Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 47 No. 01 (2026)

Loan Scope: Predicting Real Time CIBIL Impact and Recovery

Pavithra S [1], Kowsic Anand S [2], Malliboyina Kondalarao [3], Ragunath G [4]

[1, 2, 3, 4] *Department of Computer Science and Business Systems,*
KIT - Kalaignarkarunanidhi Institute of Technology, Coimbatore, India

Abstract:- This project aims to develop an intelligent system to predict an individual's future CIBIL score using machine learning models. It leverages a Random Forest Regressor trained on historical financial data, including variables such as current CIBIL scores, loan amounts, credit card usage, and repayment histories, to accurately predict changes in credit behavior. The model dynamically analyzes how new loans, late payments, and other credit activities influence the future score, making it adaptable to a wide range of user profiles. The system's architecture includes a Flask-based backend that handles user registration, login, and the fetching of user financial data based on PAN numbers from a pre-existing database. By collecting user data through a CSV file, the model uses this input to generate a prediction for future CIBIL scores. It adjusts predictions based on factors such as new loans, penalties for late payments, and rewards for timely payments, providing users with insights into how their actions could impact their financial standing. The project is designed to offer real-time insights and accurate forecasts oncredit scores, empowering users to manage their credit health proactively. This machine learning-driven approach enhances the traditional CIBIL score analysis with a data driven framework.

Keywords: *Lending Decision, Loan Portfolio, Decision Tree, Random Forest, Prediction Model, Machine Learning, Pandas, Numpy, Scikit-learn, Pickle.*

1. Introduction

In today's financial landscape, maintaining a strong CIBIL score is crucial for accessing credit and securing favorable loan terms. However, borrowers often face uncertainty regarding how new loans or changes in their repayment behavior will affect their credit score. Loan Scope aims to bridge this gap by providing users with real-time insights into how their loan decisions—such as taking out a specific loan amount can impact their CIBIL score. Loan Scope offers predictive analysis, showing users how their CIBIL score may decrease when a loan is taken and how it could recover over time with regular, on-time payments. By empowering users to see both the short-term and long-term effects of their financial decisions, this application serves as a valuable tool for responsible borrowing and better credit management.

This project aims to develop a sophisticated machine learning model to revolutionize credit risk assessment. By leveraging advanced techniques, the model will accurately predict the real-time impact of financial transactions on CIBIL scores and estimate loan recovery rates. This will enable financial institutions to make informed decisions, reduce loan defaults, optimize customer experiences, and gain a competitive edge in the market. The model will be trained on a diverse dataset, incorporating historical credit data, demographic information, behavioral data, and macroeconomic indicators. By continuously refining the model and integrating it into existing systems, financial institutions can enhance their risk management practices and achieve sustainable growth.

In today's fast-changing financial ecosystem, accurately predicting the effects of loan decisions on a borrower's credit score (CIBIL score) and estimating loan recovery likelihood is crucial for financial institutions. Conventional credit risk assessment approaches often depend on static models and historical data, which fail to fully account for the dynamic and evolving nature of creditworthiness.

This project seeks to develop an advanced machine learning model to predict real-time changes in CIBIL scores and loan recovery rates. By employing techniques such as feature engineering, model optimization, and hyperparameter tuning, the model aims to provide financial institutions with a robust tool for informed lending decisions. This innovation will help refine risk management practices, improve portfolio performance, and enhance the overall decision-making process.

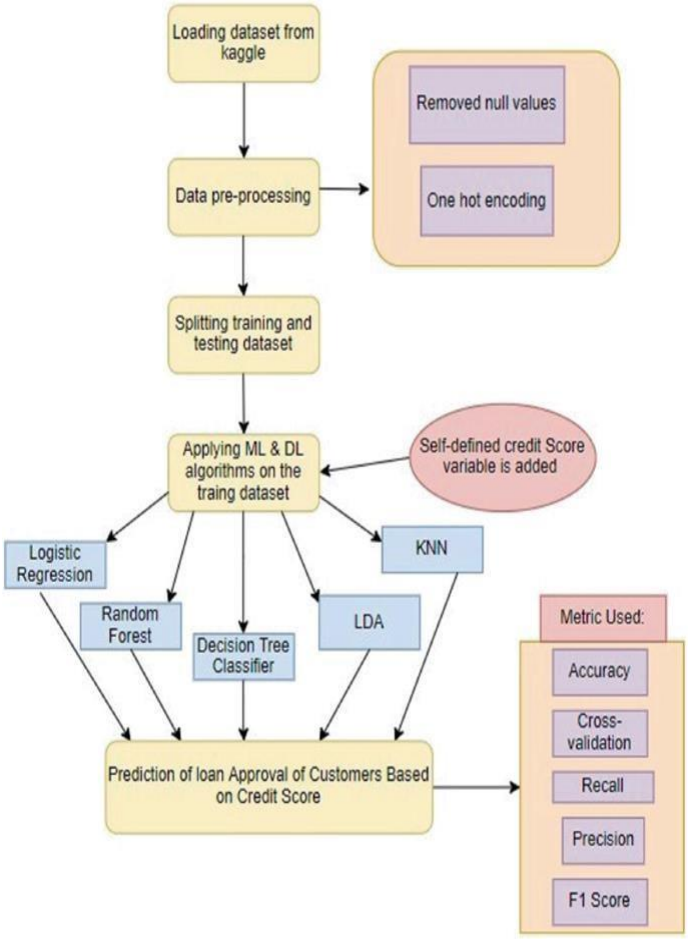

Figure 1: Prediction of Loan Approval of Customers Based on Credit Score

Traditional credit scoring systems, which rely heavily on fixed parameters and past data, often fall short in capturing the fluid nature of credit behaviour.

This project strives to bridge that gap by offering a more adaptive and dynamic solution. To address this limitation, this project aims to develop a sophisticated machine learning model capable of predicting real-time CIBIL score impact and loan recovery rates.

By leveraging advanced techniques such as feature engineering, model selection, and hyperparameter tuning, we seek to create a powerful tool that empowers financial institutions to:

- **Make informed lending decisions:** Accurately assess the creditworthiness of borrowers in real-time.
- **Optimize risk management strategies:** Identify high- risk borrowers and implement appropriate mitigation measures.
- **Enhance portfolio performance:** Reduce default rates and improve overall profitability.

This research will delve into the intricacies of credit risk modeling, exploring the potential of machine learning to revolutionize the way financial institutions evaluate loan applications and manage their portfolios.

2. Related Work

a. Machine Learning Approach for Credit Score Predictions

The proposed methodology is evaluated against existing credit score models using five widely recognized metrics: Accuracy, ROC AUC, Precision, Recall, and F1 Score. The findings demonstrate that the proposed approach surpasses traditional models in these evaluation metrics, effectively balancing predictive accuracy and computational efficiency. The conclusion highlights the importance of this approach for the banking and financial industry, offering a robust and reliable framework for assessing client creditworthiness and enhancing decision-making processes.

b. Loan Eligibility Prediction Using Credit Scores and Past Data

With the rapid growth of the banking sector, there is a critical need for an advanced system to predict loan eligibility based on credit scores and past financial behaviour. This project introduces an innovative solution that integrates multiple components powered by advanced machine learning techniques. The first component predicts optimal bank rate using algorithms like Random Forest, Gradient Boosting, and Linear Regression.

The second component focuses on assessing loan eligibility by analysing applicants' credit histories along with other essential factors such as gender, marital status, education level, income, and financial behaviour.

Additionally, a sophisticated mortgage estimation tool, built using Decision Tree Regression, is included to assist users in analysing and forecasting mortgage-related scenarios. By combining these elements with historical data, credit ratings, and modern machine learning methodologies, this research proposes a comprehensive system for accurate and efficient loan eligibility prediction, addressing the evolving needs of the banking and financial industry.

c. Identifying the Optimal Machine Learning Technique for Predicting Bank Loan Defaulters

In today’s financial environment, effectively managing credit risk and maintaining stable loan portfolios are critical challenges for banks. Accurate prediction of loan defaults plays a pivotal role in strengthening the risk management process and minimizing financial losses.

This study evaluates various machine learning techniques to identify the most effective model for predicting loan defaulters. The results reveal that for imbalanced datasets, the Random Forest algorithm delivers outstanding performance, achieving an accuracy score of 0.91. For balanced datasets, the Random Forest model surpasses all other techniques, achieving a perfect accuracy score of 1.00, making it the most reliable approach for loan default prediction. These findings underscore the significance of employing Random Forest for robust and accurate credit risk assessment.

d. Building Reliable Loan Approval Systems

This study introduces new financial indicators during the feature engineering stage to extract critical client information, enhancing both the robustness and accuracy of predictions. The dataset undergoes thorough preprocessing with cleaning, transformation, and feature engineering techniques to optimize model input. To further improve performance, three powerful ensemble methods—Random Forest, AdaBoost Classifier, and Gradient Boosting Classifier—are employed.

A comparative analysis using ROC curves and feature importance metrics highlights the superior performance of the proposed approach. The model achieves a larger area under the ROC curve and a lower false positive rate compared to the original pre-processed model. The results reveal a stronger dependency on financial features rather than personal demographic features, emphasizing the model's ability to deliver robust classification results. These findings demonstrate the potential of this advanced approach to replace existing loan approval systems in real-world applications, offering improved accuracy and reliability.

e. Analysis and Predict CIBIL Score using Machine Learning

CIBIL, India's first Credit Information Company, established in August 2000, collects and maintains records of individual loan and credit card payments. These records are used to generate Credit Information Reports (CIR) and credit scores, which aid financial institutions in evaluating and approving loan applications. However, despite a good CIBIL score, the rise in NPAs and bad loans indicates gaps in the existing scoring methodology.

This study proposes a machine learning-based platform that integrates additional essential parameters currently overlooked in CIBIL score prediction. By leveraging advanced algorithms, this approach aims to provide more comprehensive and accurate insights, thereby reducing the occurrence of bad loans and improving overall credit risk management.

3. Existing System

Existing loan approval systems primarily rely on traditional methods that are often cumbersome and time-consuming. These systems typically use limited customer data, focusing more on personal information rather than financial indicators. Machine learning has been applied in some systems, but many fail to leverage advanced techniques for feature engineering and data preprocessing. As a result, accuracy rates in predicting loan defaults remain suboptimal. Therefore, there is a pressing need for more robust, automated solutions that enhance prediction accuracy and efficiency in real-time loan assessments. Existing loan prediction systems typically rely on traditional methods of assessing a user's eligibility for loans, primarily based on static credit scores and basic financial information.

These systems evaluate a user's creditworthiness by looking at historical credit data, repayment behavior, income, and existing debts. The approval process is often rigid, and predictions regarding the impact of loans on future credit scores are limited. Most systems do not provide real-time insights into how a new loan will affect the user's CIBIL score or predict the likelihood of loan default. Users are usually not provided with personalized forecasts, leaving them unaware of the long-term consequences of their financial decisions. Furthermore, the systems lack flexibility in offering dynamic repayment options and tend to follow a one size-fits-all approach

4. Proposed System

The proposed system, Loan Scope: Predicting Real-Time CIBIL Impact and Recovery, introduces a machine learning-based approach to enhance the loan evaluation and decision making process. This system aims to provide users with personalized predictions on how taking out a loan will impact their CIBIL score and assess their risk of default. The core feature of this system is the CIBIL score prediction model, which uses regression algorithms to forecast the potential reduction in the score due to the loan. Unlike existing systems, this solution offers dynamic, scenario-based forecasting and integrates machine learning models to deliver personalized, data-driven predictions. The proposed work for the project Loan Scope: Predicting Real-Time CIBIL Impact and Recovery involves developing an advanced, machine learning-powered system that offers real-time insights into how taking a loan will affect a user's CIBIL score The system will predict both the immediate impact of a loan on the CIBIL score and potential recovery over time, based on repayment behavior

Objectives of the Proposed Work

Linear Regression Purpose: To predict the immediate reduction in a user's CIBIL score based on loan parameters such as loan amount, tenure, and interest rates.

Function: Linear Regression learns from historical data by fitting a linear equation to predict how much the CIBIL score will drop when a specific loan is taken.

Multiple Linear Regression: Purpose: When more than one independent variable (loan amount, tenure, income, etc.) affects the CIBIL score, Multiple Linear Regression is applied to account for multiple factors influencing the prediction. This provides a more accurate estimate of the CIBIL score impact.

Function: It builds a linear model with multiple features to predict how the combined effect of various loan features impacts the CIBIL score.

Logistic Regression: Purpose: This algorithm will be used to predict the probability of loan default. It is a classification algorithm that estimates the likelihood of an event (loan default) occurring based on input features such as loan amount, income, and repayment history.

Function: Logistic Regression outputs a probability score (between 0 and 1) indicating the chance that a user may default on their loan, helping to assess loan risk

It plays a central role in the application, as it estimates how much a user's CIBIL score will decrease based on their loan inputs. Using regression models such as Linear Regression and Multiple

Linear Regression, the module analyzes the relationship between loan factors and historical CIBIL score data to provide real-time predictions. This allows users to foresee the impact of taking a loan before making financial commitments.

The application features a user-friendly interface that simplifies the process of inputting loan details and quickly obtaining predictions, making financial planning more accessible. Furthermore, the system delivers personalized financial insights

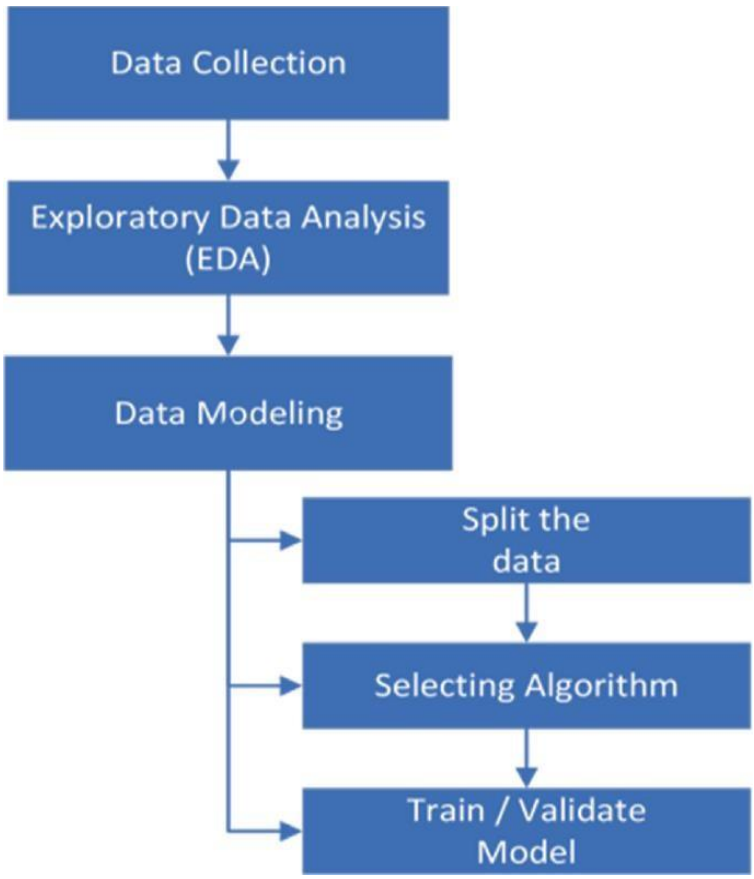

Figure 2: Block diagram

based on individual data and historical loan behavior, ensuring that users receive relevant and tailored information to manage their credit health effectively. Overall, these advantages make the proposed system a valuable tool for enhancing financial decision- making and promoting better credit management.

5. Experimental and Results

Predicting a CIBIL score based on loan amount and monthly tenure provides valuable insight for both borrowers and lenders

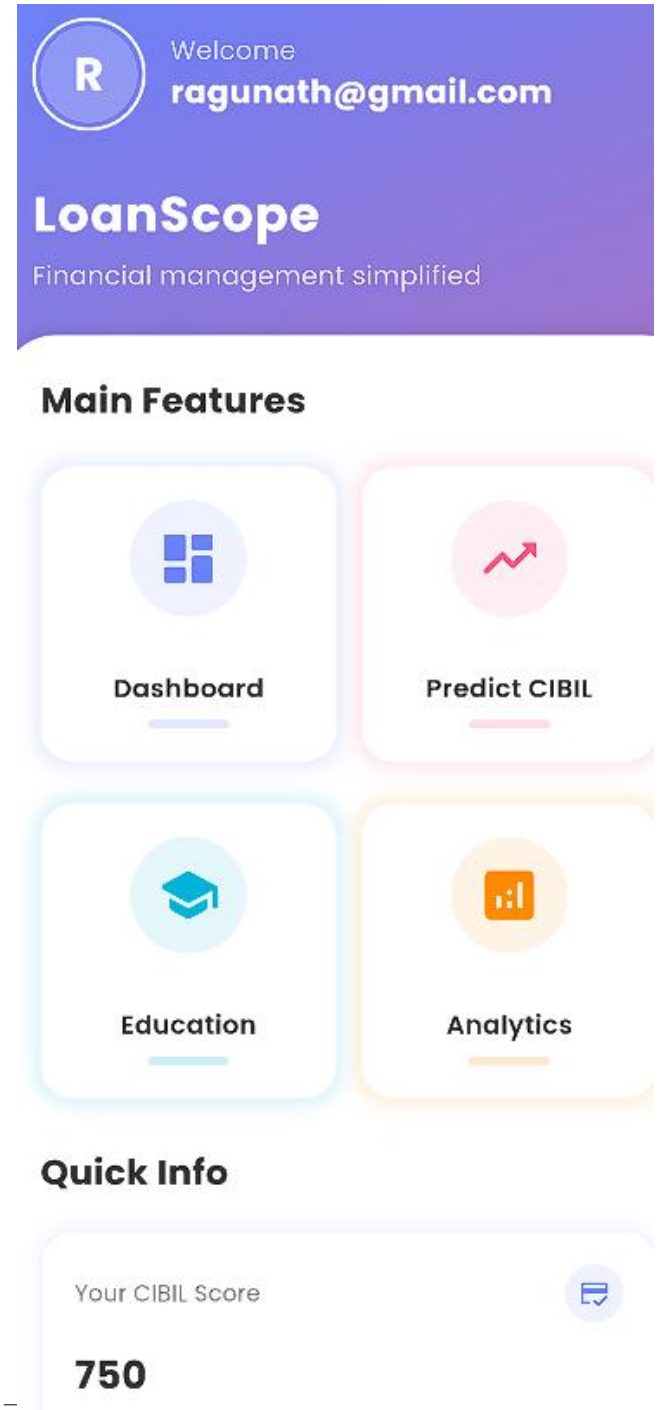

Figure 3: Home Page

To incorporate the content from the dashboard.Here's a concise way to present it: "The user interface features a dashboard accessible via a registered email, offering functionalities such as CIBIL score prediction, educational resources, analytics, and account management options including logout. This design facilitates user engagement and data-driven decision-making." This summary integrates the dashboard's features into a broader discussion of its purpose and utility in a research or technical context.

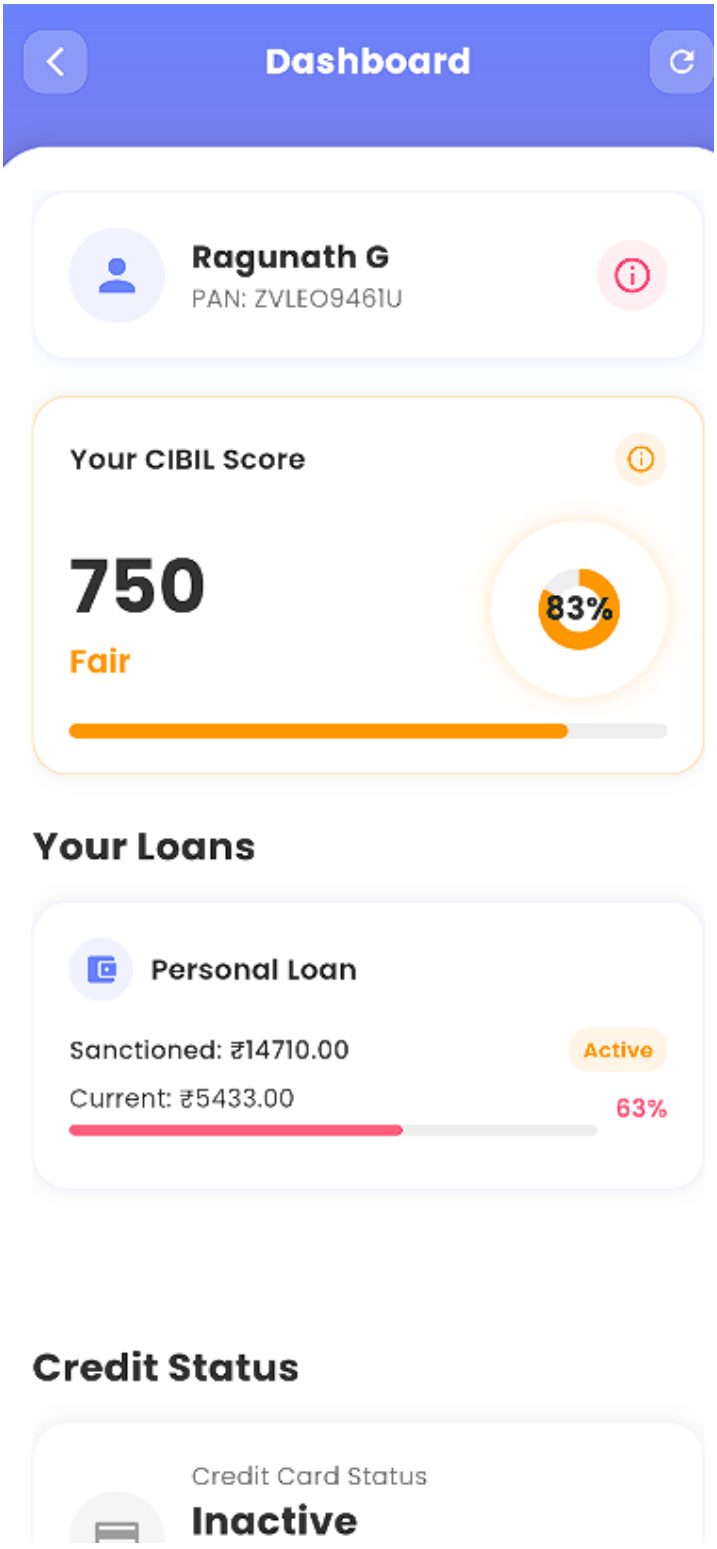

Figure 4: Dashboard

The dashboard provides a comprehensive financial overview for the user, including a CIBIL score, categorized as average. It details sanctioned and current amounts for personal and home loans, along with the status of credit card activity, which is currently inactive. This information aids users in monitoring their financial health and making informed decisions. This description contextualizes the data within the framework of financial management and user empowerment.

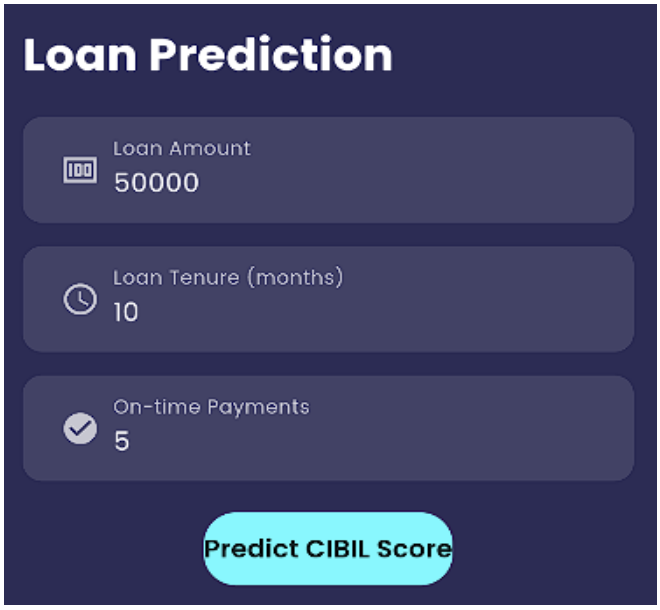

Figure 5: Loan Prediction (Input Field)

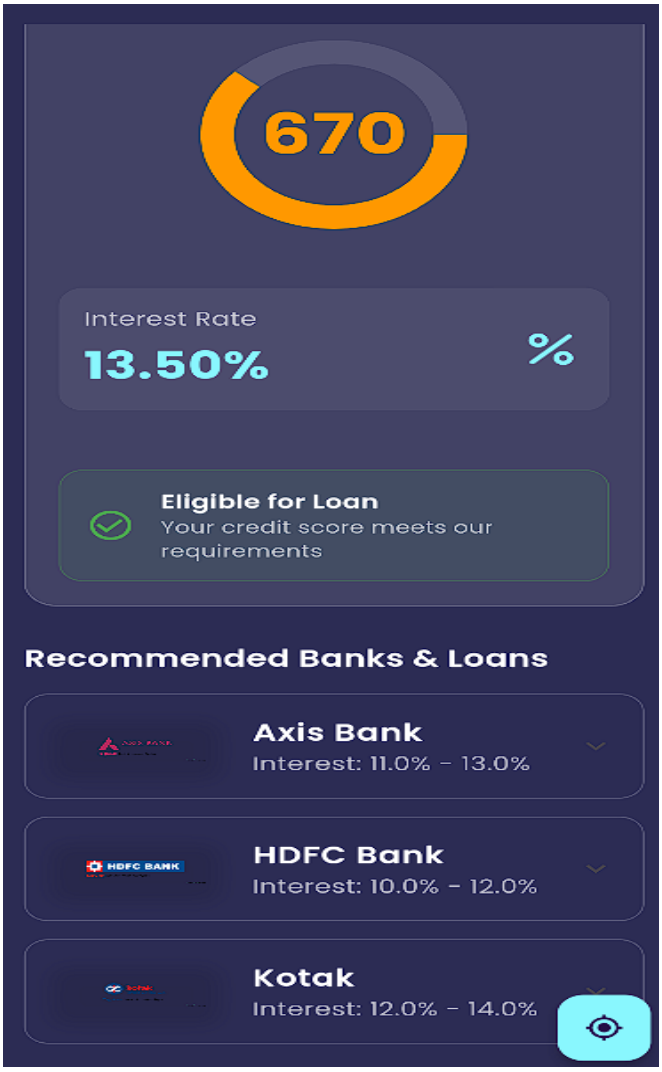

Figure 6: CIBIL Based Bank

Recommendations

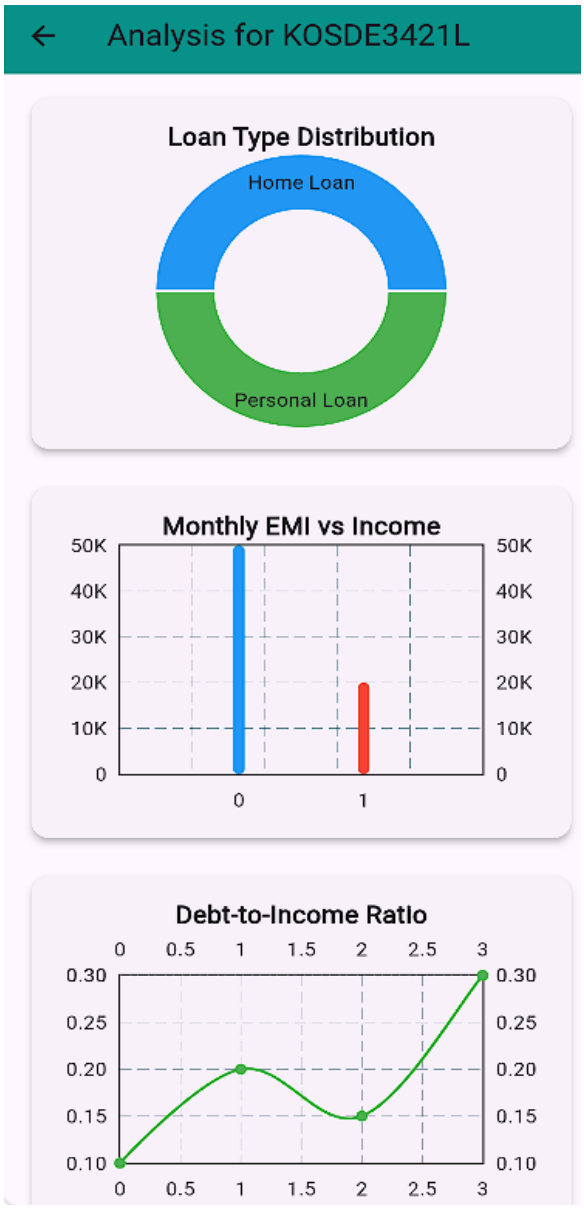

Figure 7: Analysis

6. Conclusion

In conclusion, the Loan Scope: Predicting Real-Time CIBIL Impact and Rec overy project represents a significant advancement in the way individuals and financial institutions assess the implications of loan agreements on credit health. By leveraging machine learning algorithms, this application provides users with accurate predictions of how their CIBIL scores will be affected by various loan parameters, along with insights into the likelihood of default and potential recovery of credit scores over time. The user-friendly interface ensures that individuals can easily navigate the application, input their financial data, and receive personalized predictions that empower them to make informed borrowing decisions. The insights generated by the Loan Scope application can also aid financial institutions in their lending processes. By better understanding the risk profiles of potential borrowers, lenders can make more informed decisions, tailor loan products to meet customer needs, and reduce the likelihood of defaults. This contributes to a more stable financial environment and fosters trust between borrowers and lenders. The project lays the groundwork for future enhancements, such as integrating additional features like real-time monitoring of CIBIL scores, alerts for payment reminders, and educational resources about credit

management. The scalability of the system allows for the incorporation of more sophisticated algorithms and data sources, potentially expanding its applicability to various financial products beyond personal loans, thus catering to a broader audience and addressing diverse financial needs. As a result, this project not only aids users in understanding their financial standing but also promotes responsible borrowing behavior, ultimately leading to improved credit management. With its innovative approach, the Loan Scope application stands to benefit a wide range of users by enhancing transparency in lending processes and supporting better financial literacy in the realm of credit management. In conclusion, the Loan Scope: Predicting Real-Time CIBIL Impact and Recovery project represents a significant advancement in the way individuals and financial institutions assess the implications of loan agreements on credit health. By leveraging machine learning algorithms, this application provides users with accurate predictions of how their CIBIL scores will be affected by various loan parameters, along with insights into the likelihood of default and potential recovery of credit scores over time. The user-friendly interface ensures that individuals can easily navigate the application, input their financial data, and receive personalized predictions that empower them to make informed borrowing decisions. The insights generated by the Loan Scope application can also aid financial institutions in their lending processes. By better understanding the risk profiles of potential borrowers, lenders can make more informed decisions, tailor loan products to meet customer needs, and reduce the likelihood of defaults. This contributes to a more stable financial environment and fosters trust between borrowers and lenders. The project lays the groundwork for future enhancements, such as integrating additional features like real- time monitoring of CIBIL scores, alerts for payment reminders, and educational resources about credit management. The scalability of the system allows for the incorporation of more sophisticated algorithms and data sources, potentially expanding its applicability to various financial products beyond personal loans, thus catering to a broader audience and addressing diverse financial needs. As a result, this project not only aids users in understanding their financial standing but also promotes responsible borrowing behavior, ultimately leading to improved credit management. With its innovative approach, the Loan Scope application stands to benefit a wide range of users by enhancing transparency in lending processes and supporting better financial literacy in the realm of credit management.

References

[1] Tsholofelo Mokheleli1, Tinofirei Museba, Machine Learning Approach for Credit Score Predictions, Volume 14, DOI:10.51519/journalisi.v5i2.487, Issue November 2023.

[2] Senarathna B.T.N and Weerarathna K.C.M, Comprehensive Loan Eligibility Prediction Utilizing Credit Scores and Historical Financial Behaviour for Enhanced Lending Decisions, Volume 09, DOI:10.47001/IRJIET/2023.710070, December 2023.

[3] Digambar B Uphade and Aniket A Muley, Identification of Most Preferable Machine Learning Technique for Prediction of Bank Loan Defaulters, Volume 8, DOI:10.479201/DUKGI /2024.77862, June 2024.

[4] Maryam Shoaeinaeini and Milad Shoaeinaeini, Building Reliable Loan Approval Systems: Leveraging Feature Engineering and Machine Learning, Volume 17, DOI:10.365967/JCSE/2024.88122, January 2024.

[5] Kirtee Kushwaha and Pankaj Malik, Analysis and Prediction of CIBIL Score using Machine Learning, Volume 2, DOI:10.8578998/JDKU/2021.8979, December 2021.

[6] Dr. Saravanan Laxmanan and Dr. Sankaramuthukumar, Awareness of Credit Score Mechanism in India: A Study with reference to Credit Information Bureau India Limited (CIBIL), Volume 5, DOI:10.98674/DKL/2021.87923, March 2021.

[7] Keerthana Sivamavilvelam, Leveraging Data-Driven Analytics for Optimizing Debt Collector Performance and Enhancing Recovery Efficiency, Volume7, DOI:10.365967/JCSE/2024.8812, December 2024.

[8] M Jayadev, Predictive Power of Financial Risk Factors, An Empirical Analysis of Default Companies, Volume 18, DOI:10.365967/JCSE/2024.88122, December 2024.

[9] Yi Wu, Yuwen Pan, Advanced Application Analysis of Credit Scoring in Financial Institutions Using Machine Learning Models for Improved Risk, Volume 12, DOI: 10:1155/2021/9222617, April 2021.