Bank Loan Eligibility and Credit Limits Prediction Using a Machine Learning Approach

1st Pathirathna D.N
Dept. Computer Systems Engineering
Faculty of Computing
Sri Lanka Institute of Information
Technology
Colombo, Sri Lanka
dilinikanethmini@gmail.com

2nd Pravean P.R.I
Dept. Computer Systems Engineering
Faculty of Computing
Sri Lanka Institute of Information
Technology
Colombo, Sri Lanka
praveanravi02@gmail.com

3rd Hilma M.I.F
Dept. Computer Systems Engineering
Faculty of Computing
Sri Lanka Institute of Information
Technology
Colombo, Sri Lanka
hilmailliyas2@gmail.com

4th Suranjini Silva
Dept. Computer Systems Engineering
Faculty of Computing
Sri Lanka Institute of Information
Technology
Colombo, Sri Lanka
suranjini.s@sliit.lk

5th Anuradha Jayakody
Dept. Computer Systems Engineering
Faculty of Computing
Sri Lanka Institute of Information
Technology
Colombo, Sri Lanka
anuradha.j@sliit.lk

Abstract—**Loan eligibility assessment based on various eligibility requirements in a financial institution, most typically a bank, is a critically challenging process that often leads to delays, inconsistencies, and a lack of transparency in lending decisions. This research presents a web-based machine learning-powered solution for loan eligibility prediction to address these challenges. The software system has been integrated mainly with three components. Firstly, an AI chatbot that makes decisions and provides dynamic responses based on bank-provided documents and user data regarding bank policies, loan agreements, and user-specific loan data. Secondly, a loan eligibility and eligible loan amount prediction module that evaluates identification documents, payslips, and credit information using document analysis mechanisms to determine whether a system user is eligible for a loan on behalf of the financial institution. Lastly, a financial literacy and alternative financing module assists rejected applicants by providing collateral-based loan alternatives in case the salary and bank deposit-based loan application fails. With an accuracy of 92% for the AI chatbot, 89% accuracy for the eligibility prediction model, and an average of 63% re-applications prompted with alternative financing suggestions model, this system streamlines the loan application processing, improves the overall transparency of eligibility evaluation practices, and enhances financial literacy, introducing a robust solution for modern banking.**

Keywords—loan eligibility prediction, AI chatbot, OCR-based document analysis, machine learning, collateral-based loan alternatives

I. INTRODUCTION

The financial sector is a crucial player in the world of economics that facilitates means for individuals to safekeep their financial assets and take out loans. Even though the financial sector has become more and more important over time, the traditional methods for loan processing still being used mainly within banks are often slow and rely on extensive manual verification, is sometimes subjected to human bias of the assessors, and inconvenient to both parties in a technically evolving world. Especially in the context of Sri Lanka, applicants frequently experience delays in the loan approval process when it comes to time-consuming document checks, lack of transparency, and inefficient communication with the financial institutions. Additionally, many applicants also struggle to understand the criteria for loan eligibility and ultimately end up getting rejected.

The system developed in accordance with this research introduces a web-based machine learning-powered system explicitly for predicting loan eligibility with the goal of enhancing loan accessibility, decision-making efficiency, and financial literacy. This system automates the entire loan application process, improves customer engagement, and provides alternative financing solutions when conventional loan applications fail by employing the latest technological developments such as artificial intelligence, machine learning, and optical character recognition (OCR).

When the problem itself is taken into consideration, the current loan approval process in many banks and financial institutions relies on static eligibility criteria, extensive manual processing, and limited customer guidance. This creates the key problems of inefficient loan processing resulting from manual workload, lack of financial literacy including lack of awareness of terms relevant to loans and alternative collateral-based lending within the common populace, and high rejection rates when customers fail to meet the salary-based eligibility, whereas they could have been approved under asset-based criteria. Thus, there is a need for a computerized solution capable of predicting eligibility and providing customer support to improve the decision-making accuracy, cut back on redundant manual processes, and enhance financial inclusivity.

As a solution, the presented loan eligibility prediction system has utilized the latest technologies to make three critical components available to customers.

1. AI chatbot – A smart chatbot that refers to the pre-provided document and database-based data to answer loan-related queries using a structured knowledge base. The main target of this chatbot is to assist logged-in users with personalized loan insights based on their personal financial history and assist non-registered users with general information regarding eligibility, required documentation, and bank policies.

2. Loan eligibility and lendable loan amount prediction – A component that takes user input in the form of a loan application form and validates attached documents and personal banking information to determine the financial status so that the user's eligibility for the requested loan can be determined through the formula used within the bank.

3. Financial literacy and alternative financing – A guidance system for rejected applicants that offers personalized advice based on rejection reasons and asset-based financing solutions for those who fail to meet salary-based eligibility.

This study aims to develop an AI-powered chatbot based on the context of Sri Lankan loan eligibility processes, specifically to meet the requirements determined by Ceylan Bank. The goal of this software system lies in enhancing customer engagement and loan-related assistance from a business perspective and prompt rejected applicants to explore other available alternative financing options. Furthermore, it mainly attempts to improve the loan approval efficiency and reduce the processing delays by incorporating automatic techniques that minimize the need for human intervention.

II. RELATED WORK

A loan is a gift of either money, property, or other tangible items from a third party in exchange for the early repayment of the loan amount with the inclusion of interest. Loan processing is a lengthy process when done manually, and every business that is involved with money lending regularly faces difficulty with eligibility confirmation [1]. Loan eligibility and eligible loan amount prediction is a very common real-life problem that every finance company faces in their lending operations. If the loan approval process is automated, it can save a lot of man hours and improve the speed of service to the customers. [2]. Not only that, it is seen that automating the processes can also boost the safety of the processes and minimize human error and biased interventions in the long run. It is a step towards preventing fraudulent activities, too. For instance, the primary goal in the banking sector is to place their funds in safe hands. Many banks and financial institutions now grant loans after a lengthy process of verification and validation, but there is no guarantee that the chosen applicant is the most deserving of all applicants. [3]. But an automated system evaluates candidates without bias and is transparent about the process, giving every candidate an equal opportunity to obtain a loan.

Thus, with the technological advancements in the industry, several countries and individuals have already come up with solutions that computerize at least a part of the eligibility validation process for loans. Most recently, machine learning techniques that autonomously decipher patterns and make data-driven predictions have surfaced as a promising solution that can address the current inefficiency issues [4]. For instance, there are existing studies evaluating decision tree algorithms such as ID3, C4.5, CART, CHAID, QUEST, CTREE, CRUISE, and GUIDE against parameters such as loan amount required, age, asset value, income, family details, guarantor asset value, guarantor income, already avail loan on asset to determine the feasibility of employing machine learning techniques in loan eligibility validation [5].

On that note, apart from the main loan application process based on current savings of the customer and their income, there are several other types of loans offered by financial organizations. The use of collateral is a widespread feature of credit contracts between firms and financial institutions, and collateral is an important tool for the banks to mitigate informational asymmetries such as low-risk burrowers dropping out when the interest rates are high and solve the credit-rationing problems [6]. From a banking perspective, collateral is a regular characteristic of loan contracts whose use is still not fully understood [7]. So, an opportunity to raise more awareness and proactively suggest available options for collateral-based loans would be ideal to continue the business interactions between banks and customers to provide alternatives where income-based loan applications fail.

Furthermore, even though finance is such a popular domain involved in the world economy, there is an astounding lack of awareness within the general populace, even on basic economics and finance. The studies in an international project indicate that financial literacy differs by population subgroup. Specifically, age patterns were found notable where financial knowledge follows an inverted U-shaped pattern, being lowest for the young and the older groups, but peaks in the middle of the life cycle [8]. Even though banks and financial institutions offer many services, many of these widely available financial products – student loans, mortgages, credit cards, pension accounts, annuities – have proven to be complex and difficult for financially unsophisticated customers [9]. This extends to the unawareness regarding alternative financing options, too. Wrong choices in finances are going to ultimately cost a lot, especially for the younger generations if they are not sufficiently financially literate [10]. Therefore, it is important to offer proper guidance and recommendations from an expert perspective to offer the best possible service to help individuals make the best financial decisions from the available options.

III. METHODOLOGY

This section encompasses the design and implementation of the loan eligibility and prediction system discussed in this research context by detailing the technologies, processes, and models used in loan assessment, AI chatbot interactions, and financial guidance module implementation. The methodology of this project has followed a structured data-driven approach that integrates machine learning, natural language processing (NLP), and optical character recognition (OCR) to automate the decision-making capabilities and enhance the user engagement with the system.
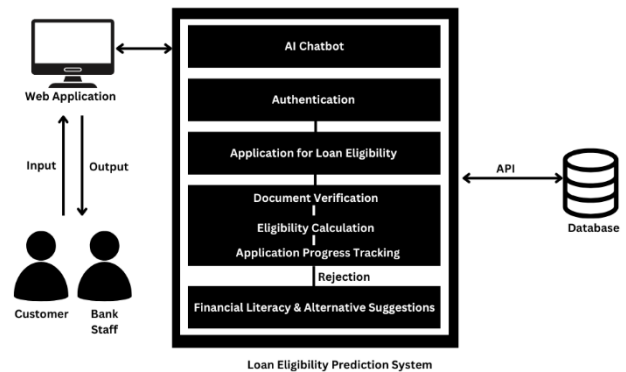
A. System Architecture

Fig. 1. High-Level Architecture Diagram

Fig. 1 represents a high-level overview of the entire system architecture. The system has been structured into three primary components, namely, the AI chatbot for loan information discussion, loan eligibility and eligible loan

amount prediction component for document analysis and application processing, and alternative financial guidance module for providing actionable recommendations for rejected applicants. The general workflow consists of the following processes;

1. User registration and authentication

2. Loan inquiry via AI chatbot

3. Loan application submission

4. Document analysis

5. Machine-learning based loan eligibility prediction and application processing

6. Financial guidance for rejected applicants

The system caters to the needs of multiple user groups, including new and existing customers, bank staff, and administrative users.
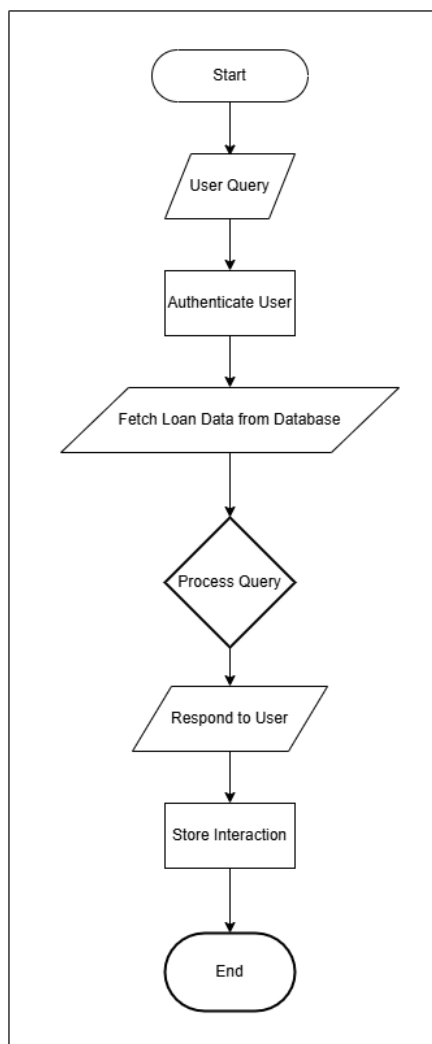
B. AI Chatbot

Fig. 2. AI Chatbot Workflow

The AI chatbot has been built by using natural language processing and retrieval-augmented generation (RAG) for generating loan-related answers. In the case of non-registered users, the chatbot provides general loan-related information by retrieving responses from the data fed into the model through properly structured financial documentation, including data related to bank loan policies, eligibility criteria,

and required documentation. On the other hand, as Fig. 2 depicts, if the user has already registered and logged into the system, following the login and OTP verification, the chatbot will act as a personalized inquiry tool that is capable of retrieving loan details from the customer details and answer queries around the up to date information about the user's already existing loans, payment dates, and interest rates. When the user performs a query, the system fetches the personal loan information of the user and stores it in a vector database for efficient retrieval.

Several questions and answers from a typical conversation with the chatbot as a user are presented below.

1. **Scenario 1 (Existing User): Requesting Loan Repayment Schedule**

User: "Can you show me my loan repayment schedule?"

Chatbot: "Here is your repayment schedule:

- April 5, 2025 – LKR 30,000

- May 5, 2025 – LKR 30,000

- June 5, 2025 – LKR 30,000

Would you like a PDF copy emailed to you?"

2. **Scenario 2 (Bank Staff): Retrieving Customer Loan History**

User: "Show me the loan history for Customer ID 7865."

Chatbot: "Customer ID 7865 has:

- Home Loan: LKR 3 Million (Outstanding: LKR 1.5 Million)

- Personal Loan: LKR 500,000 (Fully repaid in 2023)

Would you like details on repayment schedules or late payment records?"

3. **Scenario 3 (Admin Inquiry): Managing Chatbot Responses**

User (Admin): "How many unresolved queries were logged today?"

Chatbot: "There were 15 unresolved queries today. Most of them were related to loan eligibility criteria. Would you like to review them?"

User: "Yes, show me the details."

Chatbot: "Here is a summary:

- 6 inquiries about missing documents for loan applications

- 4 inquiries about delayed loan approvals

- 5 general banking inquiries outside chatbot scope

Would you like to escalate any of these to support?"

C. Loan Eligibility and Eligible Loan Amount Prediction

The loan eligibility assessment follows a completely data-driven approach that processes the financial data of the user, including OCR-extracted payslip data, CRIB (Credit Information Bureau of Sri Lanka) records, and identification documents. The demographic information (age, employment type), financial information (monthly income, savings,

existing debt, other incomes, fixed deposits), credit score (loan repayment history from CRIB verification) are all taken into account and validated with a machine learning model utilizing decision tree algorithm to determine whether the user is eligible for the requested loan or not. This follows the workflow graphically represented in Fig. 3.

OCR, Optical Character Recognition technology, has been used in extracting data from national identity cards (NICs), pay slips, and other supporting documents submitted as attachments to the loan application. The text extraction of NICs has been performed with Tesseract OCR. The payslip information has also been cross-checked with customer-provided data, and the salary details have been validated prior to confirming eligibility. If the user is eligible for the loan requested, a machine learning model that utilizes the random forest regression predicts the eligible loan amount.
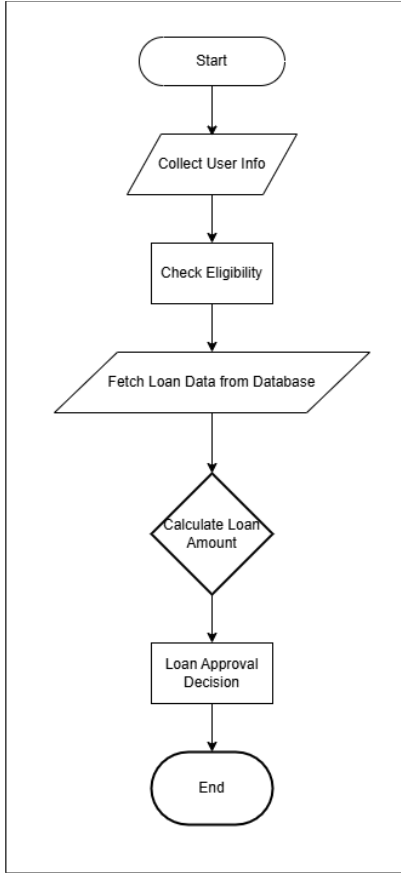

Fig. 3. Loan Eligibility Prediction Workflow

D. Loan Application Status and Approval Workflow

The profile section of the web application allows the users to track their loan application status in real time, and the approval process consists of several different steps. These include pending review (initial submission received), personal detail verification (cross-checking the banking information provided with the bank data), attachment analysis (OCR-based document analysis), CRIB verification (assessing the credit score and the past payment behavior), and bank officer verification (manual review for the final approval or rejection by an authorizes bank officer). At any stage, the application can be automatically moved to rejection based on a system-made decision if an issue is detected. In that case, the financial literacy and alternative guidance module provides solutions based on the issue that occurred.

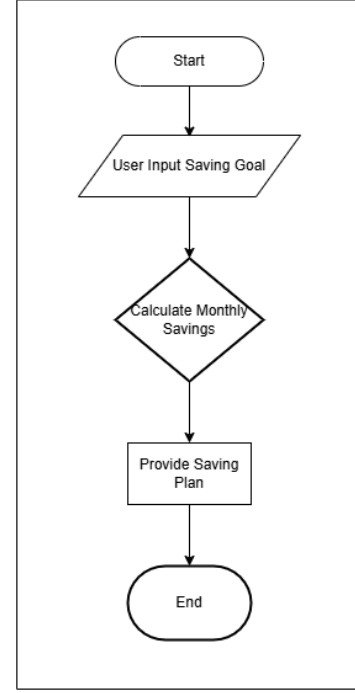### E. Financial Literacy and Alternative Financing Module

Fig. 4. Saving Plan Recommendation

In the case where an application rejection occurs based on the salary and savings-related eligibility criteria, the system offers a collateral-based solution that fits the use case. For instance, the next steps the user can take are suggested, including suggestions such as increasing the bank guarantee by increasing the fixed deposit amount, land revaluations to meet the LTV (Loan-to-Value) threshold for properties, and using vehicles or other assets as collateral. Apart from that, as Fig. 4 shows, a personalized savings plan for loan eligibility can also be suggested through the system.

Altogether, backend technologies like Python (Flask), TensorFlow, scikit-learn, and MySQL and frontend technologies such as React.js and Tailwind CSS have been used together with AI and NLP technologies such as Hugging Face Transformers (RAG) and Pinecone to develop the software solution.

IV. RESULTS & DISCUSSION

The loan eligibility prediction system has been evaluated based on its accuracy, efficiency, and user engagement across its three core components, AI chatbot, loan eligibility prediction, and financial literacy guidance in realizing the research objectives. The system has been tested against a dummy bank database for evaluating performance against real-world loan application scenarios and synthetic financial data to ensure comprehensive validation.

The AI chatbot's response system has been tested against a set of well-structured documents and a customer database using policy documents, loan agreements, and customer queries. The results have indicated a total of 92% accuracy in correctly answering loan-related questions for non-registered users and 87% accuracy when retrieving loan-specific details of authenticated users, and the response time has been recorded as under 2 seconds, ensuring quality real-time interactions.

The loan eligibility prediction component has been evaluated using standard classification metrics on a sample dataset of 1000 loan applicants. It has demonstrated a high accuracy of 87% when predicting the eligibility of loan applications after evaluating the personal data, and it later suggests a feasible loan amount based on the data.

In the context of rejected applications, it has been seen that the system provides alternative financing suggestions with high accuracy for users to opt for collateral-based solutions. From the data that has been tested against the system, common rejections have been caused due to reasons such as low salary (56%), poor CRIB history (22%), and incomplete documents (15%). 63% of the rejected users have reapplied using collateral of their choice (land, vehicles, etc.) to redo the application process with the suggestions provided through the system.

These results indicate that the system significantly contributes to improving loan accessibility, reducing processing time, and enhancing financial literacy, making it an effective solution for modern banking. This system has been deemed beneficial not just for the customers but also for banking staff, given the ease of use and automated evaluations provided with just a few simple clicks.

V. Conclusion & Future Work

This research has introduced a web-based, machine learning-powered loan eligibility system that has been carefully designed to enhance the loan processing efficiency, transparency, and financial literacy as a whole. The system has utilized AI chatbots, machine learning models, and OCR-based document analysis mechanisms to automate the eligibility assessment processes and provide alternative financing suggestions when applicants fail to meet conventional criteria. It acts as a customer support system for users to understand their loan eligibility and get needed guidance before going through the actual loan application process with a financial institution.

The software system that has been developed has demonstrated an accuracy rate of 92% for non-registered users seeking help with loan-related information in general and an accuracy of 87% in providing personalized responses for personal loan-related inquiries by authenticated users with the AI chatbot component. The loan eligibility prediction model has demonstrated intricate machine learning capabilities with both eligibility prediction and eligible loan amount prediction, going beyond the traditional rule-based systems. Furthermore, the financial guidance module has also played a crucial role in boosting financial literacy in customers, prompting 63% of the rejected applicants to reapply with collateral-based solutions, as suggested by the system.

With the incorporation of the latest technologies, this system has been capable of bridging the gap between applicants and financial institutions, fostering a more inclusive and efficient workflow by addressing common challenges faced with conventional loan processing procedures in the context of Sri Lankan banking systems.

Even though the system has displayed a significant performance and success, there are several steps that could improve the functionalities even further to bring the system to the next level. For instance, the data sources can be expanded for dynamic financial data retrieval by integrating real-time banking APIs, the chatbot can be enhanced for multilanguage support, and anomaly detection algorithms can be incorporated for potential fraud detection. These future improvements could further increase the potential for the system to grow and become a valuable asset in the financial sector with better accuracy, security, and scalability.

References

[1] A. Shaik, K. S. Asritha, N. Lahre, B. Joshua, and V. S. Harsha, "Customer Loan Eligibility Prediction using Machine Learning," Jun. 2022, [Online]. Available: https://publishoa.com

[2] M. A. Sheikh, A. K. Goel, and T. Kumar, *An Approach for Prediction of Loan Approval using Machine Learning Algorithm*. [IEEE], 2020.

[3] M. Al Mamun, A. Farjana, and M. Mamun, "Predicting Bank Loan Eligibility Using Machine Learning Models and Comparison Analysis," Jun. 2022.

[4] N. Uddin, M. K. Uddin Ahamed, M. A. Uddin, M. M. Islam, M. A. Talukder, and S. Aryal, "An ensemble machine learning based bank loan approval predictions system with a smart application," *International Journal of Cognitive Computing in Engineering*, vol. 4, pp. 327–339, Jun. 2023, doi: 10.1016/j.ijcce.2023.09.001.

[5] M. Mohankumar *et al.*, "COMPARATIVE ANALYSIS OF DECISION TREE ALGORITHMS FOR THE PREDICTION OF ELIGIBILITY OF A MAN FOR AVAILING BANK LOAN," 2016.

[6] T. Steijvers and W. Voordeckers, "Collateral and credit rationing: A review of recent empirical studies as a guide for future research," *J Econ Surv*, vol. 23, no. 5, pp. 924–946, Dec. 2009, doi: 10.1111/j.1467-6419.2009.00587.x.

[7] C. ; Kislat, L. ; Menkhoff, D. Neuberger, and Kiel, "The use of collateral in formal and informal lending Standard-Nutzungsbedingungen," Oct. 2013. [Online]. Available: https://hdl.handle.net/10419/85071

[8] A. Lusardi, O. S. Mitchell, and G. Washington, "Financial Literacy around the World: An Overview," 2011.

[9] A. S. Lusardi Olivia Mitchell *et al.*, "The Economic Importance of Financial Literacy: Theory and Evidence," 2013. [Online]. Available: http://www.nber.org/papers/w18952

[10] N. Garg and S. Singh, "Financial literacy among youth," 2018, *Emerald Group Publishing Ltd.* doi: 10.1108/IJSE-11-2016-0303.

.