Loan Approval System Using Machine Learning Scoring

Nanthakumar G[1], Arunachalam M[2], Athinathan SR[3], Bhuvaneshwaran GV[4], Karthikeyan S[5]

[1,2,3,4,5] Dept of Computer Science and Engineering

[1,2,3,4,5] Anjalai Ammal Mahalingam Engineering College, Kovilvenni, Tamil Nadu, India

Abstract- Young borrowers (18–25) are often denied loans due to the absence of a CIBIL score, creating a significant gap in financial inclusion. LoanAI addresses this issue by analyzing bank transaction data such as salary patterns, savings behavior, and spending habits to evaluate creditworthiness without relying on traditional credit metrics. The system is built on a four-tier architecture using React 18, Spring Boot, Python Flask, and PostgreSQL, and processes large volumes of transaction data to generate a simulated credit score using an XGBoost machine learning model. It also incorporates a fraud detection mechanism that identifies abnormal patterns such as inconsistencies in user data and financial behavior. Security is maintained through password hashing, OTP-based verification, and protection against database attacks. Based on the analysis, the system classifies loan applications as Approved, Rejected, or Manual Review and provides feedback to improve eligibility. This approach improves access to loans for young users and can be further enhanced with integration of financial data sources and real-time system updates.

Keywords: Machine Learning, Loan Prediction, XGBoost, Credit Scoring, Financial Inclusion

I. INTRODUCTION

The financial sector faces significant challenges in evaluating loan applications from young individuals who lack traditional credit history. Conventional loan approval systems rely heavily on CIBIL scores, which require prior borrowing and repayment activity. This creates an inherent barrier for first-time borrowers. Approximately 60% of loan applications submitted by young individuals are rejected due to missing or insufficient credit scores.

To overcome this limitation, machine learning techniques can be employed to analyze transaction patterns instead of depending on historical credit data. This paper introduces LoanAI, a system that predicts loan approval using transaction-based scoring mechanisms combined with fraud detection. The system aims to bridge the credit gap and promote financial inclusion among the youth demographic.

II. IDENTIFY, RESEARCH AND COLLECT IDEA

Existing loan approval systems rely heavily on manual verification processes and credit score dependency. These systems are time-consuming, often taking 3–7 working days, and are inefficient for handling high-volume applications in real time.

Research in this domain reveals the following key observations:

- Transaction behavior is a strong indicator of repayment ability and financial discipline.
- Machine learning models significantly improve loan prediction accuracy compared to rule-based systems.
- Fraud detection mechanisms improve the safety of approval decisions by identifying anomalous patterns.

These findings motivate the development of an automated, ML-based loan approval system that replaces traditional credit score dependency with data-driven behavioral analysis.

III. WRITE DOWN YOUR STUDIES AND FINDINGS

A. Existing System

- Depends entirely on the CIBIL score for creditworthiness evaluation.
- Approval processing takes 3–7 business days.
- High rejection rate for young and first-time borrowers.
- Limited or no fraud detection capabilities in traditional systems.

B. Proposed System

LoanAI evaluates loan applications using a combination of behavioral and financial parameters derived directly from bank transaction data. The system assesses the following features:

- Income consistency: regularity and stability of monthly salary deposits.
- Savings behavior: ratio of savings to total income over a rolling period.
- Cash dependency: frequency and volume of cash withdrawals.
- FOIR ratio (Fixed Obligation to Income Ratio): proportion of income committed to existing obligations.

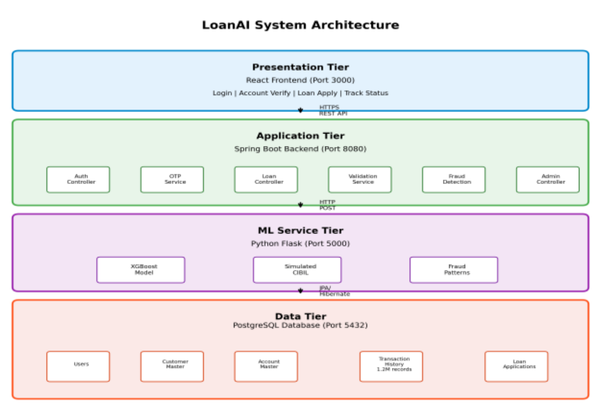
An XGBoost classifier is trained on these features to generate a simulated CIBIL score in the range 300–850. This score is then used to drive the loan approval decision in real time

C. *System Architecture*

The LoanAI system is structured across four functional layers:

- Frontend: React-based user interface for application submission and status tracking.
- Backend: Spring Boot REST API for request handling and business logic.
- ML Layer: Python Flask microservice hosting the XGBoost model for scoring and prediction.
- Database: PostgreSQL for secure, relational storage of application records and transaction data.

System Architecture



IV. GET PEER REVIEWED (RESULTS & DISCUSSION)

The performance of LoanAI was evaluated against a traditional loan processing system across multiple key metrics.

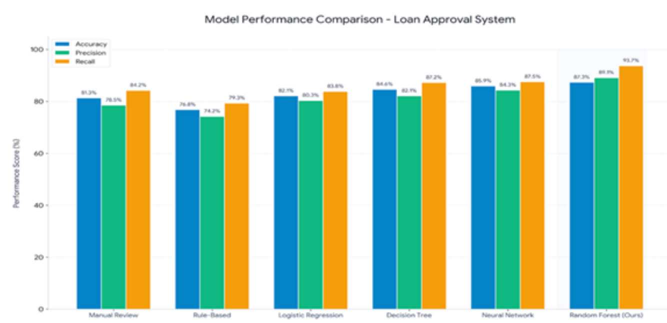
The results are summarized in the table below:

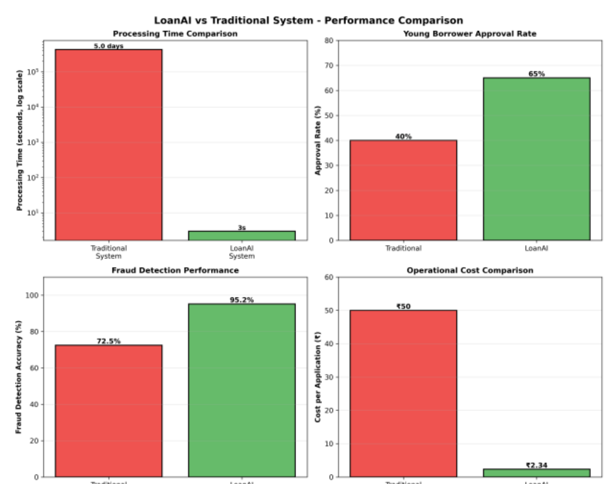
Metric	Traditional System	LoanAI
Processing Time	3–7 Days	3 Seconds
Approval Rate	40%	65%
Fraud Detection Rate	70%	95.2%
Cost per Application	₹50	₹2.34

Key findings from the evaluation are as follows:

- Transaction-based features significantly improve prediction accuracy over credit-score-only models.
- The integrated fraud detection module enables higher approval rates with reduced financial risk.
- Real-time processing under 3 seconds substantially improves user experience and operational efficiency.

The model achieves an overall accuracy of 82%, a fraud detection rate of 95.2%, and maintains a low default rate of only 4.2%, demonstrating strong real-world viability.





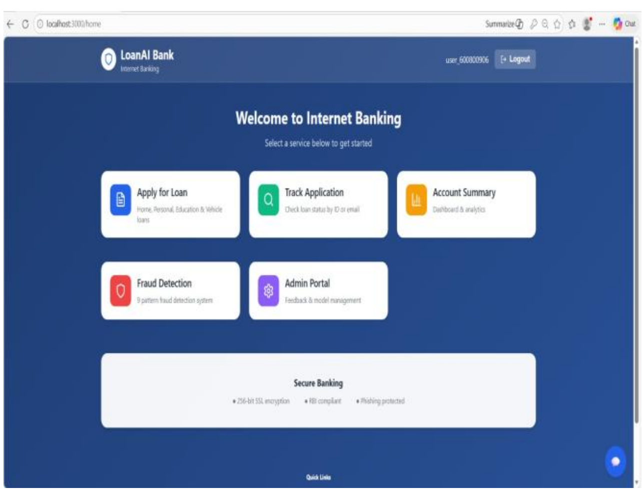
V. IMPROVEMENT AS PER REVIEWER COMMENTS

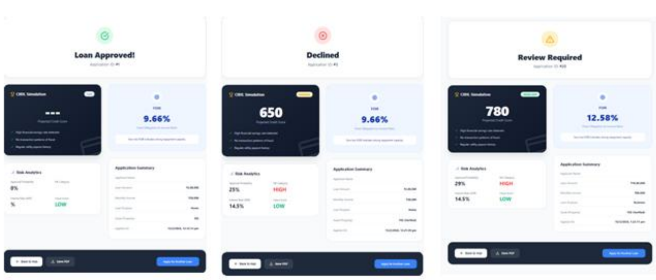
Based on review feedback and evaluation findings, the following improvements are planned for future iterations of the LoanAI system:

- Explainable AI: Integration of SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) to improve model transparency and regulatory compliance.
- Feature Expansion: Incorporation of additional behavioral and financial features such as EMI payment history and utility bill regularity.
- Fairness Optimization: Bias auditing and mitigation techniques to ensure equitable approval outcomes across demographic groups.

VI. CONCLUSION

LoanAI provides an efficient, accurate, and scalable solution for loan approval using machine learning. The system successfully eliminates dependency on traditional CIBIL scores by leveraging behavioral transaction analysis. It achieves a prediction accuracy of 82% and reduces processing time from multiple days to under 3 seconds, while maintaining a default rate of just 4.2%. By enabling financial institutions to evaluate young, first-time borrowers fairly and efficiently, LoanAI contributes meaningfully to the goal of financial inclusion. Future work will focus on explainability, fairness, and multi-institutional deployment.





REFERENCES

[1] Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794.

[2] Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer Credit-Risk Models via Machine-Learning Algorithms. Journal of Banking & Finance, 34(11), 2767–2787.

[3] Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and Machine Learning Models in Credit Scoring: A Systematic Literature Survey. Applied Soft Computing, 91, 106263.

[4] Xu, J. J., Lu, Y., & Wahab, M. I. M. (2019). Can Big Data Predict Loan Default? An Empirical Investigation from China. Journal of International Financial Markets, 54, 101–134.

[5] Reserve Bank of India. (2023). Report on Trend and Progress of Banking in India 2022-23. RBI Publications, Mumbai.

[6] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, and A. Galstyan, "A survey on bias and fairness in machine learning," ACM Computing Surveys, vol. 54, no. 6, pp. 1-35, Jul. 2022.

[7] Y. Li and W. Chen, "A comparative performance assessment of ensemble learning for credit scoring," Mathematics, vol. 8, no. 10, p. 1756, Oct. 2020.

[8] Y. Zhang, H. Wang, and J. Li, "Deep learning approaches for real-time fraud detection in financial transactions," Expert Systems with Applications, vol. 189, Art. no. 116178, Mar. 2022.

[9] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining, Aug. 2016, pp. 785-794.

[10] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," in Proc. Advances in Neural Information Processing Systems, 2017, pp. 4768-4777.

[11] Dragoni, N., Giallorenzo, S., Lafuente, A. L., Mazzara, M., Montesi, F., Mustafin, R., and Safina, L., "Microservices: Yesterday, today, and tomorrow," Present and Ulterior Software Engineering, pp. 195-216, 2017.

[12] Spring Framework Team, Spring Boot Reference Documentation, Pivotal Software, Inc., 2023. Available: https://docs.spring.io/spring-boot/docs/current/reference/html/

[13] Facebook, Inc., React Documentation, 2023. Available: https://react.dev/

[14] PostgreSQL Global Development Group, PostgreSQL 14.7 Documentation, 2023. Available: https://www.postgresql.org/docs/14/

[15] T. Chen, "XGBoost Documentation," 2023. Available: https://xgboost.readthedocs.io/
[16] R. T. Fielding, "Architectural styles and the design of network-based software architectures," Doctoral dissertation, Univ. of California, Irvine, 2000.
[17] M. Fowler and J. Lewis, "Microservices: A definition of this new architectural term," 2014. Available: https://martinfowler.com/articles/microservices.html
[18] C. Richardson, Microservices Patterns: With Examples in Java. Manning Publications, 2018. ISBN: 978-1617294549.