Team Avish

Team member 1 : AASHISH KUMAR RA

Team member 2 : AVANTIKA G

Problem Statement 2: Real-Time Credit Risk Assessment Using Alternative Data

Github link : Project Link (Credit-Risk-Analysis)

Github link : Project Link (Banking App)

Github link : Project Link (Predict loan default risk)

PROPOSED SOLUTION

Idea Solution

- **AI-powered platform** for real-time credit risk assessment, integrating traditional financial data and alternative data sources like social media, utility payments, and spending habits.
- Uses **Machine Learning** to predict creditworthiness and deliver explainable **credit risk scores** to lenders.
- Incorporates alternative data sources to improve accessibility for individuals with limited credit history.
- **Real-time dashboard** displaying risk analysis, borrower profiles, and risk heatmaps, ensuring transparency and usability.
- Data sources include:
  - Social media activity for sentiment analysis.
  - **Utility** and **rent payments** for financial stability.
  - Geolocation data and **spending patterns** for behavioral insights.
- Implements **secure Blockchain** for storing sensitive borrower and transaction data, ensuring trust and data integrity.

Voice Bot for Lenders

A real-time voice assistant to provide **credit risk details**, borrower profiles, and **scoring trends via phone calls**, supporting multiple languages.

AI-Driven Insights

Machine learning models provide actionable recommendations for improving credit **scores and identifying risk patterns.**

Blockchain-Based Smart Contracts

Immutable and secure contracts that verify alternative data inputs, **offering tamper-proof validation**.

What-If Scenarios

Enables lenders to foresee the impact of alternative data or sudden financial **changes on credit risk**, improving decision-making.

Dynamic Credit Score Optimization

A feature that provides **borrowers with actionable** insights on how to improve their credit scores by adjusting spending habits, paying off specific debts, or maintaining consistent utility payments.

5-Km Risk Detection for Lenders

Geolocation-based analysis showing credit risk hotspots within a customizable radius, helping lenders prioritize regions for offering financial products.

Problem Resolution

- **Early Detection:** Identifies risky borrowers early through AI analysis of alternative data, reducing loan defaults.
- **Enhanced Accessibility:** Alternative data integration enables individuals with limited credit history to access loans.
- **Optimized Lending Decisions:** Provides lenders with explainable credit risk models for confident decision-making.
- **Conflict-Free Integration:** Resolves discrepancies between traditional and alternative data with a robust verification pipeline.
- **Feedback Mechanism:** Public users can provide insights and feedback on creditworthiness recommendations via the platform.

TECHNICAL APPROACH

Smart Contract Deployment
Ensures secure data exchange and validation between traditional and alternative data sources.

PROVIDER

METAMASK

INFURA

IPFS/ SWARM

Block

Block

Block

ETHEREUM VIRTUAL MACHINE

Smart Contracts

Smart Contracts

Smart Contracts

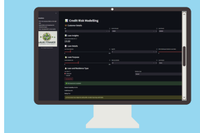
Smart Contracts

Why is it used?

Variable	Value Bucket	Scorecard Point
AGE	<22	100
AGE	22<=AGE<26	120
AGE	26<=AGE<29	185
AGE	29<=AGE<32	200
AGE	32<=AGE<36	210
AGE	36<=AGE<42	225
AGE	>=42	250
HOME STATUS	OWN	225
HOME STATUS	RENT	110
INCOME	<10000	120
INCOME	10000<=INCOME<17000	140
INCOME	17000<=INCOME<25000	180
INCOME	25000<=INCOME<35000	200
INCOME	35000<=INCOME<58000	225
INCOME	58000<=INCOME<100000	230

FRONTEND

- Borrower dashboard to track credit improvement tips and risk scores.
- Lender dashboard showcasing real-time risk assessments, "What-If" simulations, and interactive heatmaps.



METAMASK

Singer

API GATEWAY

SSL Termination

Load Balancer

Centralized access point for seamless communication between frontend, backend, and third-party data sources.

BACKEND

Authentication and Authorization Layer

(O Auth 2.0,JWT)

AI Services and Models

(Tensor flow, neo4j,LangChain)

Real Time Database PostgreSQL and Django

API (Flask)

API

contain app logic

djando

dj

URL Patterns

Serializer

Views

Models

Web Server

Request

Handshake

Websocket

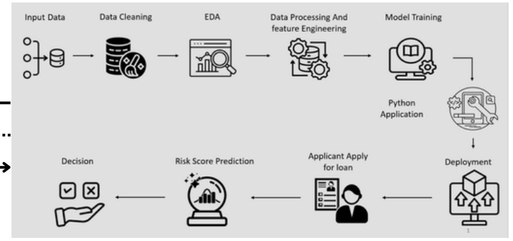
DB responds and server caches the information

CLOUD SERVICES

AWS S3

aws

Backend processes the upload request File is stored in cloud storage



CreditRisk AI (Our AI-driven insights platform) enhances lenders' ability to process and analyze complex data, including financial and alternative sources like social media, utility payments, and spending habits. It supports transparent decision-making, risk identification, and borrower engagement in real time.

Database

Text Chunking

Data Storage and Retrieval

AI Models

neo4j

Visualization Tools

Search & Retrieve details asked in by user

Crew AI (Generation of Daily Reports)

Automated Data Integration
Aggregates traditional and alternative data sources for comprehensive credit assessment.

Explainable Credit Scoring
Breaks down credit risk scores into actionable insights for lenders.

Risk Prioritization
Identifies and ranks high-risk borrowers using advanced machine learning models.

Real-Time Decision Support
Enables instant lending decisions with explainable AI outputs and simulations.

Milo(Voice assistant)

Milo

VoIP Provider

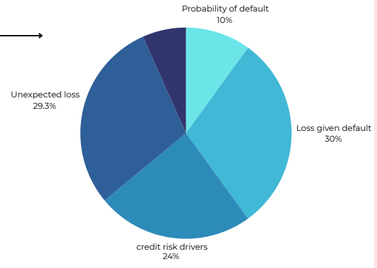
User

Process

Text to Speech

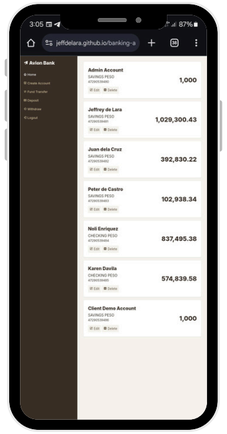
Response

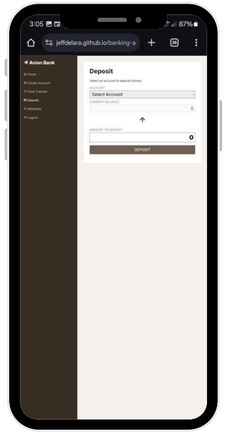
Speech to text

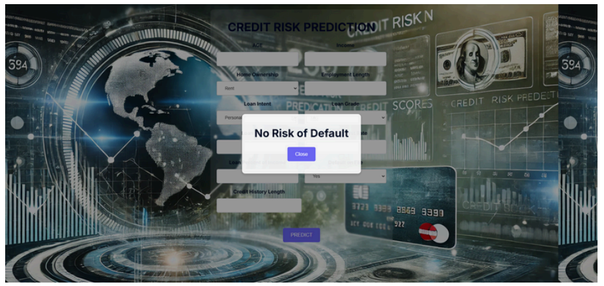


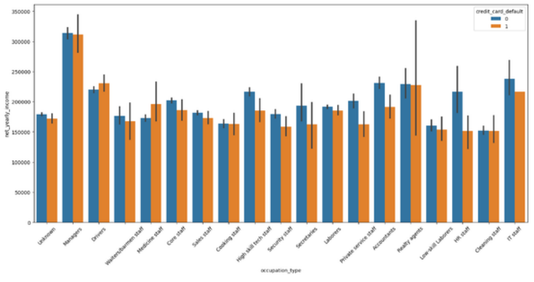
TECH STACK

Avion Bank App









FEASIBILITY AND VIABILITY

Technological Feasibility

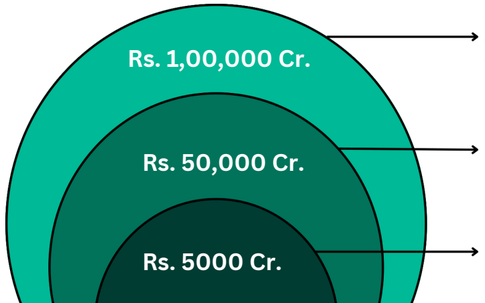
- **Integration:** Describe the key technologies, tools, and frameworks your project will use.
- **Features:** Highlight specific functionalities your system will offer (e.g., real-time credit risk scores, secure data storage).
- **Development Expertise:** Mention the technical stack or frameworks (e.g., React, Django, or blockchain solutions).

Market Feasibility

- **Urban Demand/Target Market:** Explain how your solution aligns with the needs of your target market (e.g., underserved borrowers, financial inclusion).
- **Stakeholder Engagement:** Outline the key stakeholders involved (e.g., lenders, borrowers, regulators).
- **Public Engagement:** Describe how you will engage users (e.g., through voice bots, dashboards).

Financial Feasibility

- **Subscription Model:** Explain pricing tiers or freemium models for lenders.
- **Transaction Revenue:** Highlight any additional income streams (e.g., per-transaction fees).
- **Cost Savings:** Show how your solution reduces costs (e.g., loan default risk, operational inefficiencies).



TAM (Total Addressable Market): Estimate the overall market size your solution targets. TAM = ₹20,00,000 crore × 0.05 = ₹1,00,000 crore annually.

SAM (Serviceable Addressable Market): Narrow it down to the realistic portion of the market you can capture in the medium term.SAM = ₹1,00,000 crore × 0.50 = ₹50,000 crore annually.

SOM (Serviceable Obtainable Market): Define the short-term realistic market share your solution can achieve. SOM = ₹50,000 crore × 0.10 = ₹5,000 crore annually.

Potential Challenges and Risks

Technical Challenges

- **Integration Complexity:** Combining traditional credit data with alternative data sources like social media and utility payments requires advanced integration frameworks.
- **Model Performance:** Ensuring machine learning models are accurate, unbiased, and interpretable can be challenging.

Market Adoption Risks

- **Adoption Resistance:** Lenders may be hesitant to shift from traditional credit assessment methods.
- **Competition:** Emerging fintech solutions with similar offerings pose a market threat.

Financial Risks

- **Budget Constraints:** Limited funding may hinder the development and deployment of the system.
- **Revenue Uncertainty:** Predicting returns on investment from incorporating alternative data can be uncertain.

Operational Risks

- **User Training:** Lack of adequate training for lenders on using AI-driven platforms.
- **Data Privacy Concerns:** Ensuring sensitive borrower data is securely managed and compliant with regulations.

Strategies for Overcoming Challenges

Technical Strategies

- **Robust Data Integration:** Utilize advanced middleware to seamlessly integrate traditional and alternative data sources.
- **Explainable AI Models:** Incorporate frameworks to enhance the transparency and trustworthiness of credit risk models.

Market Strategies

- **Stakeholder Engagement:** Organize workshops to demonstrate the benefits of alternative data in credit assessment.
- **Awareness Campaigns:** Promote the platform through social media and industry-specific events to build credibility.

Financial Strategies

- **Diversified Revenue Streams:** Explore partnerships, consultancy, and premium features to boost revenue.
- **Efficient Budget Allocation:** Regularly review and prioritize budget for critical development areas.

Operational Strategies

- **Comprehensive Training Programs:** Develop user-friendly guides and training modules for lenders.
- **Secure Blockchain Integration:** Ensure data privacy and compliance by leveraging blockchain for sensitive borrower data.

IMPACT AND BENEFITS

POTENTIAL IMPACT ON THE TARGET AUDIENCE

Government Departments

- **Economic Growth:** Facilitates financial inclusion, driving regional development and entrepreneurship.
- **Transparency:** Explainable AI models ensure fairness and adherence to regulatory standards.

Contractors and Media Agencies

- **Improved Loan Accessibility:** Alternative data usage helps individuals with limited credit history access loans.
- **Financial Education:** Actionable AI-driven recommendations provide guidance on improving credit scores.

Public

- **Enhanced Decision-Making:** Real-time credit risk analysis and alternative data integration support accurate and confident lending decisions.
- **Risk Mitigation:** Early identification of high-risk borrowers reduces loan defaults and financial losses.

Social Benefits	Economic Benefits	Industry Benefits	Technological Benefits
• **Increased Financial Inclusion:** Helps underserved individuals access credit products by leveraging non-traditional data sources. • **Community Trust:** Transparent credit scoring fosters trust between lenders and borrowers.	• **Cost Efficiency:** Reduces operational costs for lenders by automating credit assessment processes. • **Job Creation:** Expands lending opportunities, fostering business growth and employment.	• **Streamlined Operations:** Integration of AI models and alternative data improves efficiency in credit evaluation workflows. • **Data-Driven Insights:** Empowers financial institutions with actionable insights to improve lending strategies.	• **Integration of Advanced AI Models:** Combines sentiment analysis, behavioral data, and predictive analytics for robust credit assessments. • **Real-Time Data Processing:** Supports instant lending decisions with a secure and scalable platform.

TARGET INDUSTRY & FEASIBILITY

Weight of Evidence (WOE) and Information value (IV)

The formula for WOE is:

$$WOE = \ln\left(\frac{Distribution\ of\ Goods}{Distribution\ of\ Bads}\right)$$

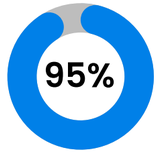
Information Value (IV)	Predictive Power
< 0.02	Useless
0.02 - 0.1	Weak predictors
0.1 - 0.3	Medium predictors
0.3 - 0.5	Strong predictors
> 0.5	Suspicious

Keywords: Credit risk analysis, logistic regression, XGBoost, machine learning.

Software Cost - Subscription Model

ITEM DESCRIPTION	PRICE
Software Development	₹ 17,000
Hosting	₹ 10000
Post Deployment Support	₹ 3000

₹ 30,000 billed annually

Progress Report : 

Completed

Conclusion:

The "**CreditSense**" project offers a revolutionary approach to **credit risk** assessment by integrating traditional financial data with alternative data sources, such as social media sentiment, utility payment records, and spending habits. By leveraging advanced machine learning models and blockchain technology, the system provides real-time, explainable credit risk scores that empower lenders to make more informed decisions while enabling financial inclusion for individuals with limited or **no credit history**.

Future Development :

The future of "CreditSense" includes enhancing AI models for personalized and explainable credit assessments, leveraging blockchain for decentralized credit histories and secure smart contracts, and integrating with financial ecosystems through **APIs** and partnerships. **A user-friendly mobile app with multi-language support** will improve accessibility, while compliance with global regulations and regular audits will ensure ethical use. Scaling the solution for global deployment through cloud-native architecture and regional adaptations will further strengthen its impact on financial inclusion and credit assessment.

RESEARCH AND REFERENCES

- Abellán, J., & Castellano, F.J. (2017). A comparative study on base classifiers in ensemble methods for credit scoring. Expert Syst. Appl., 73, 1-10.
- Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The Journal of Finance, 23(4), 589–609.
- Altman, E.I., and Sabato, G. (2007). Modelling Credit Risk for SMEs: Evidence from The US Market. ABACUS, 43 (3), 332-357.
- Breiman, Leo. 2000. Some Infinity Theory for Predictors Ensembles. Technical Report; Berkeley: UC Berkeley, vol. 577.
- Butaru, Florentin, Qingqing Chen, Brian Clark, Sanmay Das, Andrew W. Lo, and Akhtar Siddique. 2016. Risk and risk management in the credit card industry. Journal of Banking and Finance 72: 218–39.
- Chen, T.; Guestrin, C., (2016), "XGBoost: A Scalable Tree Boosting System". 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD ’16.
- Gahlaut, A., Tushar, & Singh, P.K. (2017). Prediction analysis of risky credit using Data mining classification models. 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 1-7.
- Galindo, Jorge, and Pablo Tamayo. 2000. Credit risk assessment using statistical and machine learning: Basic methodology and risk modeling applications. Computational Economics 15: 107–43.

Kindly check out our additional document that gives a detailed overview **Click here**

Click here to view the YouTube video explanation of Procura **Click here**