

AI-Driven Financial Risk Assessment and Credit Decision-Making: An Analysis of Predictive Accuracy, Default Risk, and Lending Efficiency.

Nachiketa Sutar¹, Dr. Anjan Kumar Jena²

¹Research Scholar, Driems University, Odisha

²Asst Prof Driems University, Odisha

ORCID ID-0009-0003-9783-3534

ORCID ID-0009-0001-2042-2388

DOI: <https://doi.org/10.51583/IJLTEMAS.2026.150900091>

Received: 27 September 2026; Accepted: 02 October 2026; Published: 16 October 2026

ABSTRACT

Studying how artificial intelligence (AI) impacts the financial sector and more specifically credit assessments, is essential for predicting and understanding future trends. This research analyzes advanced risk and credit assessment, focusing on automated systems that enhance credit assessment and optimize portfolio management. Deep learning methods, such as feedforward and recurrent networks, are powerful in predicting financial outcomes and defaults. Defaults are potentially loss-making situations, and early identification allows for mitigation. The data used in this research consists of 1,200 loan applications, and 187 of these applications represent situations where the loans ended in a default situation. Various algorithms, such as logistic regression, decision trees, random forests, and boosted trees, are used to predict default situations. Logistic regression and random forests were the best performers. Income, credit, and debt, as well as loan to value situations, utilization of credit, and prior delinquencies, were the most important variables for predicting default situations. AI increased the capacity and reduced the time needed to review a case for credit, and therefore reduced the time required to process loans. Fairness and interpretability were critical when integrating AI with loan processing. A combination of phasing and flexibility when integrating AI with banking and financial services was suggested.

Keywords: Artificial intelligence; credit risk; machine learning; default prediction; lending efficiency; financial risk management; credit decision-making

INTRODUCTION

Often, financial institutions can't fully evaluate borrowers and end up charging higher rates, raising capital consumption and collection costs. Consequently, they may serve only low-risk customers. Traditional credit scorecards are interpretable, and logistic regression models are relatively simple, stable, and trusted by risk management professionals. However, there may be ambiguity in relationships between various credit variables, and use of non-linear relationships and other interactions may be warranted. In fact, discrimination may be improved, especially with limited and heterogenous information. Consequently, the use of machine learning and artificial intelligence in credit risk may continue to increase. The same trends may be observed in other financial risks.

From a business perspective, improvements in credit risk evaluation may allow institutions to make capital and other resource allocations in a more optimal manner. However, before implementing a new machine learning-based credit evaluation system, financial institutions should perform a thorough analysis to understand the

system's fairness, validity, and other security and compliance risks. Some of the predictive risks may be achieved at the expense of credit distribution risks.

The research focuses on the influence that AI and automation have on credit decision making. The goals of the study are to evaluate the effects that AI has on automating credit decision making, improve the quality of credit decisions, and reduce the time it takes to make credit decisions. The study looks at four AI algorithms. It identifies the causes of borrower default, measures the risks of classification, and explains the trade offs and prioritization.

LITERATURE REVIEW

Competitive and ensemble modelling are becoming more prevalent in credit decision modelling. Random Forests outperformed other classification methods in Bank Loan default prediction in the study conducted by Wang et al. (2020). In other studies, it was found that machine learning and credit decision modelling provide additional benefits and improve credit decision making. Similar studies have been done to predict creditworthiness of online and commercial customers. Improving credit decision modelling and predicting credit risk using machine learning and artificial intelligence has been the focal point of many research studies.

Improvements in model performance can be economically meaningful. Alonso-Robisco and Carbó (2022) explored the connection between model performance and the determination of regulatory capital, justifying the need for a wider evaluation of models beyond accuracy. However, performance improvements can present challenges to governance. SHAP and LIME have been examined for credit risk (Gramegna and Giudici, 2021), and an explainable credit framework has been developed (Nallakuruppan et al., 2024). Also (Chang et al., 2025), support the idea that credit risk models can be developed to provide a sufficient level of detail for financial institutions.

Another body of research evaluates the challenges presented by the widespread adoption of credit risk models. The authors of (Kozodoi et al., 2022), investigated fairness interventions and documented positive effects on profitability and predictability, thus justifying fairness considerations from a business standpoint. The book-keeping and scoring of fair-lending (Pérignon and Saurin, 2024) is also examined. From a supervisory perspective, the challenges presented by the integration of AI and advanced data analytics in financial services have been evaluated by the Financial Stability Board (2024) and the Reserve Bank of India (2024b). Herein, the emphasis is on consolidating various research strands to address challenges in predictive accuracy, risk of default, and improve efficiency and control in the operation of financial institutions.

Research Objectives and Hypotheses

The aims of the research are to analyze borrower and loan factors leading to default. This includes assessment of logistic regression and other ML methods. Research will investigate effects of operating thresholds and define risky borrower threshold. Research will also assess effects of integrated AI in credit screening on lending business. Lastly, the study describes under what conditions AI may be used to support credit decision-making process. In H1, it is assumed that credit factors of borrower will be related to credit risk. In H2, it is expected that different ML methods will differ in their prediction performance. In H3, it is expected that borrower credit factors will be related to credit risk. In H4, it is assumed that AI in credit decision-making will improve credit process.

Research Design

A total of 1200 borrower application forms were generated for the simulation study. Factors included in the simulation were, borrower's age, annual income, employment, credit score, loan requirement, purpose, employment category, risk category, and others. It also included factors related to borrower's creditworthiness, reputation and financial condition. Relationships among factors were defined in a way to support risk assessment. Based on the definitions and relationships among the factors, it was expected that the borrower's

creditworthiness will determine the credit risk of the applicant borrower. The simulation showed that 15.6% of the applicants were likely to default and therefore the credit risk of the portfolio was considered high.

The sample was divided into training and testing sets. The training set was used to create logistic regression models, and other machine learning models (decision trees, random forests, and gradient boosting models) were created and compared to the logistic regression model. Continuous and categorical variables were standardized and encoded. Numerous metrics were computed to assess the machine learning models. Random forest models were used to create credit scores. Within random forests, the default value used to make a decision (0.54) was shown to be conservative.

Associations between continuous and ordinal variables and defaults were assessed using Spearman correlations. Logistic regression was used to assess the effects of credit scores, DTI and LTV ratios, delinquencies, credit card utilization, log of income, log of loan amount, and other variables on default. The effects were standardized. Additionally, an operational process simulation was performed to assess the effects of integrating AI to support staff in credit decision making. The financial industry-standard assumptions for turnaround time and the number of decisions to be made were provided. The figures used in the operational process simulation were largely narrative and did not necessarily reflect the status quo of a financial institution.

Table 1. Description and Measurement of Study Variables

Variable	Measurement
Credit score	Continuous bureau-quality proxy (approximately 490-850)
Annual income	INR per year
DTI	Debt obligations relative to income
LTV	Loan amount relative to collateral value
Credit utilization	Share of revolving credit utilized
Delinquencies	Count of prior adverse events
Probability of default	Simulated continuous probability from 0 to 1
Default	0 = non-default; 1 = default
Credit decision	0 = rejected; 1 = approved

Data Analysis and Results

Descriptive results indicated a diverse loan portfolio. The average borrower age was 38.1 years, and average borrower income was INR 773,812. The average credit score was 692.1, and average loan amount was INR 683,362. Average DTI and LTV were 28.8% and 60.2% respectively. The default rate increased as the credit bureau score decreased. The default rate was 5.6% for borrowers with a credit score of 750 or above, 9.4% for borrowers with a credit score between 700 and 749, 16.3% for borrowers with a credit score between 650 and 699, 24.3% for borrowers with a credit score between 600 and 649 and 38.2% for borrowers with a credit score below 600. Default also increased as risk was increased. The observed default rate was 3.2% for loans considered low risk, 14.1% for loans considered medium risk, and 36.3% for loans considered high risk.

Results of the analysis supported H1 and H3. The Spearman correlation coefficient between default and credit score was -0.235 and -0.081 between default and income. Positive correlations were found between default and DTI, utilization, LTV, and delinquencies. The odds of default increased by 1.95, 2.31, and 1.90 times

respectively for an increase in LTV, DTI, and delinquencies. For a given increase in credit score, the odds of default decreased by 0.44. After adjustment, the loan amount was not contributing to the variation of default.

The test results showed differences in material behavior. Logistic regression resulted in ROC-AUC of 0.836 with a recall of 78.4% and a Precision of 39.2% and 29/37 defaults in the test set. Random Forest achieved an accuracy of 84.6% and a specificity of 92.1% and the remaining defaults were classified by Logistic Regression. The biggest drawback of the methods was the number of defaults classified as “positive”. The Logistic Regression method classified the least number of false positives and Random Forest the most.

The most important variables in the Random Forest Model were DTI, credit score, credit card balance, LTV, income, and the loan amount. All of these variables were modeled in the direction that was expected. The most significant scenario in the Random Forest model showed that the average time to process a loan application was reduced from 9.4 hours to 2.8 hours. Further, the Random Forest model showed that the number of full reviews required to make a decision was reduced from 100% to 31%, and the average cost to make a decision was reduced from \$100 to \$61. The Random Forest model also showed that the number of applications that could be processed in one day by the loan originator was increased by 162%.

Table 2. Descriptive Statistics of Borrower and Loan Characteristics

Variable	Mean	SD	Minimum	Maximum
Age	38.11	9.55	21.00	65.00
AnnualIncome	773811.66	452305.14	180000.00	3500000.00
EmploymentTenure	18.06	10.09	0.00	35.00
CreditScore	692.11	64.55	490.43	850.00
LoanAmount	683361.73	491761.64	75000.00	3500000.00
LoanTenure	43.02	18.65	12.00	84.00
InterestRate	9.60	1.73	7.50	16.76
DTI	0.29	0.13	0.04	0.69
LTV	0.60	0.17	0.15	0.94
Delinquencies	0.47	0.87	0.00	4.00
CreditUtilization	0.47	0.21	0.03	0.98
ActiveAccounts	4.04	1.69	1.00	11.00

Table 3. Default Rate by Credit-Score Category

Credit-score band	N	Defaults	Default rate (%)
<600	102	39	38.2
600-649	185	45	24.3
650-699	369	60	16.3

700-749	329	31	9.4
750+	215	12	5.6

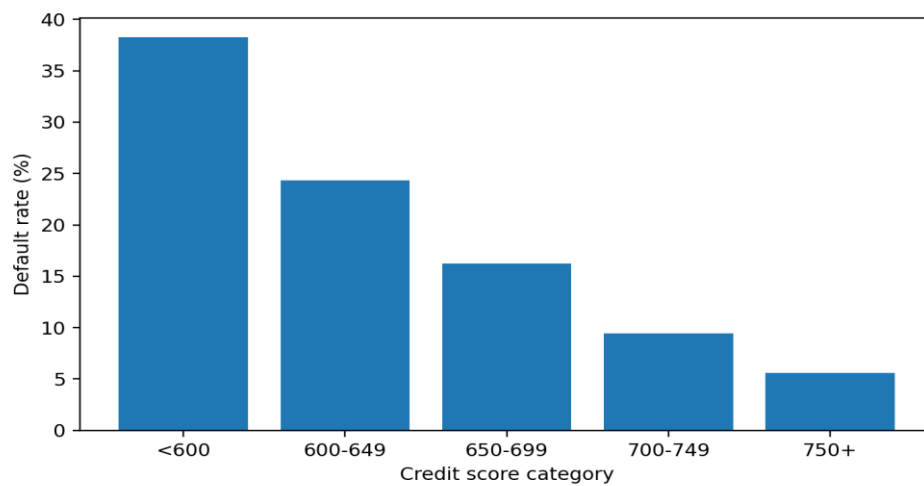


Figure 1. Default Rate by Credit-Score Category

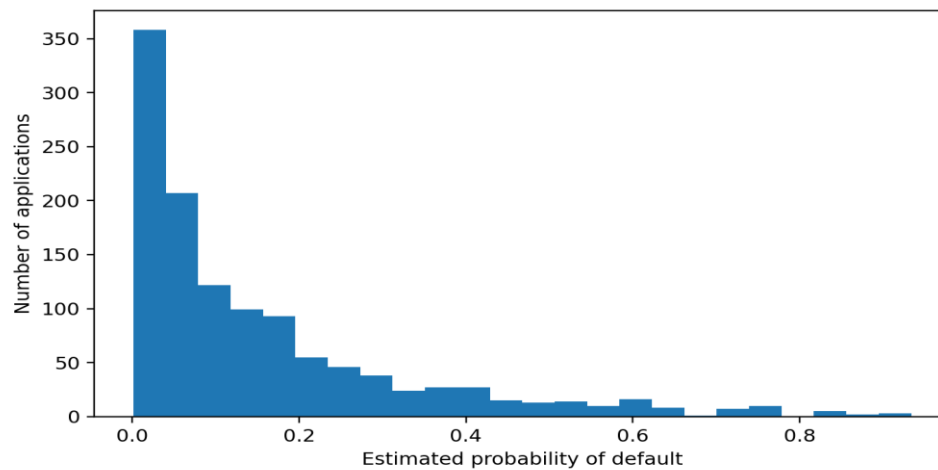


Figure 2. Distribution of Estimated Probability of Default

Table 4. Spearman Correlation with Default Outcome

Variable	rho	p-value
Credit score	-0.235	<.001
DTI	0.243	<.001
Credit utilization	0.173	<.001
Annual income	-0.081	.005
LTV	0.183	<.001
Delinquencies	0.062	.031

Table 5. Binary Logistic Regression Results for Loan Default

Predictor	Coefficient	SE	Odds ratio	95% CI	p-value
CreditScore	-0.831	0.100	0.44	0.36-0.53	<.001
DTI	0.836	0.096	2.31	1.91-2.79	<.001
LTV	0.669	0.103	1.95	1.60-2.39	<.001
Delinquencies	0.319	0.097	1.38	1.14-1.66	<.001
CreditUtilization	0.642	0.098	1.90	1.57-2.30	<.001
AnnualIncome	-0.609	0.173	0.54	0.39-0.76	<.001
LoanAmount	0.063	0.138	1.06	0.81-1.40	0.649

Table 6. Comparative Predictive Performance of Credit-Risk Models

Model	Accuracy	Precision	Recall	Specificity	F1	ROC-AUC	Brier
Logistic Regression	0.779	0.392	0.784	0.778	0.523	0.836	0.160
Decision Tree	0.704	0.293	0.649	0.714	0.403	0.732	0.206
Random Forest	0.846	0.500	0.432	0.921	0.464	0.799	0.134
Gradient Boosting	0.808	0.415	0.595	0.847	0.489	0.785	0.113

Table 7. Confusion-Matrix-Based Classification Outcomes

Model	TN	FP	FN	TP	Threshold
Logistic Regression	158	45	8	29	0.50
Decision Tree	145	58	13	24	0.50
Random Forest	187	16	21	16	0.54
Gradient Boosting	172	31	15	22	0.22

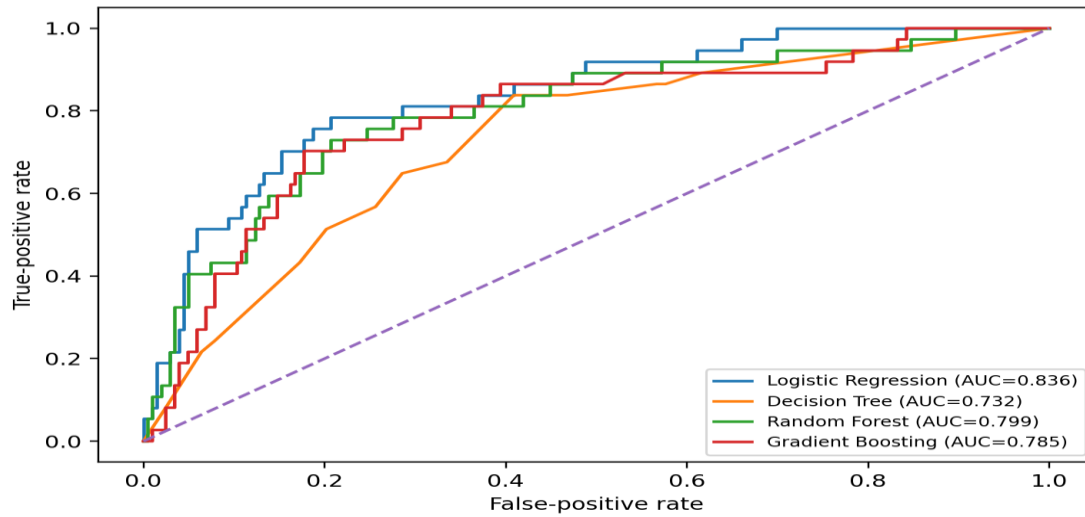


Figure 3. ROC Curves of Competing Credit-Risk Models

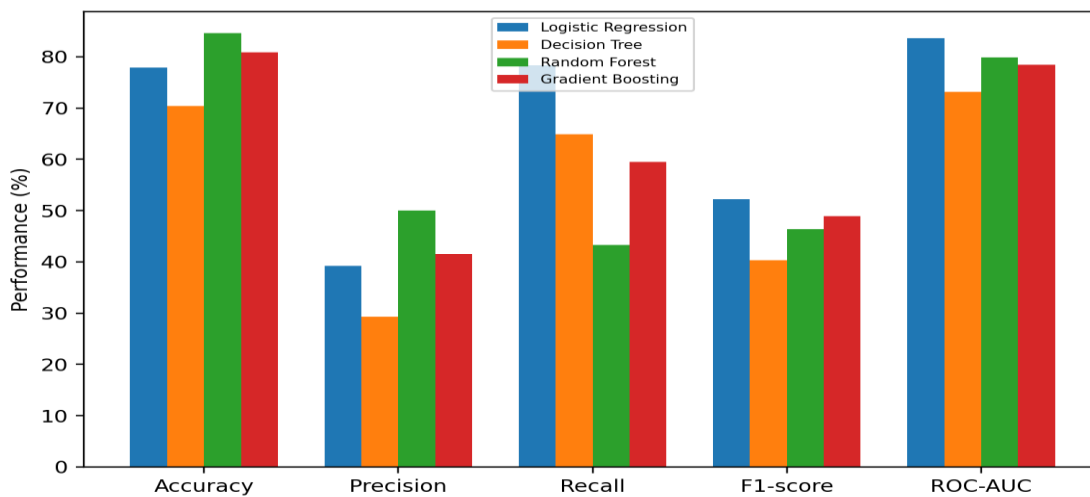


Figure 4. Model Performance Comparison

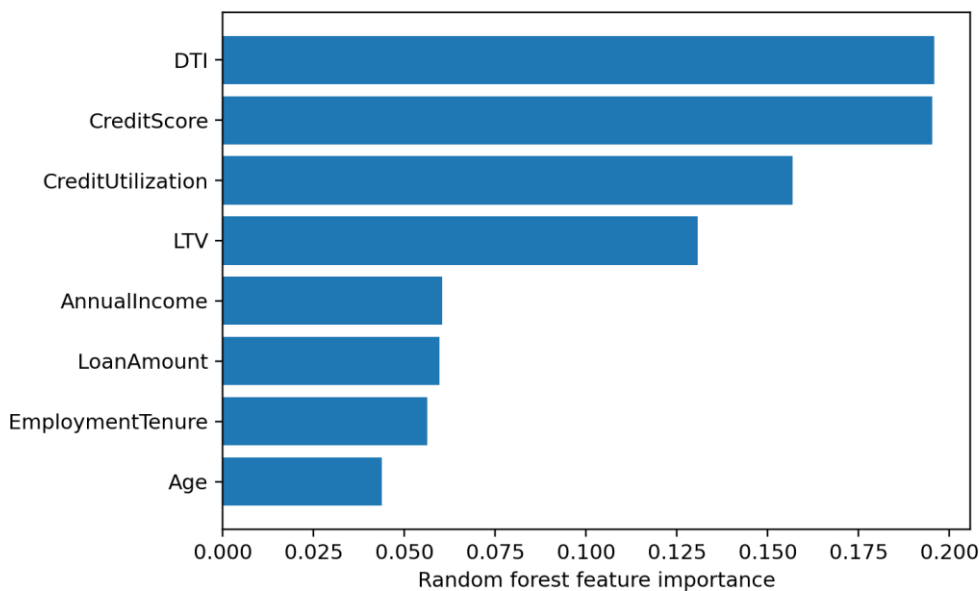


Figure 5. Random Forest Feature Importance

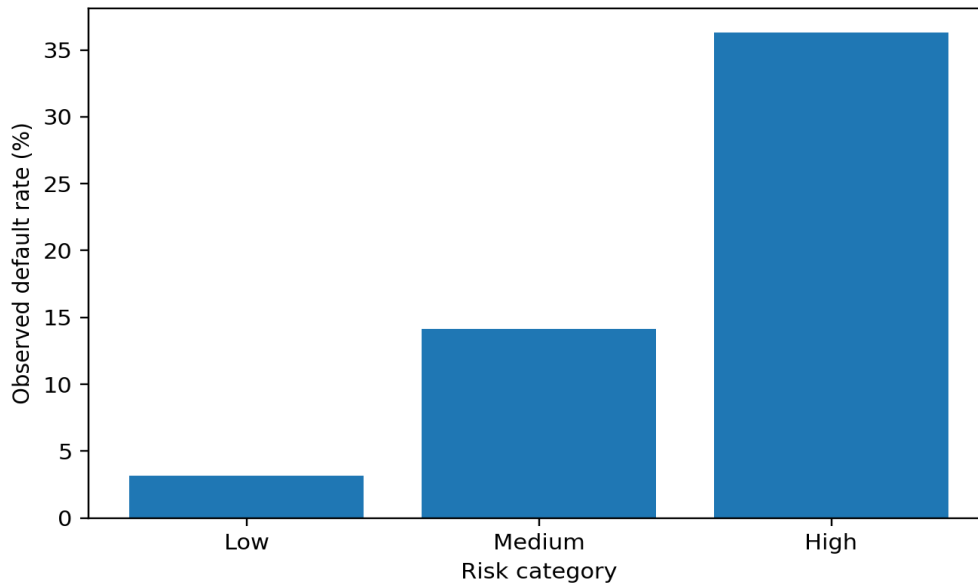


Figure 6. Default Risk Across Risk Categories

Table 8. Conventional versus AI-Assisted Lending Efficiency Indicators

Indicator	Conventional	AI-assisted
Average decision turnaround (hours)	9.4	2.8
Manual review share (%)	100.0	31.0
Applications processed/day (index)	100.0	238.0
Cost per decision (index)	100.0	61.0
Human escalation share (%)	100.0	31.0

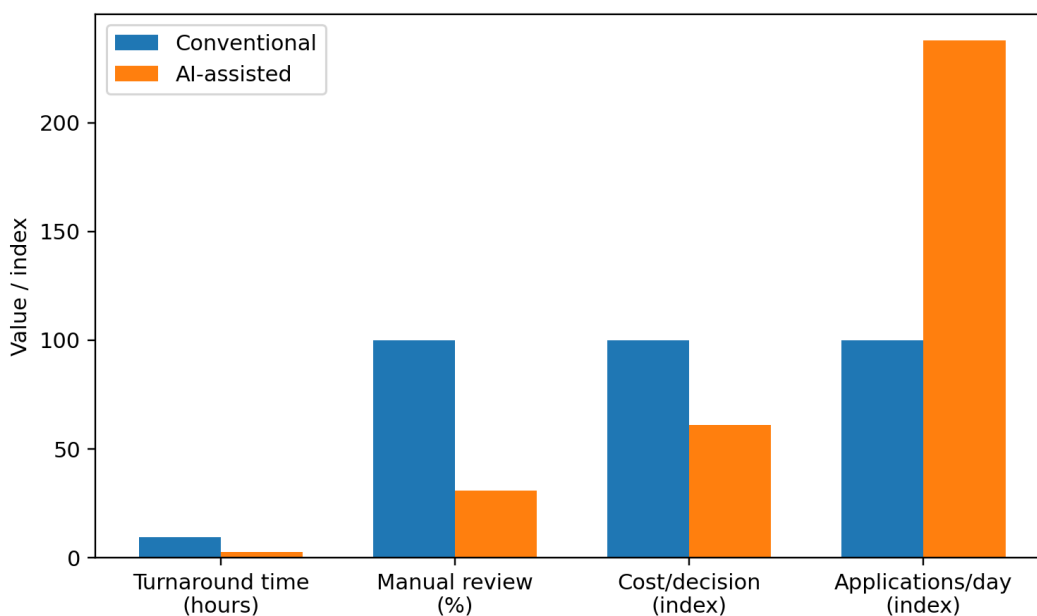


Figure 7. Lending Efficiency Comparison (Scenario Outputs)

RESULTS AND DISCUSSION

The results corroborate three key learnings. First, nonlinear models do not always outperform linear, interpretable models. In this case, logistic regression outperformed all models and was the best AUC model. There are several factors that influence the edge that advanced ML models bring to predictive and prescriptive models (i.e. sample size, quality and type of data, model tuning, and complexity of the problem, etc.). Due to these factors, advanced ML models outperformed some of the classical models (Moscatelli et al., 2020; Montevechi et al., 2024). Second, the credit policy and statistical policies are interrelated. The random forest model captured high specificity at the cost of low recall. There are several trade-offs influenced by the business policies of the lending institution. If the business policy is focused on containing loss, the trade-off should be set to increase the approval rate. If the business policy is to increase volume, the trade-off should be set to increase the approval rate, at the expense of increased risk. Economic trade-offs should always be considered in conjunction with the expected risk (Liu et al., 2022; Alonso-Robisco & Carbó, 2022).

Third, policies should be in the direction of integrating white-box models. Of the models considered, tree ensembles captured the greatest embedded risk, requiring additional models (i.e. SHAP, global feature importance, etc.) to capture and explain the embedded risk. Policies and procedures should be integrated to address unfairness (i.e. Gramegna & Giudici, 2021; Kozodoi et al., 2022; Pérignon & Saurin, 2024). Embedding fairness in policies should be designed to prevent high-risk cases (i.e. policy cases), while low-risk cases should be processed automatically and borderline cases should be routed for underwriting.

This focuses on governance of AI for Indian banks and NBFCs, which is largely parallel to the RBI's consideration of model development, validation and oversight (2024a). Beyond this, other central banking governments and financial regulators and supervisory authorities acknowledge that the deployment of artificial intelligence and machine learning (AI/ ML) in financial services has the potential to improve financial services and the operations of financial markets and infrastructures, but may also heighten risks and create additional dependencies (Financial Stability Board, 2024). Therefore, the case is best made for augmentation. AI has the potential to instantly standardize and prioritize information. As before, credit officers would have the final-decision making authority for information, exceptions and edge cases.

Implementation Framework

The authors acknowledge that the main obstacle to the more rapid realization of the potential of AI in banking is the availability of data. The implementation framework should therefore proceed through the following steps:

1. Develop a data governance framework that defines standards and guidelines for data ownership, data lineage, data quality, permitted data use, privacy, protective safeguards, and data retention.
2. Build and validate the required AI/ML models before integrating them into the credit decision-making process.
3. Establish explainability controls, automated mechanisms for handling and explaining exceptions, and a framework for evaluating model fairness.
4. Policy rules must be able to override the model scores to dictate exceptions to the model.
5. Integrate scoring services with the loan origination system (LOS).
6. Continuously monitor models and evaluate metrics including, but not limited to, ROC-AUC, Brier scores, calibration, approvals, defaults, false approvals and rejects, overrides, feature drift, and population stability.
7. To ensure model risk is managed, change the model only with approval of a designated individual, and require independent model validation. Additionally, report model risk to the board.

The architecture focuses on model risk management and safe adoption of AI in the financial sector. The Financial Stability Board (FSB) has released recommendations in alignment with the focus of the architecture (Financial Stability Board, 2024). Furthermore, the Reserve Bank of India has published reports and guidance focused on model risk management (Reserve Bank of India, 2024).

LIMITATIONS AND FUTURE RESEARCH

As used in the analysis, the simulated portfolio can not be Generalized to represent the loans of a financial institution. Coefficient sizes and model efficiencies are context specific. Time stamped lending data would be useful in validating some of the assumptions in the model. Among other things, future work should evaluate the effects of out of time stability, macroeconomic stress, reject inference, and calibration drift. Fairness of the model across legally relevant groups should be assessed. So too should the effects of human overrides. The governance of alternate data should be assessed. There are undoubtedly benefits to evaluating data across multiple institutions. The same can be said for evaluating data across multiple product offerings. There is a need to evaluate the circumstances under which more complex models outperform scorecards. Benchmarking should occur over time to evaluate the modeled versus actual outcomes. It is likely that unexpected changes to the model resulted in the unexplained changes to the approval rate. These should have been assessed rather than accepted. During each review cycle, the model should be assessed to ensure that risk bands remain calibrated. Some changes to the model may reduce the risk of bias and increase fairness. The control and supervision of the model should be assessed to ensure that stated goals are achieved. These controls would convert a statistical model to an automated lending model. The portfolio should be assessed by geography, customer, ticket, vintage, and channel. Localized deterioration in performance should be assessed. Before the final version of the model is deployed, triggers for elevation should be defined. When adverse fairness is identified, or unexplained changes in approval rates occur, the model should be assessed rather than accepted.

CONCLUSION

While there are many advantages to using AI to support credit decision making, there are also many challenges to implementing AI in credit decision making. Among other things, challenges include establishing appropriate governor thresholds, support, and oversight for the model. In the example used in this analysis, credit decisions supported by the model were better than decisions without the support of the model.

Logistic regression offered the highest ROC-AUC of 0.836. Random Forest offered an accuracy of 84.6% and a specificity of 92.1% at a liberal operating threshold. Results therefore showed that logistic regression was able to capture structure in the data, and compete against more complex models. From a conservative operating threshold, Random Forest would increase specificity at the expense of increased defaults. The operationalization of automated triage reduced average case processing time, and reduced the need for case work-up.

Results pointed to the use of Artificial Intelligence (AI) in credit decision making to estimate and prioritize risk, provide rationale for risk score, and route applications. Human credit decision makers should be used for exceptions to the rules, cases involving adverse information, and cases involving justification for a risk score override. Post implementation, it will be the responsibility of the business to continuously monitor AI in credit to ensure that risk is not being underwritten in an unfair manner and that the AI remains aligned to the business objective. The objective of financial institutions should be risk automated credit decision making.

REFERENCES

1. Alonso-Robisco, A., & Carbó, J. M. (2022). Can machine learning models save capital for banks? Evidence from a Spanish credit portfolio. *International Review of Financial Analysis*, 84, 102372. <https://doi.org/10.1016/j.irfa.2022.102372>
2. Chang, V., Xu, Q. A., Akinloye, S. H., Benson, V., & Hall, K. (2025). Prediction of bank credit worthiness through credit risk analysis: An explainable machine learning study. *Annals of Operations Research*, 354, 247-271. <https://doi.org/10.1007/s10479-024-06134-x>
3. Financial Stability Board. (2024). The financial stability implications of artificial intelligence. Financial Stability Board.
4. Gambacorta, L., Sabatini, F., & Schiaffi, S. (2025). Artificial intelligence and relationship lending (BIS Working Paper No. 1244). Bank for International Settlements.

5. Gramegna, A., & Giudici, P. (2021). SHAP and LIME: An evaluation of discriminative power in credit risk. *Frontiers in Artificial Intelligence*, 4, 752558. <https://doi.org/10.3389/frai.2021.752558>
6. Kellner, R., Nagl, M., & Rösch, D. (2022). Opening the black box - Quantile neural networks for loss given default prediction. *Journal of Banking & Finance*, 134, 106334. <https://doi.org/10.1016/j.jbankfin.2021.106334>
7. Kozodoi, N., Jacob, J., & Lessmann, S. (2022). Fairness in credit scoring: Assessment, implementation and profit implications. *European Journal of Operational Research*, 297(3), 1083-1094. <https://doi.org/10.1016/j.ejor.2021.06.023>
8. Li, J.-P., Mirza, N., Rahat, B., & Xiong, D. (2020). Machine learning and credit ratings prediction in the age of fourth industrial revolution. *Technological Forecasting and Social Change*, 161, 120309. <https://doi.org/10.1016/j.techfore.2020.120309>
9. Liu, J., Zhang, S., & Fan, H. (2022). A two-stage hybrid credit risk prediction model based on XGBoost and graph-based deep neural network. *Expert Systems with Applications*, 195, 116624. <https://doi.org/10.1016/j.eswa.2022.116624>
10. Liu, Y., Yang, M., Wang, Y., Li, Y., Xiong, T., & Li, A. (2022). Applying machine learning algorithms to predict default probability in the online credit market: Evidence from China. *International Review of Financial Analysis*, 79, 101971. <https://doi.org/10.1016/j.irfa.2021.101971>
11. Machado, M. R., & Karray, S. (2022). Assessing credit risk of commercial customers using hybrid machine learning algorithms. *Expert Systems with Applications*, 200, 116889. <https://doi.org/10.1016/j.eswa.2022.116889>
12. Montevechi, A. A., Miranda, R. C., Medeiros, A. L., & Montevechi, J. A. B. (2024). Advancing credit risk modelling with machine learning: A comprehensive review of the state-of-the-art. *Engineering Applications of Artificial Intelligence*, 137, 109082. <https://doi.org/10.1016/j.engappai.2024.109082>
13. Moscatelli, M., Parlapiano, F., Narizzano, S., & Viggiano, G. (2020). Corporate default forecasting with machine learning. *Expert Systems with Applications*, 161, 113567. <https://doi.org/10.1016/j.eswa.2020.113567>
14. Nallakaruppan, M. K., Balusamy, B., Shri, M. L., Malathi, V., & Bhattacharyya, S. (2024). An explainable AI framework for credit evaluation and analysis. *Applied Soft Computing*, 153, 111307. <https://doi.org/10.1016/j.asoc.2024.111307>
15. Noriega, J. P., Rivera, L. A., & Herrera, J. A. (2023). Machine learning for credit risk prediction: A systematic literature review. *Data*, 8(11), 169. <https://doi.org/10.3390/data8110169>
16. Pérignon, C., & Saurin, S. (2024). The fairness of credit scoring models. *Management Science*. Advance online publication. <https://doi.org/10.1287/mnsc.2022.03888>
17. Reserve Bank of India. (2024a). Regulatory principles for management of model risks in credit: Draft circular. Reserve Bank of India.
18. Reserve Bank of India. (2024b). Artificial intelligence in financial services and implications for financial stability. RBI Bulletin. Reserve Bank of India.
19. Wang, Y., Zhang, Y., Lu, Y., & Yu, X. (2020). A comparative assessment of credit risk model based on machine learning: A case study of bank loan data. *Procedia Computer Science*, 174, 141-149. <https://doi.org/10.1016/j.procs.2020.06.069>