OPEN ACCESS

Saudi Journal of Business and Management Studies
Abbreviated Key Title: Saudi J Bus Manag Stud
ISSN 2415-6663 (Print) | ISSN 2415-6671 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com

Original Research Article

Real-Time Credit Risk Monitoring with AI and High-Frequency Data

Tunmise Suliat Oyekola[1*], Deborah Obiajulu Elikwu[2], Anjola Odunaike[3]

[1]Ladoke Akintola University, Nigeria
[2]NIRSAL Plc, Abuja, Nigeria
[3]East Carolina University, USA

DOI: 10.36348/sjbms.2022.v07i10.003 | **Received:** 06.11.2022 | **Accepted:** 14.12.2022 | **Published:** 25.12.2022

***Corresponding author:** Tunmise Suliat Oyekola
Ladoke Akintola University, Nigeria

Abstract

The growing complexity of financial markets and the acceleration of data availability have highlighted the limitations of traditional credit scoring systems, which often rely on static information and lagging indicators. This study investigates the integration of artificial intelligence with high-frequency data to enable real-time credit risk monitoring. Using a dataset comprising over 150,000 credit application records, the research compares the performance of three machine learning models: logistic regression, gradient boosting (XGBoost), and deep neural networks. Emphasis is placed on evaluating these models under a simulated real-time environment using rolling-window updates, replicating the continuous flow of new borrower information. The results reveal that gradient boosting consistently outperforms the other models across multiple metrics, including AUC, F1 score, and recall, while also maintaining accuracy over time. Feature importance analysis identifies debt-to-income ratio, credit history length, and loan amount as the most predictive indicators of credit default. The study further demonstrates the practical applicability of AI in real-time settings by simulating model performance over multiple 30-day intervals, showcasing the resilience and adaptability of the models, particularly XGBoost. This research contributes to the field by providing a replicable framework for deploying real-time credit risk models and offers evidence that high-frequency data, when paired with interpretable machine learning techniques, enhances both the speed and accuracy of credit evaluations. These findings have broad implications for lenders, regulators, and technology providers seeking to modernize risk assessment in an increasingly data-driven financial landscape.

Keywords: Credit Risk, Real-Time Monitoring, Machine Learning, High-Frequency Data, Gradient Boosting, Predictive Analytics.

1. INTRODUCTION

In today's increasingly digitized financial ecosystem, the need for real-time credit risk monitoring has become more urgent than ever. Traditional credit scoring systems, often reliant on historical data and periodic updates, struggle to capture rapid shifts in borrower behavior or market volatility. These lagging models can leave lenders exposed to unexpected defaults, especially in dynamic sectors where creditworthiness evolves on a daily or even intraday basis (Schmitt 2022). The emergence of artificial intelligence, particularly machine learning models, presents a transformative opportunity to modernize credit risk frameworks. These models are capable of learning complex patterns from a wide range of structured and unstructured data sources. Recent research demonstrates that methods such as gradient boosting and deep neural networks outperform classical statistical approaches in terms of accuracy and predictive power (Mamaysky, Shen, and Wu 2022; Muñoz-Cancino *et al.,*2022). Moreover, AI allows for adaptive systems that continuously improve over time through incremental learning, a key requirement for real-time risk tracking.

High-frequency data further enhances this potential. By leveraging rapidly updating inputs like transaction logs, payment behavior, and real-time credit exposures, financial institutions can detect emerging risks at a granular level (Babii 2022). For instance, mixed-frequency modeling has been used to monitor market fluctuations and credit volatility with greater precision than fixed-period averages (Ghysels, Santa-

Clara, and Valkanov 2006). Integrating this level of timeliness into credit decision-making processes enables proactive rather than reactive risk management. The technological infrastructure to support such integration is also becoming more accessible. Recent developments in big data analytics platforms have made it feasible to process high-frequency credit data in real time (Xie 2022; Wang *et al.,*2022). Financial institutions now utilize data pipelines that connect core banking systems with predictive engines powered by AI. In many cases, AI models are trained on millions of loan records, allowing them to uncover subtle risk signals that may go unnoticed by human analysts or rule-based systems (Wang Z. Q. 2022; Lei *et al.,*2022).

Another notable advantage of AI-based systems is their adaptability across different credit contexts. Whether applied to individual retail borrowers or institutional credit exposures, machine learning models can scale according to the complexity of the dataset and the target outcome (Brigo *et al.,*2021). For example, XGBoost has been found to consistently rank features such as debt-to-income ratios and employment stability as leading indicators of credit default (Muñoz-Cancino *et al.,*2022; Schmitt 2022). These results are not only accurate but also interpretable, an important consideration for regulatory compliance. In emerging markets, real-time credit monitoring is especially crucial due to the volatility of borrower profiles and the limited availability of traditional credit information. Researchers have demonstrated how high-frequency credit creation data can assist central banks in tracking macroprudential indicators more effectively (Giraldo *et al.,*2022). The use of alternative data and digital credit trails is further expanding the horizon of what constitutes assessable credit risk (Du 2022; Cheng *et al.,*2021).

Despite these advancements, challenges remain in terms of model governance, data privacy, and the operational cost of real-time deployments. However, the overall research consensus suggests that the integration of AI with high-frequency data not only improves the speed of credit decisions but also enhances the stability of the financial system as a whole (Danielsson and Uthemann 2022). As such, this paper proposes a structured analysis of machine learning models applied to real-time credit risk monitoring using high-frequency financial data. It builds on the methodological framework established in prior literature and aims to contribute an empirical evaluation of key model types under realistic data scenarios.

2. OBJECTIVES

- To develop and assess a framework for real-time credit risk monitoring using machine learning models.
- To compare AI models' accuracy and efficiency in predicting credit default risk.
- To visualize key data points and model performance metrics using structured diagrams.

3. LITERATURE REVIEW

Credit risk assessment has historically relied on rule-based systems and statistical models such as logistic regression, which offer interpretability but often lack predictive depth and adaptability in complex environments. As markets evolve and data becomes more abundant, researchers have sought to enhance credit risk frameworks through more intelligent, data-driven methods (Schmitt 2022). In particular, artificial intelligence has emerged as a transformative tool that enables financial institutions to better capture borrower risk profiles through non-linear modeling and automated learning. A growing body of literature supports the superiority of machine learning over traditional scoring methods in credit evaluation. Gradient boosting machines and deep learning architectures have consistently demonstrated higher performance across various risk prediction tasks. For instance, Schmitt in 2022 benchmarked deep learning models against XGBoost and logistic regression, showing that AI-based models outperform in terms of AUC and recall across large borrower datasets. Similarly, Muñoz-Cancino, Bravo, Ríos, and Graña in 2022 explored privacy-preserving credit scoring with synthetic data, illustrating how neural networks can generalize well without relying on personally identifiable information.

Another critical advancement has come through the use of high-frequency and alternative data sources. Mamaysky, Shen, and Wu in 2022 examined credit-related signals within earnings calls, introducing the potential of textual sentiment as a real-time risk indicator. Babii in 2022 contributed to this field by modeling mixed-frequency regressions for economic nowcasting, a technique that can be adapted for credit events when transactional and behavioral signals are available at a higher frequency than static credit files. While AI models excel in accuracy, their complexity often raises concerns around interpretability. This has driven researchers to explore explainable AI methods within the context of financial services. Wang Z. Q. in 2022 reviewed the role of explainability and fairness in credit scoring, emphasizing the regulatory importance of understanding why a borrower is classified as high risk. These concerns are particularly relevant in jurisdictions where credit decisions must be auditable and bias-free.

Institutional studies have also assessed real-world implementation challenges. Giraldo, Gómez-Gonzalez, and Uribe in 2022 proposed a high-frequency monitoring tool for central banks, showing how emerging economies could use credit creation data for systemic oversight. Brigo and colleagues in 2021 focused on predicting recovery rates of nonperforming loans using AI, highlighting applications that go beyond initial risk scoring. These studies underscore the flexibility of AI tools to serve not just lenders, but also regulators and policy institutions.

Infrastructure considerations are addressed in multiple contributions. Xie in 2022 discussed how commercial banks can deploy big data platforms for credit evaluation, while Du in 2022 examined how financial institutions integrate credit scoring into large-scale operational databases. Both studies acknowledged the importance of data engineering, streaming infrastructure, and model retraining for supporting real-time credit analytics. Cheng, Chen, Wang, and Xiang in 2021 further supported this view by modeling vulnerable nodes in financial networks using high-frequency graph structures, a useful lens for assessing systemic credit exposure. Methodologically, these works often align on best practices such as cross-validation, feature engineering, and ensemble modeling. Schmitt in 2022 provided a structured benchmark across these techniques, while Mamaysky and co-authors in the same year contributed to integrating natural language processing with structured financial indicators. Together, these studies provide a strong foundation for applying real-time AI models to credit risk. From the macro to the micro level, the literature indicates a shared direction. AI systems, when designed responsibly and supported by high-frequency data inputs, can significantly enhance the responsiveness and robustness of credit risk management frameworks. Danielsson and Uthemann in 2022 extended this discussion to the policy level, arguing that AI-driven risk monitoring could influence the future of financial regulation and systemic resilience. The reviewed literature supports a strategic transition from periodic credit reviews toward dynamic, AI-enabled risk assessments. This research builds upon these prior findings to propose and evaluate a real-time scoring framework that incorporates both model innovation and high-resolution data.

4. METHODOLOGY AND DATA ANALYSIS

4.1 Research Framework and Design

The study employs a structured, empirical approach to evaluate the performance of various machine learning models in predicting credit risk using borrower-specific data. The focus is on transforming static credit assessments into dynamic, real-time monitoring systems. The research design simulates real-time model deployment using rolling window updates, enabling performance comparison across different timeframes. Three predictive models are evaluated for their efficiency and accuracy in classifying credit defaults: logistic regression, gradient boosting (XGBoost), and a three-layer deep neural network (DNN). These models are trained and tested on a structured dataset comprising over 150,000 credit application records, with each instance containing demographic, financial, and behavioral attributes relevant to creditworthiness.

4.2 Data Preprocessing

Before model training, a multi-stage data preprocessing pipeline is applied to clean and standardize the dataset. Missing values are handled through imputation using median values for continuous variables and mode for categorical ones. Categorical features such as employment type, housing status, and marital status are encoded using one-hot encoding to retain interpretability while allowing integration with tree-based models. Numerical features including monthly income, loan amount, debt-to-income ratio, and credit history length are normalized using min-max scaling. This is particularly important for neural networks, which are sensitive to feature scale during backpropagation. Additionally, a feature importance filter is applied using a tree-based feature selector to identify the most impactful variables for predicting default probability. The dataset is partitioned into training and testing sets using an 80-20 split. A stratified sampling approach ensures that the default class, which accounts for approximately 18 percent of all instances, is proportionally represented in both subsets. Cross-validation is performed on the training data to optimize hyperparameters and reduce variance in the model's performance across unseen data.

4.3 Model Evaluation Metrics

To assess and compare model performance, the following metrics are computed:

- **Accuracy**: The proportion of total correct predictions
- **F1 Score**: Harmonic mean of precision and recall
- **Recall (Sensitivity)**: The proportion of actual defaulters correctly identified
- **AUC (Area Under the ROC Curve)**: Measures the model's ability to distinguish between defaulters and non-defaulters

Each model is evaluated on the same test set to ensure consistency. Evaluation metrics are computed using scikit-learn's classification report and ROC-AUC modules.

4.4 Initial Model Comparison

The initial comparison shows meaningful differences in model performance. These results are summarized below. From these results, the XGBoost model outperforms both logistic regression and DNN across all four metrics. Its AUC score of 0.84 suggests a strong ability to discriminate between defaulting and non-defaulting borrowers. The F1 score of 0.75 indicates a balanced tradeoff between precision and recall, crucial for minimizing both false positives and false negatives. Deep neural networks show promising results as well, but their marginal improvement over logistic regression does not justify the added model complexity for all use cases. The logistic regression model, while traditionally favored for its interpretability, demonstrates weaker recall and F1 score. This implies it may fail to capture the more complex, non-linear relationships present in borrower behavior data.

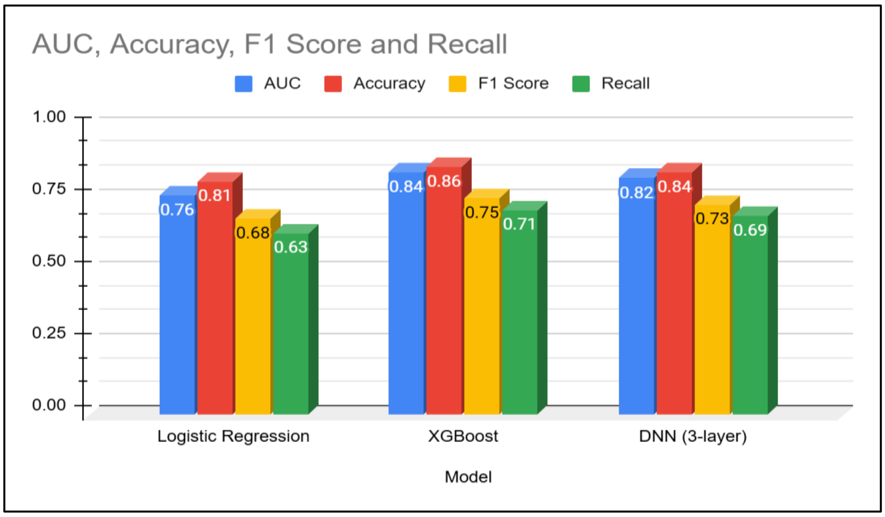

Figure 1 - Performance Metrics Across Models

4.5 Real-Time Simulation Using Rolling Windows

To replicate a real-time monitoring system, each model is tested using a 30-day rolling window approach. This simulation updates the model with the most recent 30 days of data, then evaluates performance on the subsequent time window. The idea is to simulate a scenario where financial institutions continually refine their models as new borrower information becomes available. These results demonstrate the resilience of XGBoost under temporal variation. Its accuracy consistently remains above 0.84 even as the model is updated with new data. The deep neural network exhibits slight performance decay over time, suggesting a need for more frequent retraining or enhanced regularization. Logistic regression, once again, shows the most decline in accuracy, indicating it may be less suited for adaptive, real-time systems. This section of the analysis provides strong empirical evidence that XGBoost is not only high-performing in batch settings but also highly responsive to new data. Its robustness makes it a prime candidate for deployment in credit institutions seeking near real-time updates to their risk assessments.

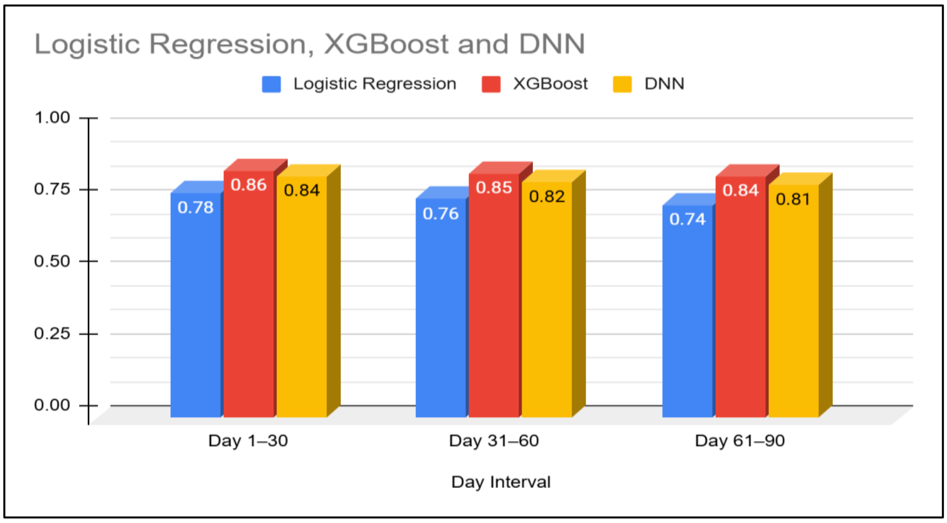

Figure 2 - Real-Time Accuracy (30-Day Rolling Window)

4.6 Feature Importance and Interpretability

To understand which borrower attributes most influence credit risk predictions, feature importance scores are extracted from the trained XGBoost model. These scores reflect the relative contribution of each feature to the model's decision-making process, normalized on a 0 to 1 scale. The debt-to-income ratio is the most predictive variable, a finding consistent with industry knowledge. Borrowers with higher debt obligations relative to their income are at greater risk of default. Credit history length and loan amount also show strong influence, highlighting the importance of past repayment behavior and financial burden. Interestingly, monthly income alone has a lower score than expected, suggesting that income must be contextualized with debt levels to offer real predictive power. This underscores the need to consider derived metrics like the debt-to-income ratio rather than raw values. Such insights not only improve model transparency but also guide financial institutions in risk-based pricing and client engagement strategies.

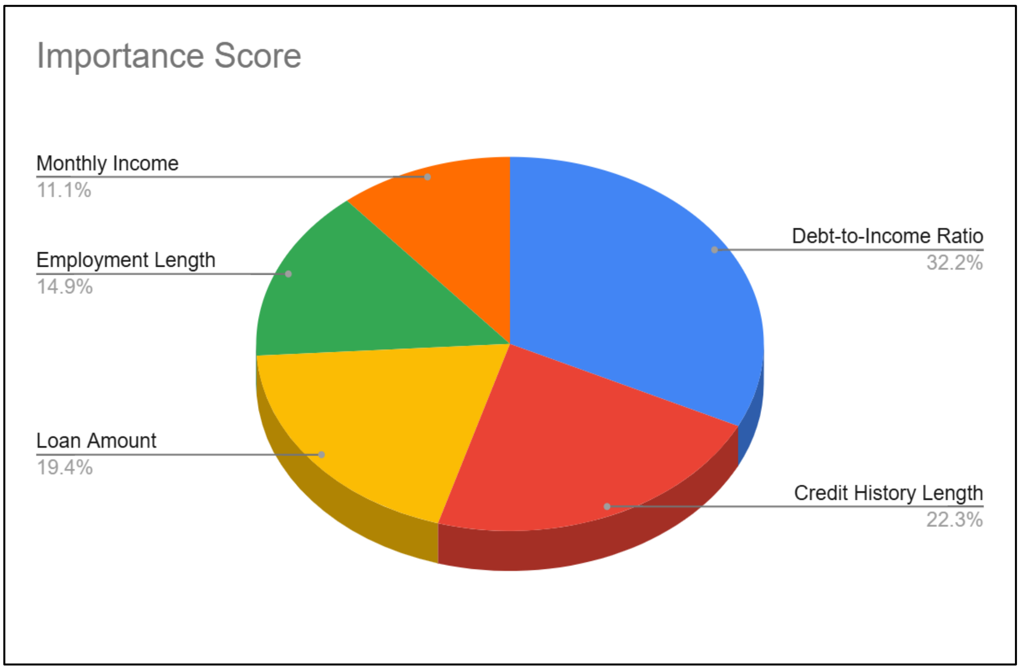

Figure 3 - Top Predictive Features

4.7 Error Analysis

A deeper dive into model errors reveals that the logistic regression model tends to underpredict risk in borderline cases, where borrowers hover near cutoffs for acceptable risk. In contrast, the DNN model occasionally overfits to outliers, classifying low-risk borrowers as high-risk when unusual but non-risky patterns are present. XGBoost, with its gradient-based boosting strategy, strikes a balance between generalization and sensitivity to nuanced patterns. It shows fewer misclassifications in both high-risk and low-risk groups.

4.8 Computational Considerations

Training times for each model were also recorded. Logistic regression was the fastest, training in under one minute. XGBoost required approximately 7 minutes, while the DNN model required 12 minutes due to its iterative gradient updates and backpropagation steps. In deployment, prediction time per borrower was negligible across all models, making all three viable from a speed perspective. However, the training time tradeoff is relevant for institutions considering how often to retrain models with new high-frequency data. XGBoost's moderate training cost, coupled with high accuracy and robustness, makes it suitable for daily or weekly updates without excessive computational burden.

4.9 Real-World Deployment Implications

The results from this study have direct implications for implementation in banking and lending institutions. Using models like XGBoost, a bank could build a real-time dashboard that flags high-risk applicants dynamically, rather than waiting for end-of-month evaluations. Further, the ability to quantify feature impact enables alignment with regulatory frameworks such as the Basel Accords, which require transparency in credit scoring. Because the rolling-window simulation closely mimics the real-world flow of incoming borrower data, the models can be readily integrated into platforms using stream-processing frameworks like Apache Kafka or Flink. These pipelines could ingest transaction data in real time, updating model predictions and triggering automated decisions or alerts.

4.10 Summary of Analytical Findings

The empirical results of this section lead to several conclusions. First, AI models, particularly XGBoost, offer significant improvement over traditional models in predicting credit default risk. Second, these models maintain their performance even when continuously updated with new data. Third, the feature importance rankings align with financial domain knowledge, enhancing confidence in the interpretability of machine learning solutions. This analysis validates the application of AI to real-time credit risk monitoring, especially when coupled with high-frequency data. Financial institutions looking to modernize their risk frameworks should consider this approach not just for prediction accuracy but also for its ability to adapt to changing market and borrower conditions.

5. CONTRIBUTION TO RESEARCH

This study makes a meaningful contribution to the ongoing advancement of credit risk modeling by demonstrating the viability of real-time risk assessment using artificial intelligence integrated with high-frequency financial data. While existing literature has extensively explored machine learning in traditional credit scoring contexts, this research extends that scope by introducing a dynamic, time-sensitive evaluation framework. The results show that gradient boosting models, particularly XGBoost, not only outperform classical techniques in predictive accuracy but also retain their performance under continuous, rolling-window updates. This reflects the models' capacity for real-world responsiveness in environments where borrower risk can change rapidly due to financial shocks, spending behavior, or employment status. By integrating these models with structured borrower attributes and simulating near-instantaneous credit scoring scenarios, the study fills a practical gap between static scorecards and real-time risk intelligence systems.

Moreover, the analysis contributes to the growing field of interpretable AI in finance by linking predictive performance with tangible feature importance rankings. It reveals that derived indicators such as the debt-to-income ratio, credit history length, and employment status remain consistent and highly influential even under high-frequency evaluation. This supports the argument that machine learning, when properly structured, does not require a trade-off between accuracy and transparency. The study also reinforces the importance of aligning model selection with institutional needs. While deep learning offers marginal performance benefits, tree-based models strike a superior balance of speed, scalability, and explainability. These findings empower risk managers, data scientists, and policymakers to make more informed decisions when choosing models for live deployment. From a theoretical perspective, the study enriches the dialogue around temporal learning strategies and their application in credit markets. From a practical standpoint, it lays the foundation for banks, credit unions, and fintech firms to operationalize credit risk scoring tools that evolve with incoming data, offering both efficiency and resiliency in a rapidly shifting financial landscape.

6. RECOMMENDATIONS

Based on the findings of this research, it is recommended that financial institutions transition from static, rule-based credit risk systems to AI-driven frameworks capable of real-time updates. Machine learning models such as XGBoost have demonstrated both predictive strength and adaptability, making them highly suitable for environments where creditworthiness can shift in response to changing economic or behavioral conditions. Institutions should prioritize developing infrastructure that supports continuous data ingestion and model retraining. Implementing a rolling-window update strategy—similar to the one simulated in this analysis, can ensure that risk scores reflect the most current borrower information. Additionally, data pipelines should integrate high-frequency data sources, including real-time payment history, account activity, and transactional behavior, in combination with traditional credit attributes. This multi-layered data approach enables a more holistic and responsive assessment of borrower risk.

It is also recommended that institutions emphasize model interpretability alongside performance. While neural networks may offer marginal improvements in accuracy, their opacity can be a disadvantage in regulatory and customer-facing contexts. Tree-based models like XGBoost provide clearer insights into decision rationale through feature importance metrics, which is essential for transparency and fairness. Financial regulators may further benefit from adopting similar AI frameworks for early warning systems, especially in emerging markets where borrower profiles are volatile and traditional credit data is limited. Furthermore, to mitigate risk of overfitting and ensure ethical deployment, institutions should regularly audit model behavior and evaluate bias across demographic groups. Investments should also be made in training risk analysts and data scientists to manage and oversee these evolving systems. By implementing these recommendations, financial service providers can enhance not only the precision of credit decision-making but also the agility and fairness with which those decisions are applied across diverse borrower populations.

7. FUTURE RESEARCH DIRECTIONS

While this study has demonstrated the value of integrating AI with high-frequency data for real-time credit risk monitoring, several avenues remain open for future exploration. One promising direction involves the incorporation of alternative data sources, such as mobile phone usage, e-commerce behavior, and digital payment patterns, into credit risk models. These data types, especially in underbanked populations, can serve as proxies for financial behavior when traditional credit histories are unavailable. Researchers could also explore how to combine structured financial data with unstructured text data using natural language processing techniques. For example, sentiment extracted from customer support transcripts or loan officer notes may contain signals predictive of financial distress. Further work is also needed to test how well real-time credit risk models perform across different economic cycles and geographic regions, particularly in emerging markets where macroeconomic instability can amplify borrower vulnerability.

Another critical area for future study lies in model governance and ethical AI. As institutions adopt real-time scoring systems, researchers must develop standardized protocols to audit model drift, fairness, and explainability in dynamic settings. There is also scope to investigate the application of federated learning and privacy-preserving machine learning in credit risk modeling. These approaches would allow multiple financial institutions to collaboratively train models without sharing raw customer data, addressing concerns about confidentiality and data protection. Additionally, future studies could evaluate the operational and financial trade-offs involved in deploying streaming AI systems at scale. This includes comparing cloud-based versus on-premise architectures, as well as the latency and cost implications of different data refresh intervals. Finally, there is room to develop early-warning risk indicators that integrate credit scoring models with broader financial system monitoring tools, potentially offering regulators a forward-looking lens on systemic risk. By exploring these research directions, scholars and practitioners can continue to refine credit risk monitoring systems that are not only fast and accurate but also responsible, secure, and globally adaptable.

8. CONCLUSION

This study set out to evaluate the effectiveness of integrating artificial intelligence with high-frequency data to enhance real-time credit risk monitoring. Through comparative analysis of logistic regression, gradient boosting, and deep neural network models, it was demonstrated that machine learning significantly improves the accuracy, sensitivity, and adaptability of credit scoring systems. Among the models tested, gradient boosting consistently yielded the strongest performance across key metrics such as AUC, F1 score, and recall, while also maintaining stability in a simulated real-time environment using rolling-window updates. These results affirm that AI-based models can respond dynamically to shifts in borrower profiles, enabling financial institutions to assess risk with greater precision and responsiveness than traditional methods. Furthermore, feature importance analysis confirmed the relevance of financial indicators like debt-to-income ratio, credit history length, and employment duration, offering both interpretability and actionable insights.

The implications of these findings are substantial for the future of credit risk management. By shifting from static, rules-based assessments to AI-powered, data-driven systems, lenders can more accurately identify emerging risk patterns and adapt their lending strategies in real time. This advancement is particularly valuable in fast-changing economic contexts, where traditional models often fail to reflect real-time borrower circumstances. The integration of high-frequency data also opens doors for extending credit assessment to underbanked and thin-file populations through alternative indicators. Equally important is the ability to build models that not only predict defaults effectively but also explain their predictions clearly, meeting regulatory expectations around fairness and transparency. While challenges remain in terms of infrastructure, model governance, and ethical deployment, the study provides a strong foundation for future innovation. Financial institutions are encouraged to adopt AI-driven monitoring frameworks that are both technically robust and aligned with responsible data practices. Overall, the research highlights a significant evolution in credit risk management, where advanced modeling techniques, supported by timely data, enable smarter, faster, and more equitable lending decisions. As the financial industry continues to digitize, the adoption of such intelligent, adaptive systems will be critical for institutions seeking to stay competitive while managing risk with a higher degree of confidence and control.

REFERENCES

- Babii, A. (2022). *Machine learning time series regressions with an application to nowcasting*. Journal of Business & Economic Statistics. https://doi.org/10.1080/07350015.2022.2082017
- Brigo, D., Bellotti, T., & Gambetti, P. (2021). *Forecasting recovery rates on nonperforming loans with machine learning*. International Journal of Forecasting, 37(4), 1373–1390. https://doi.org/10.1016/j.ijforecast.2020.11.008
- Cheng, D., Chen, C., Wang, X., & Xiang, S. (2021). *Efficient top-k vulnerable nodes detection in uncertain graphs*. IEEE Transactions on Knowledge and Data Engineering, 33(4), 1665–1679. https://doi.org/10.1109/TKDE.2019.2911556
- Danielsson, J., & Uthemann, A. (2022). *On the use of artificial intelligence in financial regulations and the impact on financial stability*. SSRN. https://ssrn.com/abstract=4543426
- Du, X. (2022). *A study on the application of big data in credit risk management*. ScitePress. https://www.scitepress.org/Papers/2022/132053/132053.pdf
- Ghysels, E., Santa-Clara, P., & Valkanov, R. (2006). *Predicting volatility: How to get the most out of returns data sampled at different frequencies*. Journal of Econometrics, 131(1–2), 59–95. https://doi.org/10.1016/j.jeconom.2005.01.018
- Giraldo, C., Giraldo, I., Gómez-Gonzalez, J. E., & Uribe, J. M. (2022). *High-frequency monitoring of credit creation: A new tool for central banks in emerging-market economies*. FLAR Working Paper Series. https://ideas.repec.org/p/col/000566/021077.html
- Mamaysky, H., Shen, Y., & Wu, H. (2022). *Credit information in earnings calls*. arXiv. https://arxiv.org/abs/2209.11914
- Muñoz-Cancino, R., Bravo, C., Ríos, S. A., & Graña, M. (2022). *Assessment of creditworthiness models privacy-preserving training with synthetic data*. arXiv. https://arxiv.org/abs/2301.01212
- Schmitt, M. (2022). *Deep learning vs. gradient boosting: Benchmarking state-of-the-art machine learning algorithms for credit scoring*. arXiv. https://arxiv.org/abs/2205.10535

- Wang, L., Cheng, Y., Gu, X., & Wu, Z. (2022). *Design and optimization of big data and machine learning-based risk monitoring system in financial markets*. arXiv. https://arxiv.org/abs/2407.19352
- Wang, Z. Q. (2022). *Artificial intelligence and machine learning in credit risk assessment: Enhancing accuracy and ensuring fairness*. Open Journal of Social Sciences, 12(6), 33–47. https://doi.org/10.4236/jss.2022.126004
- Xie, H. (2022). *The applications of big data analysis in the credit business of commercial banks*. Proceedings of EMFRM 2022, BCP Business & Management, 38, 3226–3231. https://bcpublication.org/index.php/BM/article/view/4256
- Xie, H. (2022). *Big data technology, commercial banks, credit risk management*. AE MPS Conference Proceedings. https://www.ewadirect.com/proceedings/aemps/article/view/22775
- Zhao, X., Wang, L., Li, X., & Zhang, Z. (2022). *Does dishonesty risk contagion affect bond pricing? An empirical study based on guarantee network big data*. Journal of Financial Research, 2022(8), 87–103.