

AI-Powered Loan Eligibility Advisor System



Model Research & Roadmap

Project Title: *AI-Powered Voice-Enabled Loan Advisory Chatbot*



Project Overview

Loan approval processes are often **slow, inconsistent, and difficult to understand**. This project aims to build a **global AI chatbot** that:

- Predicts loan eligibility using **ML models**.
 - Accepts **payslip uploads & annual income** as core eligibility parameters.
 - Communicates with users via **text and speech** (speech-to-text + text-to-speech).
 - Provides **explainable advisory reports** and **personalized guidance**.
 - Scales globally with **cloud deployment + database integration**.
-



Research & Literature Insights

1. Loan Eligibility Prediction

- Logistic Regression, Random Forest, and Gradient Boosting are widely used for **loan approval classification**.
- Debt-to-Income (DTI) and Loan-to-Income (LTI) ratios are strong predictors of creditworthiness.

2. Document Verification

- Payslip parsing via **OCR** (Google Document AI, AWS Textract, Tesseract) is standard for extracting structured salary data.
- Fraud detection research uses **image tampering detection** and **pattern analysis** in financial docs.

3. Conversational AI in Finance

- **Rasa, LangChain, and GPT-based bots** are widely researched for financial advisory.
- **Explainability (XAI)** is important to gain trust (SHAP/LIME for ML interpretability).

4. Voice-Enabled Chatbots

- STT: Whisper, Google Speech-to-Text, AWS Transcribe.
- TTS: Amazon Polly, Google WaveNet, Azure Neural TTS.
- Studies show voice assistants improve accessibility and user engagement.

Final System Architecture

Components

1. Frontend (User Interface)

- React.js / Next.js (Web)
- Flutter (Mobile App)
- Features:
 - Chatbot (text + speech)
 - Payslip upload
 - Loan calculators

- Dashboard & reports

2. Backend (Core Services)

- FastAPI / Django REST Framework
- Modules:
 - **STT** (Whisper/Google/AWS)
 - **NLP Chatbot** (LangChain + GPT API / Rasa)
 - **ML Loan Eligibility Model** (XGBoost + Business Rules)
 - **OCR Service** (AWS Textract/Google DocAI for payslip parsing)
 - **TTS** (Amazon Polly / Google TTS)

3. Database (Cloud-hosted)

- **PostgreSQL/MySQL** → structured user & loan data
- **MongoDB** → chatbot logs & unstructured queries

4. Cloud & Deployment

- **AWS/GCP/Azure**
- **Docker + Kubernetes** for containerization & scaling
- **CloudFront/Cloudflare** for global low-latency access

Data Flow

1. User speaks → **STT** → **Text**
2. Text → **Chatbot NLP + Loan Eligibility ML**
3. User uploads payslip → **OCR** → **Structured Income Data**

4. Eligibility Engine checks:

- Payslip income
- Annual income
- Credit score
- DTI/LTI ratios
- Loan amount requested

5. Response generated → **Text** → **TTS** → **Voice Reply**

6. Report stored in DB + downloadable PDF summary.



Loan Eligibility Rules + ML

- **Hard Conditions (Rules Engine):**

- Payslip mandatory for salaried applicants.
- Annual income \geq minimum threshold per loan type.
- DTI \leq 40%.

- **ML Prediction:**

- Predict loan approval probability (e.g., XGBoost model).
- Explain eligibility via SHAP values (XAI).



Roadmap

Phase 1: Planning & Requirements

- Finalize features & compliance requirements (RBI, GDPR, etc.).
- Define loan types & thresholds.

Phase 2: Data & Model Development

- Collect loan datasets (public + synthetic).
- Create payslip dataset for OCR training.
- Train ML model (XGBoost + Explainability).

Phase 3: Chatbot + Voice Integration

- Build chatbot flow in Rasa / LangChain + GPT API.
- Integrate **STT (Whisper/Google)** + **TTS (Polly/Google TTS)**.

Phase 4: Backend & DB

- Set up FastAPI backend with APIs.
- Database schema for users, loans, payslips.
- OCR pipeline for payslip parsing.

Phase 5: Frontend Development

- React/Next.js web app with chatbot UI.
- Upload interface for payslips.
- Voice-enabled chat window.

Phase 6: Deployment & Scaling

- Containerize (Docker).

- Deploy on AWS/GCP/Azure.
- Use Kubernetes for scaling.
- Global access via CDN.

Phase 7: Security & Compliance

- Encrypt data (at rest + in transit).
- Implement role-based access.
- Ensure compliance with financial regulations.

Phase 8: Monitoring & Improvements

- Add fraud detection for payslips.
- Add multilingual voice support.
- Continuous retraining of ML model with new data.



Innovative Features (Add-ons)

- **Explainable AI Reports** → Show “why” user is eligible/not.
 - **Multilingual Voice Chatbot** → English, Hindi, Spanish, Mandarin.
 - **WhatsApp/Telegram Integration.**
 - **Gamified Financial Health Score** → Tips to improve loan eligibility.
 - **AI Fraud Detection** → Prevent fake payslips & manipulated income docs.
-