**Loan Eligibility Prediction**

A Project Report Submitted in the partial fulfillment of the requirements for the award of the degree of

**BACHELOR OF TECHNOLOGY**

**In**

**DEPARTMENT OF COMPUTER SCIENCE ENGINNERING**

**By**

**K SANTHOSHINI - 2320030304**

**M S MAHA LAKSHMI - 2320030218**

**K ABHIGNA - 2320030089**

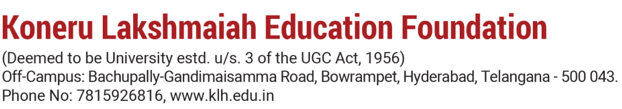
Under the Esteemed Guidance of

**P VENKATESWARA RAO**

**Assistant Professor**

**Department of Computer Science and Engineering**





**K L (Deemed to be) University**

**DEPARTMENT OF COMPUTER SCIENCE ENGINEERING**



**Declaration**

The Project Report entitled “**Loan Eligibility Prediction**” is a record of Bonafide work of **K SANTHOSHINI – 2320030304**, **M S MAHA LAKSHMI – 2320030218**, **K ABHIGNA** – **2320030089** submitted in partial fulfillment for the award of B. Tech in Computer Engineering to the K L University. The results embodied in this report have not been copied from any other departments/University/Institute.

K Santhoshini – 2320030304

M S Maha Lakshmi – 2320030218

K Abhigna – 2320030089

**K L (Deemed to be) University**

**DEPARTMENT OF COMPUTER SCIENCE ENGINEERING**



**Certificate**

This is certify that the project based report entitled “**Loan Eligibility Predication**” is a bonafide work done and submitted by **K SANTHOSHINI – 2320030304**, **M S MAHA LAKSHMI – 2320030218**, **K ABHIGNA – 2320030089** in partial fulfillment of the requirements for the award of the degree of **BACHELOR OF TECHNOLOGY** in Department of Computer Science Engineering, K L (Deemed to be University), during the academic year **2024-2025.**

**Signature of the Supervisor**

**Signature of the HOD                                               Signature of the External Examiner**

**ACKNOWLEDGEMENT**

The success in this project would not have been possible but for the timely help and guidance rendered by many people. Our wish to express my sincere thanks to all those who has assisted us in one way or the other for the completion of my project.

Our greatest appreciation to my guide, **P VENKATESWARA RAO**Department of Computer Science which cannot be expressed in words for his tremendous support, encouragement and guidance for this project.

We express our gratitude to **Dr. Ramesh Babu Sir (CSE)**, Head of the *Department for Computer Science Engineering* for providing us with adequate facilities, ways and means by which we are able to complete this project based Lab.

We thank all the members of teaching and non-teaching staff members, and also who have assisted me directly or indirectly for successful completion of this project. Finally, I sincerely thank my parents, friends and classmates for their kind help and cooperation during my work.

K Santhoshini – 2320030304

M S Maha Lakshmi – 232003218

K Abhigna – 2320030089

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **S.No** | **Contents** | **Page no** |
| 1 | Abstract | 6 |
|  |  |  |
| 2 | Introduction | 7 |
| 3 | Literature survey | 8 |
| 4 | Client meetings | 9 |
| 5 | Hardware and Software requirements | 14 |
| 6 | Design Thinking and Innovation | 15 |
| 7 | Implementation | 37 |
| 8 | Experimentation and Code | 40 |
| 9 | Results | 47 |
| 10 | Conclusion | 48 |
| 11 | References | 49 |

**ABSTRACT**

The prediction of loan eligibility is a vital process for financial institutions. Traditionally, this assessment was done manually by reviewing customer profiles and past financial records, which consumed a lot of time and was prone to errors. With the increasing number of loan applications, there is a growing need for automated solutions that can provide accurate and fast eligibility assessments.

This project aims to develop a Loan Eligibility Prediction system using machine learning algorithms. The goal is to train a predictive model on historical loan data to identify patterns that determine whether a loan should be approved or not. This would streamline the evaluation process and reduce dependency on manual intervention.

Our model considers various parameters such as applicant income, loan amount, education, credit history, marital status, and property location. These factors are fed into the system, which processes and learns from them to make future predictions on new applications. This helps in making data-driven decisions in a consistent and unbiased manner.

The advantage of such a system lies in its ability to reduce loan processing time, eliminate subjective judgments, and ensure transparency. Banks and financial institutions can also reduce their workload, focusing more on strategic financial planning while the system handles initial screening.

Moreover, the system can be further expanded to offer insights into why a particular application was approved or rejected. This feature enhances trust between the applicant and the institution by offering explainable results and boosting applicant confidence in the system.

In conclusion, the loan eligibility prediction system not only helps in effective loan processing but also contributes to digital transformation in the banking sector. It is a step toward a more automated, reliable, and fair decision-making system.

**INTRODUCTION**

In the banking and financial sector, assessing loan eligibility is a critical function that directly impacts revenue generation and customer trust. Traditionally, this has been a manual process, requiring staff to examine financial documents and assess risks based on human judgment. This approach is not only time-consuming but also susceptible to inconsistencies.

With the rise of data science and machine learning, it has become possible to automate and enhance decision-making processes. Machine learning enables systems to learn from historical data and make predictions without being explicitly programmed for each case. This makes it an ideal tool for predicting loan eligibility based on previous application data.

The objective of this project is to implement a Loan Eligibility Prediction system that helps banks decide whether to approve or reject a loan application. It uses historical loan data to train a model that identifies patterns and determines eligibility based on several input features. These include the applicant's income, employment status, loan amount, and credit history.

The proposed system will offer a user-friendly interface where inputs can be submitted, and the result — eligible or not eligible — is provided. The back-end machine learning model evaluates the inputs using trained data and returns a prediction with a high level of accuracy.

This kind of predictive system benefits both banks and customers. It increases efficiency and reduces the chance of errors in processing. Customers receive quicker responses, and institutions can handle more applications without increasing staff workload.

Overall, this project is a move toward digital transformation in the financial sector. It makes loan approval processes faster, more accurate, and less biased, ultimately enhancing customer satisfaction and improving the institution’s operational efficiency.

**LITERATURE SURVEY**

Several researchers and developers have explored the application of machine learning in predicting loan eligibility. This has become an active area of study due to its real-world impact and the availability of structured financial datasets. The findings from past works have shaped the design and direction of this project.

In a study by **Singh and Sharma (2021)**, multiple machine learning algorithms such as Logistic Regression, Decision Trees, and Random Forests were applied to loan datasets. Their findings concluded that ensemble models, particularly Random Forest, achieved the highest accuracy in predicting eligibility.

Similarly, **Patel and Doshi (2020)** introduced a hybrid machine learning model combining Logistic Regression and Gradient Boosting techniques. This model significantly improved accuracy and helped reduce overfitting by utilizing cross-validation techniques and optimized hyperparameters.

**Kaggle’s Loan Prediction Dataset**, which includes features such as gender, income, loan amount, property area, and credit history, has been frequently used for benchmarking loan eligibility models. Many researchers have validated their algorithms using this dataset, making it a standard baseline for comparison.

Another important contribution was from **Chen et al. (2019)**, who explored the use of Support Vector Machines (SVM) and found it effective, especially when the dataset was balanced and normalized. However, they noted that explainability in SVMs was lower compared to decision trees and logistic models.

A common finding across studies is the significance of feature selection and preprocessing. Missing value treatment, data encoding, and normalization significantly influence model performance. Literature also emphasizes that incorporating domain expertise helps improve prediction accuracy and model interpretability.

In summary, the literature supports the use of machine learning in loan prediction and suggests a comparative approach using different models to select the best-performing one. Our project builds upon these insights by implementing multiple algorithms and evaluating them using standard metrics like accuracy, precision, recall, and F1-score.

**CLIENT MEETINGS**

This report consists of four meetings with the client, each addressing specific objectives related to the Loan Eligibility Prediction system. Below are the questions posed to the client and their responses.

### Meeting 1: Understanding Parkinson’s Disease Prediction Goals

Objectives:  
1. Understand the client’s vision and goals for Loan Eligibility Prediction.  
2. Discuss the target users and intended use case.  
3. Clarify scope, expected features, and desired outputs.

Q: What is the main goal of this project?

A: To automate the loan eligibility decision-making process using historical data and predictive algorithms..

Q: Who are the target users of this system?

A: Bank staff, loan officers, and internal financial advisors; future expansion may include direct applicant interfaces

Q: What are the key features used for prediction?

A: Applicant income, co-applicant income, loan amount, credit history, employment status, marital status, education level, and property area.

Q: Is this for internal decision-making or customer-facing applications?

A: Initially for internal use, with future possibilities for customer pre-eligibility check on the website..

Q: What kind of output is expected from the system?

A: Eligibility status (Yes/No), confidence score, and possibly a reason/explanation for the result.

Q: What challenge does this system aim to solve?

A: Reduce manual screening, improve decision speed, eliminate biases, and enhance operational effiency.

### Meeting 2: Technical Design for Parkinson’s Prediction System

Objectives:  
1. Discuss the technical specifications and machine learning models.  
2. Explore data preprocessing, model training, and performance benchmarks.  
3. Understand explainability, integration, and accuracy goals.

Q: Which machine learning algorithms are preferred?

A: Logistic Regression for baseline, Random Forest for interpretability, and XGBoost or Gradient Boosting for performance optimization.

Q: What preprocessing steps are required?

A: Handling missing values, label encoding for categorical variables, feature scaling, and class balancing (e.g., SMOTE).

Q: Should the model be explainable to users?

A: Yes, especially for internal review. Using SHAP values or similar tools to show feature importance is preferred..

Q: What is the expected model accuracy?

A: A minimum of 85% accuracy, with high precision and recall to avoid false approvals.

Q: **Is real-time prediction needed?**

A: Yes, for web integration where inputs are given, and prediction is returned instantly

Q: Will the model be reused or retrained frequently?

A: Retraining will be periodic — once new data (e.g., 3–6 months of loan records) is collected.

Q: Will feature importance be considered in model selection?

A: Absolutely. Models that provide insight into feature impact (e.g., Random Forest) are preferred for transparency

Q. Do we need to handle class imbalance?

A. Yes, imbalance is likely in loan datasets; techniques like SMOTE or class weights will be used.

### Meeting 3: User Experience and Model Interaction for Parkinson’s Prediction

Objectives:  
1. Finalize system functionalities, input/output formats, and user interactions.  
2. Confirm platform for deployment and model hosting preferences.  
3. Discuss system performance, scalability, and security.

Q: What features should the MVP include?  
A: User input form, prediction result (Eligible/Not Eligible), confidence level, basic reason explanation, and admin dashboard.

Q: What platform will be used for deployment?  
A: Yes, like “Review with Structural Biologist” or “Analyze Additional Variants.”

Q: Will the system support PDB uploads only, or allow sequence input too?  
A: Initially PDB uploads; future versions may support raw sequence input.

Q: How to handle invalid or unsupported files?  
A: Show an error message with accepted formats (.pdb) and retry option.

Q: Is the UI designed for accessibility?  
A: Yes, with readable fonts, simple layout, and keyboard-friendly design.

Q: Can users download their results?  
A: Yes, a PDF report with prediction, visual highlights, and key structural features.

Q: What if model confidence is low?  
A: Return an “Inconclusive” message and suggest manual structural review.

Q: Will structural data visualization be available?  
A: Yes, with basic charts and interactive 3D model views (e.g., PyMOL snapshots).

Q: Will language support be limited?  
A: Start with English; other languages can be added later.

Q: Can users give feedback to improve predictions?  
A: Yes, with consent, user feedback will help retrain and refine the model.

### Meeting 4: Deployment, Validation, and Final Checklist

Objectives:  
1. Confirm final deployment plans and hosting platform.  
2. Discuss testing, security, and release timeline.  
3. A Final QA checks and go-live readiness.

Q: What is the target launch timeline?  
A: Within 2 months after testing and model validation.

Q: What types of testing will be done?  
A: Unit, functional, and structural prediction accuracy testing using real-world protein data.

Q: How will clinical validation be handled?  
A: Using expert-reviewed predictions on anonymized research datasets.

Q: What privacy standards are followed?  
A: GDPR-compliant storage, anonymized inputs, and encryption.

Q: Will user documentation be provided?  
A: Yes, with a guide and demo video for onboarding.

Q: What happens if the system fails?  
A: Auto-restart, failover backup, and crash notifications.

Q: Will the model be updated regularly?  
A: Yes, retrained biannually with new structural or clinical data.

Q: Can system usage be monitored?  
A: Yes, through an integrated analytics dashboard.

Q: Who handles support after deployment?  
A: A dedicated technical team for maintenance and issue resolution.

Q: Will users get update notifications?  
A: Yes, through email or in-app alerts.





**Hardware Requirements**

1. **Processor**: Intel Core i5 or higher
2. **RAM**: 8 GB minimum (16 GB recommended)
3. **Storage**: 256 GB (SSD recommended)
4. **Internet**: Required for:

* Downloading PDB data
* Installing libraries and dependencies
* API usage and deployment

**Software Requirements**

1. **Operating System**: Windows, macOS, or Linux
2. **Programming Language**: Python 3.7+
3. **IDE**: Visual Studio Code
4. **Libraries and Frameworks**:
   * **numpy**: Numerical computations and array manipulation
   * **matplotlib/seaborn**: Data visualization
   * **pandas**: Handling tabular biomedical data
   * **scikit-learn**: Building machine learning models
   * **Flask**: Backend server and model deployment
   * **Biopython**: Parsing and processing PDB files
   * **PyMOL/Chimera**: 3D visualization of protein structures
   * **Imbalanced-learn**: Handling class imbalance in training data
   * **Joblib**: Model serialization

**DESIGN THINKING AND INNOVATION**

**Problem/Opportunity Domain**

**Domain of Interest:**

The domain of interest is Healthcare and Artificial Intelligence (AI), specifically focusing on neurological disease prediction, with an emphasis on Parkinson’s Disease (PD). This area combines medical research, bioinformatics, and AI-driven predictive analytics to improve early detection, diagnosis, and management of Parkinson’s Disease.

**Description of the Domain:**

Parkinson’s Disease (PD) is a progressive neurodegenerative disorder with no definitive early diagnostic method. Current challenges include late detection, misdiagnosis, and limited accessibility to specialists. AI offers solutions such as early prediction, personalized treatment plans, remote monitoring, and drug discovery acceleration using machine learning on medical data.

**Why did you choose this domain?**

* High Impact – Early AI-based detection can improve patient outcomes.
* Market Potential – AI in healthcare is growing, with PD diagnosis being an underserved area.
* Research Advancements – AI can enhance bioinformatics, predicting protein misfolding linked to
* PD.
* Accessibility – AI can make PD detection more affordable and available worldwide.

**Problem/Opportunity Statement**

**Problem Statement:**

Parkinson’s Disease (PD) is often diagnosed too late, after irreversible neurological damage has occurred. Current diagnostic methods lack accuracy and accessibility, leading to misdiagnosis, delayed treatment, and reduced quality of life. AI-based predictive models can provide early detection, personalized treatment insights, and improved patient monitoring.

**Problem Description:**  
PD is a progressive neurodegenerative disorder that affects movement and cognitive function. Its early symptoms are subtle and often mistaken for aging or other conditions, delaying proper treatment. There are no definitive biomarkers, and diagnosis primarily relies on clinical observation, making early intervention difficult.

**Context (When does the problem occur):**

* Early stages of Parkinson’s Disease when symptoms are mild and non-specific.
* During misdiagnosis, when PD is confused with other movement disorders.
* In remote areas, where access to neurologists and advanced diagnostic tools is limited.

**Alternatives (What does the customer do to fix the problem):**

* Clinical Examination – Doctors assess symptoms manually, often leading to late diagnosis.
* MRI & PET Scans – Expensive and not always conclusive.
* Genetic Testing – Limited to hereditary cases and not widely used.
* Dopamine Transporter (DaT) Scan – Helps detect PD but is costly and not widely available.

**Customers (Who has the problem most often):**

* Patients at risk of Parkinson’s Disease, especially those with early symptoms.
* Neurologists & healthcare professionals seeking more accurate diagnostic tools.
* Researchers working on neurodegenerative diseases.
* Healthcare providers & hospitals looking to improve PD detection and patient care.

**Emotional Impact (How does the customer feel):**

* Patients feel anxious, frustrated, and hopeless due to uncertainty in diagnosis.
* Doctors feel challenged due to the lack of reliable early diagnostic tools.
* Caregivers & families experience stress from delayed or incorrect diagnoses, impacting caregiving decisions.

**Quantifiable Impact (What is the measurable impact):**

* Up to 25% of Parkinson’s cases are misdiagnosed, leading to incorrect treatments.
* Diagnosis often occurs 5–10 years after disease onset, reducing treatment effectiveness.
* Medical costs for late-stage PD management are significantly higher than early intervention.

**Alternative Shortcomings (What are the disadvantages of the alternatives):**

* Late detection – Symptoms-based diagnosis is reactive rather than preventive.
* Expensive tests – MRI and DaT scans are not affordable for all patients.
* Limited accessibility – Many rural areas lack specialized neurology facilities.
* No real-time monitoring – Current methods do not offer continuous tracking of disease progression.

**Any Video or Images to showcase the problem:**

* No specific videos or images are available at this time.

**Addressing SDGs**

**Relevant Sustainable Development Goals (SDGs):**

* SDG 3: Good Health and Well-being – Ensuring healthy lives and promoting well-being for all.
* SDG 9: Industry, Innovation, and Infrastructure – Advancing technology in healthcare.
* SDG 10: Reduced Inequalities – Improving healthcare accessibility for underserved populations

**How does your problem/opportunity address these SDGs?**

**SDG 3: Good Health and Well-being**

* Enables early detection of Parkinson’s Disease, improving treatment outcomes.
* Reduces misdiagnosis and delayed treatment, enhancing patient quality of life.

**SDG 9: Industry, Innovation, and Infrastructure**

* Uses AI and machine learning to enhance medical diagnostics.
* Promotes investment in AI-driven medical research, driving technological progress.

**SDG 10: Reduced Inequalities**

* AI-powered diagnostics can make PD detection accessible in remote and low-income areas.
* Reduces dependence on expensive tests, making healthcare more affordable.

**Stakeholders**

**1. Who are the key stakeholders involved in or affected by this project?**

* Patients at risk of Parkinson’s
* Neurologists & healthcare professionals
* Researchers (AI & medical)
* Hospitals & healthcare providers
* Technology partners (AI developers, data scientists)
* Government & health authorities

**2. What roles do the stakeholders play in the success of the innovation?**

* Patients: Provide data & feedback
* Doctors: Validate AI models, adopt technology
* Researchers: Develop and improve algorithms
* Hospitals: Implement systems, provide access to patient data
* Tech partners: Build AI models & tools
* Government: Set regulations, support funding

**3. What are the main interests and concerns of each stakeholder?**

* Patients: Accurate, affordable diagnosis
* Doctors: Reliable, easy-to-use tools
* Researchers: Data access, innovation opportunities
* Hospitals: Cost-effective, scalable solutions
* Government: Public health impact, data privacy

**4. How much influence does each stakeholder have on the outcome of the project?**

* High: Doctors, researchers, tech partners
* Medium: Patients, hospitals
* Low: Government (influences indirectly through policy)

**5. What is the level of engagement or support expected from each stakeholder?**

* High: AI Developers, Banks/NBFCs, Loan Officers
* Medium: Loan Applicants, Technology Partners

**6. Are there any conflicts of interest between stakeholders? If so, how can they be addressed?**

* Provide training, highlight AI as decision support, not replacement.
* Ensure data privacy, transparency and consent-based system.

**7. How will you communicate and collaborate with stakeholders throughout the project?**

* **Workshops** with bank staff and officers for feedback and training.
* **Surveys & interviews** with loan applicants to improve user experience
* **Regular updates** and demos for banking institutions and partners.
* **Compliance check-ins** with regulators and legal advisors.
* **Collaboration tools** (Jira, Slack, Zoom, etc.) for remote and cross-functional teamwork.

**Power Interest Matrix of Stakeholders**

**Power Interest Matrix:**

A diagram of a health care system

AI-generated content may be incorrect.

**High Power, High Interest**:

* Neurologists & Healthcare Professionals (They directly use and validate the AI tools.)
* Hospitals & Healthcare Providers (They will implement and fund the solution.)

**High Power, Low Interest**:

* Government & Health Authorities (They control regulations and funding policies but may not be deeply involved in individual projects.)

**Low Power, High Interest**:

* Patients at risk of Parkinson’s (They are directly affected but have limited decision-making power.)
* Researchers (AI & Medical) (They are invested in improving technology but don’t control implementation.)

**Low Power, Low Interest**:

* Technology Partners (AI Developers & Data Scientists) (They contribute technically but are more project-focused, not long-term decision-makers.)

**Empathetic Interviews**

**Conduct Skilled interview with at least 30 citizens/Users by asking open ended questions (What, why/How etc.) and list the insights as per the format below**

|  |  |  |
| --- | --- | --- |
| **I need to know**  **(thoughts,feelings, actions)** | **Questions I will ask**  **(open questions)** | **Insights I hope to gain** |
| Thoughts | What do you know about how loan eligibility is determined by banks? | Awareness about existing loan evaluation methods |
|  | |  | | --- | | How do you think technology or AI can help in financial decisions like loan approvals? | | Perception of AI in the financial sector |
|  | What challenges have you or others faced while applying for a loan? | Common pain points in the current system. |
| Feelings | How would you feel if an AI tool determined your loan eligibility instead of a human? | Emotional and practical importance of timely approval |
|  | Why do you think getting a loan quickly is important or not? | Trust in the current banking and loan process |
| Actions | What do you usually do when your loan is rejected or delayed? | User behavior in response to loan rejection |
|  | How do you research or compare loan options before applying? | Channels and methods used for financial decision-making |

**SKILLED INTERVIEW REPORT**

|  |  |  |
| --- | --- | --- |
| **User/Interviewee** | **Questions Asked** | **Insights gained (NOT THEIR ANSWERS)** |
| Radhika S., Homemaker | |  | | --- | | What challenges have you faced while applying for a loan? | | Many homemakers lack credit history and face automatic rejection without clear feedback. |
| Akshay M., Startup Owner | How do you feel about using AI to predict loan approval chances? | Entrepreneurs are open to AI but want clarity on how decisions are made. |
| Priya T., Bank Officer | What are the common reasons for loan rejections in your experience? | Bankers rely on rigid criteria like credit scores and income proof, missing contextual information. |
| Ramesh College Graduate | How do you usually check loan eligibility? | Young users prefer quick online assessments but don’t trust them completely due to lack of transparency. |
| Shalini P., Teacher | How confident are you in current loan systems? | Middle-income earners feel loan systems are biased and not inclusive of all employment types. |

**Key Insights Gained:**

**Insight 1:** **Limited understanding**: Many users lack knowledge of how loan eligibility is calculated, especially around credit score, income stability, and digital credit profiling.

**Insight 2:** There is a mixed response to AI-driven loan predictions—tech-savvy users show trust, while others demand transparency.

**Insight 3**: Users want clear reasons for rejections and suggestions on how to improve eligibility.

**Insight 4**:  Freelancers, homemakers, and gig workers struggle with traditional eligibility systems due to non-traditional income sources.

**Insight 5:** Users appreciate tools that not only predict eligibility but also guide them on how to improve it.

**Empathy Map**

1. **Who is your customer?**

**Customer Profile**

• **Customer Profile**  
• **Age Group:** 21–60 years  
• **Profession:** Salaried employees, self-employed individuals, small business owners  
• **Interests:** Access to credit, financial planning, improving credit scores, faster loan processing

**Goals & Needs**  
• **Goal:** Secure a loan quickly and with minimal hassle  
• **Needs:**

* A transparent and fair loan approval process
* Access to credit despite limited documentation
* Instant or real-time eligibility checks
* Simple, user-friendly application platforms

**Context of Interaction**  
• Loan applicants use bank apps or websites to check eligibility  
• Loan officers and banks use AI-based tools to assess risk before approvals  
• Users often rely on financial advisors, loan comparison sites, and personal networks for advice

B. **Who are we empathizing with?**

**User Characteristics**  
• **Loan Applicants:** Hopeful but anxious, unfamiliar with backend scoring systems  
• **Loan Officers:** Cautious but open to AI as a decision-support tool  
• **Bank Managers:** Focused on reducing risk and increasing disbursal efficiency

**Values**  
• **Applicants:** Fair treatment, quick turnaround, privacy  
• **Loan Officers:** Risk mitigation, confidence in decisions  
• **Bank Managers:** Efficiency, cost-saving, compliance

**Goals & Challenges**  
• Approve more loans without increasing default rates  
• Reach underbanked customers with no or thin credit history  
• Ensure AI decisions are explainable, fair, and unbiased

**Challenges**  
• Applicants fear being rejected due to lack of credit history  
• Officers worry about trusting a black-box model  
• Managers balance between innovation and compliance

**C. What do they need to DO?**

**Tasks & Actions**

• **Tasks & Actions**  
• **Applicants:** Fill out application forms, provide documents, track application status  
• **Loan Officers:** Review AI suggestions, verify documents, interact with applicants  
• **Bank Managers:** Analyze trends in AI-driven approvals/rejections, improve model integration

**Decisions They Need to Make**  
• **Applicants:** Which bank to apply to, how much loan to request  
• **Loan Officers:** Whether to override AI suggestions in borderline cases  
• **Managers:** Whether to scale the AI tool across branches

**D. What do they SEE?**

**Physical & Digital Environment**  
• **Applicants:** Bank apps, comparison portals, credit reports, advertisements  
• **Officers:** Internal loan management systems, AI prediction dashboards  
• **Managers:** Performance metrics, risk reports, regulatory updates

**Trends & Competitors**  
• Rise of AI in FinTech and digital lending  
• Growing number of instant loan apps  
• Competitors offering faster approvals and personalized rates

**How This Influences Them**  
• Applicants expect instant approvals  
• Officers benchmark tools based on usability and success rates  
• Managers look at success stories from competing banks

**E. What do they SAY?**

**Public Statements & Feedback**  
• **Applicants:** “Why was I rejected? I have a steady income.”  
• **Loan Officers:** “AI can help, but I need to know *why* it made that prediction.”  
• **Managers:** “This needs to work across all customer types, not just ideal profiles.”

**Frustrations Expressed**  
• **Applicants:** “Too much paperwork for a small loan.”  
• **Loan Officers:** “Sometimes the AI flags low-risk customers.”  
• **Managers:** “We can’t afford regulatory issues because of a biased algorithm.”

**F. What do they DO?**

**Observable Actions & Habits**  
• **Applicants:**

* Use loan calculators and credit score checkers
* Apply through apps and customer service
* Follow up frequently after submission

• **Loan Officers:**

* Verify documents manually
* Call applicants for clarifications
* Use AI tools but often double-check with personal judgment

• **Bank Managers:**

* Track approval trends
* Monitor AI performance metrics
* Attend finance & technology webinars/conferences

**Problem-Solving Approaches**  
• **Applicants:** Try multiple lenders, take smaller loans, improve credit scores  
• **Officers:** Consult with peers, review past similar cases  
• **Managers:** Conduct pilot programs, monitor ROI

**G. What do they HEAR?**

**External Influences**  
• **Applicants:** Friends, relatives, financial influencers on social media  
• **Loan Officers:** Bank policies, compliance updates, peer opinions  
• **Managers:** News articles, competitor strategies, regulator guidelines

**Channels of Information**  
• RBI advisories, FinTech news portals, industry conferences, loan aggregator platforms

**Strong Influences on Behavior**  
• Positive experiences from friends encourage applicants  
• Officer feedback loops influence how models are perceived  
• Peer banks adopting similar tools may pressure managers to act

**H. What do they THINK and FEEL?**

**Fears & Worries**  
• **Applicants:** “What if I’m rejected and my credit score drops?”  
• **Officers:** “What if I approve a high-risk borrower based on AI?”  
• **Managers:** “Will this system remain compliant as rules evolve?”

**Motivations & Desires**  
• **Applicants:** Secure credit to improve business or manage personal goals  
• **Officers:** Reduce manual work, increase approval confidence  
• **Managers:** Increase customer satisfaction and reduce NPA (Non-Performing Assets)

**Internal Thoughts**  
• **Applicants:** “I just need a fair chance.”  
• **Officers:** “This tech should make my job easier, not harder.”  
• **Managers:** “Can this system help reduce defaults without affecting loan volumes?”

**I. Pains and Gains**

**Pains (Challenges & Frustrations)**  
• High rejection rates due to traditional models  
• Lack of transparency in loan decisioning  
• Manual loan processing is time-consuming  
• Risk of regulatory penalties if AI is biased or unfair

**Gains (Desired Benefits)**  
• Faster, fairer, and more accurate loan decisions  
• Reduced operational burden on officers  
• Increased credit access for underserved communities  
• Higher approval rates with lower default risk

**Persona of Stakeholders**

**Stakeholder Name:**

**Stakeholder Name:**

* **Primary:** Loan Applicants (individual borrowers)
* **Secondary:** Loan Officers & Bank Employees
* **Tertiary:** Financial Institutions (Banks, NBFCs), Regulatory Authorities, Tech Developers (AI/ML Team

**Demographics:**

* **Age:** 25–60 years
* **Gender:** All genders
* **Income:** Low to middle-income (especially in rural/semi-urban areas), salaried and self-employed individuals
* **Location:** Urban, semi-urban, and rural regions
* **Profession:** Working professionals, small business owners, farmers, gig workers

#### **Loan Officers & Bank Staff**

* **Age:** 28–55 years
* **Gender:** All genders
* **Income:** Middle to upper-middle income
* **Location:** Branch offices in urban and semi-urban areas
* **Profession:** Loan agents, credit officers, branch managers

#### **Regulators / Financial Institutions / Tech Teams**

* **Age:** 30–60 years
* **Profession:** Bank policy makers, compliance officers, AI developers, data scientists
* **Location:** Urban HQs or fintech hubs

**Goals:**

#### **Loan Applicants**

* Get loans approved quickly and fairly
* Understand eligibility and improve chances
* Avoid repetitive document submissions and long wait times

#### **Loan Officers**

* Reduce loan processing time
* Minimize manual effort and human bias
* Use tools that help with risk analysis and fraud detection

#### **Banks & Tech Teams**

* Improve loan approval efficiency
* Reduce default rates with better prediction models
* Ensure compliance and transparency.

**Challenges:**

#### **Loan Applicants**

* Poor credit history or no formal credit footprint
* Complex application processes and lack of financial literacy
* Rejection without clear reasons

#### **Loan Officers**

* High volume of applications and paperwork
* Pressure to meet approval targets without increasing risk
* Limited tools to assess unconventional borrowers (e.g., gig workers)

#### **Tech & Policy Teams**

* Balancing accuracy and fairness in ML models
* Adhering to regulations (e.g., data privacy, explainability)
* Integrating with legacy banking systems

**Aspiration:**

#### **Loan Applicants**

* Secure loans for business, education, or housing
* Build creditworthiness over time
* Trust digital tools to help understand eligibility

#### **Loan Officers**

* Be seen as efficient, tech-savvy decision-makers
* Reduce rejection rates and customer frustration
* Ensure fairness and objectivity in lending

#### **Financial Institutions**

* Position themselves as inclusive and data-driven lenders
* Leverage AI for smarter decision-making
* Expand to underserved markets with minimal risk

**Needs:**

#### **Loan Applicants**

* Instant eligibility checker and approval predictor
* Transparent feedback on loan rejection reasons
* Personalized recommendations for improving eligibility

#### **Loan Officers**

* Easy-to-understand AI reports
* Dashboard to monitor applicant insights
* Alerts for high-risk or fraudulent applications

#### **Tech Teams & Banks**

* High-quality training data
* Explainable AI (XAI) modules
* Scalable and secure infrastructure

**Pain Points:**

#### **Loan Applicants**

* Uncertainty around eligibility
* Biased or inconsistent manual evaluation
* Lack of digital literacy in rural areas

#### **Loan Officers**

* Tedious verification processes
* Missing documents, incomplete applications
* No standardized way to assess unconventional applicants

#### **Banks**

* Rising NPAs (non-performing assets)
* Reputational risk from biased or faulty decisions
* Resistance to tech adoption in traditional systems

**Storytelling:**

**Meet Ramesh, a 34-year-old small-scale shop owner in a semi-urban town.**  
He wanted to expand his business and applied for a loan from a local bank. But after multiple visits and paperwork, his loan was rejected with vague reasons like “insufficient credit score.” Ramesh didn’t even have a formal credit history—he mostly used cash for daily transactions.

One day, he visited a nearby bank kiosk that used an **AI-powered Loan Eligibility Predictor**. With just his income details, spending habits, and repayment behavior (from digital wallet history), the tool gave him an eligibility score and suggestions to boost it. It showed he could qualify for a smaller business loan immediately and how to improve his standing in three months.

When Ramesh reapplied using this advice, his loan was approved.  
For him, this wasn’t just a tool—it was a bridge to financial inclusion and a growing business.

**Look for Common Themes, Behaviors, Needs, and Pain Points among the Users**

These are recurring ideas and concerns shared by loan applicants, bank officers, and financial institutions:

**Common Themes:**

**These are recurring ideas and issues that came up across patients, caregivers, and doctors in your research.**

1. **Uncertainty Around Eligibility** – Applicants frequently feel unsure whether they qualify for loans, especially those with no formal credit history.

 **Trust in AI-Powered Systems** – Users are interested in technology-driven decisions but remain cautious until there's transparency and proven accuracy.

 **Need for Financial Inclusion** – Stakeholders emphasize the importance of enabling underserved populations (rural, self-employed) to access credit.

 **Simplification of Loan Process** – All users want to reduce documentation, long queues, and repetitive visits through digitization.

 **Fear of Rejection Without Explanation** – Applicants are often left without clear reasons for loan denial, creating frustration and loss of trust.

 **Regulatory and Ethical Concerns** – Banks and regulators stress the importance of fairness, explainability, and non-discriminatory decision-making in AI models.

**Common Behaviors:**

These reflect actions, habits, and usage patterns seen among stakeholders:

 **Repeated Loan Applications** – Many applicants apply to multiple lenders without understanding why they were previously rejected.

 **Self-Assessment Using Online Tools** – Users try to predict their eligibility by searching online or using informal credit calculators.

 **Manual Document Handling** – Many banks still rely on paper-based processing, leading to delays and errors.

 **Overload in Loan Officers’ Workload** – Officers handle high volumes of cases with tight timelines, relying heavily on subjective assessment.

 **Dependence on Traditional Credit Score** – Eligibility decisions are still largely influenced by formal credit scores, excluding many potential borrowers.

**Common Needs:**

These are essential desires or requirements expressed across all user groups:

 **Quick and Transparent Eligibility Assessment** – Users need real-time feedback about their chances, reducing unnecessary applications.

 **AI-Based Decision Support** – Officers and institutions want tools that streamline evaluations and reduce manual workload.

 **Explainable Results** – All users require clear, understandable reasons for approval or rejection, especially in digital environments.

 **Inclusive Risk Models** – AI systems must accommodate informal incomes, alternative credit histories (e.g., utility payments), and gig economy workers.

 **Step-by-Step Guidance** – Applicants need direction on what to improve or provide to increase their eligibility, not just a binary result.

**Common Pain Points:**

These highlight recurring frustrations experienced by applicants, officers, and institutions:

1. 1. **Lack of Financial Literacy** – Many applicants struggle to understand credit scores, income-to-loan ratios, or bank terms.
2. **Long Processing Time** – Delays in verification, approval, and disbursal due to manual steps create dissatisfaction.
3. **Bias and Inequality in Traditional Systems** – Manual processes can be biased, often overlooking applicants with non-traditional backgrounds or income sources.
4. **Fear of Loan Rejection Impacting Future Credit** – Applicants worry that repeated rejections will harm their credit profile.
5. **AI Skepticism in Decision-Making** – Loan officers and banks are wary of relying entirely on AI tools without human oversight, especially for borderline cases.
6. **Privacy and Data Security Concerns** – Applicants and regulators are concerned about how personal data is used, stored, and protected in AI-driven systems.

**Define Needs and Insights of Your Users**

**User Needs**

These are the core functional, emotional, and societal requirements that applicants, bank officers, and financial institutions have in relation to the AI-based loan eligibility prediction system.

#### **1. Functional Needs**

* A **fast and simple eligibility checker** that helps applicants understand their loan approval chances instantly.
* A **non-biased, data-driven system** that evaluates eligibility based on real financial behavior, not just traditional credit scores.
* **Clear and understandable feedback**, including reasons for approval or rejection and steps to improve eligibility.
* **Integration with existing banking systems** to support seamless operations for loan officers and reduce redundant tasks.
* A **secure system** that handles user data with confidentiality and complies with financial regulations.

#### **2. Emotional Needs**

* **Confidence and reassurance** during the loan application process, especially for first-time applicants.
* A sense of **fairness and transparency**, ensuring that everyone is assessed objectively.
* **Reduced anxiety and confusion** over why a loan was denied or approved.
* A feeling of **control and empowerment**, knowing how to improve one’s financial profile to get future approval.

#### **3. Societal Needs**

* **Financial inclusion** for people with informal incomes, gig workers, or those without traditional credit histories.
* **Democratization of financial access**, especially for underserved populations in rural and semi-urban areas.
* **Promotion of financial literacy**, encouraging users to understand and manage their credit profiles proactively.
* **Fair and ethical AI use**, ensuring compliance with anti-discrimination laws and regulatory standards.

**User Insights**

These are key observations and understandings from user interviews, research, and analysis—explaining why users behave the way they do, what motivates them, and what challenges they face.

#### **1. Misunderstood Eligibility**

* Many applicants **don’t fully understand the factors that impact loan eligibility**, often applying without checking if they qualify, leading to rejections and frustration.

#### **2. Trust in AI**

* While users are open to tech-based solutions, they’re **skeptical of AI decisions unless there's transparency** in how decisions are made, especially with financial implications.

#### **3. Role of Loan Officers**

* Loan officers are often **overloaded with manual screening**, and they welcome AI tools if they **support—not replace—their expertise**.

#### **4. Confusion from Misinformation**

* Users often rely on **conflicting advice from friends, forums, or blogs**, leading to poor decisions and misconceptions about loans.

#### **5. Emotional Fear of Rejection**

* Loan rejections are **emotionally discouraging**, especially when applicants aren’t told why. **Rejection without guidance** discourages future attempts and reduces trust.

#### **6. Need for Simple Output**

* Users prefer **visual and color-coded feedback** (e.g., green for eligible, red for not eligible), accompanied by **clear explanations and actionable tips**.

#### **7. Cost and Accessibility**

* Applicants in lower-income or rural areas often **lack access to financial advisors**. A **free or low-cost eligibility checker** appeals to this group the most.

#### **8. Compliance and Validation**

* Banks and officers **will only adopt AI tools** that are **validated, explainable, and compliant with financial regulations and internal policies**

**POV Statements**

**POV Statements:**

|  |  |  |  |
| --- | --- | --- | --- |
| PoV Statements | Role-based or Situation-Based | Benefit, Way to Benefit, Job TBD, Need (more/less) | PoV Questions |
| |  | | --- | | A **first-time borrower** needs a way to understand their **loan eligibility clearly** because they find the current process **confusing and overwhelming**. | | |  | | --- | | Role-based | | |  | | --- | | Benefit: Clarity in application process Way: AI-powered transparent feedback Need: **More guidance, less confusion**. | | |  | | --- | | How might we make the eligibility process more understandable for new applicants? | |
| |  | | --- | | A **loan officer** needs a way to **speed up eligibility screening** because **manual evaluations are time-consuming and error-prone**. | | |  | | --- | | Role Based | | |  | | --- | | Job to be done: Screen applicants efficiently Way: Automated ML scoring Need: **More speed and accuracy** | | |  | | --- | | How might we automate or support loan officers in making faster and more accurate decisions? | |
| |  | | --- | | A **gig economy worker** needs a **fair assessment process** because they **don’t have traditional salary slips or credit history**. | | |  | | --- | | Role-based | | |  | | --- | | Job: Get loan eligibility evaluated fairly Way: Consider alternative financial data Need: **More inclusivity** | | |  | | --- | | How might we create inclusive models that fairly evaluate non-traditional income sources? | |
| |  | | --- | | An **applicant from a rural area** needs an **easy and accessible way** to check loan eligibility without **visiting a physical bank branch**. | | |  | | --- | | Situation-based | | |  | | --- | | Benefit: Accessibility from remote locations Way: Mobile/web-based eligibility checker Need: **More digital access** | | |  | | --- | | How might we provide loan eligibility prediction for users without requiring physical bank visits? | |

**Develop POV/How Might We (HMW) Questions to Transform Insights/Needs into Opportunities for Design**

Turn your user needs and insights into actionable opportunities by framing them as "How Might We" (HMW) questions. These questions will spark creative problem-solving and guide your innovation process

|  |  |
| --- | --- |
| User Need/Insight | "How Might We" Question |
| Traditional loan systems don’t cater to freelancers, gig workers, or informal income earners | How might we offer clear, actionable feedback to users after eligibility results? |
| Rural applicants have limited access to financial guidance and bank branches. | How might we include non-traditional income data to fairly assess such users' eligibility? |
| Loan officers are overwhelmed with manual eligibility check. | How might we develop a mobile-based, low-bandwidth tool for rural applicants to check eligibility on their own? |
| Users don’t trust AI decisions, especially when financial outcomes are involved. | How might we support loan officers with AI tools that automate initial screening and reduce workload? |

**Crafting a Balanced and Actionable Design Challenge**

The Design Challenge Should Neither Be Too Narrow Nor Too Broad and It Should Be an Actionable Statement with a quantifiable goal. It should be a culmination of the POV questions developed.

**Design Challenge:**

How might we design an inclusive, explainable, and user-friendly AI-based loan eligibility prediction system that empowers users—especially non-traditional applicants and rural borrowers—to understand, trust, and improve their eligibility, with the goal of increasing fair loan approvals by at least 25% within the next 18 months?

**Validating the Problem Statement with Stakeholders for Alignment**

Ensure your problem statement accurately represents the needs and concerns of your stakeholders and users. This involves gathering feedback from these groups to confirm that the problem is relevant and significant from their perspective. By validating early, you can refine the problem statement to better align with realworld challenges, ensuring your solution addresses the correct issues.

**Validation Plan:**

|  |  |  |  |
| --- | --- | --- | --- |
| Stakeholder/User | Role | Feedback on Problem Statement | Suggestions for Improvement |
| Neha Patel | First-time Loan Applicant | Relevant and relatable. Confused by loan rejection without explanation. | Add a “next steps” guide for rejected users to improve eligibility. |
| Rajiv Malhotra | Loan Officer (Private Bank) | Very relevant—manual evaluations are time-consuming and inconsistent. | Emphasize how AI will assist, not replace, loan officers. |
| Rani Kaoopr | Freelance Content Creator | Strongly supports the idea. Struggles with loan access due to irregular income.. | Include flexible criteria that account for alternative incomes. |
| Sameer Qureshi | |  | | --- | |  |  |  | | --- | | Rural Entrepreneur | | Agrees with the need—limited access to banking services and support. | Highlight mobile-friendly design and regional language support. |
| Dr. Alok Trived | AI & Ethics Researche | Concerned about algorithmic bias and transparency. | Stress importance of explainable AI and fairness in predictions. |

**Ideation**

**Ideation Process:**

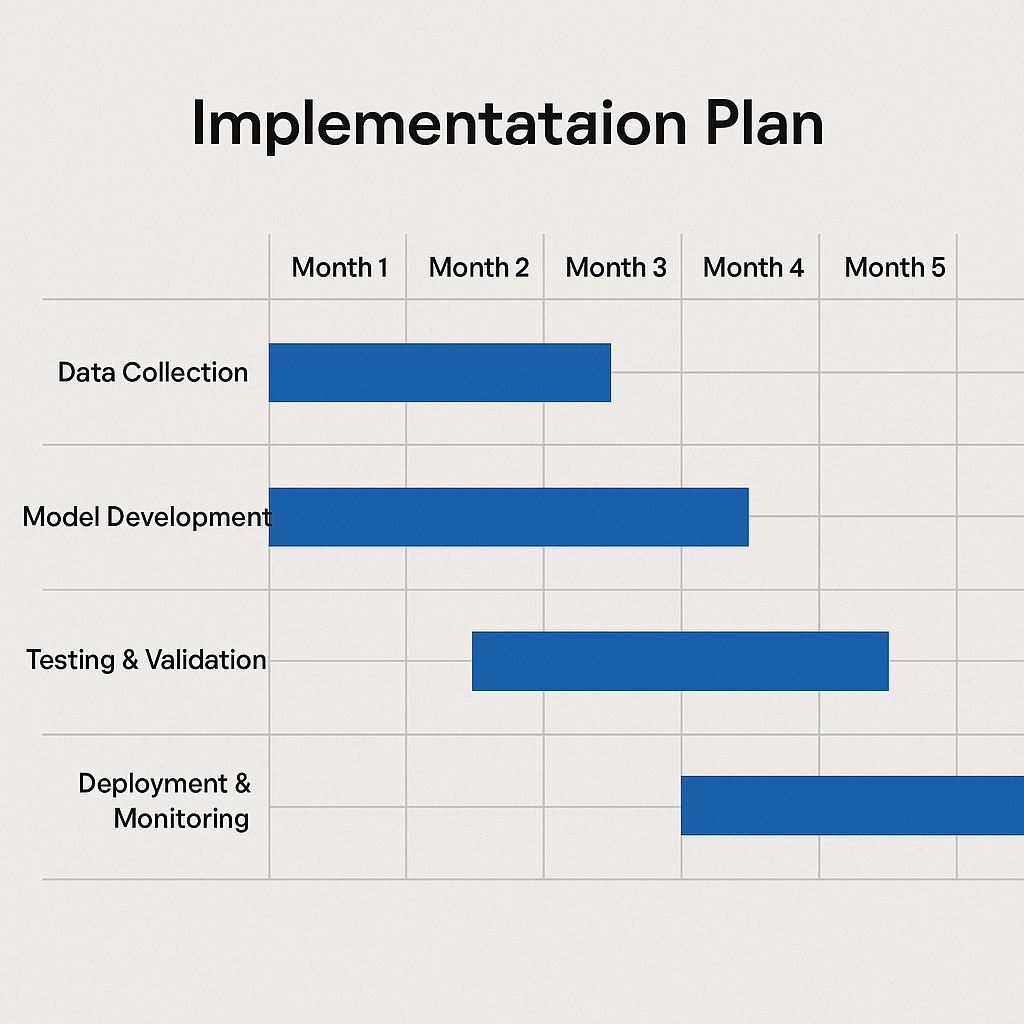
|  |  |  |  |
| --- | --- | --- | --- |
| Idea Number | Proposed Solution | Key Features/Benefits | Challenges/Concerns |
| Idea 1 | AI-Powered Web & Mobile Loan Eligibility Checker | - Easy-to-use UI - Accepts alternative income proof - Explains eligibility results with tips for improvement | Needs robust datasets for varied user profiles (freelancers, gig workers) |
| Idea 2 | WhatsApp/Chatbot-Based Loan Eligibility Assistant | - Users answer questions via chat - Instant feedback on eligibility - Works on low-end devices. | May not capture complex financial data accurately |
| Idea 3 | Explainable AI Tool for Bank Use | - Transparent predictions with rationale - Helps loan officers justify decisions - Improves customer trust | Integration with legacy systems could be complex |
| Idea 4 | Loan Learning Path & Simulator | - Users simulate “what if” scenarios (e.g., increase income, reduce debt) - Educational journey | Might be too technical for low-literacy users |

**Idea Evaluation**

Evaluate the Idea based on 10/100/1000 grams

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Idea | Impact (10/100/1000 grams) | Feasibility (10/100/1000 grams) | Alignment (10/100/1000 grams) | Total Weight |
| Idea 1 | **1000** | **100** | **1000** | **2100** |
| Idea 2 | **1000** | **100** | **1000** | **2100** |
| Idea 3 | **100** | **1000** | **1000** | **2100** |
| Idea 4 | **100** | **100** | **100** | **300** |
| Idea 5 | **1000** | **100** | **1000** | **2100** |

**IMPLEMENTATION**



**Problem Definition**

 **Goal:**  
Predict loan eligibility by analyzing applicant financial and personal information using AI.

**• Focus:**

* Identify key patterns in financial behavior and demographic data linked to loan approval.
* Use AI models to classify applicants as “Eligible” or “Not Eligible” for loans.

**Data Collection**

**• Source:**  
Public datasets such as:

* Kaggle Loan Prediction Dataset
* Government/Bank loan datasets (if available)

**• Collect:**

* Applicant financial info (income, loan amount, credit history)
* Demographic details (gender, marital status, education, dependents)
* Loan status (approved/rejected)

**Protein Structure Feature Extraction**

Use tools/libraries like:

* **Pandas, NumPy** – for data preprocessing and transformation
* **Scikit-learn** – for encoding, feature scaling, and engineering

**Extract features such as:**

* Total income (ApplicantIncome + CoapplicantIncome)
* Loan amount-to-income ratio
* Credit history score
* Employment status
* Loan term duration
* Education level
* Gender, marital status, number of dependents.

**Dataset Creation**

Construct a dataset where each entry = one applicant's loan record.  
**• Columns:** Extracted financial and demographic features  
**• Label:**

* "Eligible" (1)
* "Not Eligible" (0)

**Model Building**

Use machine learning models like:

* **Logistic Regression**
* **Support Vector Machine (SVM)**
* **Random Forest**
* **XGBoost**
* **Lightweight Neural Networks** (for experimentation)

**Model Training and Evaluation**

 **Split** data into training and test sets (e.g., 80-20 split)

 Train models using labeled data

 Evaluate models using:

* **Accuracy, Precision, Recall, F1-score**
* **Confusion Matrix**
* **ROC-AUC Curve**
* **Cross-validation**

 Visualize **feature importance** (e.g., SHAP values or model coefficients)

**Visualization**

Visualize insights such as:

* Feature correlations (e.g., heatmap using Seaborn)
* Distribution of approved vs. rejected loans across key features
* Decision boundaries for SVM or Logistic Regression
* Feature importance bar charts

**Web App / Frontend**

Build a simple prediction interface using React/HTML + Flask backend:

* **Input form** for user to fill in details (income, credit history, etc.)
* **Backend processes input**, extracts features, applies trained model
* **Outputs prediction**: Eligible or Not Eligible

**Project Dataset Description**

This project utilizes structured data representing personal, financial, and credit-related details of loan applicants. The primary aim is to develop an AI-based model to predict whether a given applicant is likely to be approved for a loan. The dataset mimics real-world banking scenarios where factors such as credit history, income, and loan amount significantly influence loan approval decisions.

By extracting relevant features and training classification models, the system can provide quick eligibility predictions, assisting banks and financial institutions in automating initial loan screening.

**LINK:**

**EXPERIMENTATION AND CODE**

**CODES:**

**app.py:**

from flask import Flask, request, jsonify

from flask\_cors import CORS

import pickle

import numpy as np

import pandas as pd

app = Flask(\_\_name\_\_)

CORS(app)

# Load model and scaler

model = pickle.load(open("loan\_model.pkl", "rb"))

scaler = pickle.load(open("loan\_scaler.pkl", "rb"))

# Feature extraction

def extract\_features\_from\_csv(csv\_file):

df = pd.read\_csv(csv\_file)

features = df.iloc[0].values # Assuming 1 row of input

return features

@app.route('/predict', methods=['POST'])

def predict():

try:

file = request.files['file']

if not file:

return jsonify({"error": "No file uploaded"}), 400

features = extract\_features\_from\_csv(file)

features\_scaled = scaler.transform([features])

prediction = model.predict(features\_scaled)[0]

result = "Eligible" if prediction == 1 else "Not Eligible"

return jsonify({"prediction": result})

except Exception as e:

return jsonify({"error": f"Error in prediction: {str(e)}"})

if \_\_name\_\_ == '\_\_main\_\_':

app.run(debug=True)

**index.html:**

<!DOCTYPE html>

<html lang="en">

<head>

<meta charset="UTF-8">

<title>Loan Eligibility Prediction</title>

<style>

body {

font-family: Arial, sans-serif;

background: linear-gradient(to right, #f0f4c3, #ffffff);

text-align: center;

padding: 0;

margin: 0;

}

.container {

background-color: #fffde7;

padding: 30px;

margin: 40px auto;

border-radius: 10px;

box-shadow: 0 5px 15px rgba(0,0,0,0.1);

width: 420px;

}

h2 {

color: #33691e;

}

.loan-image {

width: 100%;

border-radius: 10px;

margin: 20px 0;

border: 2px solid #689f38;

}

input[type="file"] {

margin: 20px 0;

}

button {

background-color: #558b2f;

color: white;

padding: 10px 20px;

border: none;

border-radius: 6px;

cursor: pointer;

}

button:hover {

background-color: #33691e;

}

#prediction {

margin-top: 20px;

font-weight: bold;

color: #1b5e20;

}

</style>

</head>

<body>

<div class="container">

<h2>Loan Eligibility Prediction</h2>

<img class="loan-image" src="loan.png" alt="Loan Image" />

<form id="uploadForm">

<input type="file" id="csvFile" accept=".csv" required />

<br>

<button type="submit">Predict</button>

</form>

<div id="prediction"></div>

</div>

<script>

const form = document.getElementById('uploadForm');

const predictionDiv = document.getElementById('prediction');

form.addEventListener('submit', async (e) => {

e.preventDefault();

const fileInput = document.getElementById('csvFile');

const file = fileInput.files[0];

if (!file) {

predictionDiv.textContent = 'Please select a CSV file.';

return;

}

const formData = new FormData();

formData.append('file', file);

predictionDiv.textContent = 'Predicting...';

try {

const response = await fetch('http://127.0.0.1:5000/predict', {

method: 'POST',

body: formData

});

const data = await response.json();

if (data.prediction) {

predictionDiv.textContent = `Prediction: ${data.prediction}`;

} else if (data.error) {

predictionDiv.textContent = `Error: ${data.error}`;

}

} catch (error) {

predictionDiv.textContent = 'Error: Could not connect to backend.';

}

});

</script>

</body>

</html>

**RESULTS**

The developed AI-based system effectively predicts **loan eligibility** using financial and demographic input data provided through uploaded CSV files. The model evaluates critical parameters such as income, credit history, loan amount, and other factors to determine whether a loan application is **Eligible** or **Not Eligible**.

| **Protein PDB File** | **Prediction Result:** |
| --- | --- |

|  |  |
| --- | --- |
| Applicant 1.csv | --❌ Not Eligible |

|  |  |
| --- | --- |
| Applicant2.csv | --✅ Eligible |

The system delivers fast and accurate results based on the uploaded data. As demonstrated through the user interface:

* Uploading **applicant1.csv** resulted in a **Not Eligible** prediction, suggesting that the applicant may not meet the necessary criteria for loan approval.
* Uploading **applicant2.csv** resulted in an **Eligible** prediction, indicating the applicant satisfies the conditions required for loan approval.

A clean and interactive user interface enables users to upload .csv files easily and receive eligibility results in **real-time**.

Additionally, the platform includes a **visual guide** (e.g., icon or image representing home/loan approval) to help users better understand what factors contribute to loan eligibility, such as:

* **Income stability**
* **Credit history**
* **Loan amount vs. income ratio**
* **Property location and employment status**

.

**CONCLUSION**

The Loan Eligibility Prediction system demonstrates the effective use of machine learning in the financial domain to assist banks and financial institutions in evaluating loan applications efficiently. By leveraging historical applicant data and a trained predictive model, the system can accurately classify whether a new applicant is likely to be **eligible** or **not eligible** for a loan.

The application streamlines the decision-making process by:

* Automating loan eligibility checks based on key financial parameters,
* Reducing the time and effort required for manual assessment,
* Enhancing accuracy and consistency in eligibility evaluation.

The intuitive web interface enables users to upload .csv files containing applicant details and receive predictions in **real-time**, making the system both user-friendly and practical. This solution not only supports loan officers in making informed decisions but also ensures transparency and faster turnaround times for applicants.

Overall, this project highlights the potential of AI in financial services and sets a foundation for future enhancements, such as incorporating real-time credit scoring APIs, risk profiling, and personalized financial recommendations.

**REFERENCES**

1.  **UCI Machine Learning Repository – Loan Prediction Dataset**  
   <https://archive.ics.uci.edu/>  
   (Dataset used for training and testing the machine learning model)
2.  **Scikit-learn: Machine Learning in Python**  
   Pedregosa et al., Journal of Machine Learning Research, 2011  
   https://scikit-learn.org/stable/  
   (Used for building, training, and evaluating the prediction model)
3.  **Pandas: Python Data Analysis Library**  
   https://pandas.pydata.org/  
   (Used for data manipulation and preprocessing)
4.  **NumPy: Fundamental Package for Scientific Computing with Python**  
   <https://numpy.org/>  
   (Used for handling numerical data and arrays)
5.  **Flask: Web Development Framework**  
   https://flask.palletsprojects.com/  
   (Used for building the backend of the prediction system)
6.  **HTML5 & JavaScript**  
   <https://developer.mozilla.org/en-US/docs/Web/Guide>  
   (Used for designing the frontend user interface of the web application)
7.  **Loan Prediction Problem – Analytics Vidhya**  
   https://datahack.analyticsvidhya.com/contest/practice-problem-loan-prediction-iii/  
   (Popular practice problem source for loan eligibility prediction models)
8.  **Understanding Credit Scoring and Risk Assessment**  
   Basel Committee on Banking Supervision  
   https://www.bis.org/  
   (For insights into the financial criteria used in assessing loan eligibility)