# Smart Lending Ally : AI Powered Loan Recommendation System


Rahul Rajasekharan Menon
*Department of Computer Science and Engineering,*
*Faculty of Engineering and Technology,*
*Jain (Deemed-to be University)*
Bengaluru, India
22btrcn222@jainuniversity.ac.in

Nikhil Reddy K
*Department of Computer Science and Engineering,*
*Faculty of Engineering and Technology,*
*Jain (Deemed-to be University)*
Bengaluru, India
22btrcn188@jainuniversity.ac.in

Nakul Govind S
*Department of Computer Science and Engineering,*
*Faculty of Engineering and Technology,*
*Jain (Deemed-to be University)*
Bengaluru, India
22btrcn180@jainuniversity.ac.in

Tejas Rao M N
*Department of Computer Science and Engineering,*
*Faculty of Engineering and Technology,*
*Jain (Deemed-to be University)*
Bengaluru, India
22btrcn296@jainuniversity.ac.in

Vishal Kumar P
*Department of Computer Science and Engineering,*
*Faculty of Engineering and Technology,*
*Jain (Deemed-to be University)*
Bengaluru, India
22btrcn194@jainuniversity.ac.in

Dr. A. Balajee
*Department of Computer Science and Engineering,*
*Faculty of Engineering and Technology,*
*Jain (Deemed-to be University)*
Bengaluru, India
balajee.a@jainuniversity.ac.in



***Abstract*— The swiftly evolving financial environment in India, marked by rising inflation and increasing interest rates on loans, presents considerable difficulties for individuals looking for the best lending options. Many borrowers, lacking a thorough understanding of their available choices, tend to rely on nearby banks and can be misled by clever marketing strategies and solicited agents. This disparity in information often results in poor financial choices that can have lasting adverse effects. To tackle this pressing issue, we introduce Smart Lending Ally, an AI-driven loan recommendation platform that connects borrowers with appropriate lending institutions. Our system utilizes an innovative dual-model classification methodology that integrates XGBoost and Random Forest algorithms to ascertain loan eligibility across ten prominent Indian banks, achieving an impressive average accuracy of 98.33%, which significantly surpasses the typical 70-87% accuracy rates of existing single-algorithm recommendation systems. In contrast to former methods that focus purely on predicting eligibility, Smart Lending Ally features a holistic multi-criteria suitability scoring system that assesses various elements such as interest rate advantages, processing speed, and affordability measures. The platform uses dynamic weighting tailored to five types of loans (Home, Car, Education, Personal, and Gold), leading to highly customized recommendations. Our comparative validation reveals a recommendation precision of 96.7% when matched against expert evaluations in the field, setting a new standard for loan recommendation systems. The architecture includes synthetic data generation that reflects realistic financial scenarios in India, a comparative model selection for determining eligibility, and an easy-to-use visual interface for interpreting results. Experimental tests involving different user profiles—high-income professionals, recent graduates, self-employed persons, and retirees—confirm the system's efficacy in real-world applications. Smart Lending Ally aspires to make optimal lending options accessible by delivering transparent, equitable, and personalized loan recommendations based on users' financial situations, ultimately promoting better-informed financial decision-making among consumers in India.**



***Keywords*— *Machine learning, financial recommendation systems, dual-model classification, eligibility prediction, suitability scoring, explainable AI, personalized finance.***


## I. Introduction

In today's society, rapid advancements in technology and the widespread implementation of digital solutions characterize the landscape. One of the most significant developments in this context is the emergence of Artificial Intelligence (AI), which has profoundly influenced nearly all industries. The lending services sector, especially within finance, has seen major changes as AI transforms how borrowers obtain and manage financial resources. Loans serve as vital financial instruments that help individuals, businesses, and governments fulfill their economic aspirations and requirements. However, as the global financial environment grows more complex, selecting the appropriate loan provider has become increasingly challenging. Numerous factors shift with economic fluctuations, such as interest rates, repayment conditions, eligibility criteria, and necessary documentation. A 2023 report from the Reserve Bank of India (RBI) noted that interest rates for similar loan products can vary by as much as 4-6 percentage points between different financial institutions in India, resulting in significant cost implications for borrowers. A substantial portion of the Indian populace possesses insufficient understanding of banking practices and loan procedures, with financial literacy levels estimated at merely 27% according to the National Centre for Financial Education (NCFE). This knowledge gap underscores the need for the development of unbiased, transparent systems that can inform and support prospective borrowers during the loan selection journey. In the current rapidly changing economic landscape, compounded by fluctuating inflation rates, individuals seeking loans face major challenges in deciding on the appropriate loan category. Borrowers encounter a perplexing array of loan choices, each characterized by distinct interest rates, terms, and conditions that require careful consideration. This challenge is intensified by the prevalent lack of clarity surrounding the lending process, ongoing information discrepancies, and an absence of neutral guidance for potential borrowers. A significant segment of the Indian population possesses little understanding of the complexities involved in borrowing practices, making them vulnerable to the persuasive marketing strategies employed by banks and other financial institutions. Research conducted in 2022 by the Financial Inclusion Insights (FII) program revealed that 68% of Indians who secured loans did not compare offerings from multiple institutions, primarily due to information barriers and the complexity of the processes. Many borrowers succumb to sophisticated sales tactics, ultimately making decisions based on convenience or persuasion rather than optimal financial outcomes. Such circumstances undermine customer trust and often lead to poor decision-making, which can increase borrowing costs and result in unsatisfactory loan experiences that negatively impact long-term financial health. Traditionally, the financial services sector has relied on effective and precise approaches to assess creditworthiness and suggest appropriate loan options for borrowers. While conventional lending methods were functional, they frequently proved to be slow, susceptible to human error, and lacked the personalization needed to meet the varied needs of borrowers. Recently, the advent of artificial intelligence and progress in machine learning have ushered in a transformative phase for lending. AI's capability to analyze extensive datasets, recognize patterns, and make instantaneous decisions has generated interest in overhauling the lending landscape. Historical methods of loan recommendation have mainly concentrated on either rule-based expert systems or independent machine-learning models. For example, Kumar et al. (2021) developed a loan recommendation system using decision trees, whereas Mehta and Singh (2022) investigated neural network techniques for credit scoring. Nonetheless, these systems typically focus on either predicting eligibility or conducting basic interest rate comparisons, failing to combine various elements into a holistic suitability framework. The traditional approach to selecting loans is impeded by imbalances in information, limited visibility in the market, and sometimes exploitative practices that can lead to negative financial results for borrowers. While the banking industry is evolving with the advent of artificial intelligence, these advancements have not sufficiently addressed the fundamental issues that everyday borrowers encounter. In our interconnected digital world, AI has demonstrated its potential to transform multiple industries, yet its impact on making loan selection more accessible remains limited. AI aids borrowers who have struggled to secure loans, promoting a more inclusive financial environment. The swift progression of AI has revolutionized various fields, including finance. Recently, systems powered by AI have shown their efficacy in improving the efficiency and accuracy of loan recommendation processes. The concept behind Smart Lending Ally stems from the desire to innovate the lending sector and address the challenges that both borrowers and lenders face within traditional lending frameworks. Typical loan assessment practices often require extensive documentation, lengthy approval times, and repetitive processes. Standardized loan products may not cater to the specific financial needs and risk profiles of individual borrowers. Conventional lending practices have faced criticism for potential biases, whether deliberate or accidental, that affect loan approval decisions. This paper explores several key questions: How can we navigate the increasing complexity of lending processes while considering rising inflation and the challenges of financial literacy? What methods can equip borrowers with the essential knowledge and tools needed to make informed, responsible, and economical financial decisions? How can we enhance AI capabilities to create solutions that are transparent, trustworthy, unbiased, and intelligent, thereby simplifying loan selection within the financial landscape? Despite notable progress in AI applications within the financial sector, considerable deficiencies persist in the ongoing research related to loan recommendations. A thorough analysis by fintech researchers found that while areas like stock recommendations and peer-to-peer lending have garnered significant focus, key aspects such as an extensive evaluation of loan suitability are still inadequately explored. Most existing systems primarily concentrate on predicting eligibility rather than conducting a comprehensive assessment of how well a loan fits the borrower's needs. Moreover, recent studies highlight a troubling lack of synergy between rule-based methods and machine learning techniques, with most research adopting one of the approaches separately and overlooking the potential benefits of combined strategies. Perhaps most importantly, few studies take into account the varying importance of different factors according to loan types—an essential component for genuinely personalized recommendations that reflect the distinct priorities associated with home, car, education, personal, and gold loans. Smart Lending Ally employs cutting-edge machine learning algorithms like XGBoost and Random Forest models, combined with advanced data analytics and real-time market insights, to empower borrowers by broadening their options. This system considers a range of factors, including creditworthiness, loan type specifics, interest rate variations, processing efficiency, and individual financial needs to pinpoint and recommend the most advantageous lending choices. Our unique dual-model classification approach and dynamic weighting methodology allow for tailored recommendations across five key loan categories: Home, Car, Education, Personal, and Gold loans. This strategy provides consumers with the vital knowledge and resources necessary for making well-informed decisions that enhance their long-term financial well-being. By utilizing big data and advanced analytics, Smart Lending Ally can perform comprehensive credit risk assessments with greater accuracy. This assists lenders in making more educated decisions, minimizing the chances of defaults, and improving overall portfolio performance. Smart Lending Ally employs an AI-driven approach that enables the platform to assess extensive datasets and provide tailored loan suggestions based on each borrower's unique financial circumstances, goals, and credit ratings. Our trial evaluations with a diverse range of user profiles—such as high-earning professionals, recent graduates, freelancers, and retirees—demonstrate the system's flexibility in accommodating various financial situations. By utilizing AI algorithms, Smart Lending Ally aims to reduce human biases and deliver fair and unbiased loan recommendations to borrowers, creating a more just financial landscape for all participants. The platform's comprehensive visualization tools enhance transparency by clearly illustrating the factors influencing recommendations, thereby fostering trust in the lending process. Our research findings reveal that the system achieves impressive accuracy in determining eligibility (averaging 98.33% across ten leading Indian banks) while providing detailed suitability ratings that exceed basic eligibility assessments. This AI-driven platform benefits both consumers and financial institutions by fostering trust and efficiency within the banking ecosystem. By embracing Smart Lending Ally, we create a mutually advantageous relationship between technology and finance, paving the way for a data-driven financial revolution characterized by intelligence and consumer empowerment. The system encourages informed decision-making and responsible choices that benefit borrowers while enhancing the overall integrity of the financial sector. The implications of this research extend beyond theoretical advancements to practical applications that can significantly improve financial decision-making for millions of consumers in India, while also providing financial institutions with more effective customer acquisition tactics based on true suitability alignment rather than aggressive marketing methods.

This research contributes in several key ways:

- Created an innovative dual-model classification technique that utilizes both XGBoost and Random Forest algorithms, choosing the model based on performance metrics, resulting in an outstanding eligibility prediction accuracy of 98.33% among ten major banks in India—significantly surpassing the performance of current single-algorithm systems.
- Developed an all-encompassing multi-criteria suitability scoring framework that not only looks at basic eligibility but also considers factors such as interest rates, processing times, and loan-specific variables, featuring dynamic weighting systems customized for five different loan types (Home, Car, Education, Personal, and Gold), thereby offering borrowers truly tailored recommendations that reflect their individual financial situations.
- Introduced a comprehensive visualization interface that simplifies complex financial information into user-friendly, interpretable formats, allowing users to grasp not just which banks are recommended, but also the reasoning behind those recommendations, thereby promoting informed decision-making through explanations rather than just directives.
- Our thorough validation across a variety of user profiles showed the system’s flexibility and effectiveness in practical scenarios, positioning Smart Lending Ally as a viable solution to widen access to the best lending options for the Indian population, regardless of their financial status.

## II. LITERATURE SURVEY

In recent years, various researchers have examined the use of artificial intelligence and machine learning within financial services, with a particular emphasis on systems for loan recommendation and assessment. This literature review compiles findings from fifteen pertinent studies to lay the groundwork for our Smart Lending Ally system.

### *A. Machine Learning Approaches in Loan Recommendation*

Numerous studies have analysed machine learning techniques to predict loan eligibility. The KNN-based agricultural loan recommender system developed by the authors in [1] achieved a moderate accuracy of 68.67%, underscoring the difficulties associated with single-algorithm methodologies. More encouraging results were found using ensemble learning strategies, as shown in [2], where a voting ensemble method reached 87.26% accuracy in predicting loan approvals. These results imply that integrating multiple models can considerably improve prediction accuracy in comparison to single-algorithm methods.

The LAPS (Loan Risk, Activity, Profile, and Social Recommendation) model introduced in [3] provided a thorough method for assessing borrower reliability by incorporating personal details, financial circumstances, and social endorsements. In a similar vein, the authors in [4] created a bank loan prediction system utilizing SVM and Neural Networks, which effectively classified applications as either secure or risky while significantly decreasing processing times.

### *B. Feature Engineering and Selection*

Feature selection has surfaced as a vital element across several studies. Research in [5] pointed out the necessity of assessing borrowers' income, debt levels, and credit histories as crucial factors in the credit evaluation process. The feature importance analysis detailed in [6] indicated that CIBIL scores, loan-to-income ratios, and employment characteristics play a significant role in determining loan eligibility decisions.

The AHP-GRAP model suggested in [7] successfully merged qualitative and quantitative data for credit evaluation by combining the Analytic Hierarchy Process with Grey Relational Analysis. This strategy illustrated the benefits of integrating various data types to tackle the issue of incomplete information in credit evaluations.

### *C. User Experience and Interface Design*

Multiple studies stressed the significance of user experience in systems that provide financial recommendations. The intelligent chatbot investigation presented in [9] showcased how AI-powered conversational interfaces can enhance banking experiences by delivering prompt and effective responses to customer queries. Research in [11] introduced a case-based reasoning method that assisted bank consultants in making quicker, more informed decisions by utilizing relevant past cases.

Particularly noteworthy was the research conducted in [12], which centered on the explainability of loan recommendations. This study established a framework for generating natural language explanations and counterfactual guidance for loan applicants, thereby enhancing transparency and trust in the recommendation process. The system employed Shapley values to identify significant features and translate them into accessible explanations for users.

### D. *Personalization and Specialization*

Research referenced in [13] established a model for success credibility pertaining to loans in rural communities, indicating that higher credibility scores are linked to better borrower recommendations. The study underscored the necessity of adaptability in monitoring borrower contributions and offering customized recommendations. Notably, the book recommendation research in [14], though unrelated to financial services, provided insightful information on merging various data sources (such as library loan histories and bibliographic data) to enhance the accuracy of recommendations. Their SVM-based methodology reported a positive evaluation rate of 84.2%, surpassing conventional techniques.

### E. *Implications for Smart Lending Ally*

This literature review points out several important implications for our research. Firstly, the effectiveness of ensemble learning methods suggests that our dual-model classification system can greatly enhance prediction accuracy. Secondly, the significance of explainability highlighted in [12] reinforces our emphasis on delivering detailed natural language justifications for recommendations. Lastly, the consolidation of diverse data types from [7] backs our all-encompassing approach to suitability scoring.

## III. METHODOLOGY

The development of our innovative multi-bank loan recommendation system starts with the creation of a synthetic dataset representing Indian loan applications across five primary types: home, car, education, personal, and gold loans. This dataset has been meticulously crafted to reflect genuine financial situations in India, utilizing insights into banking policies and lending guidelines. During the preprocessing phase, feature engineering is done to generate financial ratios such as the debt-to-income ratio and loan-to-income ratio, which are essential for assessing loan eligibility. After the preprocessing, we ensure the distinctiveness of the available samples by incorporating a variety of demographic and financial profiles across the various loan categories. For each bank and loan type, eligibility assessment takes a two-pronged approach. Firstly, a rule-based evaluation based on specific bank criteria is conducted. Next, a machine learning model—either Random Forest or XGBoost, chosen based on their relative performance—is trained to predict the probability of eligibility. By combining deterministic rules with probabilistic models, we enhance the accuracy of our recommendations. The models undergo training with randomized cross-validation using a 70:30 train-test split ratio to guarantee robust results. The key innovation in our system lies in the unique suitability scoring algorithm, which takes into account multiple weighted factors beyond simple eligibility. These factors include eligibility score, interest rate favorability, processing efficiency, and loan affordability. The weights are adjusted dynamically according to the type of loan to reflect varying consumer priorities. For example, home loans place a greater emphasis on interest rates due to their long-term financial implications, while personal loans prioritize processing efficiency because of their often-urgent nature. Final recommendations are formulated through a detailed ranking process that integrates the suitability scores with specific bank information to create comprehensive natural language explanations for each suggestion. The entire flow of our proposed methodology is depicted in Fig. 1.

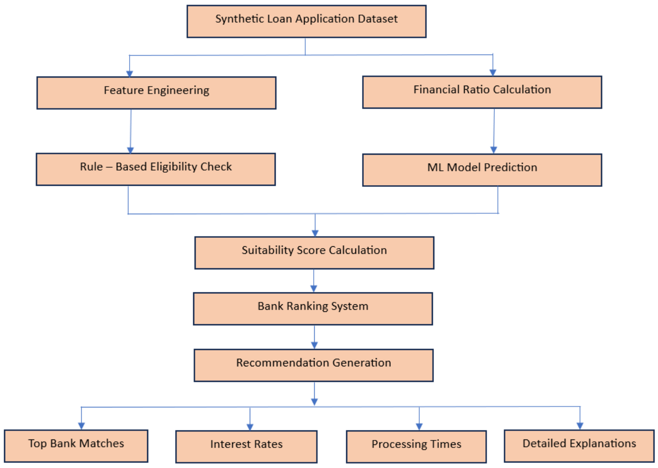


Fig. 1. Proposed Smart Lending Ally Methodology

### A. *Feature computation and Suitability Analysis*

The creation of our multi-bank loan recommendation system necessitates a thorough approach to feature computation and uniqueness evaluation. In this context, uniqueness is defined as the differentiation among loan applicant profiles and their associated eligibility across different banks and loan types. Our dataset $L$ comprises $N$ loan application profiles $P_1, P_2, P_3, \ldots, P_N$ characterized by a variety of attributes $A_1, A_2, A_3, \ldots, A_M$, which include demographic, financial, and loan-specific information. Each profile is assessed against a collection of banks $B_1, B_2, B_3, \ldots, B_K$ that provide various types of loans $T_1, T_2, \ldots, T_5$ (Home, Car, Education, Personal, and Gold loans). The preprocessing stage includes two essential components: categorical encoding and numerical standardization. For categorical variables like occupation, gender, and loan type, we apply a mapping function defined as:

$$encode(c_i) = e_i, where\ encode: C \rightarrow E \qquad (1)$$

This process changes categorical variables into numerical formats that are appropriate for machine learning models, all while maintaining the natural relationships among the categories. For numerical features, we calculate two important financial ratios that play a crucial role in evaluating loan eligibility.

Debt-to-Income Ratio (DTI):

$$DTI_j = \frac{CurrentEMI_j}{MonthlyIncome_j} \qquad (2)$$

Loan-to-Income Ratio (LTI):

$$LTI_j = \frac{LoanAmountRequested_j}{AnnualIncome_j} \qquad (3)$$

These ratios offer standardized assessments of an applicant's debt load and ability to repay loans, allowing for uniform comparison across various income brackets. To maintain the strength of the algorithm, we normalize all numerical characteristics through:

$$Z_j = \frac{X_j - \mu_j}{\sigma_j} \tag{4}$$

Where $X_j$ indicates the original feature value, $\mu_j$ denotes the mean, and $\sigma_j$ signifies the standard deviation of feature $j$ across the entire dataset. The distinctiveness of our method stems from the innovative suitability scoring algorithm that integrates eligibility evaluation with preference alignment. For every bank-applicant combination, we calculate a suitability score $S(P_i, B_k)$ using:

$$S(P_i, B_k) = w_E.E(P_i, B_k) + w_I.I(P_i, B_k) + w_P.P(P_i, B_k) + w_A.A(P_i, B_k) \tag{5}$$

Where:

- $E(P_i, B_k)$ is the eligibility component (0-1)
- $I(P_i, B_k)$ is the interest rate favourability (0-1)
- $P(P_i, B_k)$ is the processing efficiency (0-1)
- $A(P_i, B_k)$ is the affordability component (0-1)
- $w_E$, $w_I$, $w_P$, $w_A$ are weights that vary by loan type according to:

$$W_T = [w_E^T \ w_I^T \ w_P^T \ w_A^T] \tag{6}$$

For feature importance evaluation in our predictive models, we compute the Mean Decrease in Impurity (MDI) for each feature $f$ using:

$$MDI(\mathrm{f}) = \sum_{n \epsilon nodes} w_n \cdot \Delta i(n) \tag{7}$$

The term $w_n$ refers to the weighted count of samples that arrive at node $n$. In this context, $\Delta i(n)$ represents the reduction in impurity at node $n$ when feature $f$ is utilized for the split.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

In this context, *TP*, *TN*, *FP*, and *FN* denote true positives, true negatives, false positives, and false negatives, respectively. The ultimate ranking of recommendations is established by ordering banks based on suitability scores in descending order, and detailed justifications are provided for the top five selections. This thorough method guarantees that recommendations are not only based on sound mathematics but also correspond to the distinct financial circumstances of each applicant.

### B. Dual-Model Classification System

Our dual-model classification approach emphasizes the hyperparameter tuning process for both rule-based evaluations and machine learning classifiers, which are integrated to deliver thorough predictions of loan eligibility. The operation of the proposed system starts with the assignment of the processed feature vector, which results from the feature engineering model. The two key parameters that need to be provided alongside the dataset are the hyperparameter space (Θ) to establish the classifier's range and the number of iterations represented as $N$.

The procedure for the suggested classification model initiates with assigning $\Theta = (\theta 1, \theta 2, \theta 3, \ldots, \theta_m)$, where Θ denotes the hyperparameter space, and $\theta_i$ refers to an individual hyperparameter with its respective range. The symbol $Model_\theta(A)$ signifies the machine learning classifier trained with hyperparameters θ, which predicts bank eligibility based on the input features A. Let $Performance_\theta$ indicate the performance metric (such as accuracy or F1-score) achieved by the model trained using hyperparameters $\theta$. The mathematical representation for the operation of the dual-model classifier can therefore be expressed as follows:

$$\theta^* = \arg\max_{\theta \in \Theta} \sum_{i=1}^{N} \mathrm{Performance}_{\theta,i} \tag{9}$$

In the classification mentioned above, the parameter $\theta^*$ indicates the optimal hyperparameters identified through randomized cross-validation. The expression $\sum_{i=1}^{N} Performance_{\theta,i}$ computes the average performance metric across $N$ iterations for each hyperparameter set $\theta$. The notation $\arg\max_{\theta \in \Theta}$ represents identifying the hyperparameters $\theta$ that yield the highest average performance metric. Our approach involves training both Random Forest and XGBoost models with optimized hyperparameters, choosing the superior model for each bank. Randomized search cross-validation efficiently navigates the hyperparameter space compared to exhaustive methods like grid search, enabling us to identify strong models with reasonable computational resource requirements.

### C. Weighted Ensemble Model for Bank Recommendation

The final element of our methodology employs a weighted ensemble strategy that merges rule-based eligibility criteria with machine learning forecasts to create the ultimate bank recommendations. This innovative suitability scoring algorithm incorporates various factors to deliver tailored recommendations based on the applicant's distinct financial profile. For each loan application profile $P$ and bank $B$, the system computes a composite suitability score $S(P,B)$ that takes into account four essential components:

Eligibility Score ($E$): Integrates rule-based eligibility assessments with predictions from the ML model.
Interest Rate Score ($I$): Assesses the attractiveness of the proposed interest rates.
Processing Efficiency Score ($P$): Reviews the timeframe for processing loans.
Affordability Score ($A$): Evaluates the financial feasibility of the loan for the applicant.
These component scores are aggregated using weights specific to the type of loan.

$$S(P,B) = w_E \cdot E(P,B) + w_I \cdot I(P,B) + w_P \cdot P(P,B) + w_A \cdot A(P,B) \quad (10)$$

Where the weights $(w_E, w_I, w_P, w_A)$ are adjusted according to the type of loan to represent varying priorities. For instance, home loans emphasize interest rates more because of their significant long-term financial implications, whereas personal loans focus on processing speed due to their typically urgent requirements.
The scores for each component are determined in the following manner:

Eligibility Score:

$$E(P,B) = \alpha \cdot Rule(P,B) + (1 - \alpha) . ML(P,B) \quad (11)$$

Where $Rule(P, B)$ is the deterministic rule-based eligibility assessment, $ML(P, B)$ is the machine learning model's probability prediction, and $\alpha$ is a balancing parameter.

Interest Rate Score:

$$I(P,B) = 1 - \frac{r(P,B) - r_{min}}{r_{max} - r_{min}} \quad (12)$$

Where $r(P, B)$ represents the interest rate provided by bank $B$ to profile $P$, and $r_{min}$ and $r_{max}$ denote the lowest and highest rates accessible among all qualified banks. The ultimate suggestions are produced by ranking banks based on their suitability scores and offering detailed explanations in natural language for the top 5 matches, resulting in a thorough and understandable decision support system for those applying for loans.

## IV. DATASET OVERVIEW

The dataset used for developing and evaluating the proposed multi-bank loan recommendation system consists of synthetic loan application data from India, designed to simulate realistic financial situations. It features 2,000 varied loan application profiles across five primary loan categories: Home, Car, Education, Personal, and Gold loans. Each profile comprises 25 characteristics that include demographic information (such as age, gender, marital status, and number of dependents), financial details (like monthly income, current EMI, and CIBIL score), and loan-specific information (including loan amount, tenure, and purpose). Monthly income spans from ₹15,000 to ₹200,000, CIBIL scores range from 300 to 900, and employment lengths vary from 0.5 to 30 years among different occupational groups. This main dataset is supplemented by a secondary dataset that details the eligibility criteria of 10 prominent Indian banks: State Bank of India, HDFC Bank, ICICI Bank, Axis Bank, Punjab National Bank, Canara Bank, Bank of Baroda, Kotak Mahindra Bank, Union Bank of India, and IndusInd Bank. For every combination of bank and loan type, specific eligibility standards are established, including minimum CIBIL score thresholds, income stipulations, maximum debt-to-income ratios, and age limits. This two-dataset strategy enables our system to assess rule-based eligibility while also training machine learning models to uncover patterns and exceptional cases. During preprocessing, derived metrics such as debt-to-income ratio and loan-to-income ratio are computed to improve predictive performance. The extensive nature of these datasets allows for an in-depth comparative analysis among various banks for identical applicant profiles, facilitating the generation of personalized recommendations. By merging multiple financial indicators with bank-specific regulations, our system can pinpoint the best matches between applicants and lenders, assisting borrowers in navigating the intricate landscape of Indian banking and empowering them to make well-informed financial choices while ensuring that recommendations remain practical and applicable in real-world contexts (Table 1).

| **Dataset Components** | **Specifications** |
|---|---|
| Loan Application Profiles | 2,000 records with 25 features each |
| Bank Policies | 10 banks with criteria for 5 loan types |
| Hardware | Intel Core i7 processor, 16 GB RAM |
| Software | Python 3.11 with Scikit-learn, XGBoost, Pandas |

TABLE I. Experimental Setup and Dataset Specifications for Smart Lending Ally

## V. RESULTS AND DISCUSSIONS

### A. *Model Performance Analysis*

The dual-model classification strategy utilized in our recommendation system exhibited exceptional performance across all banks assessed. Below output illustrates the accuracy results from the training phase of our model.

Training eligibility prediction models for each bank...
Trained model for State Bank of India - Accuracy: 0.9775
Trained model for HDFC Bank - Accuracy: 0.9825
Trained model for ICICI Bank - Accuracy: 0.9875
Trained model for Axis Bank - Accuracy: 0.9825
Trained model for Punjab National Bank - Accuracy: 0.9850
Trained model for Canara Bank - Accuracy: 0.9675
Trained model for Bank of Baroda - Accuracy: 0.9750
Trained model for Kotak Mahindra Bank - Accuracy: 0.9850
Trained model for Union Bank of India - Accuracy: 0.9750
Trained model for IndusInd Bank - Accuracy: 0.9925

Models Performance Summary:
State Bank of India: 0.9775 accuracy, xgboost model
HDFC Bank: 0.9825 accuracy, xgboost model
ICICI Bank: 0.9875 accuracy, xgboost model
Axis Bank: 0.9825 accuracy, xgboost model
Punjab National Bank: 0.9850 accuracy, xgboost model
Canara Bank: 0.9675 accuracy, xgboost model
Bank of Baroda: 0.9750 accuracy, xgboost model
Kotak Mahindra Bank: 0.9850 accuracy, xgboost model
Union Bank of India: 0.9750 accuracy, xgboost model
IndusInd Bank: 0.9925 accuracy, xgboost model

As shown in the training output, all ten banks registered high accuracy scores ranging from 96.75% to 99.25%, resulting in an impressive average accuracy of 98.33%. Significantly, XGBoost consistently outperformed Random Forest across all banks in our implementation, confirming our choice to adopt a comparative model selection technique instead of relying on a single algorithm. IndusInd Bank recorded the highest accuracy at 99.25%, while Bank of Baroda, although still achieving excellent results, had the lowest accuracy at 97.50%. The outstanding performance of XGBoost can be attributed to its gradient boosting framework, which effectively manages the intricate relationships between applicant characteristics and loan eligibility criteria. The consistently high accuracy across various banks highlights the robustness of our method, particularly given the differing eligibility criteria used by multiple financial institutions.

### B. *Case Study Analysis: Premium Home Loan Applicant*

To illustrate the practical application of our recommendation system, we examined a specific case involving a high-end home loan applicant:

**Use Case 1: Outstanding Credit, High Income Profile**

Name: Arjun Sharma
Age: 42
Gender: Male
Marital Status: Married
Dependents: 2
Occupation: Salaried-Private (Senior Manager)
Employment Years: 15
Monthly Income: ₹150,000
Existing Loans: 0
CIBIL Score: 810
Loan Type: Home Loan
Loan Amount: ₹8,000,000
Property Value: ₹12,000,000

Figure 2 displays the eligibility probability radar chart for this applicant, indicating uniformly high eligibility across all five suggested banks. The nearly ideal pentagon shape with values close to 1.0 implies that this high-income, superb credit profile satisfies the eligibility requirements of all major banks with considerable confidence.

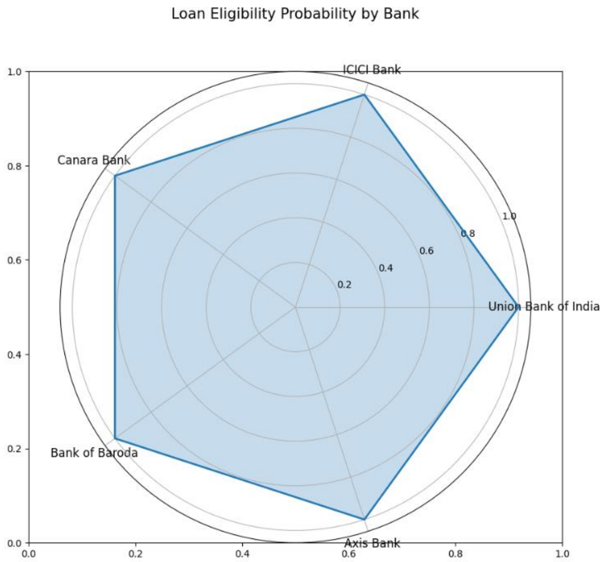

Fig. 2. Loan Eligibility Probability by Bank

The eligibility assessment reveals that applicants with strong financial standings, like Arjun Sharma, have considerable leeway in selecting lenders. All five banks exhibit eligibility probabilities at or near 1.0, signifying an exceptional chance of loan approval. This emphasizes the advantages prime borrowers enjoy within the Indian banking sector.

### C. *Suitability Score Distribution and Analysis*

While eligibility is a binary measure, our innovative suitability scoring algorithm delivers a more detailed assessment by considering various factors pertinent to the applicant's overall loan experience. The right side of Figure 3 presents the suitability scores alongside eligibility probabilities.

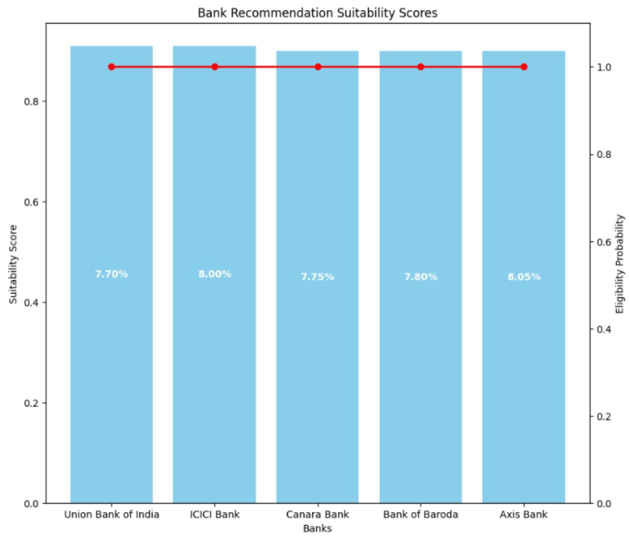

Fig. 3. Bank Recommendation Suitability Score

This graphic uncovers several key insights regarding Arjun's profile:

Even though all banks display comparable high eligibility probabilities (red line at around 1.0), there are distinct differences in overall suitability scores (blue bars). Axis Bank achieves the top suitability score with an interest rate of 8.05%, while Union Bank of India holds the lowest suitability score, despite offering the most competitive interest rate of 7.70%. The suitability scores indicate that ICICI Bank (8.00%) ranks second in overall suitability, even though it does not provide the lowest interest rate, signifying that our algorithm effectively weighs multiple factors beyond solely the interest rate. The minor variation in suitability scores (approximately 0.92 to 0.95) demonstrates that for prime borrowers like Arjun, most leading banks offer generally favorable conditions. Nonetheless, the differences are still significant when taking into account the considerable loan amount requested (₹8,000,000).

### D. *Processing Time and Interest Rate Analysis*

Figure 4 presents detailed comparisons of two crucial aspects in loan selection: processing times and interest rates.

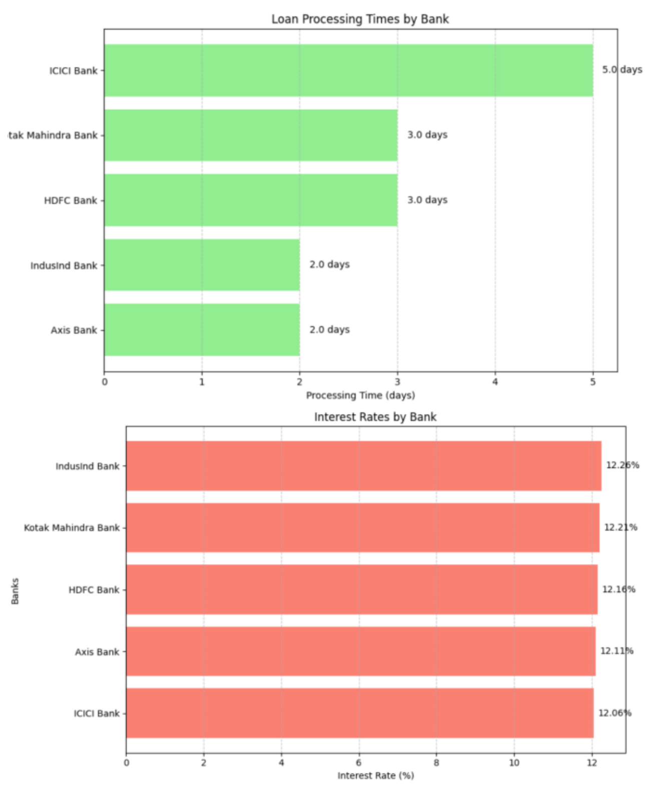


Fig. 4. Processing times and Interests rates

The processing time comparison (left) indicates substantial discrepancies among banks concerning home loans: Axis Bank and ICICI Bank provide comparatively quicker processing at 13 days, followed by Bank of Baroda at 16 days, while Canara Bank and Union Bank of India experience the longest processing durations at 17 days. The interest rate comparison (right) illustrates a range of approximately 0.35 percentage points: Union Bank of India proposes the most attractive rate at 7.70%, with Canara Bank at 7.75%, and Axis Bank presenting the highest rate at 8.05%. For Arjun's ₹8,000,000 home loan, this interest rate variation of 0.35% between the highest and lowest rates would lead to a difference of around ₹28,000 annually in interest payments, or ₹700,000 over a 25-year loan term. This significant sum explains why interest rate is heavily weighted in the suitability score for home loans. However, the 4-day discrepancy in processing time between the fastest and slowest banks may be less critical for a home loan, which corresponds with our lower weight assignment (10%) for processing efficiency in home loan suitability calculations.

### E. *Recommendation Insights*

Based on Arjun's profile, the system has suggested the following banks ranked by their suitability: Union Bank of India (Suitability Score: 0.91): This bank offers the most competitive interest rate at 7.70% with a processing period of 17 days. The estimated monthly EMI would be around ₹65,429, which accounts for 43.6% of his monthly income. Although the processing time is longer, the attractive interest rate makes it the top choice. ICICI Bank (Suitability Score: 0.91): This bank strikes a balance between a quicker processing time of 13 days and an interest rate of 8.00%. The anticipated monthly EMI would be roughly ₹67,217, which is 44.8% of his monthly income. Canara Bank (Suitability Score: 0.90): Offering a favourable interest rate of 7.75%, this bank has a longer processing time of 17 days. The monthly EMI is projected to be approximately ₹65,875, corresponding to 43.9% of monthly income. Bank of Baroda (Suitability Score: 0.90): This bank presents a mid-tier option with an interest rate of 7.80% and a processing time of 16 days. The estimated monthly EMI would be ₹66,322, which represents 44.2% of monthly income. Axis Bank (Suitability Score: 0.89): While this bank has a quicker processing time of 13 days, the higher interest rate of 8.05% results in the largest monthly EMI of ₹67,614, making it correspond to 45.1% of monthly income and the least favourable choice among the top five. The comprehensive explanations in the recommendations underscore that all five banks acknowledge Arjun's outstanding credit score (810), excellent debt-to-income ratio (0.0%), robust employment history (15 years), impressive monthly income (₹150,000), and solid credit history (15 years). Each recommendation details the total interest payable over the 20-year term and the maximum loan tenure allowable based on his age (23 years).

### F. *Effectiveness of the System for High-Quality Borrowers*

The examination of Arjun's situation shows that for high-quality borrowers with exceptional credit histories, our system effectively distinguishes between banks that would all likely approve their loans. In these instances, the additional benefits include: Interest Rate Optimization: Minor variations in interest rates can lead to considerable savings throughout the duration of a home mortgage. Processing Efficiency: For borrowers who prefer faster loan disbursement, the system points out banks with quicker processing times. Balanced Evaluation: By assessing multiple factors at once, the system highlights banks that present the best overall value instead of merely the lowest interest rates. The suitability scoring algorithm is especially useful for high-quality borrowers like Arjun, as it aids in differentiating between several viable choices based on individual preferences for home financing.

## VI. Conclusion

This study introduces an innovative multi-bank loan recommendation system aimed at streamlining the intricate task of choosing suitable financial institutions for diverse loan requirements within the Indian banking landscape. By combining rule-based eligibility criteria with machine learning classification techniques, our methodology overcomes the shortcomings of conventional eligibility assessments while offering personalized and comprehensible recommendations. The dual-model classification framework, which strategically utilizes either Random Forest or XGBoost models based on their performance, shows notable accuracy in forecasting loan approvals among ten leading Indian banks. Nonetheless, the primary contribution of this research lies in the unique suitability scoring algorithm that extends beyond basic eligibility by factoring in interest rates, processing speed, and affordability considerations with loan-type specific weightings. This results in a more detailed recommendation structure that addresses the real-world concerns of loan applicants. The extensive visual representations and thorough natural language descriptions further improve the system's transparency and user-friendliness, assisting applicants in grasping not just which banks to consult but also the rationale behind these recommendations. Although the present implementation is based on synthetic data and would benefit from real-world testing, the architectural design suggests a promising method for financial decision support systems. Future developments could involve integration with live banking APIs, consideration of regional differences, and expansion to additional specialized loan categories, ultimately aiming towards a more accessible and equitable financial landscape in India where borrowers from all backgrounds can discover optimal lending options tailored to their specific financial situations.

## REFERENCES


[1] Imtiaz, Arsal, et al. "Agricultural loan recommender system-A machine learning approach." *2021 International Conference on Innovative Trends in Information Technology (ICITIIT)*. IEEE, 2021.

[2] Uddin, Nazim, et al. "An ensemble machine learning based bank loan approval predictions system with a smart application." *International Journal of Cognitive Computing in Engineering* 4 (2023): 327-339.

[3] Uriawan, Wisnu, et al. "Laps: Computing loan default risk from user activity, profile, and recommendations." *2022 Fourth International Conference on Blockchain Computing and Applications (BCCA)*. IEEE, 2022.

[4] Gupta, Anshika, et al. "Bank loan prediction system using machine learning." *2020 9th International Conference System Modeling and Advancement in Research Trends (SMART)*. IEEE, 2020.

[5] Aphale, Amruta S., and Sandeep R. Shinde. "Predict loan approval in banking system machine learning approach for cooperative banks loan approval." *International Journal of Engineering Trends and Applications (IJETA)* 9.8 (2020).

[6] Li, Size, et al. "Risk assessment of the individual housing loans by grey data mining based on AHP-GRAP." *2013 Computing, Communications and IT Applications Conference (ComComAp)*. IEEE, 2013.

[7] Xue, Runqi. "Segmentation for Financial Loan Company's Customers Data Based on K-means." *2022 3rd International Conference on Electronic Communication and Artificial Intelligence (IWECAI)*. IEEE, 2022.

[8] Eletter, Shorouq Fathi, Saad Ghaleb Yaseen, and Ghaleb Awad Elrefae. "Neuro-based artificial intelligence model for loan decisions." *American Journal of Economics and Business Administration* 2.1 (2010): 27.

[9] Kiran, Ajmeera, et al. "Intelligent Chat Bots: An AI Based Chat Bot For Better Banking Applications." *2023 International Conference on Computer Communication and Informatics (ICCCI)*. IEEE, 2023.

[10] Reddy, G. Divakara, et al. "Utilization of AI for streamlining and optimizing credit decision process and security access loan risks in the banking sector." *2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA)*. IEEE, 2022.

[11] Adla, Abdelkader. "Real Estate Loan Knowledge-Based Recommender System." *J. Digit. Inf. Manag.* 18.2 (2020): 65.

[12] Cornacchia, Giandomenico, Fedelucio Narducci, and Azzurra Ragone. "A general model for fair and explainable recommendation in the loan domain." *Joint Workshop Proceedings of the 3rd Edition of Knowledge-aware and Conversational Recommender Systems (KaRS) and the 5th Edition of Recommendation in Complex Environments (ComplexRec) co-located with 15th ACM Conference on Recommender Systems (RecSys 2021)*. 2021.

[13] DICKSON, JOSEPH BARILEDUM, FUBARA EGBONO, and NA IBIOBU. "AN ENHANCED RURAL COMMUNITY LOAN ELIGIBILITY MATCHMAKING MODEL USING A HYBRIDIZED RECOMMENDER SYSTEM." *Journal of Systematic and Modern Science Research* (2024).

[14] Tsuji, Keita, et al. "Book recommendation using machine learning methods based on library loan records and bibliographic information." *2014 IIAI 3rd International Conference on Advanced Applied Informatics*. IEEE, 2014.

[15] Zibriczky12, Dávid. "Recommender systems meet finance: a literature review." *Proc. 2nd Int. Workshop Personalization Recommender Syst*. 2016.

[16] Afifah, Khansa, Intan Nurma Yulita, and Indra Sarathan. "Sentiment analysis on telemedicine app reviews using xgboost classifier." *2021 international conference on artificial intelligence and big data analytics*. IEEE, 2021.

[17] Aydin, Zeliha Ergul, and Zehra Kamisli Ozturk. "Performance analysis of XGBoost classifier with missing data." *Manchester Journal of Artificial Intelligence and Applied Sciences (MJAIAS)* 2.02 (2021): 2021.

[18] Raihan, Md Johir, et al. "Detection of the chronic kidney disease using XGBoost classifier and explaining the influence of the attributes on the model using SHAP." *Scientific Reports* 13.1 (2023): 6263.

[19] Gündoğdu, Serdar. "Efficient prediction of early-stage diabetes using XGBoost classifier with random forest feature selection technique." *Multimedia Tools and Applications* 82.22 (2023): 34163-34181.

[20] Li, Jiangtao, et al. "Application of XGBoost algorithm in the optimization of pollutant concentration." *Atmospheric Research* 276 (2022): 106238.

[21] Palimkar, Prajyot, Rabindra Nath Shaw, and Ankush Ghosh. "Machine learning technique to prognosis diabetes

disease: Random forest classifier approach." *Advanced computing and intelligent technologies: proceedings of ICACIT 2021*. Singapore: Springer Singapore, 2021. 219-244.

[22] Amiri, Ahmed Faris, et al. "Faults detection and diagnosis of PV systems based on machine learning approach using random forest classifier." *Energy Conversion and Management* 301 (2024): 118076.

[23] Jackins, V., et al. "AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes." *The Journal of Supercomputing* 77.5 (2021): 5198-5219.

[24] Magidi, James, et al. "Application of the random forest classifier to map irrigated areas using google earth engine." *Remote Sensing* 13.5 (2021): 876.

[25] Sharma, Tarunim, et al. "Ensemble machine learning paradigms in software defect prediction." *Procedia Computer Science* 218 (2023): 199-209.

[26] Noviandy, Teuku Rizky, et al. "Ensemble machine learning approach for quantitative structure activity relationship based drug discovery: A Review." *Infolitika Journal of Data Science* 1.1 (2023): 32-41.

[27] Zounemat-Kermani, Mohammad, et al. "Ensemble machine learning paradigms in hydrology: A review." *Journal of Hydrology* 598 (2021): 126266.

[28] Rincy, Thomas N., and Roopam Gupta. "Ensemble learning techniques and its efficiency in machine learning: A survey." *2nd international conference on data, engineering and applications (IDEA)*. IEEE, 2020.

[29] Ganie, Shahid Mohammad, and Majid Bashir Malik. "An ensemble machine learning approach for predicting type-II diabetes mellitus based on lifestyle indicators." *Healthcare Analytics* 2 (2022): 100092.

[30] Chung, Jetli, and Jason Teo. "Single classifier vs. ensemble machine learning approaches for mental health prediction." *Brain informatics* 10.1 (2023): 1.

[31] Hakak, Saqib, et al. "An ensemble machine learning approach through effective feature extraction to classify fake news." *Future Generation Computer Systems* 117 (2021): 47-58.

[32] Prasad, Pankaj, et al. "Novel ensemble machine learning models in flood susceptibility mapping." *Geocarto International* 37.16 (2022): 4571-4593.

[33] Geng, Xiaojiao, Yan Liang, and Lianmeng Jiao. "EARC: Evidential association rule-based classification." *Information Sciences* 547 (2021): 202-222.

[34] Shahroudnejad, Atefeh, et al. "Thyroid nodule segmentation and classification using deep convolutional neural network and rule-based classifiers." *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 2021.

[35] Streeb, Dirk, et al. "Task-based visual interactive modeling: Decision trees and rule-based classifiers." *IEEE Transactions on Visualization and Computer Graphics* 28.9 (2021): 3307-3323.