasks, such as data manipulation, file I/O, network programming, and more. Python has a number of popular third-party libraries and frameworks, including NumPy, Pandas, Matplotlib, Flask, Django, and many others, that extend its capabilities for specific use cases. Overall, Python is a versatile and powerful programming language that is widely used in a variety of fields due to its ease of use, readability, and rich set of libraries and frameworks.

Flask is a lightweight web framework for Python that allows developers to quickly build web applications. It was first released in 2010 and has since become one of the most popular web frameworks for Python.Flask is designed to be simple and flexible, making it easy for developers to get started with building web applications. It provides a set of tools and libraries that allow developers to handle HTTP requests and responses, manage sessions and cookies, and interact with databases.

Operating System :

Windows

Windows is a popular operating system developed by Microsoft. It was first introduced in 1985 and has since become one of the most widely used operating systems in the world, powering millions of personal computers and servers. There have been several versions of Windows released over the years, with the most recent being Windows 11. Windows is known for its user-friendly interface, compatibility with a wide range of software and hardware, and its ability to run on various types of devices such as desktops, laptops, tablets, and even smartphones. Windows also includes a variety of built-in features and tools such as the Windows File Explorer, Task Manager, Control Panel, and Command Prompt. Additionally, Windows supports a wide range of third-party software and applications, making it a versatile operating system that can be customized to suit a user’s specific needs.

CHAPTER 7

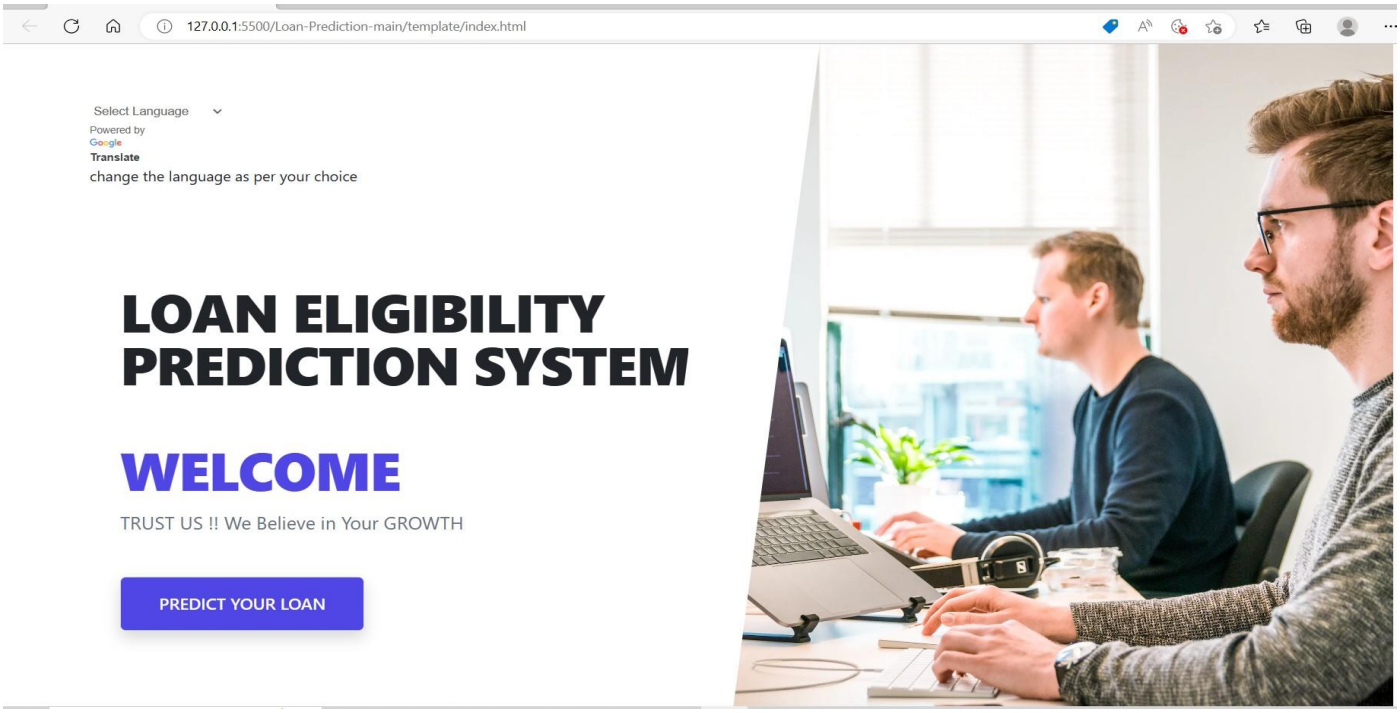
PROJECT SCHEDULING

Sr. No	Group Member	Time duration	Work to be done
1	Om Chavan Sampada Mahadik Parthavi Khatu	2nd week of January	Designing phase of user Interface.
		3rd week of January	Implementation and testing of design.
2	Om Chavan Sampada Mahadik Parthavi Khatu	1st and 2nd week of February	Training and Testing dataset.
3	Om Chavan Sampada Mahadik Parthavi Khatu	By the end of march month	Final testing of application and resolving issues if any.

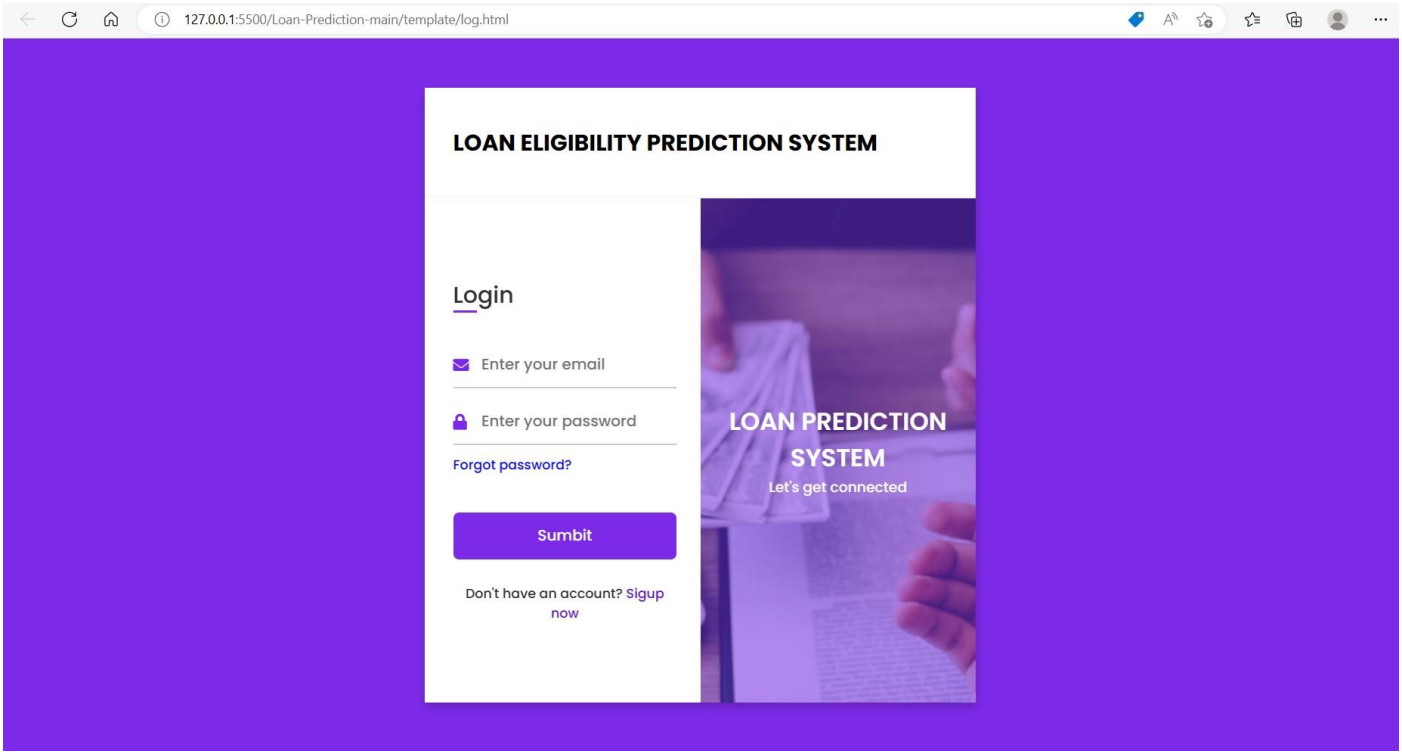
CHAPTER 9

RESULT AND DISCUSSION

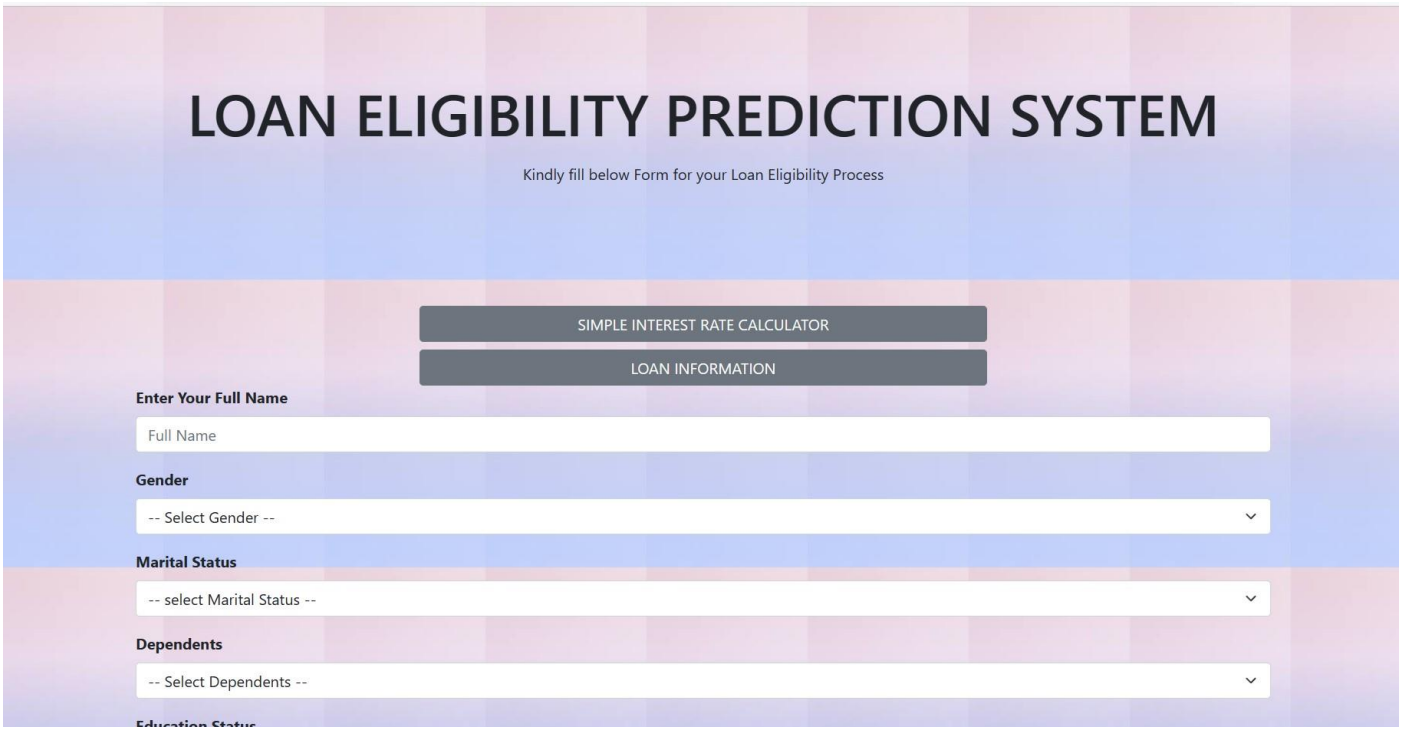
Home Page: This is the home page of the loan prediction system.



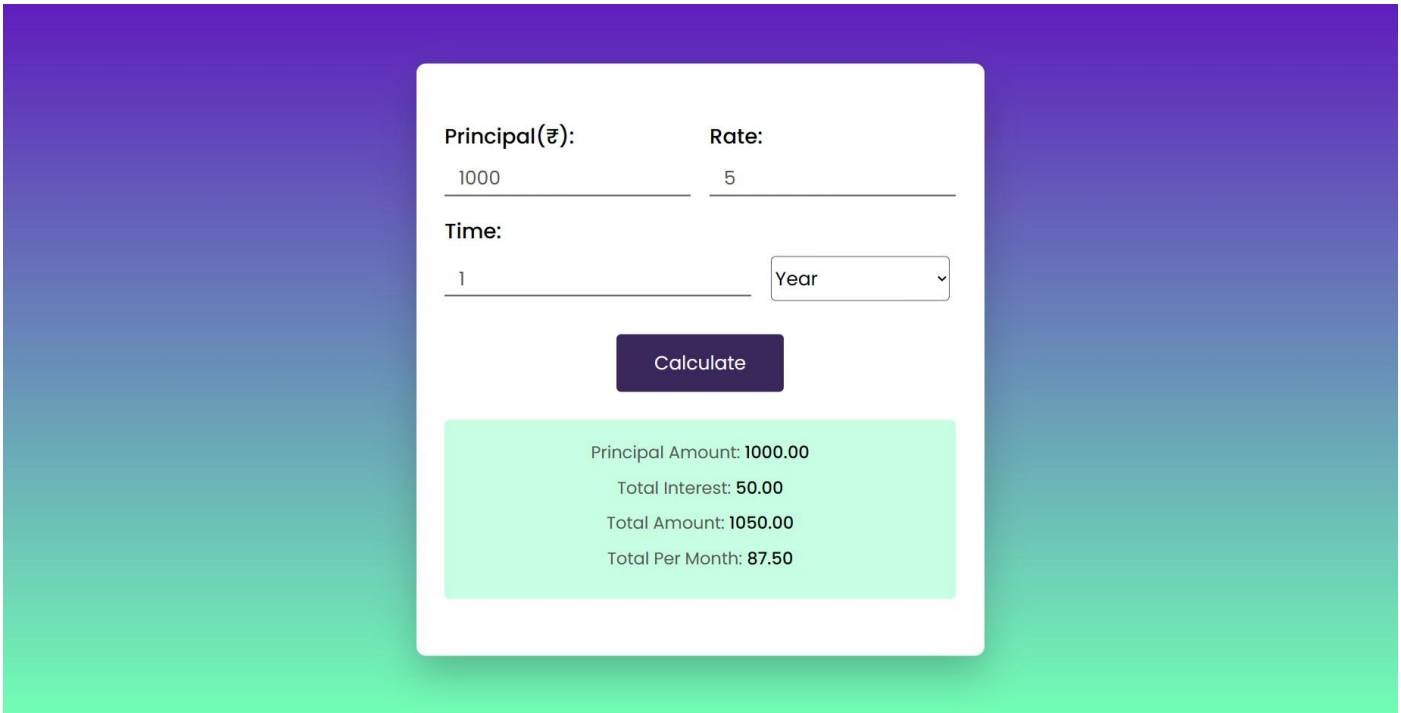
Login Page: Here user can Login and Register to the application.



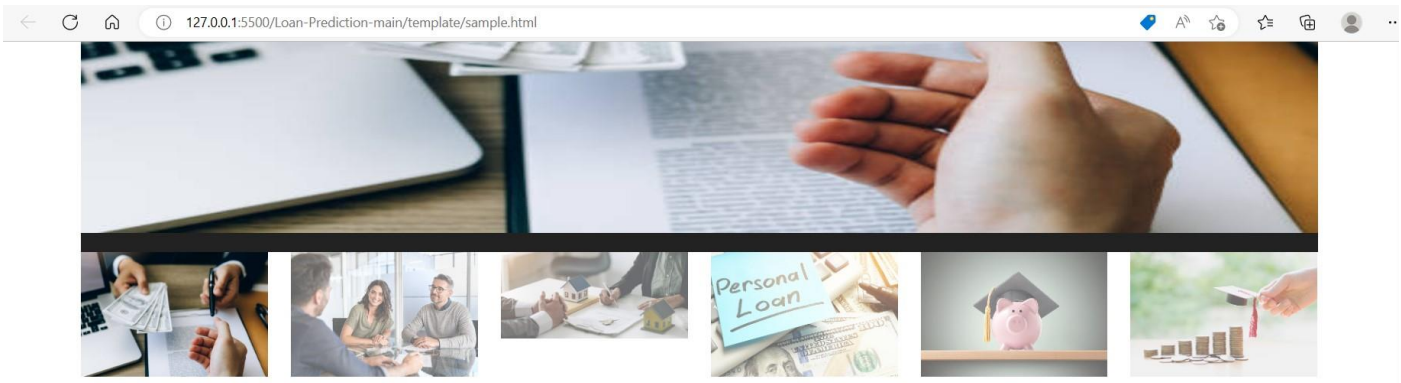
Application Form : This is the personal information application form which is to be filled by applicant for predicting his/her eligibility status.



Simple interest rate calculator: This the simple interest rate calculator page which is helpful for the user to check the amount of loan by calculating it within few clicks.



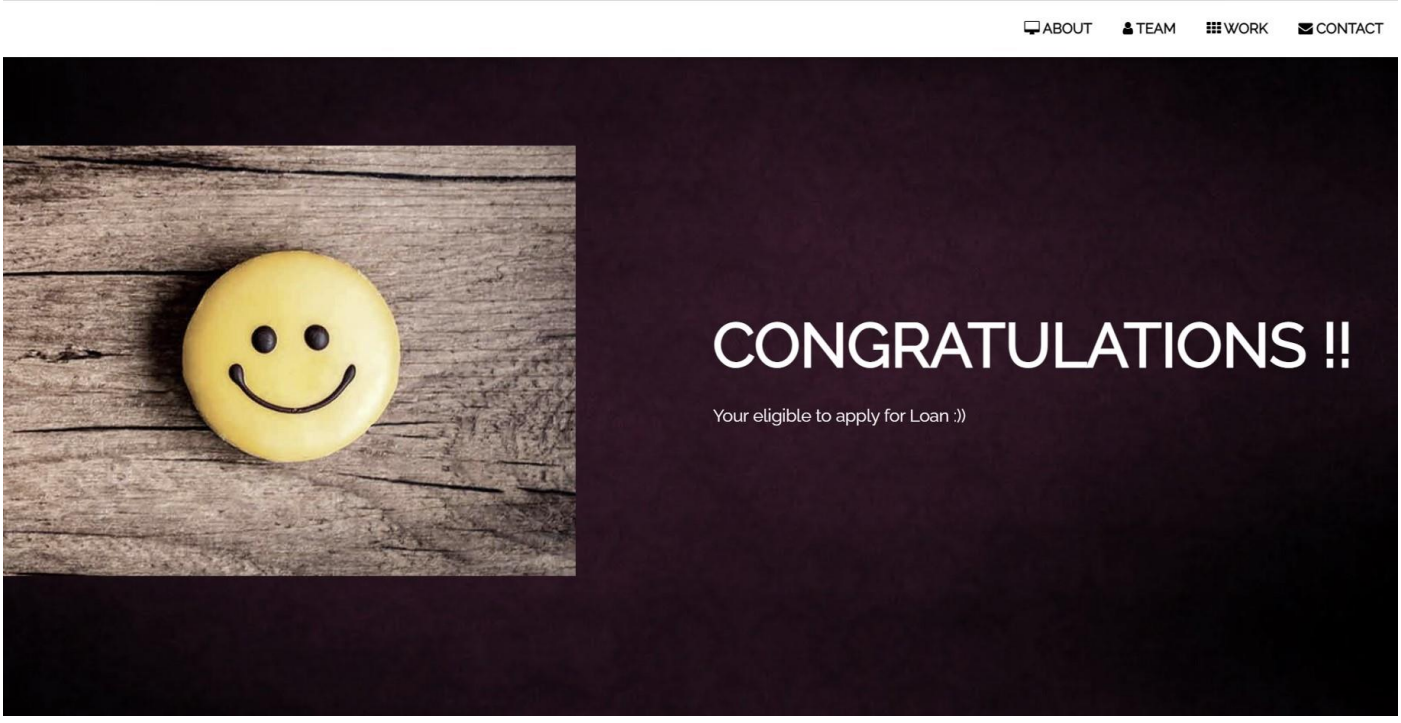
Loans information section: This is the loans information section which gives a basic idea regarding the types of loans.



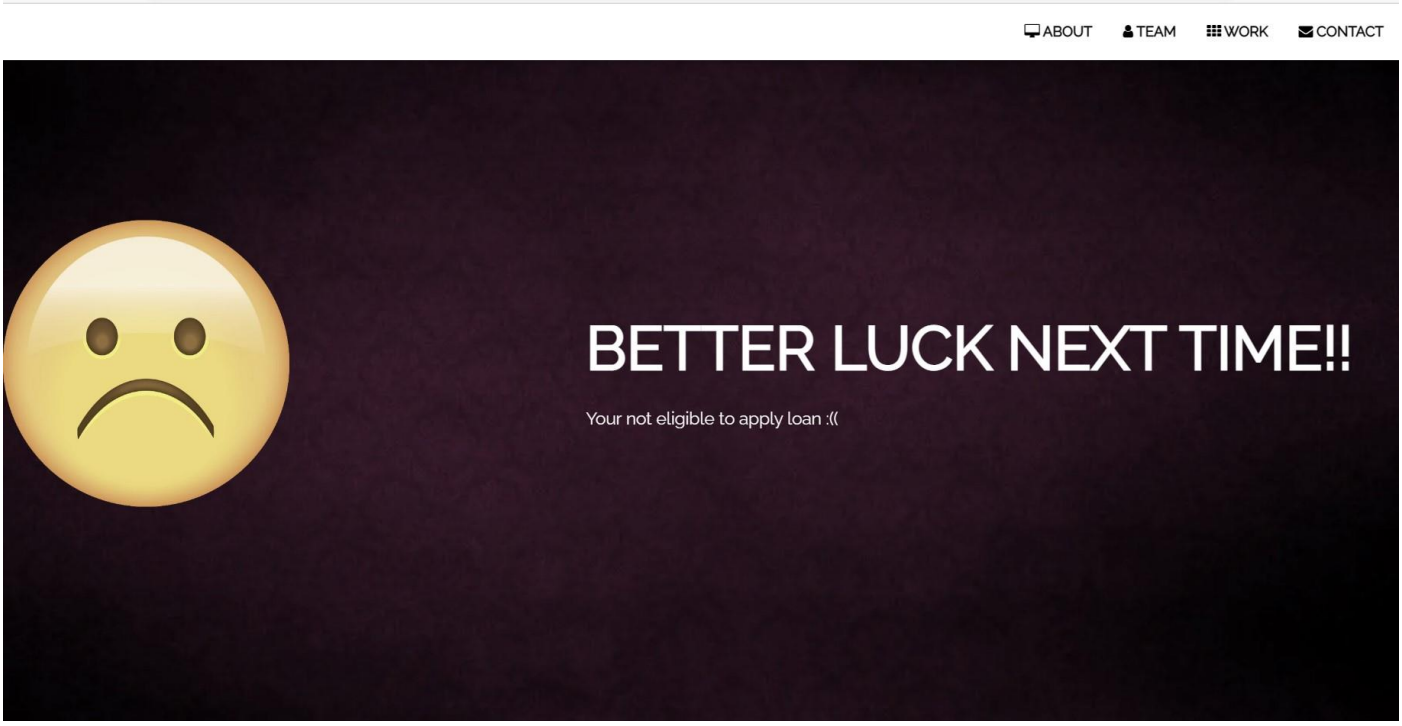
LOAN INFORMATION

No.	TYPES	DESCRIPTION	RATE OF INTEREST
1.	BUISNESS LOAN	A business loan is a loan specifically intended for business purposes. As with all loans, it involves the creation of a debt, which will be repaid with added interest.	11.9 p.a.
2.	PERSONAL LOAN	A personal loan (also known as a consumer loan) describes any situation in which an individual borrows money for personal need, including making investments in a company. All personal loans have three common elements: Evidence of the debt (promissory note) An amount borrowed (principal)	10.25 p.a
3.	EDUCATION LOAN	An education loan is a sum of money borrowed to finance post-secondary education or higher education-related expenses. Education loans are intended to cover the cost of tuition, books and supplies, and living expenses while the borrower is in the process of pursuing a degree.	8.95 p.a

Eligibility Page: This two are the eligibility result pages which user will receive after filling the form and clicking on predict button. This is the output if applicant is eligible for loan.



This the output if applicant is not eligible for the loan.



CHAPTER 10

CONCLUSION

The loan prediction system project is a valuable tool for financial institutions to evaluate the creditworthiness of potential borrowers. By using machine learning algorithms to analyze various factors such as credit history, employment status, income level, and loan amount, the system can provide accurate predictions on whether a borrower is likely to repay their loan on time or default. The loan prediction system not only helps financial institutions mitigate risks and make informed lending decisions but also provides borrowers with fair and transparent loan evaluation processes. Moreover, the project can be further enhanced with the incorporation of new features such as real-time data analysis, automated loan processing, and personalized loan recommendations. Overall, the loan prediction system project has the potential to revolutionize the lending industry by improving efficiency, reducing costs, and promoting responsible lending practices.

FUTURE SCOPE

The loan prediction system has a wide range of future scope for development and improvement. Some of the potential areas of future scope are, The loan prediction system can incorporate advanced machine learning algorithms such as deep learning and reinforcement learning to improve the accuracy of predictions. The loan prediction system can be developed to provide personalized loan recommendations based on the borrower's financial profile, risk appetite, and repayment capacity. This can help borrowers to choose the most suitable loan product and improve their chances of loan approval. The loan prediction system can be automated to streamline the loan processing workflow. This can reduce the time and resources required for loan processing and improve the efficiency of the lending process. In summary, the loan prediction system has a vast potential for future development and enhancement, with the integration of advanced machine learning algorithms, alternative data sources, blockchain technology, and personalized loan recommendations. These advancements can help to improve the accuracy of loan predictions, streamline loan processing, and provide borrowers with more personalized loan recommendations.

REFERENCES

1. S. Sreesouthry, A. Ayubkhan, M. Mohamed Rizwan, D. Lokesh, K. Prithivi Raj (2021). Loan Prediction Using Logistic Regression in Machine Learning. Annals of the Romanian Society for Cell Biology. http://annalsofrscb.ro/index.php/journal/article/view/2818

2. T Sunitha, M Chandravallika, M Ranganayak (2020), "Predicting loan status using Logistic Regression and Binary Tree (ICICNIS). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3769854

3. Veeraballi Nagajyothi, "Loan approval prediction using KNN, decision Tree and Naïve Bayes models"(2020), International Journal of Engineering in Computer (IJECS) . Science. https://www.computersciencejournals.com/ijecs/archives/2020.v2.i1.A.30

4. Pidikiti Supriya, MyneediPavani, Nagarapu Saisushma, NamburiVimala Kumari, k Vikash (2019), "Loan Prediction by using Machine Learning Models", International Journal of Engineering and Techniques. http://www.ijetjournal.org/volume5/issue2/IJET-V5I2P28.pdf

5. Nikhil Madane, SiddharthNanda,"Loan Prediction using Naïve bayes", Journal of the Gujrat Research History,Volume 21 Issue 14s, December 2019. https://www.irjmets.com/uploadedfiles/paper/issue_3_march_2022/20103/final/fin_irjmets1648135217.pdf

6. Om Prakash Yadav,Chandan Soni, Saikumar Kandakatla, Shantanu Swanth (2019), "Loan Prediction using Machine learning",International Journal of Information and computer S2019cience . https://www.researchgate.net/publication/357449126_THE_LOAN_PREDICTION_USING_MACHINE_LEARNING