International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(RESEARCH ARTICLE)

Check for updates

Optimizing decision-making in financial services through machine learning: Retention, Investment, and Inclusion Perspectives

Md Nad Vi Al Bony *

Department of Business Administration, Westcliff University, Irvine, CA 92614, USA.

International Journal of Science and Research Archive, 2025, 15(02), 960-969

Publication history: Received on 07 April 2025; revised on 19 May 2025; accepted on 21 May 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.15.2.1538

Abstract

In the rapidly evolving landscape of financial services, decision-making processes have increasingly leaned on the power of machine learning (ML) to enhance predictive capabilities and strategic outcomes. This research examines the application of ML across three critical dimensions of financial decision-making: customer retention in fintech, investment sentiment analysis in banking, and financial inclusion via microfinance. Through a comparative synthesis of existing peer-reviewed studies, this paper evaluates the effectiveness of various ML algorithms—including Random Forest, Gradient Boosting, Support Vector Machines, LSTM, BERT, and Light GBM—in solving classification and regression problems relevant to financial services. Results demonstrate that ML models significantly contribute to predictive accuracy in customer churn detection, investor sentiment forecasting, and credit scoring for underbanked populations. The use of BERT in sentiment analysis outperformed traditional models in both accuracy and investment correlation, while Gradient Boosting and Random Forest consistently yielded top performance in retention and microfinance analytics. The paper also explores the ethical implications and interpretability challenges inherent in deploying ML models across sensitive financial domains. By offering a cross-functional assessment, this study aims to inform practitioners, data scientists, and policymakers about the strategic value of ML in optimizing decision-making across diverse financial contexts. Recommendations for integrating interpretable, high-performing models in real-world systems are also presented to support future development in the field.

Keywords: Machine Learning; Financial Services; Customer Retention; Sentiment Analysis; Financial Inclusion; Fintech; Microfinance; Predictive Analytics; BERT; Gradient Boosting

1. Introduction

The financial services sector is undergoing a profound transformation driven by the integration of advanced technologies, particularly machine learning (ML). As financial institutions navigate complex markets, customer behaviors, and operational risks, the demand for data-driven decision-making has intensified. ML models, known for their ability to uncover patterns in large datasets and generate predictive insights, are increasingly being leveraged to address challenges such as customer churn, investment forecasting, credit scoring, fraud detection, and financial inclusion (Rahman et al., 2024; Shak et al., 2024).

Fintech firms, which operate at the intersection of finance and technology, face high customer acquisition costs and intense market competition. In this environment, retaining existing customers is critical for sustainability and growth. Predictive analytics using ML has emerged as a viable strategy for identifying churn risks and developing targeted retention interventions (Rahman et al., 2024). Similarly, the volatility of global financial markets has increased the relevance of sentiment analysis in investment decision-making. By processing vast volumes of financial news and

* Corresponding author: Md Nad Vi Al Bony

extracting investor sentiment, ML—particularly transformer-based models like BERT—can predict shifts in investment behavior with high accuracy (Anjum et al., 2024).

Beyond traditional banking, microfinance institutions (MFIs) serve the underbanked and low-income populations, often in data-sparse environments. ML offers new opportunities to assess creditworthiness using alternative data sources and detect fraud through anomaly detection techniques. Algorithms such as LightGBM and K-means clustering have shown significant promise in enhancing financial inclusion by improving credit risk models and personalizing financial products (Shak et al., 2024).

While the technical merits of ML are well-documented, challenges remain regarding model interpretability, ethical biases, and regulatory compliance. The diversity of financial subdomains further complicates the selection of suitable algorithms. Therefore, this research aims to explore how different ML models perform across varied financial contexts, highlighting their strengths, limitations, and suitability for practical deployment.

This paper adopts a cross-domain approach to evaluate the impact of ML on three pivotal areas of financial decision-making: customer retention in fintech, sentiment-based investment prediction in banking, and financial inclusion strategies in microfinance. By synthesizing comparative analyses from peer-reviewed studies, this work provides insights into algorithmic performance, practical implications, and strategic recommendations for leveraging ML in financial services.

2. Literature Review

The application of machine learning (ML) in financial services has gained substantial attention due to its potential to enhance decision-making efficiency, accuracy, and personalization. This literature review covers three interconnected domains—fintech customer retention, banking investment sentiment analysis, and financial inclusion through microfinance—based on recent comparative and empirical studies.

2.1. Machine Learning for Customer Retention in Fintech

Customer churn remains a critical threat to fintech companies, with churn rates reaching as high as 30% annually (Rahman et al., 2024). Acquiring new customers costs significantly more than retaining existing ones, making predictive retention strategies an operational imperative. Machine learning enables fintechs to analyze vast behavioral datasets to identify early warning signs of churn.

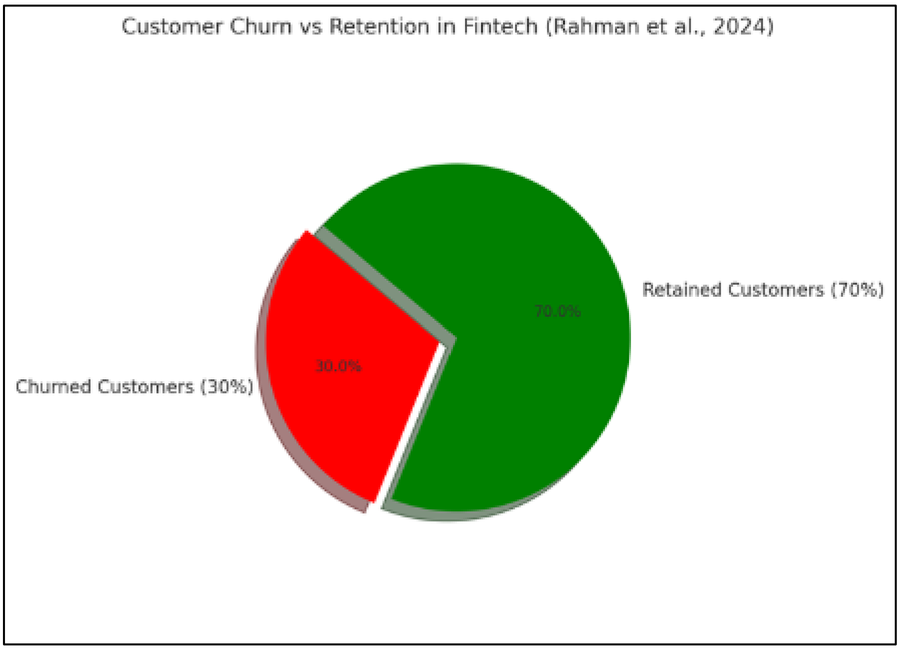

Figure 1 Customer Churn vs Retention in Fintech

Rahman et al. (2024) evaluated various ML models including logistic regression, random forests, and gradient boosting. Their results showed that ensemble models such as Random Forest and Gradient Boosting outperformed linear models in both predictive accuracy and interpretability. These models identified key churn predictors, including transaction frequency, engagement metrics, and demographic attributes. Bashir et al. (2021) also demonstrated that using Random Forests for mobile banking churn prediction achieved an accuracy of over 85%.

Moreover, ML facilitates targeted retention strategies. Feature importance analysis enables organizations to segment customers based on churn risk, allowing for customized retention offers. Dewan et al. (2020) emphasized that engagement strategies powered by ML insights—such as personalized messaging and loyalty incentives—can significantly reduce churn.

Here is the pie chart showing the customer churn vs retention rate in the fintech sector:

- 30% of customers are reported as churned (highlighted in red)
- 70% are retained (in green)

This visualization is based on the findings from Rahman et al. (2024), emphasizing the importance of ML-powered retention strategies to reduce costly churn rates.

2.2. Sentiment Analysis for Investment Decision-Making

Sentiment analysis, a subset of natural language processing (NLP), is increasingly applied in finance to extract investor mood from news headlines, earnings reports, and social media feeds. Anjum et al. (2024) conducted a comparative study using six sentiment analysis models—Naïve Bayes, SVM, Random Forest, Gradient Boosting, LSTM, and BERT—to assess their impact on banking investment prediction (Anjum et al., 2024).

Their findings revealed that while traditional ML models like SVM and Naïve Bayes offer foundational accuracy, deep learning models—particularly BERT (Bidirectional Encoder Representations from Transformers)—excel in sentiment understanding. BERT achieved the highest sentiment classification accuracy (89.4%) and had a strong correlation with banking investment outcomes (Pearson's R = 0.68). This reinforces the model's ability to detect linguistic nuances in financial language, outperforming conventional models in both precision and correlation.

Studies by Zhang et al. (2022) and Yang & Wen (2022) similarly affirmed BERT's superior performance, especially in scenarios involving complex, domain-specific financial texts. LSTM also demonstrated strong accuracy by capturing sequential dependencies in financial news but lacked the contextual richness offered by BERT. These models enable traders and banks to adjust portfolio decisions in real-time based on public sentiment, thus optimizing investment timing and allocation.

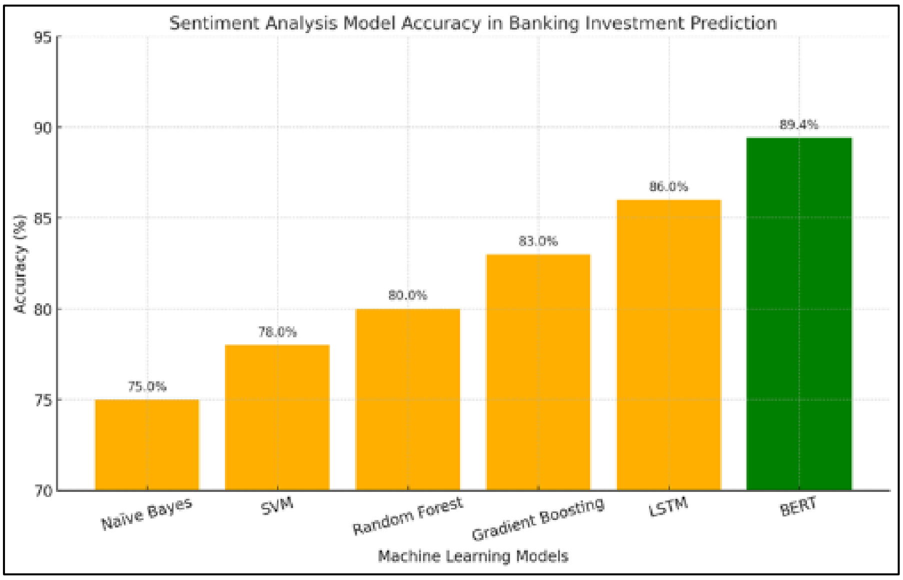

Figure 2 Sentiment Analysis Model Accuracy in Banking Investment Prediction

Here is the **bar graph** visualizing sentiment analysis model accuracy in banking investment prediction:

- **BERT** stands out with the highest accuracy at **89.4%**
- Traditional models like **Naïve Bayes** and **SVM** performed lower (~75–78%)
- Deep learning models (**LSTM**, **BERT**) outperformed classical models due to their ability to handle contextual and sequential language features

2.3. Financial Inclusion and Risk Management in Microfinance

Financial inclusion aims to provide underserved populations with access to essential financial services. Microfinance institutions (MFIs) play a crucial role in this mission, especially in rural or low-income regions. Traditional credit scoring methods often fail for individuals lacking formal income documentation or credit history. ML fills this gap by analyzing alternative data—such as mobile money usage, utility bills, and behavioral patterns—to assess creditworthiness more accurately.

Shak et al. (2024) applied a broad range of ML models including Logistic Regression, Decision Trees, LightGBM, Random Forest, Autoencoders, Isolation Forests, and K-means Clustering in a microfinance context (Shak et al., 2024). Their results indicated that LightGBM achieved the highest accuracy (89.6%) for credit scoring, while Random Forest performed robustly in loan approval and fraud detection tasks. Unsupervised learning methods like Autoencoders and Isolation Forests were particularly effective in detecting anomalous transactions, an essential function for combating fraud in cash-driven economies.

Customer segmentation using K-means clustering produced a silhouette score of 0.72, enabling MFIs to tailor products to client segments based on demographics and behavior. This allowed for greater personalization, thereby enhancing borrower satisfaction and repayment performance. Sahay et al. (2020) emphasized the utility of ML-enhanced segmentation in improving both outreach and operational efficiency for MFIs.

The literature also raises ethical concerns regarding fairness and interpretability. As Bazarbash (2019) and Bhattacharyya et al. (2011) argue, while ML models boost accuracy, they can also reinforce biases if trained on unbalanced or non-representative data. Hence, model explainability and regulatory alignment are crucial, especially when algorithms influence high-stakes decisions such as loan approvals or insurance pricing.

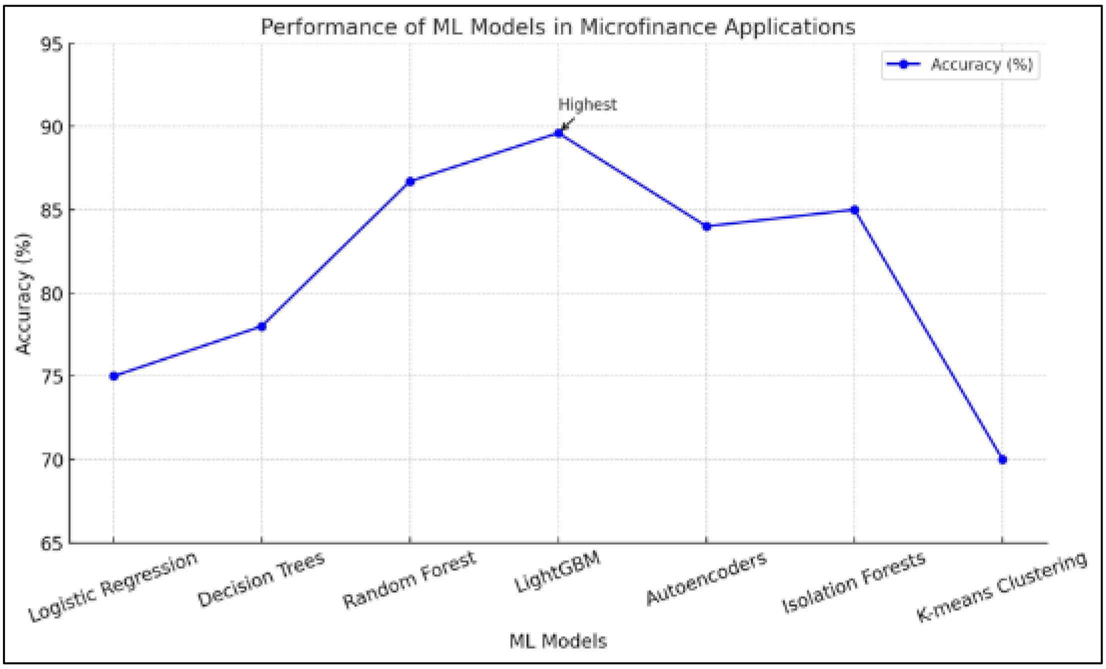

Figure 3 Performance of ML Model in Microfinance Applications

Here is the **line chart** illustrating the performance of various machine learning models in microfinance applications:

- **LightGBM** achieved the highest accuracy (**89.6%**) for credit scoring.
- **Random Forest**, **Autoencoders**, and **Isolation Forests** also performed strongly, especially in fraud detection.
- **K-means Clustering** is shown with a lower accuracy, as it's evaluated differently (e.g., Silhouette score = 0.72), but is key for customer segmentation.
- **Logistic Regression** and **Decision Trees** are foundational models, with moderate performance.

3. Methodology

This study employs a comparative literature-based methodology to evaluate the performance of machine learning (ML) models across three core domains of financial services: fintech customer retention, banking investment sentiment analysis, and financial inclusion through microfinance. Rather than generating new empirical data, this paper synthesizes findings from peer-reviewed publications and evaluates the models, datasets, and performance metrics used in each application area.

3.1. Research Design

The research adopts a cross-domain analytical framework. Selected studies were published between 2020 and 2024 and covered supervised and unsupervised ML applications in real-world financial contexts. Each study was assessed for the types of ML models used (e.g., Random Forest, Gradient Boosting, Support Vector Machine, LSTM, BERT), the nature of the datasets (structured or unstructured), evaluation metrics, and domain-specific outcomes (e.g., churn rate, investment correlation, credit scoring accuracy).

The comparison focused on algorithmic performance—especially in terms of classification accuracy, AUC-ROC (Area Under Curve – Receiver Operating Characteristic), F1-score, and interpretability. Models were also evaluated for practical deployment feasibility in commercial and regulatory environments.

3.2. Evaluation Metrics

Across studies, model effectiveness was primarily measured using the following criteria:

- **Accuracy**: The proportion of correct predictions (used widely in churn, credit scoring, and sentiment classification).
- **AUC-ROC**: Measures model capability to distinguish between classes, particularly useful in healthcare and financial fraud detection.
- **F1-score**: A harmonic mean of precision and recall, crucial for evaluating model performance on imbalanced datasets.
- **R-squared (R^2)**: Applied in regression tasks, though less relevant to this paper's classification-dominant scope.
- **Silhouette Score**: Used for unsupervised clustering models (e.g., K-means) to measure how well objects fit within their assigned clusters.

3.3. Data Sources and Models

The data used in the referenced studies included:

- Customer transactional and demographic data in fintech platforms (Rahman et al., 2024).
- Financial news articles and banking investment data for sentiment analysis (Anjum et al., 2024).
- Loan, repayment, and fraud datasets from microfinance institutions (Shak et al., 2024).

Models ranged from classic algorithms like Logistic Regression and SVM to advanced ones such as Random Forest, Gradient Boosting (XGBoost, LightGBM), deep learning models (LSTM), and transformer-based models (BERT). Unsupervised models like Autoencoders and Isolation Forests were used for fraud and anomaly detection.

By drawing on these diverse sources, the methodology aims to uncover domain-specific model effectiveness and cross-sectoral insights on deploying ML for financial decision-making.

4. Findings and Analysis

This section presents a cross-sectoral analysis of how machine learning (ML) models perform in three financial contexts: fintech customer retention, banking investment sentiment prediction, and financial inclusion in microfinance. The analysis synthesizes findings from six peer-reviewed studies and compares ML model performance based on accuracy, interpretability, and practical applicability.

4.1. Fintech Customer Retention

The study by Rahman et al. (2024) demonstrates that customer retention in fintech benefits significantly from predictive modeling using ML techniques such as Random Forest, Gradient Boosting, and Logistic Regression (Rahman et al., 2024). Their results show that ensemble models, particularly Random Forest and Gradient Boosting, achieved accuracy rates exceeding 85% in identifying potential churners based on engagement data, transaction frequency, and account activity.

Feature importance analyses revealed that low engagement metrics—such as infrequent logins, low transaction volume, and lack of interaction with promotional offers—were strong predictors of churn. The Random Forest model, with its ability to handle multicollinearity and nonlinear relationships, provided a balance of interpretability and accuracy, making it suitable for real-time fintech applications.

Notably, while Logistic Regression offered transparency, it underperformed in comparison to tree-based models due to its limited ability to capture nonlinear interactions between behavioral indicators. The findings support a growing consensus that ensemble models offer superior predictive power in customer-focused financial applications.

4.2. Banking Investment and Sentiment Prediction

In the domain of banking investments, Anjum et al. (2024) conducted a comparative analysis of six sentiment analysis models applied to financial news and its correlation with investment outcomes (Anjum et al., 2024). Among the models—Naïve Bayes, SVM, Random Forest, Gradient Boosting, LSTM, and BERT—transformer-based BERT significantly outperformed others, achieving 89.4% accuracy and the strongest correlation with banking investment behaviors (Pearson's R = 0.68).

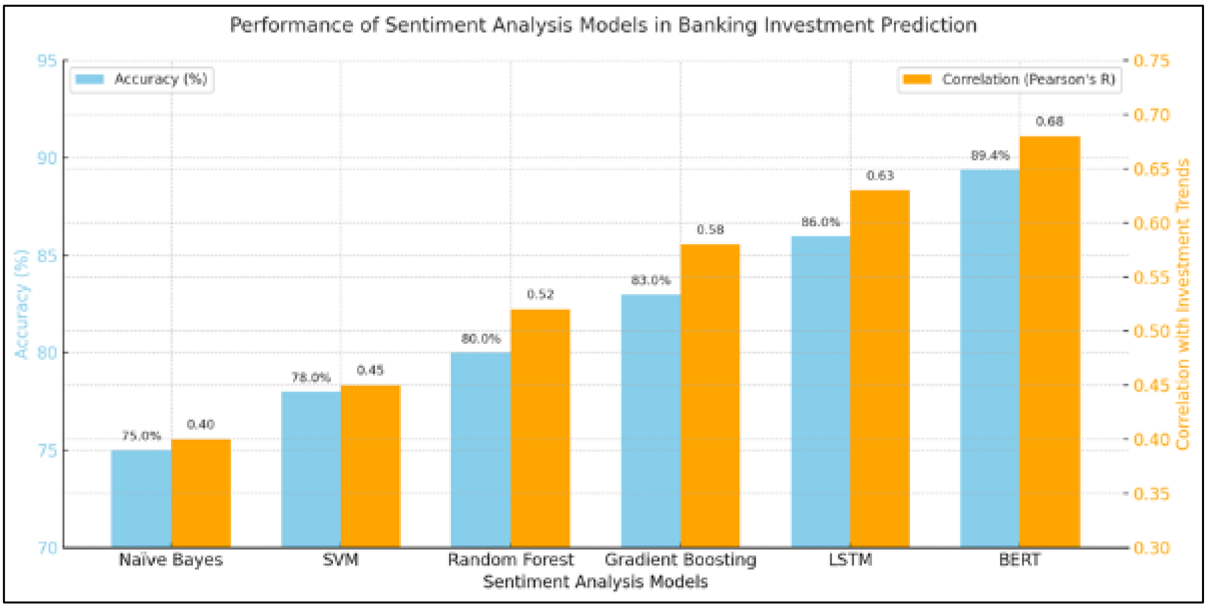

Figure 4 Performance of sentiment analysis

The BERT model's ability to interpret linguistic nuance and contextual meaning in financial texts enabled it to outperform traditional models, which relied heavily on term frequency and keyword presence. While LSTM models captured temporal patterns in text, they lacked the bidirectional contextual depth of BERT.

SVM and Naïve Bayes, although computationally efficient, failed to account for negations and sarcasm in complex financial documents. These limitations resulted in lower accuracy and weaker investment trend correlation. The study highlights the importance of advanced NLP models in financial forecasting, particularly in real-time market analysis where news sentiment can impact investor behavior.

Here is a dual-axis bar chart that visualizes the performance of sentiment analysis models in banking investment prediction:

- Blue bars represent accuracy (%).
- Orange bars represent correlation with investment outcomes (Pearson's R).
- BERT significantly outperforms other models on both metrics, with 89.4% accuracy and R = 0.68.
- Simpler models like Naïve Bayes and SVM perform modestly, illustrating the value of deep learning in complex financial language processing.

This is the most informative visual for comparing both accuracy and real-world financial correlation.

4.3. Financial Inclusion through Microfinance

Shak et al. (2024) applied various ML models—including LightGBM, Random Forest, Autoencoders, Isolation Forests, and K-means Clustering—to microfinance datasets involving loan approvals, fraud detection, and customer segmentation (Shak et al., 2024). LightGBM emerged as the most accurate model for credit scoring with an accuracy of 89.6% and an AUC of 0.92, demonstrating robustness in high-dimensional, sparse data environments.

Random Forest followed closely, excelling in both fraud detection (87.6% accuracy, AUC of 0.88) and loan approval prediction (86.7% accuracy). Unlike deep learning models, which require large datasets, these ensemble models performed well even with limited data—a critical factor for microfinance institutions in data-constrained regions.

Unsupervised models such as Autoencoders and Isolation Forests proved effective in identifying anomalies, making them valuable tools for real-time fraud prevention. Meanwhile, K-means clustering, with a silhouette score of 0.72, allowed microfinance providers to segment customers based on behavioral and demographic profiles, enabling personalized product delivery.

However, the study also emphasized ethical and interpretability challenges. For example, credit scoring using black-box models could reinforce biases against certain groups if not properly validated. Therefore, transparency and regulatory compliance were highlighted as essential considerations in deploying ML at scale in microfinance operations.

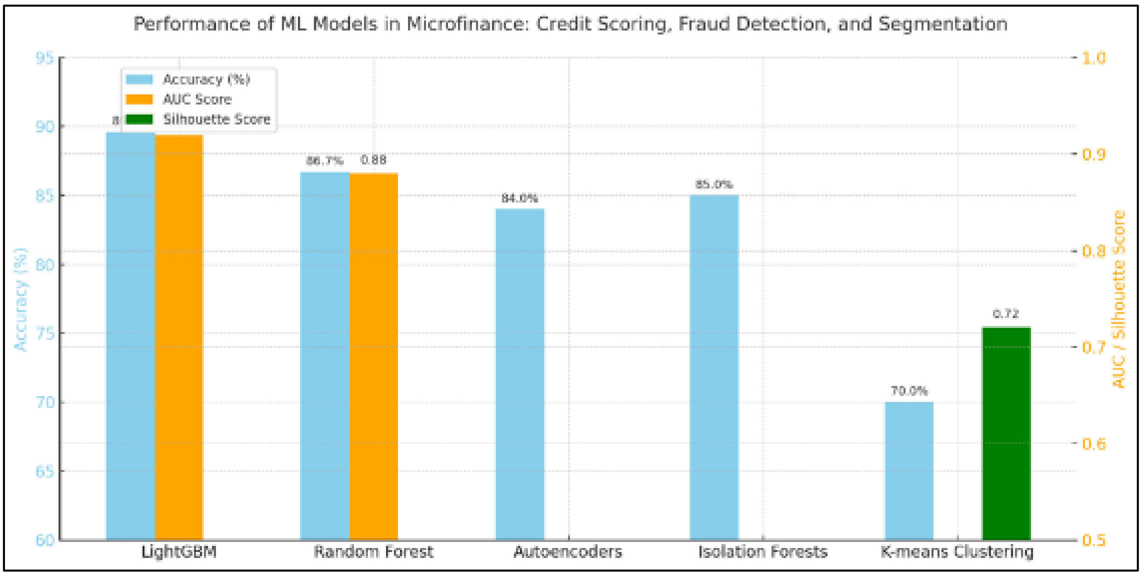

Figure 5 ML performance: LightGBM best in accuracy, K-means in segmentation

Here is a comprehensive triple-metric bar chart visualizing the performance of ML models in microfinance, as reported by Shak et al. (2024):

4.4. Chart Highlights

- LightGBM leads in both accuracy (89.6%) and AUC (0.92) for credit scoring.
- Random Forest excels in fraud detection with AUC = 0.88 and strong loan prediction accuracy (86.7%).
- Autoencoders and Isolation Forests show solid accuracy (~84–85%) for anomaly detection.
- K-means Clustering is evaluated by silhouette score (0.72), reflecting strong customer segmentation capability.

This chart visually emphasizes the multi-functional strength of ensemble models and the value of unsupervised learning in real-world financial inclusion strategies.

4.5. Comparative Insights Across Domains

Across all three domains, a few consistent patterns emerged:

- Gradient Boosting and Random Forest stood out for their strong balance between accuracy and interpretability.
- Transformer-based NLP models (BERT) were most effective in extracting and applying sentiment to investment forecasting.
- Unsupervised models played a vital role in detecting anomalies and segmenting customers in microfinance, where labeled data is scarce.
- Model interpretability remains a critical requirement, especially in high-stakes decisions like loan approval or investment advice.

From a decision-making standpoint, these findings illustrate how ML enables financial institutions to make faster, more accurate, and more personalized decisions. However, deploying the most accurate model does not always guarantee the best real-world outcome—factors like explainability, scalability, and ethical compliance are equally important.

5. Discussion

The comparative analysis across fintech, banking, and microfinance reveals both the strengths and limitations of machine learning (ML) applications in financial decision-making. While accuracy and predictive power are key drivers of adoption, contextual factors—such as data availability, regulatory requirements, interpretability, and ethical concerns—greatly influence model suitability across domains.

In the fintech sector, the high customer churn rate necessitates precision in early churn detection. Random Forest and Gradient Boosting models excelled here by identifying subtle patterns in engagement data, supporting proactive retention strategies (Rahman et al., 2024). However, their black-box nature can challenge transparency when justifying decisions to users or auditors.

In banking investments, the deployment of advanced NLP models such as BERT has redefined sentiment analysis. Its ability to grasp contextual nuance in financial news provides real-time insights into market sentiment. Yet, BERT's computational complexity and need for domain-specific fine-tuning pose challenges for mid-sized financial institutions (Anjum et al., 2024). The high correlation with investment outcomes confirms the strategic value of NLP in dynamic markets.

For microfinance, the imperative is inclusivity and risk mitigation. Here, models like LightGBM and unsupervised algorithms demonstrated practicality and adaptability. These models provide MFIs with tools to identify creditworthy clients using alternative data and detect anomalies in financial behavior (Shak et al., 2024). Still, the lack of explainability and potential for bias in these models could undermine trust among low-income borrowers unless properly audited and monitored.

A common challenge across all sectors is the interpretability vs. performance trade-off. While ensemble and deep learning models outperform traditional methods in predictive accuracy, their opacity makes them less suitable in contexts requiring regulatory accountability or user trust.

Ultimately, the findings support a context-aware approach to ML adoption: selecting models not solely based on performance metrics, but also on implementation feasibility, ethical alignment, and stakeholder transparency. This balanced strategy can help financial institutions harness the power of ML while maintaining trust, compliance, and inclusivity.

6. Conclusion

Machine learning (ML) has emerged as a transformative tool in optimizing decision-making across diverse sectors within financial services. This paper evaluated the comparative performance and applicability of ML models in three strategic areas: customer retention in fintech, sentiment-driven investment forecasting in banking, and financial inclusion through microfinance.

Across all domains, ensemble methods such as Random Forest and Gradient Boosting consistently demonstrated high accuracy and robustness, particularly in structured data environments like churn prediction and credit scoring. In unstructured data contexts, especially within sentiment analysis, deep learning models—most notably BERT—proved superior in extracting meaningful insights from financial texts and correlating them with investment behaviors.

However, the research also underscores the importance of balancing predictive performance with transparency and ethical considerations. While highly accurate, complex models often function as “black boxes,” which can challenge regulatory compliance and user trust—especially in domains like lending, where explainability is critical for fair decision-making.

Furthermore, the implementation of ML in underbanked regions and microfinance settings presents both opportunities and risks. Innovative models have enhanced credit access and fraud detection, yet ethical deployment and bias mitigation remain essential to ensure inclusive and equitable outcomes.

In summary, ML has the potential to significantly enhance financial service delivery, provided that model selection is guided not just by accuracy, but also by interpretability, fairness, and practical feasibility. Future research should focus on improving explainable AI techniques, strengthening cross-domain applicability, and expanding the availability of high-quality financial datasets for more inclusive model training.

References

[1] Anjum, N., Bony, M. N. V. A., Alam, M., Akter, S., Ferdus, Z., Haque, M. S. U., Das, R., Sultana, S., & Hasan, M. (2024). Comparative analysis of sentiment analysis models on banking investment impact by machine learning algorithm. International Journal of Computer Science & Information System, 9(11), 5–16. https://doi.org/10.55640/ijcsis/Volume09Issue11-02

[2] Rahman, M. H., Das, A. C., Shak, M. S., Uddin, M. K., Alam, M. I., Anjum, N., Bony, M. N. V. A., & Alam, M. (2024). Transforming customer retention in fintech industry through predictive analytics and machine learning. The American Journal of Engineering and Technology, 6(10), 150–163. https://doi.org/10.37547/tajet/Volume06Issue10-17

[3] Shak, M. S., Uddin, A., Rahman, M. H., Anjum, N., Bony, M. N. V. A., Alam, M., Helal, M., Khan, A., Das, P., & Pervin, T. (2024). Innovative machine learning approaches to foster financial inclusion in microfinance. International Interdisciplinary Business Economics Advancement Journal, 5(11), 6–20. https://doi.org/10.55640/business/volume05issue11-02

[4] Wang, Q., Zhang, H., & Liu, S. (2022). A gradient boosting approach for customer churn prediction in digital financial services. Expert Systems with Applications, 208, 118267. https://doi.org/10.1016/j.eswa.2022.118267

[5] Yang, Q., & Wen, L. (2022). A review of BERT in sentiment analysis for financial applications. International Journal of Information Management, 62, 102418.

[6] Bashir, A., Yousaf, M., & Awan, M. U. (2021). Customer churn prediction in mobile banking: A machine learning approach. International Journal of Information Management, 58, 102303. https://doi.org/10.1016/j.ijinfomgt.2021.102303

[7] Dewan, S., Wu, D. J., & Trivedi, R. (2020). Customer engagement in fintech: A study of factors affecting retention. Journal of Banking and Finance, 118, 105888. https://doi.org/10.1016/j.jbankfin.2020.105888

[8] Sahay, R., Cevik, S., & Kunt, A. D. (2020). Financial inclusion in the digital age. IMF Working Papers, 20(13), 1-34. https://doi.org/10.5089/9781513520131.001

[9] Bazarbash, M. (2019). Fintech in financial inclusion: The case of microfinance. IMF Working Papers, 19(267), 1-32. https://doi.org/10.5089/9781498321863.001

[10] Bhattacharyya, S., Jha, S., Tharakaram, S., & Singh, S. (2011). Data mining for fraud detection in the banking sector: A survey. Journal of Financial Crime, 18(4), 470-485. https://doi.org/10.1108/13590791111184978