Forecasting of Customer's Eligibility for The Bank Loan and, Also Predict Credit Limit, Using ML Approach

Project Id: Research Project - (24-25J-268)

Requirements Specification (SRS)

Student ID	Name
IT21164644	D N Pathirathna
IT21142178	Hilma M.I.F
IT21191060	P.R.I.Pravean

Supervised by

Suranjini Silva

B.Sc. (Hons) in Information Technology Specializing in Information Systems Engineering
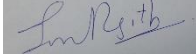
Department of Computer Systems Engineering

Sri Lanka Institute of Information Technology
Sri Lanka

08 December 2024

Declaration

We declare that this is our own work, and this proposal does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or institute of higher learning, and to the best of our knowledge and belief, it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

Name	Student Id	Signature
D N PATHIRATHNA	IT21164644	
Hilma M.I.F	IT21142178	
P.R.I.Pravean	IT21191060	

I.   Table of Contents

Table of Contents

Table of Figure

List of Tables

Acronyms

ML – Machine Learning
AI – Artificial intelligence
CRIB - Credit Information Bureau of Sri Lanka

1.Knowledge gap (Problem) with novel and creative solution

It's common knowledge that the lending business generates billions of dollars annually. Lenders profit from interest payments (Murfin and Petersen, 2014), so they provide competitive rates, while borrowers look for the most favorable conditions. Borrowers are often expected to disclose their income, assets, as well as credit score when applying for a loan. Lenders and borrowers alike may find this procedure to be tedious and aggravating. This essay will go over the usage of artificial intelligence and machine learning to gauge loan qualifications. As data scientists, it's crucial to have a firm grasp of the difficulties associated with determining loan eligibility (Azzini et al., 2019) , as well as the means by which these obstacles might be overcome via the development of machine learning models. We'll show how machine learning may be utilized to benefit both lenders and applicants in the loan application process (Padimi, .. and Ningombam, 2022).

Eligibility for a loan refers to the simple procedure of figuring out whether a possible borrower satisfies the requirements for that loan. This may depend on things including a person's salary, length of work, as well as credit history. Eligibility for loans is crucial in guaranteeing that borrowers will be able to pay back their debt. The default might have serious implications, such as a drop-in credit score and trouble getting new funding in the future. Lenders thus use extreme caution when determining loan suitability. Borrowers may increase their chances of loan approval by learning more about the criteria used by lenders. Lenders will typically employ a combination of criteria to establish a borrower's suitability for a loan, since there are many factors that might influence this decision. Copying the borrower's credit report is a typical practice. A borrower's current debt load and payment history may be seen in this document. Lenders will also research the borrower's credit history and other public sources (Barron, 2000).

Customers who are likely to default on investment and business loans may be identified with the use of feasibility analysis, financial documents, and research into the company's and borrower's reputations in the marketplace. In order to make sound lending decisions, banks offering private loans must collect detailed information on prospective borrowers and keep close tabs on their behavior throughout the life of the loan. Banking industry Score Company includes details on the credit history of the clients and moreover helps to prevent default customers via the use of a scoring system (Gallati, 2003). Due to data scarcity and the lack of reliable methods for interpreting client credit histories, this is easier said than done.In Sri Lanka, a customer must wait several days to see if he or she eligible for a loan. As a result, they are not actual idea about they are eligible or not for the loan process also they didn't have actual idea about loan amount the bank will allocate for them.

Nowadays Banking sector collaborate with IT industry and they experienced new technology for their services (Ho et al., 2019). According to the ML and data science, banking sector can implement loan eligibility predictor for their customer. Another problem addressed in this project is the lack of an advanced, AI-powered solution for loan eligibility assessment in the financial sector, particularly within emerging markets. Present systems frequently lack efficiency, personalization, and the necessary capabilities to navigate the varied and ever-changing financial environments present in emerging markets. Financial institutions often depend on traditional techniques that do not offer immediate, precise evaluations, resulting in inefficiencies and possible discrepancies in loan authorizations. The deficiency in current systems prevents financial institutions from efficiently serving their customers and restricts credit access for individuals, especially those with non-traditional financial histories.

The aim of this research project is to create an AI assistant chatbot tailored to tackle these obstacles through offering a more comprehensive, clear, and effective loan eligibility evaluation procedure. The plan entails using sophisticated data analysis, machine learning methods, and intuitive interfaces to guarantee precise, unbiased, and personalized loan eligibility evaluations for each customer. This project seeks to improve financial inclusion and enhance the efficiency of loan processing in emerging markets by tackling the drawbacks of existing systems

In Scope

- Identify the most suitable algorithms and train with the date set and identify best model for the prediction.
- Predict loan amount after consideration of input parameters and using ML algorithms.
- Change the model and prediction approach with economic patterns.

- All the input parameters are highly restricted and consider privacy high improperness

Out Scope

- Eligible customer can apply loan though the system
- The proposed system can implement or other organizations to consider their employees financial situations.
- Introduce loan guarantors to the system and identify guarantors who are eligible or not (Bloomenthal, 2022).
- CRIB (Credit Information Bureau of Sri Lanka) integrates with the system.
- Implement centralized system and connected all the financial organizations to the system and full fill the customer requirement according to their financial situation.

"The project's aim is to design, develop and enhance forecasting of customer's eligibility for the bank loan and, also Predict credit limit, using ML approach"

The interactive tools and calculators in this system provide real-time assessments of the user's financial health, foreseeing potential gaps in their financial literacy and giving advice tailored to these gaps. This approach does not only qualify customers for loans but also makes them realize the consequences of their decisions on borrowing. Therefore, this system will facilitate responsible and informed borrowing by ensuring that each customer has appropriate financing options that best meet their financial goals and hence no potential mismanagement of debt leading to financial distress[2]. Through KYC data, financial institutions will be better placed to provide customized advice in view of every customer's financial situation to set feasible financial targets and ensure prudent debt management and sensible borrowing decisions. Further, interactive financial education that is inbuilt in the lending platform may raise customers' levels of financial literacy with core concepts, thus better arming them to tackle the pitfalls associated with personal finance. The tools may provide real-time financial decision support through budgeting evaluations and loan prediction calculators, allowing the customer to make informed decisions on the spot.

1. Developing detail-oriented educational modules on the concepts of personal finance, which involve budgeting, saving, investing, and repayment planning among others.
2. Use KYC data to power simulation-based, calculator-based, and other interactive tools to dynamic, context-driven suggestions in real-time about finances.
3. Collate the resources for education, individual counseling, and predictive analytics into a system that seamlessly transitions into the loan eligibility system.

2. Existing systems and related work

Citation	Description	Improvements	Limitations
Loan Prediction Using Machine Learning and Its Deployment on Web Application (Sujatha et al., 2021)	The loan prediction is developed using machine learning algorithms such as logistic regression. The Python programming language is used or the implementation of the code which has been developed in Collab and the html pages are developed for deployment of website using Visual Studio code.	Logistic Regression and Decision tree algorithms have been used to predict the results The python is used to implement the code in Collab. Flask and Heroku app are used to deploy the project. In order to enhance the operating speed of the proposed system.	This work can be extended in order to improve the focus where the high accuracy can be obtained. System only predicts the eligibility of the customer. In this paper predicting the loans according to the dataset taken. Different countries have different attributes as their priority, those attributes can be included in the dataset while training according to their bank's choice
Loan Prediction System Using Decision Tree and Random Forest Algorithms (Chaudhary, 2020)	Logistic Regression, Decision Tree, and Random Forest is applied to this prediction s. Through this system, one can predict if that specific candidate is protected and the entire interaction of approval of highlights is robotized by the ML method.	The outcome of the research is in 79.45% of accuracy achieved. Used two Algorithms to get an accurate prediction.	Consider few utilizes models and also dataset Outdated with considering new patterns. Required larger training sets to give accurate results Only

An Empirical Study on Loan Prediction Using Logistic Regression and Decision Tree (Aditya Sobika et al., 2021)	Logistic Regression, and Decision tree examination used to predict eligibility of loan	According to the paper researchers applied a novel approach for credit scoring by combining genetic programming with deep neural networks.	Required larger training sets to give accurate results there are less than 10 explanatory attributes the accuracy of SVM is reduced.
Prediction of Modernized Loan Approval System Based on Machine Learning Approach (Singh et al., 2021)	In this research paper, using three Machine Learning algorithms. prediction of Data set. XGBoost – XGBoost is a Decision tree based open-source software library. Random forest Decision tree.	proposed model will characterize the behavior of customers on the Basis of their record.	Only show client is eligible or not. According to the paper in some situations like client going through some disaster algorithm cannot predict the appropriate result. Consider few utilizes models and also dataset Outdated with considering new patterns.
Prediction of loan status in commercial bank using machine learning classifier (Arutjothi and Senthamarai, 2017)	According to the paper K-NN credit scoring model was built with the programming language R.	Credit scoring is widely analyzed using classification techniques. Feature selection techniques are used to remove the irrelevant attributes.	Loan applicant should be default or valid customer. Only show client is eligible or not.
Customer Loan Eligibility Prediction using Machine Learning Algorithms in Banking Sector (Kumar et al., 2022)	Feature selection method is applied to select more appropriate features to predict the customer loan eligibility according to the research Decision tree, random forest, support vector machine, k-nearest neighbor, and hybrid model with decision tree and adaboost combination are implemented in this work.	Python PYSPARK machine learning package is used to train the model. Used three Algorithms to get an accurate prediction.	Required design an innovative model for customer loan eligibility prediction over banking sector. The model's results are analyzed in terms of accuracy parameter, and performance is represented in graphical format.

Table 1 : Existing Technologies
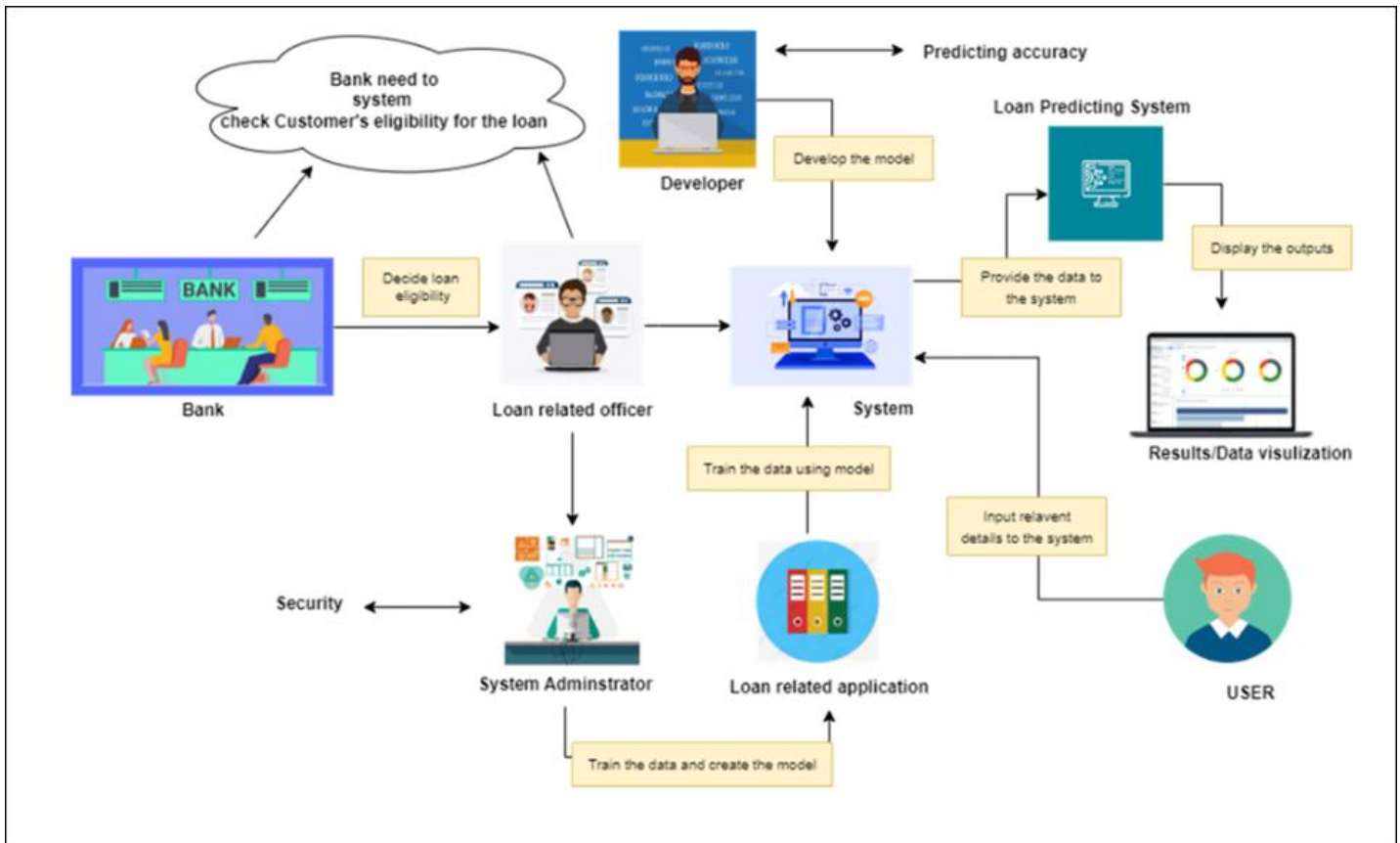
3. System Design

Figure 1 : High level system diagram

The QuickLoan system is an advanced solution aimed at optimizing the loan approval process for banks through the use of machine learning and artificial intelligence. It comprises three key components such as Loan Eligibility Prediction, Financial Advisor, and AI Assistant. Together, these components improve the accuracy of loan assessments, explore alternative approval methods for rejected applications, and provide quick assistance to both bank staff and customers.

The Loan Eligibility Prediction component is the core of the system. A machine learning model, developed by the system's developer using historical loan data, forms the backbone of this functionality. Loan-related officers input customer details into the system, which then evaluates the information to determine loan eligibility with a high level of predictive accuracy. The results are displayed through a user-friendly data visualization interface, allowing both officers and customers to clearly understand the decision.

The second component, the Financial Advisor, is designed to minimize loan rejection rates. When a loan application is declined by the prediction system, the Financial Advisor evaluates additional data, such as the customer's KYC (Know Your Customer) information, to identify alternative ways to approve the loan. By recommending personalized financial solutions, this component make sures to give better customer retention and satisfaction.

The AI Assistant serves as an intelligent support tool for two user groups such bank staff and customers. For bank staff, the assistant retrieves detailed customer profiles, loan histories, and repayment plans, allowing staff

to access the required information without manual searches. For customers, the AI Assistant provides instant answers to queries about the loan process, upcoming payment schedules, and repayment plans, providing a smooth customer experience.

A System Administrator plays an important role in maintaining the overall system functionality. This includes managing the security of the platform, as well as overseeing the Loan-related Application, which handles the training and updating of the machine learning model. Security protocols are available throughout the system to protect customer data and compliance with banking regulations.

In terms of data flow, the system begins with customer and loan-related data input by loan officers. This data is processed by the machine learning model, which generates predictions on loan eligibility. If a loan is rejected, the Financial Advisor steps in to evaluate alternative solutions. The outcomes are visualized through intuitive dashboards, making them easily accessible to both customers and staff.

4.Functional Requirements / Non-Functional Requirements

<u>**Functional Requirements**</u>

1. Loan Eligibility Prediction

- Description: Utilize machine learning to analyze users' data to forecast eligibility for a loan and its maximum amount.
- Inputs: User financial profiles, collateral values, and historical economic trends.
- Outputs: A prediction of loan eligibility and the recommended loan amount.
- Use Case: Somebody applies for a loan, and the system predicts eligibility while suggesting how to improve the chances of getting approved.

2. Loan Status Notification and Updates

- They collect, analyze, and ultimately use the KYC data to build strong financial profiles of the users.
- Description: Inform users of application status and necessary steps to move forward.
- Functionality:
  Real-time alerts on loan application progress.
- Messages of documents missing or actions require.

3. KYC Data Integration

- They collect, analyze, and ultimately use the KYC data to build strong financial profiles of the users.
- Inputs: User demographics, employment status, income sources, liabilities, assets, and past financial transactions.
  Outputs: A full financial profile for loan eligibility assessment and personalized financial recommendations.
- Application   Scenario:
  A user uploads their KYC information, including employment and financial history.
- It analyzes this data in order to forecast loan eligibility and offers recommendations for improving financial well-being.

4. Collateral Valuation Module

- Description: The system shall assess collateral options offered by the user, such as property, savings,

gold, or any other asset.

- Inputs: Info regarding assets, land value, savings accounts, physical assets, etc. Outputs: The projected value of collateral and this appropriateness for satisfying loan criteria.
- · Application Scenario:
- Somebody provides their asset information. It computes the collateral value and suggests any other collateral that may be required.

5. AI Support for End User Interactions

- It is essential to record user interactions with the system to guarantee traceability and support prospective improvements.
- Outputs: A secure, time-stamped record of user actions.

Non-Functional Requirements

1. Scalability
The system should be able to handle increased users' demands without performance degradation, ensuring smooth operation with up to 100,000 concurrent users.

2. Usability
The interface should be user-friendly with intuitive navigation and multilingual support to handle diverse user bases.

3.Dependability
To Ensure on time and uptime and smooth operation with automated failover mechanisms in the event of system failure.

4. Efficacy
Loan eligibility predictions should be generated within 2 seconds of input submission. Security and Privacy.

5. Encryption

The user's data should be encrypted both in transit and at rest.

6.Personalization

 Allow financial institutions to revise their models and predictive algorithms, with respect to economic trends and market movements.

 7. Accessibility

 The platform should be accessible on different devices such desktops following the Web Content Accessibility Guidelines (WCAG).

5. External Interfaces

Login page

Figure 2 : Login Page

Registration Page

Figure 3 : Registration page

Application Form

Figure 4 : Application

Figure 5 : Application

6.Application of key pillars in the specialized area of knowledge

The Loan Eligibility and Prediction System is a cutting-edge technology that automates the loan application process with machine learning (ML) and artificial intelligence (AI). This technology not only determines loan eligibility, but also forecasts results based on a thorough examination of applicants' financial information. By including an AI chatbot, the solution improves the client experience by offering real-time assistance and guidance during the loan application process. Customers can communicate with the chatbot to get fast answers to their questions and help navigating the application process, significantly reducing wait times and increasing satisfaction.

In addition to serving clients, the system offers a bank staff login feature that provides loan officers and financial analysts with crucial data and insights. This feature allows bank employees to assess applications, monitor loan statuses, and make informed decisions based on real-time statistics. The automated nature of the system streamlines operations, allowing employees to concentrate on higher-value tasks rather than human data entry.

Another important component of the system is the job of the financial advisor, which is specifically created to assist consumers who may be rejected throughout the loan application process. This role is critical to the bank's administrative operations since it provides individualized help and guidance to rejected applicants. The financial advisor can investigate the reasons for rejection and advise clients on how to improve their financial profiles, boosting their chances of future acceptance. This proactive strategy not only increases customer loyalty, but it also strengthens the bank's reputation for assisting clients with their financial journeys.

Overall, the Loan Eligibility and Prediction System, powered by ML and an AI chatbot, revolutionizes the loan application experience. By automating eligibility assessments and providing real-time customer support, it significantly enhances user engagement and operational efficiency. Meanwhile, the financial advisor feature ensures that rejected applicants receive the necessary guidance to improve their financial standings, ultimately contributing to a more inclusive banking environment. This comprehensive system positions banks to better serve their customers while optimizing internal processes and decision-making capabilities.

7. Application of technologies in the relevant key pillar/area

Software Requirements

Software Requirement	Justification
Operating System (macOS /Windows 10 /Windows 11)	The OS is required to manage all processes and able to run required programs such as IDEs, Libraries, and Tools
Python/Angular/.NET/Gemini API	Python will be used as the primary programming language to implement and develop prediction model using ML and data science Angular will be used as the primary programming language to implement front end. .NET will be used as the primary programming language to implement Backend end
PyCharm IDE / Jupiter IDE	IDEs used for implement Code for the systems

IntelliJ IDEA/ Visual Studio Code/Note Pad	
Google Drive/Git Hub/Git Lab	Will be used for Document and Code backups
SQL Database - MYSQL	SQL Database will be used as the primary Data base management system.
MS Office/ Mendeley/ Google doc	A documentation, reporting and researching tools use manage documentation and backup purpose.
Figma/Draw. Io	A designing tool used to create diagrams, wire frames and data visualization charts.
Google Collab	Google Collab will be used as cloud development environment to build, train & test ML & Deep Learning models.

Table 2 : Software requirement

Hardware Requirements

Hardware Requirement	Justification
Core i5 Processor or above	To perform long running, Intensive processing power.
16GB RAM	To manage data-sets & Setup local development environments.
Disk space 100GB or above	To store datasets, files, applications, IDE and frameworks necessary for development purposes.
Screen resolution 1360*768	To visualized data and display outputs

Table 3 : Hardware requirements

8. Limitations

Description:	Implications	Approach to Mitigate
Memory Constraints	The system deals with huge data, including KYC information, financial records, historical trends, and real-time contributions of users. The in-memory management and storage of this huge data for fast access could become a big challenge.	High memory use at peak times can slow down system performance. Resource-intensive operations, such as training machine learning models, may require external infrastructure. Use distributed data storage systems, for example, cloud-based databases and Optimize memory usage by efficient data compression and processing
Performance Constraints	The potential delays in providing real-time financial advice or loan eligibility predictions during peak usage times. Latency is higher when interfacing with external systems, like financial institutions or credit bureaus.	Implement caching mechanisms for frequently accessed data. Implement load balancing and horizontal scaling to effectively allocate system workload.
Data Integrity and Accessibility	The system is heavily dependent on accurate and comprehensive user data, which includes KYC details, income records, and asset valuations. Prediction accuracy may be affected by data gaps, inaccuracies, or inconsistencies.	Implement data validation checks and imputation methods to handle missing data. Leverage alternative data sources, such as mobile money transaction records or community-based credit scoring.
Reliance on Economic Cycles	This may impact the predictive accuracy of the system due to unforeseen economic changes such as inflation, recession, or sudden regulatory changes. Loan predictions and financial advice may not reflect the latest economic conditions. Challenges in relevance in dynamic markets	Regularly retrain predictive models with updated economic data. Integrate scenario-based simulations to dynamically tweak guidance.

| Low financial literacy of users | It assumes users have a basic understanding of financial concepts, which is not always the case. This might limit the effectiveness of the learning modules. Hard to get users with limited literacy or language barriers engaged. Lower acceptance among the less educated strata of society. | We provide multilingual support and simplified explanations of financial terms. Use gamification to make learning more engaging and accessible |

Table 4 : Limitations

9. Reference

1. R. Manglani and A. Bokhare, "Logistic Regression Model for Loan Prediction: A Machine Learning Approach," 2021 Emerging Trends in Industry 4.0 (ETI 4.0), May 2021, doi: https://doi.org/10.1109/eti4.051663.2021.9619201.

2. C. N. Sujatha, A. Gudipalli, B. Pushyami, N. Karthik, and B. N. Sanjana, "Loan Prediction Using Machine Learning and Its Deployement On Web Application," IEEE Xplore, Nov. 01, 2021. https://ieeexplore.ieee.org/abstract/document/9696448

3. K. Aditya Sobika, M. Jamuna, H. Limitha, M. Saroja, and S. Shivappriya, "An Empirical Study on Loan Prediction Using Logistic Regression and Decision Tree," 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Oct. 2021, doi: https://doi.org/10.1109/icaeca52838.2021.9675505.

4. G. Arutjothi and C. Senthamarai, "Prediction of loan status in commercial bank using machine learning classifier," IEEE Xplore, Dec. 01, 2017. https://ieeexplore.ieee.org/abstract/document/8389442

5. Ch. Naveen Kumar, D. Keerthana, M. Kavitha, and M. Kalyani, "Customer Loan Eligibility Prediction using Machine Learning Algorithms in Banking Sector," 2022 7th International Conference on Communication and Electronics Systems (ICCES), Jun. 2022, doi: https://doi.org/10.1109/icces54183.2022.9835725.

6. Azzini, S. Marrara, Amir Topalović, Mirjiana Pejic Bach, and M. J. Rattigan, "Opportunities and Risks for Data Science in Organizations: Banking, Finance, and Policy - Special Session Overview," Oct. 2018, doi: https://doi.org/10.1109/dsaa.2018.00078.

7. Padimi, Venkata Sravan, and D. D. Ningombam, "Applying Machine Learning Techniques To Maximize The Performance of Loan Default Prediction," Journal of Neutrosophic and Fuzzy Systems, pp. 44–56, 2022, doi: https://doi.org/10.54216/jnfs.020204.

8. Alrabiah and M. M. Elgammal, "Optimal regulation of banking system's advanced credit risk management by unified computational representation of business processes across the entire banking system," Cogent Economics & Finance, vol. 6, no. 1, p. 1486685, Jan. 2018, doi: https://doi.org/10.1080/23322039.2018.1486685.

9. Natasha Robinson and Nidhi Sindhwani, "Loan Default Prediction Using Machine Learning," [online]https://ieeexplore.ieee.org/document/10522232, [Accessed 14 May 2024]

10. H. K. Sarisa, V. Khurana, V. C. Koti, and N. Garg, "Loan prediction using machine learning," [online] https://ieeexplore.ieee.org/abstract/document/10466049 [Accessed 20 May 2024

10. Conclusion

In summary, the Loan Eligibility and Prediction System is a game-changing breakthrough in the banking and financial services industries. This solution automates the loan eligibility evaluation process using advanced machine learning and artificial intelligence technologies, delivering real-time support via an AI chatbot. This not only improves the customer experience by providing fast support, but it also streamlines processes for bank employees, freeing them to focus on more strategic responsibilities.

The addition of a dedicated financial advisor feature improves the system's functionality by giving targeted assistance to consumers who may receive loan denials. This proactive strategy helps applicants understand the reasons for their refusal and provides practical information to enhance their financial profiles, establishing an inclusive and customer-empowering culture.

Overall, the use of this system enables financial institutions to better meet the requirements of their customers while streamlining internal processes and decision-making capacities. By embracing these new solutions, banks may increase customer happiness, operational efficiency, and, ultimately, generate growth in an increasingly competitive market. The Loan Eligibility and Prediction System not only meets modern banking requirements, but it also establishes an innovative standard for customer service and engagement in the financial industry.