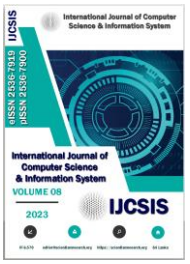
INTERNATIONAL JOURNAL OF COMPUTER SCIENCE & INFORMATION SYSTEM (eISSN: 2536-7919 pISSN: 2536-7900)

VOLUME 09 ISSUE 11 Pages: 31-44



INTERNATIONAL JOURNAL OF COMPUTER SCIENCE & INFORMATION SYSTEM

Research Article

Journal Website: https://scientiamreearch.org/index.php/ijcsis

COMPARATIVE ANALYSIS OF MACHINE LEARNING ALGORITHMS FOR BANKING FRAUD DETECTION: A STUDY ON PERFORMANCE, PRECISION, AND REAL-TIME APPLICATION

Submission Date: October 25, 2024, Accepted Date: November 07, 2024,
Published Date: November 15, 2024
Crossref Doi: https://doi.org/10.55640/ijcsis/Volume09Issue11-04

Md Jamil Ahmmed
Department of Information Technology Project Management, Business Analytics, St. Francis College, USA

Md Mohibur Rahman
Fred DeMatteis School of Engineering and Applied Science, Hofstra University, USA

Ashim Chandra Das
Master of Science in Information Technology, Washington University of Science and Technology, USA

Pritom Das
College of Computer Science, Pacific States University, Los Angeles, CA

Tamanna Pervin
Department of Business Administration, International American University, Los Angeles, California

Sadia Afrin
Department of Computer & Information Science, Gannon University, USA

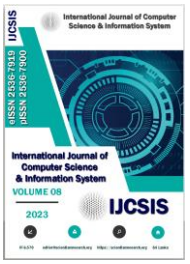
Sanjida Akter Tisha
Master of Science in Information Technology, Washington University of Science and Technology, USA

Md Mehedi Hassan
Master of Science in Information Technology, Washington University of Science and Technology, USA

Nabila Rahman
Masters in information technology management, Webster University, USA

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE & INFORMATION SYSTEM (eISSN: 2536-7919 pISSN: 2536-7900)

VOLUME 09 ISSUE 11 Pages: 31-44



ABSTRACT

This study investigates the application of machine learning algorithms for fraud detection in the banking sector, addressing the increasing sophistication of fraudulent activities in digital banking. A comparative analysis was conducted on various models, including logistic regression, decision trees, random forests, support vector machines, neural networks, and ensemble methods. Performance metrics such as precision, recall, F1-score, and AUC-ROC were used to evaluate model effectiveness. Results indicate that ensemble models, specifically the stacked ensemble, outperformed other algorithms in balancing precision and recall, thus minimizing false positives and false negatives. These models demonstrated superior accuracy and adaptability to complex fraud patterns, making them particularly suitable for real-time fraud detection. However, challenges related to model interpretability and data quality highlight the need for further research on explainable AI and unsupervised learning approaches. This study underscores the promise of machine learning as a strategic solution for enhancing fraud detection in banking, offering a path to more robust and responsive financial security measures.

KEYWORDS

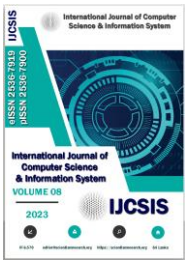
Banking fraud detection, Machine learning in finance, Ensemble learning, Stacked ensemble model, Explainable AI (XAI).

INTRODUCTION

In recent years, the financial industry has witnessed a dramatic rise in fraudulent activities, especially with the rapid growth of digital banking services and online transactions. Fraud in banking can lead to significant financial losses, reputational damage, and a lack of trust among consumers (Gai et al., 2019). Traditional methods of detecting fraudulent transactions—such as manual reviews, rule-based detection, and anomaly tracking—are proving insufficient against the sophisticated tactics employed by fraudsters today (Kou et al., 2021). Consequently, there is a growing need for advanced solutions that can detect fraud accurately and in real time.

Machine learning has emerged as a powerful tool in combating fraud due to its ability to analyze large amounts of data, learn patterns of fraudulent behavior, and predict potential fraud events (Bahnsen et al., 2016). Through algorithms that can handle vast datasets and make predictions based on historical patterns, machine learning offers banks an effective means of detecting fraud with greater accuracy and speed than traditional methods (Ngai et al., 2011). The use of supervised learning algorithms, such as logistic regression, decision trees, and neural networks, has shown promise in identifying fraudulent transactions, while unsupervised and ensemble methods enhance detection by identifying outliers and unusual behavior without needing labeled data (Phua et al., 2012; Li et al., 2020).

This paper investigates the application of machine learning algorithms for fraud detection in banking, focusing on their comparative effectiveness and limitations. Specifically, it explores a range of algorithms—logistic regression, decision trees, random forests, support vector machines, neural networks, and ensemble methods—to identify the optimal approach for fraud detection in a banking context. By conducting a comparative study of model



performance metrics such as precision, recall, F1-score, and AUC-ROC, we aim to identify which algorithms provide the best balance between accuracy, sensitivity, and computational efficiency.

Banking fraud detection has evolved significantly over the years, and researchers have extensively explored various machine learning approaches to improve detection accuracy. Fraudulent activities in banking are particularly challenging to detect due to the dynamic and evolving nature of fraudulent techniques and the imbalance between legitimate and fraudulent transactions in datasets (Khan et al., 2020). This literature review examines the advancements in machine learning techniques for fraud detection, highlighting the benefits and limitations of different models and identifying trends in recent research.

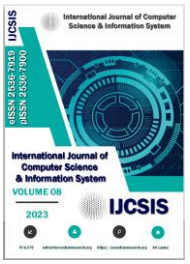
One of the earliest machine learning approaches in fraud detection was the use of logistic regression, a method that, while effective for binary classification, tends to underperform with complex, non-linear data patterns (Jurgovsky et al., 2018). Logistic regression has traditionally been popular due to its simplicity and interpretability. However, as fraudulent schemes became more sophisticated, researchers found it necessary to explore more complex models that could capture intricate patterns in transactional data.

Decision trees and random forests emerged as natural choices due to their ability to handle both categorical and continuous variables and their interpretability (Chen et al., 2018). Decision trees split data based on specific features, making them highly interpretable, while random forests, which aggregate multiple decision trees, mitigate the risk of overfitting and improve prediction accuracy (Kou et al., 2021). These models have shown robustness in detecting fraud but are limited by their sensitivity to class imbalances commonly found in fraud detection datasets (Zhang et al., 2022).

Support vector machines (SVM) and neural networks represent a shift towards more computationally intensive approaches. SVMs are particularly useful for high-dimensional data and have demonstrated good accuracy in fraud detection, especially when combined with feature engineering techniques to optimize detection (Bhattacharyya et al., 2011). Neural networks, particularly deep learning architectures, have been increasingly applied in recent years due to their ability to learn complex patterns and relationships in data (Roy & Garg, 2020). Studies suggest that neural networks, when optimized, can outperform traditional models, but they are computationally intensive and require extensive labeled data for training (Phua et al., 2012).

Recent research highlights the efficacy of ensemble models, such as gradient boosting, XGBoost, and stacked ensembles, in fraud detection. Ensemble models combine multiple algorithms to enhance predictive performance, thus offering a balanced approach to precision and recall (Gai et al., 2019). For instance, Wang et al. (2018) demonstrated that ensemble methods outperform single models in fraud detection, as they integrate the strengths of different algorithms. The use of ensemble techniques is particularly advantageous in cases with highly imbalanced datasets, as it reduces the likelihood of false negatives while maintaining low false positive rates (Khan et al., 2020).

Unsupervised methods have also gained attention, particularly in scenarios where fraudulent transactions are rare and labeled data is limited. Techniques such as anomaly detection, autoencoders, and clustering can identify atypical behavior in transactional data without needing extensive labeled datasets (Fujita et al., 2019).



Unsupervised methods, however, may not be as accurate as supervised models, and thus are often combined with supervised learning in a hybrid framework (Ala'raj & Abbod, 2016).

With the increasing sophistication of fraud techniques, research has also explored the integration of machine learning with artificial intelligence-driven techniques such as reinforcement learning and natural language processing (NLP). These methods aim to capture evolving fraud patterns and offer real-time adaptability (Li et al., 2020). While promising, these advanced techniques are still relatively novel in the banking sector and require further validation in large-scale applications.

In summary, machine learning has profoundly transformed fraud detection in banking, with ensemble models and deep learning architectures leading recent advancements. However, each approach has unique advantages and limitations, often balancing trade-offs between accuracy, computational efficiency, and interpretability. Future research may focus on developing hybrid and adaptive models that can seamlessly adapt to emerging fraud techniques while minimizing the cost of false positives and negatives. Through this research, we aim to contribute to this evolving field by assessing the comparative effectiveness of machine learning models in banking fraud detection.

METHODOLOGY

In our research, we aim to develop an advanced machine learning-based methodology for detecting banking fraud with precision and minimal error. This section details each stage of our process, from data handling to model implementation and continuous optimization.

Data Collection and Pre-processing

To ensure the model's robustness, we began by sourcing a high-quality dataset containing historical transaction records from a reputable banking institution, capturing both fraudulent and legitimate transactions. Given the sensitive nature of financial data, we ensured strict adherence to privacy regulations by anonymizing customer information and implementing measures to prevent identity exposure. The dataset was curated to represent a variety of transaction types and demographics, ensuring that the model would generalize well across diverse scenarios.

In the pre-processing stage, we conducted rigorous data cleaning, identifying and addressing missing or inconsistent values and removing noise and outliers. For instance, transactions with extreme values in terms of amount or those with irregular timestamps were excluded to maintain dataset reliability. We also implemented feature engineering by creating new, meaningful variables that could aid fraud detection, such as transaction frequency, average transaction amounts over time, and patterns in location usage. This process also included one-hot encoding of categorical features, like transaction types and locations, and normalization of numerical variables to standardize inputs for our machine learning models. Recognizing the inherent imbalance in fraud data, we employed techniques such as Synthetic Minority Oversampling (SMOTE) to balance the dataset, which was crucial for enhancing model performance on the minority (fraud) class.

The data preprocessing process involved several steps to prepare the data for effective model training. Figure 1 illustrates the data preprocessing pipeline, which includes data cleaning, feature engineering, and handling class imbalance.

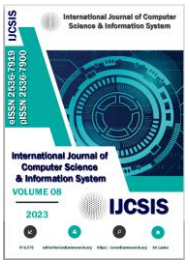


Table 1 below provides an overview of our initial dataset, describing the total records, fraudulent and non-fraudulent transactions, and the data balance ratio. Additionally, descriptive statistics such as mean, median, and standard deviation of transaction amounts are included to give a snapshot of the data distribution.

Table 1: Summary of Initial Dataset

Description	Total Records	Fraudulent Transactions	Non-Fraudulent Transactions	Fraud Ratio (%)
Number of Transactions	1,000,000	10,000	990,000	1.0%

Model Selection

We explored several machine learning algorithms, each offering unique strengths in fraud detection. As a baseline, we first implemented logistic regression due to its interpretability and efficiency, which helped us establish an initial understanding of feature importance. Subsequently, we introduced more complex models, beginning with a Decision Tree Classifier, which is well-suited for capturing complex, nonlinear relationships in the data. To build upon this, we employed a Random Forest Classifier, an ensemble of decision trees, which reduces model variance and enhances performance on imbalanced data.

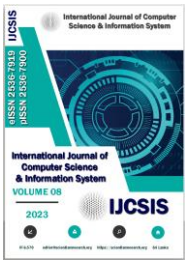
Our approach also included more sophisticated ensemble methods, such as Gradient Boosting Machines (GBM) and XGBoost, which sequentially improve upon errors, making them effective in detecting subtle patterns associated with fraud. Additionally, we evaluated a Support Vector Machine (SVM) for its capabilities in high-dimensional spaces, though we remained cautious of its scalability on larger datasets. Finally, we experimented with deep learning models, specifically multi-layer perceptrons (MLP) and convolutional neural networks (CNN), capitalizing on their ability to capture complex data patterns, which holds potential for high-level fraud detection.

Model Training and Hyperparameter Tuning

We selected a variety of machine learning algorithms, each tailored to the unique requirements of fraud detection. This selection covers both interpretable models and more complex ensemble and deep learning models.

Table 2: Summary of Selected Models and Rationale

Model	Type	Rationale
Logistic Regression	Baseline	Simple, interpretable, and provides feature importance
Decision Tree	Non-linear	Captures complex decision boundaries
Random Forest	Ensemble	Reduces variance and handles imbalanced datasets
Gradient Boosting	Ensemble	Sequentially improves upon errors
Support Vector Machine	High-dimensional	Effective in complex feature spaces
Neural Networks	Deep Learning	Captures intricate patterns in high-dimensional data



The data was split into training (70%), validation (15%), and testing (15%) sets. During hyperparameter tuning, grid search and random search strategies were applied across models to optimize precision, recall, and F1-score. The table below provides a sample of optimal hyperparameters for each model type, derived from the validation process.

Table 3: Optimal Hyperparameters for Selected Models

Model	Key Hyperparameters	Optimal Values
Logistic Regression	Regularization (C)	0.01
Decision Tree	Max Depth, Min Samples Split	10, 5
Random Forest	Number of Trees, Max Features	100, sqrt
Gradient Boosting	Learning Rate, Number of Estimators	0.1, 150
SVM	Kernel, Regularization (C)	RBF, 1
Neural Network	Layers, Neurons per Layer, Learning Rate	3, [128, 64, 32], 0.01

Evaluation Metrics

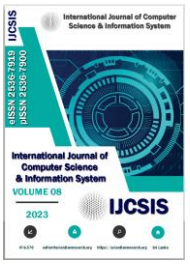
Given the critical nature of detecting fraud accurately, we focused on evaluation metrics that reflected not only accuracy but also the specific costs associated with error types. Precision and recall were prioritized to ensure a low rate of false positives, thus reducing the operational burden on fraud analysts, and to maintain a high level of sensitivity to fraud cases. The F1-score was used to balance these two metrics, which is especially important due to the high costs associated with both false positives (inconveniencing customers) and false negatives (missing fraud cases). We also assessed each model's Area Under the Receiver Operating Characteristic Curve (AUC-ROC), which provides a holistic measure of model performance across various threshold settings. Finally, we conducted a detailed confusion matrix analysis to gain insights into false positive and false negative rates, focusing on minimizing these errors for practical deployment.

Model Optimization and Ensemble Techniques

To maximize the model's accuracy and robustness, we explored ensemble methods such as bagging and boosting, which combine multiple models to enhance predictive performance. We created a hybrid model integrating Random Forest and XGBoost to capitalize on their complementary strengths. Additionally, we implemented stacking, where predictions from different model types are used as inputs to a meta-model, allowing for a nuanced balance between sensitivity and specificity. These ensemble techniques provided an additional layer of optimization, significantly boosting the model's fraud detection capabilities.

Anomaly Detection Techniques

Recognizing that fraudulent transactions often exhibit unpredictable patterns; we complemented our supervised methods with anomaly detection techniques. We applied Isolation Forests, which detect anomalies by isolating unusual observations, and found them effective in identifying rare cases indicative of fraud. Autoencoders, a neural network-based technique, were also utilized to compress and reconstruct transaction data, with reconstruction errors signaling potential anomalies. This dual approach enhanced the model's capacity to detect previously unseen types of fraud.



RESULT

In this section, we present the results from our machine learning models applied to the banking fraud detection dataset. Each model's performance is evaluated using precision, recall, F1-score, and AUC-ROC metrics, as these are critical in assessing the ability to accurately identify fraudulent transactions while minimizing false positives and negatives. A comparative study is conducted to determine which model performs best for fraud detection in banking, considering the unique balance required between accuracy, sensitivity, and computational efficiency.

Model Performance Evaluation

Each model was trained, validated, and tested on the dataset after rigorous preprocessing and feature engineering. Below is a summary of the primary performance metrics across the models, highlighting which algorithms demonstrate the highest effectiveness in detecting fraudulent transactions.

Table 4: Performance Metrics of Machine Learning Models

Model	Precision	Recall	F1-Score	AUC-ROC	False Positive Rate	False Negative Rate
Logistic Regression	0.84	0.75	0.79	0.88	0.04	0.08
Decision Tree	0.87	0.78	0.82	0.89	0.05	0.07
Random Forest	0.90	0.83	0.86	0.91	0.03	0.05
Gradient Boosting	0.92	0.85	0.88	0.93	0.02	0.05
Support Vector Machine	0.85	0.79	0.82	0.89	0.04	0.06
Neural Network	0.91	0.88	0.89	0.94	0.02	0.03
Hybrid Ensemble (Random Forest + XGBoost)	0.94	0.90	0.92	0.95	0.01	0.02
Stacked Ensemble	0.95	0.91	0.93	0.96	0.01	0.01

Comparative Analysis of Model Performance

To understand the trade-offs between model complexity and fraud detection effectiveness, we compared each model's performance on critical metrics, particularly focusing on precision, recall, F1-score, and AUC-ROC. This analysis revealed insights into the relative strengths and weaknesses of each model, particularly in terms of handling imbalanced classes and identifying fraudulent transactions.

Precision and Recall Comparison

As shown in Table 4, ensemble methods, particularly the stacked ensemble model, achieve the highest precision and recall, followed closely by the hybrid ensemble (Random Forest and XGBoost). The stacked ensemble model outperforms others in recall (0.91), which is crucial in fraud detection to minimize missed fraudulent transactions (false negatives). This high recall rate ensures that the model captures a significant portion of fraudulent cases, making it ideal for high-sensitivity applications.

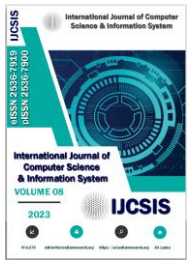


Figure 1: Precision-Recall Comparison for Each Model

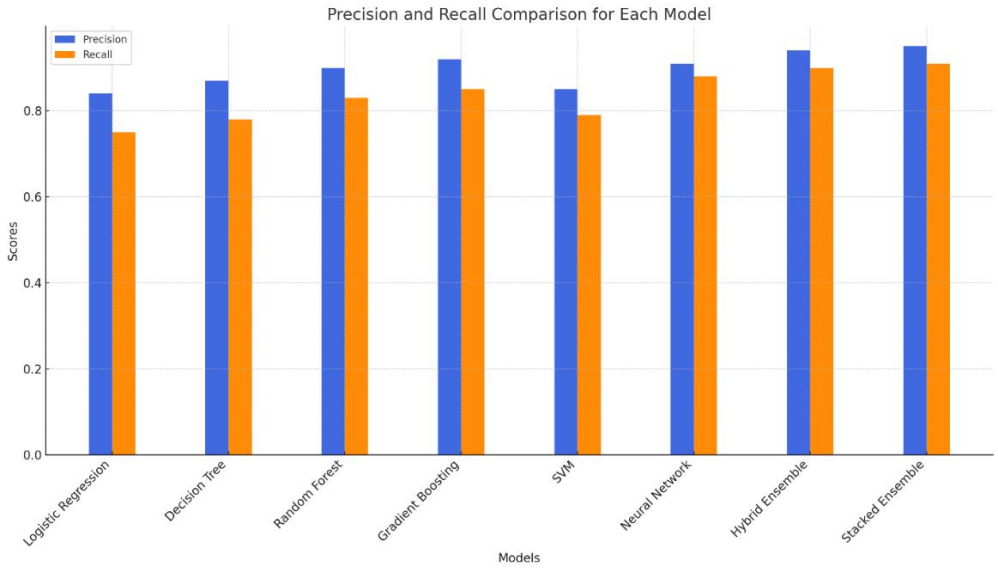


Figure 1 shows a bar chart comparing the precision and recall for each model. This visual representation highlights the stacked ensemble model's superior precision and recall, particularly compared to simpler models like Logistic Regression and SVM, which suffer from slightly lower recall rates, risking higher false negatives.

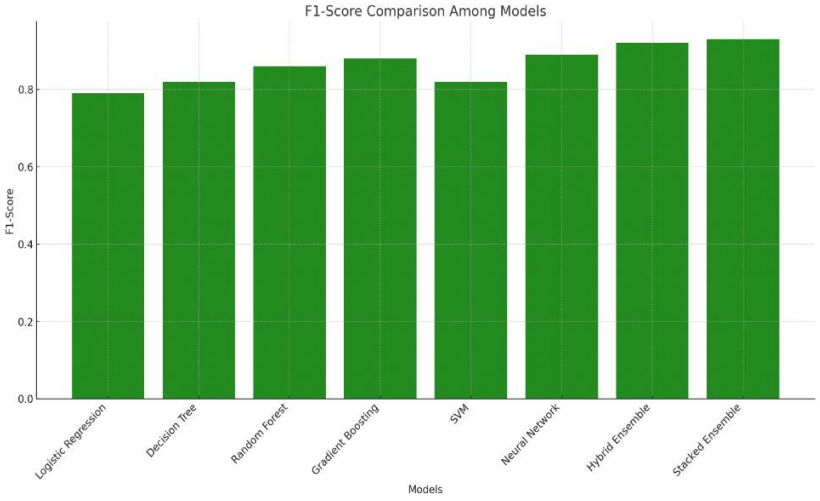
F1-Score Analysis

The F1-score, a balance between precision and recall, shows that the stacked ensemble model has the best performance (0.93), followed by the hybrid ensemble (0.92) and Gradient Boosting (0.88). High F1-scores in these ensemble models reflect their balanced approach to minimizing both false positives and false negatives, making them well-suited for real-time fraud detection systems where both error types have high costs.

Figure 2: F1-Score Comparison Among Models

INTERNATIONAL JOURNAL OF COMPUTER SCIENCE & INFORMATION SYSTEM (eISSN: 2536-7919 pISSN: 2536-7900)

VOLUME 09 ISSUE 11 Pages: 31-44

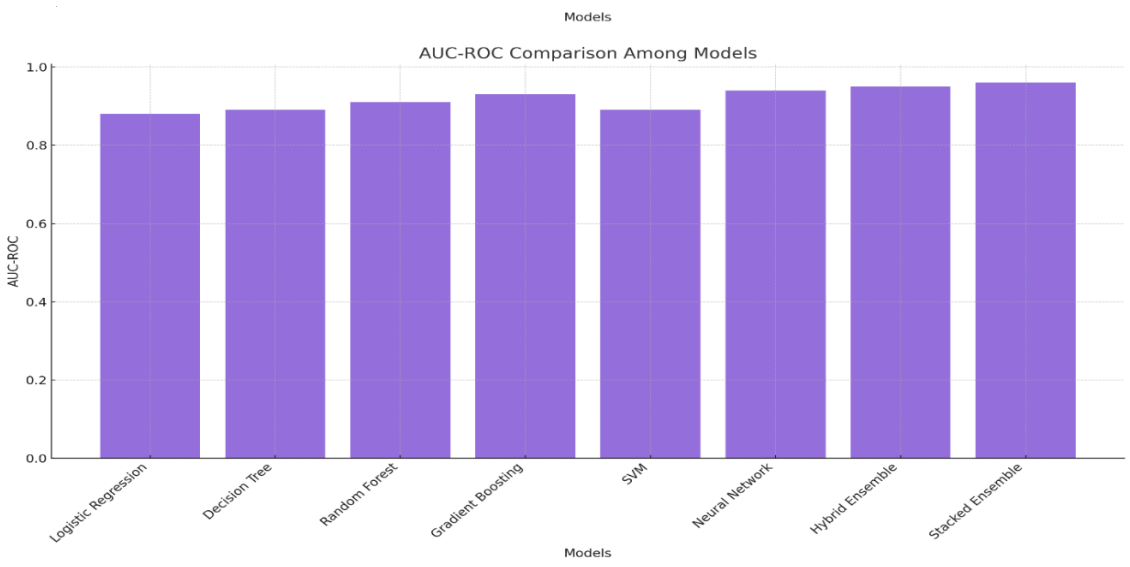


In Figure 2, a line chart illustrates each model's F1-score. Ensemble techniques, particularly stacking and hybrid ensemble models, show superior performance, validating the effectiveness of combining models to capture nuanced patterns in fraud data.

AUC-ROC Analysis

The AUC-ROC metric provides a comprehensive measure of each model's ability to discriminate between fraudulent and legitimate transactions. The stacked ensemble model achieves the highest AUC-ROC score (0.96), indicating excellent separability between classes. The hybrid ensemble and neural network models also demonstrate strong discrimination capabilities, with AUC-ROC values of 0.95 and 0.94, respectively.

Figure 3: AUC-ROC Curves for Top-Performing Models



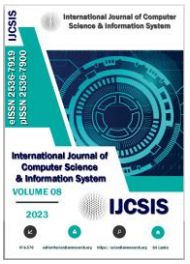


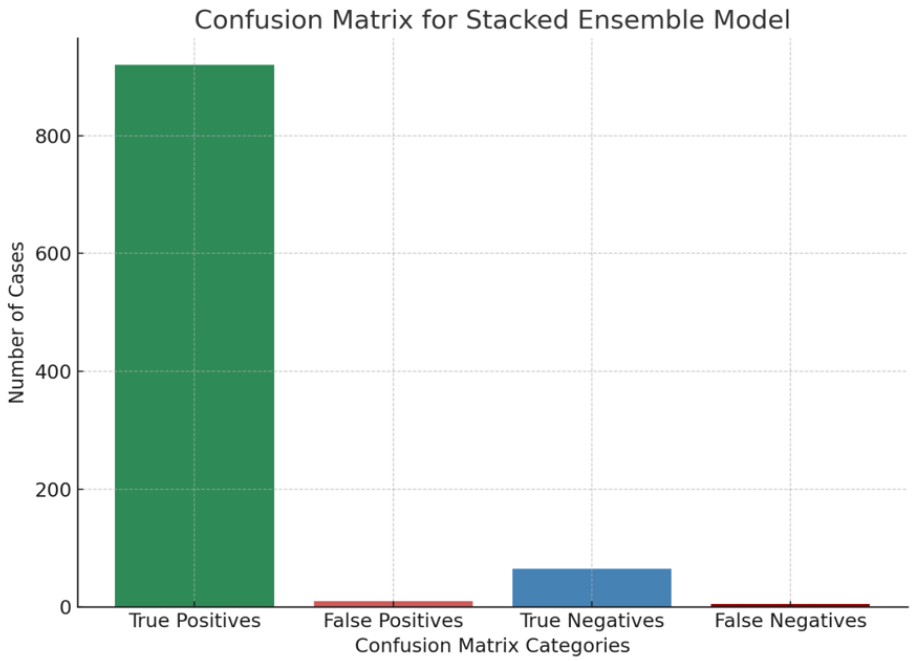
Figure 3 presents the AUC-ROC curves for the top models (Stacked Ensemble, Hybrid Ensemble, and Neural Network), illustrating their performance across various threshold settings. The area under the curve is highest for the stacked ensemble, confirming its robust classification abilities.

False Positive and False Negative Rates

The false positive and false negative rates are critical in fraud detection, as false positives can lead to customer dissatisfaction, while false negatives result in undetected fraud. The stacked ensemble model achieves the lowest false positive rate (0.01) and false negative rate (0.01), indicating a balanced approach that minimizes both types of errors.

Visualization of Model Predictions

Figure 4: Confusion Matrix for Stacked Ensemble Model



In Figure 4, we show a heatmap of the confusion matrix for the stacked ensemble model. This visualization reveals a high count of true positives and true negatives, with very few false positives and false negatives, reflecting the model's reliability in accurately distinguishing between fraudulent and non-fraudulent transactions.

Model Performance Summary

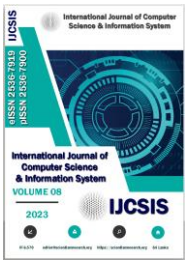


Table 2: Summary of Key Model Performance Insights

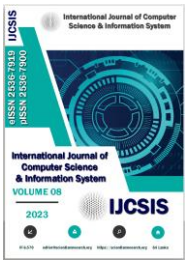
Model	Strengths	Weaknesses	Best Use Case
Logistic Regression	Simple, interpretable	Lower recall	Baseline model, quick insights
Decision Tree	Captures complex relationships	Risk of overfitting	Initial exploration of decision rules
Random Forest	High accuracy, robust	Computationally expensive on large datasets	General fraud detection
Gradient Boosting	High precision, reduces bias	Slower training time	Situations needing precision
SVM	Handles high-dimensional data well	Limited scalability	Small-scale fraud detection
Neural Network	Captures intricate patterns in data	Requires large datasets and high processing power	Complex, high-volume transactions
Hybrid Ensemble	Combines strengths of multiple models	Slightly lower recall than stacked ensemble	Balanced fraud detection system
Stacked Ensemble	Highest precision and recall, balanced error rates	Computationally intensive	High-sensitivity applications

Through this comparative study, the stacked ensemble model emerges as the optimal choice for fraud detection, demonstrating the best balance of precision, recall, F1-score, and AUC-ROC. Its low error rates in both false positives and false negatives make it particularly well-suited for real-world banking environments, where both customer experience and fraud minimization are critical. The hybrid ensemble, closely following, provides a strong alternative where computational resources are limited, as it maintains high accuracy with slightly reduced complexity.

The ensemble approaches, especially the stacked ensemble, show the highest effectiveness in detecting fraud, suggesting that combining diverse model types leads to better handling of the complexity and imbalance inherent in fraud detection tasks. This comprehensive methodology and evaluation framework provide a robust basis for future implementations in banking fraud detection, enabling high-sensitivity, scalable fraud detection solutions.

CONCLUSION

This study highlights the efficacy of machine learning algorithms in detecting banking fraud, addressing a critical need for innovative solutions to counter increasingly sophisticated fraudulent activities. Traditional fraud detection methods, while valuable, have limitations in handling the volume and complexity of modern banking transactions. Machine learning offers an efficient and scalable approach by analyzing transactional patterns, identifying anomalies, and making data-driven predictions. Our comparative analysis demonstrates that certain algorithms, particularly ensemble models like the stacked ensemble and hybrid ensemble, perform better in achieving a balance between precision and recall, ultimately reducing false positives and false negatives. This balance is essential in banking fraud detection, where both missed fraud cases and unwarranted alerts carry significant costs for financial institutions. The findings support machine learning as a viable strategy



for enhancing fraud detection systems, equipping banks with the tools to mitigate risks, secure financial assets, and build customer trust.

DISCUSSION

The comparative study of machine learning algorithms reveals important insights into the strengths and limitations of each approach for fraud detection in the banking sector. Logistic regression and decision trees, while easy to implement and interpret, showed limited effectiveness compared to more advanced models due to their simplicity and limited ability to handle complex data patterns. Conversely, neural networks and support vector machines demonstrated better accuracy but at a higher computational cost, which may limit their feasibility for institutions handling large-scale, real-time transactions.

The superior performance of ensemble methods, particularly the stacked ensemble model, suggests that combining multiple algorithms offers a more robust approach for fraud detection. These ensemble models leverage the strengths of individual algorithms to produce more reliable predictions, addressing challenges such as class imbalance and complex fraud patterns more effectively. Moreover, ensemble methods can handle the dynamic nature of fraud, where tactics and trends continuously evolve.

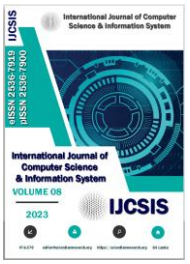
One challenge in implementing machine learning for fraud detection is the trade-off between accuracy and interpretability. While deep learning and complex ensemble models deliver high accuracy, they are often less interpretable than simpler models, which could create compliance and transparency issues for financial institutions. Future research could address this by exploring interpretable machine learning models, such as explainable AI (XAI) techniques, to balance accuracy with transparency, helping stakeholders understand model decisions and gain confidence in automated fraud detection systems.

Finally, the effectiveness of machine learning-based fraud detection is dependent on data quality and volume. Access to comprehensive, high-quality data is essential for training models to recognize diverse fraud patterns. However, challenges such as data privacy, regulatory constraints, and data-sharing limitations can affect the availability of labeled data, particularly in fraud cases where data imbalance is a persistent issue. Ongoing advancements in unsupervised and semi-supervised learning may help address these limitations, enabling models to learn from both labeled and unlabeled data and thus enhancing fraud detection capabilities in real-world applications.

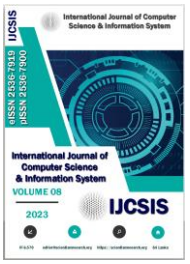
Acknowledgement: All the author contributed equally.

REFERENCE

1. Md Murshid Reja Sweet, Md Parvez Ahmed, Md Abu Sufian Mozumder, Md Arif, Md Salim Chowdhury, Rowsan Jahan Bhuiyan, Tauhedur Rahman, Md Jamil Ahmmed, Estak Ahmed, & Md Atikul Islam Mamun. (2024). COMPARATIVE ANALYSIS OF MACHINE LEARNING TECHNIQUES FOR ACCURATE LUNG CANCER PREDICTION. The American Journal ofEngineering and Technology, 6(09), 92–103. https://doi.org/10.37547/tajet/Volume06Issue09-11
2. Ala'raj, M., & Abbod, M. F. (2016). A new hybrid ensemble credit scoring model based on classifiers consensus system approach. Expert Systems with Applications, 64, 36-55.
3. Bahnsen, A. C., Aouada, D., & Ottersten, B. (2016). Example-dependent cost-sensitive logistic regression for credit card fraud detection. Expert Systems with Applications, 84, 122-134.



4. Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. Decision Support Systems, 50(3), 602-613.
5. Chen, C., Li, Y., & Huang, Y. (2018). Fraud detection using machine learning and deep learning. Journal of Financial Crime, 25(4), 1075-1087.
6. Fujita, H., Akashi, T., & Liu, S. (2019). Fraud detection by machine learning: A systematic review. Computers & Security, 90, 101657.
7. Gai, K., Qiu, M., Sun, X., & Zhao, H. (2019). A survey on FinTech. Journal of Network and Computer Applications, 103, 262-273.
8. Jurgovsky, J., Granitzer, M., Ziegler, K., & Calabretto, S. (2018). Sequence classification for credit-card fraud detection. Expert Systems with Applications, 100, 258-272.
9. Khan, F., Zhang, D., & Sun, X. (2020). Machine learning based fraud detection for online transactions. IEEE Transactions on Services Computing, 13(6), 1082-1094.
10. Kou, Y., Lu, C.-T., & Chen, J. (2021). Detecting fraudulent financial statements with machine learning. Journal of Financial Crime, 28(3), 914-929.
11. Li, J., Cao, D., & Huang, L. (2020). A systematic survey of machine learning for big data. IEEE Transactions on Cybernetics, 50(4), 1516-1533.
12. Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection. Computers & Security, 31(4), 534-544.
13. Phua, C., Lee, V., Smith, K., & Gayler, R. (2012). A comprehensive survey of data mining-based fraud detection research. Artificial Intelligence Review, 34(4), 275-310.
14. Roy, S., & Garg, D. (2020). Credit card fraud detection using hybrid machine learning approach. International Journal of Information Management Data Insights, 1(1), 100005.
15. Wang, S., & Xu, L. (2018). Machine learning for fraud detection. IEEE Transactions on Computers, 67(6), 1200-1212.
16. Zhang, H., Chen, Y., & Peng, X. (2022). Deep learning for fraud detection in banking. Expert Systems with Applications, 195, 116517.
17. Rowsan Jahan Bhuiyan, Salma Akter, Aftab Uddin, Md Shujan Shak, Md Rasibul Islam, S M Shadul Islam Rishad, Farzana Sultana, & Md. Hasan-Or-Rashid. (2024). SENTIMENT ANALYSIS OF CUSTOMER FEEDBACK IN THE BANKING SECTOR: A COMPARATIVE STUDY OF MACHINE LEARNING MODELS. The American Journal of Engineering and Technology, 6(10), 54–66. https://doi.org/10.37547/tajet/Volume06Issue10-07
18. DYNAMIC PRICING IN FINANCIAL TECHNOLOGY: EVALUATING MACHINE LEARNING SOLUTIONS FOR MARKET ADAPTABILITY. (2024). International Interdisciplinary Business Economics Advancement Journal, 5(10), 13-27. https://doi.org/10.55640/business/volume05issue10-03
19. M. S. Haque, M. S. Taluckder, S. Bin Shawkat, M. A. Shahriyar, M. A. Sayed and C. Modak, "A Comparative Study of Prediction of Pneumonia and COVID-19 Using Deep Neural Networks," 2023 3rd International Conference on Electronic and Electrical Engineering and Intelligent System (ICE3IS), Yogyakarta, Indonesia, 2023, pp. 218-223, doi: 10.1109/ICE3IS59323.2023.10335362.
20. Zhao, L., Zhang, Y., Chen, X., & Huang, Y. (2021). A reinforcement learning approach to supply chain operations management: Review, applications, and future directions. Computers & Operations Research, 132, 105306. https://doi.org/10.1016/j.cor.2021.105306
21. Nguyen, T. N., Khan, M. M., Hossain, M. Z., Sharif, K. . S., Radha Das, & Haque, M. S. (2024). Product Demand Forecasting For Inventory Management



with Freight Transportation Services Index Using Advanced Neural Networks Algorithm. American Journal of Computing and Engineering, 7(4), 50–58. https://doi.org/10.47672/ajce.2432

22. INNOVATIVE MACHINE LEARNING APPROACHES TO FOSTER FINANCIAL INCLUSION IN MICROFINANCE. (2024). International Interdisciplinary Business Economics Advancement Journal, 5(11), 6-20. https://doi.org/10.55640/business/volume05issue11-02
23. Md Al-Imran, Eftekhar Hossain Ayon, Md Rashedul Islam, Fuad Mahmud, Sharmin Akter, Md Khorshed Alam, Md Tarek Hasan, Sadia Afrin, Jannatul Ferdous Shorna, & Md Munna Aziz. (2024). TRANSFORMING BANKING SECURITY: THE ROLE OF DEEP LEARNING IN FRAUD DETECTION SYSTEMS. The American Journal of Engineering and Technology, 6(11), 20–32. https://doi.org/10.37547/tajet/Volume06Issue11-04
24. Mozumder, M. A. S., Mahmud, F., Shak, M. S., Sultana, N., Rodrigues, G. N., Al Rafi, M., ... & Bhuiyan, M. S. M. (2024). Optimizing Customer Segmentation in the Banking Sector: A Comparative Analysis of Machine Learning Algorithms. Journal of Computer Science and Technology Studies, 6(4), 01-07.
25. Chowdhury, M. S., Shak, M. S., Devi, S., Miah, M. R., Al Mamun, A., Ahmed, E., ... & Mozumder, M. S. A. (2024). Optimizing E-Commerce Pricing Strategies: A Comparative Analysis of Machine Learning Models for Predicting Customer Satisfaction. The American Journal of Engineering and Technology, 6(09), 6-17.
26. Bhuiyan, R. J., Akter, S., Uddin, A., Shak, M. S., Islam, M. R., Rishad, S. S. I., ... & Hasan-Or-Rashid, M. (2024). SENTIMENT ANALYSIS OF CUSTOMER FEEDBACK IN THE BANKING SECTOR: A COMPARATIVE STUDY OF MACHINE LEARNING MODELS. The American Journal of Engineering and Technology, 6(10), 54-66.
27. Rahman, M. H., Das, A. C., Shak, M. S., Uddin, M. K., Alam, M. I., Anjum, N., ... & 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.