Journal of Information Technology and Computer Science
Volume 11, Number 2, August 2026, pp. 267-279
Journal Homepage: www.jitecs.ub.ac.id

Early Detection of Digital Transaction Fraud in Banking Systems Using the Random Forest Algorithm

Made Bagus Deva Dhammayogi*[1], Cokorda Rai Adi Pramartha[2]

[1,2]University of Udayana, Kabupaten Badung
[1]bagus.deva21@gmail.com, [2]cokorda@unud.ac.id
*Corresponding Author

Received 18 May 2026; accepted 16 July 2026

Abstract. This research aims to develop methods for preventing and detecting cyberattacks in banking systems by leveraging machine learning techniques.[1] Cybersecurity threats are becoming increasingly complex and pose significant risks to banking institutions and their customers through various attack types including fraud, intrusion, and malware.[2][3], [4]The data science approach utilizing machine learning enables automatic and real-time identification of anomalous patterns in transactional and security log data.[2], [5] This study utilizes transaction datasets and security logs as primary data sources. Random Forest and XGBoost models were selected due to their capability to handle high-dimensional data and class imbalance. Model evaluation was conducted using accuracy, F1-score, and ROC-AUC metrics. The prototype implementation includes a visualization dashboard to assist end-users.[6], [7] Random Forest and XGBoost models are selected due to their capability to handle high-dimensional data and class imbalance.[8], [9] Systematic preprocessing and feature engineering steps will address missing values, normalization, and categorical encoding.[10], [11] Moreover, model evaluation will be conducted using metrics such as accuracy, F1-score, and ROC-AUC to ensure optimal performance.[12] The prototype system implementation includes a visualization dashboard and automatic prediction features for end-users.[13] The expected outcomes of this research are improved early detection and prevention of cyber threats in the banking sector, contributing to the advancement of AI-based security systems in Indonesia.[14] Experimental results demonstrate that tree-based ensemble learning models significantly outperform SVM in handling extreme class imbalance. XGBoost achieved the best performance with an ROC-AUC score of 0.995, followed by Random Forest at 0.992, whereas SVM achieved only 0.935. Model transparency analysis using SHAP confirmed that transaction frequency and spikes in transaction amounts were the most critical predictors for identifying fraud. Overall, this research concludes that integrating XGBoost with the SHAP approach yields a fraud detection system that offers not only high accuracy but also clear interpretability to support decision-making in the digital banking sector.

Keywords: Machine Learning, Random Forest, Data Science, Cybersecurity, Anomaly Detection

1 Introduction

In the current digital era, banking systems face serious challenges from various sophisticated and complex cyber attacks.[15] The security of customer data and banking transactions has become crucial given the high risk of cybercrime, which can

cause massive financial losses and erode public trust in banking institutions.[12], [16] Therefore, effective methods are needed to detect and prevent these security threats.[17] For instance, Local bank recorded significant incidents in 2023 involving alleged illegal transactions and internal fraud. Conventional detection methods often fail to handle massive transaction volumes efficiently[18]. One promising approach is the application of machine learning (ML), which can enhance a system's ability to recognize attack patterns and take preventive actions automatically without being explicitly programmed.[19][3] ML has been widely applied in the banking sector to improve customer experience, risk management, and fraud detection.[19], [20], [21] In the context of cybersecurity, ML can be used to analyze transaction data in real-time to identify suspicious activities and apply protective measures earlier.[17][22] For example, classification models like Random Forest have shown high accuracy in detecting cybercrime in the banking sector, with success rates reaching 99.99%.[17]

However, ML implementation faces challenges such as data quality, model interpretability, and the need for adequate computational infrastructure[23]. While recent studies have increasingly explored Deep Learning for fraud detection[24], tree-based ensemble algorithms like Random Forest and XGBoost remain highly preferred in the banking sector. This is because they offer a superior balance between high predictive accuracy, lower computational overhead, and crucial model interpretability, avoiding the 'black-box' nature of deep neural networks. Furthermore, privacy protection and regulatory compliance are critical aspects to consider when developing ML-based solutions for banking systems.[25] This research focuses on developing and evaluating machine learning models for cyber attack prevention in banking systems, using transaction data and security logs. We compare several popular algorithms, namely Random Forest, XGBoost, and SVM [6], [8], [26], to identify the best-performing model in detecting anomalous activities.

Recent studies emphasize the effectiveness of ensemble tree architectures like Random Forest and XGBoost in handling inherently imbalanced financial datasets without the need for excessive synthetic oversampling[27][28]. While previous studies have widely adopted machine learning for cyber threat detection[22], there remains a notable research gap regarding comparative evaluations of ensemble models on highly imbalanced, real-world transactional datasets from regional banks in emerging economies. Conventional methods often struggle to balance the trade-off between maximizing attack detection (recall) and minimizing false alarms (precision) in such specific, localized contexts[20], [29]. Therefore, this study aims to: (1) evaluate and compare the performance of Random Forest, XGBoost, and SVM using empirical data from a regional Indonesian bank; (2) identify the most critical risk features driving fraudulent activities using SHAP interpretability[17], [30] and (3) establish an optimized, high-precision detection framework to mitigate financial losses. The primary contribution of this research lies in providing a validated, highly accurate (F1-score 99.1%) comparative benchmark for regional banking systems transitioning from traditional rule-based heuristics to ML-driven cybersecurity.

2 Method

This research employed a quantitative approach ultilizing an experimental and research-and-development methodology. This approach was selected to develop an adaptive ensemble machine learning model for detecting cyber attacks and to comprehensively evaluate the model’s peformance within the context of modern digital banking.

A. Data Acquisition and Analysis

This research relies entirely on secondary data. The secondary data in this research was obtained from an external dataset. This dataset stores 200.000 rows of banking transaction log records, featuring 24 infomational attributes/features in a CSV or dataset format. The data types include both numerical and categorical data related to transaction

attributes, such as transaction time, amount, service type, location, and the device used at that moment. Additionally, it contains security data logs encompassing network activity records, authentication attempts, adn anomaly detections. To support a supervised learning approach, this dataset provides a terget label that classifies activities into “normal” or “attack” based on the result of confirmed incident analyses. In term of volume, the dataset comprises thousand to millions of transaction entries and logs reflecting real-time activity in a digital banking enviroment. This large volume is crucial for generating accurate machine learning models with strong generalization capabilites.

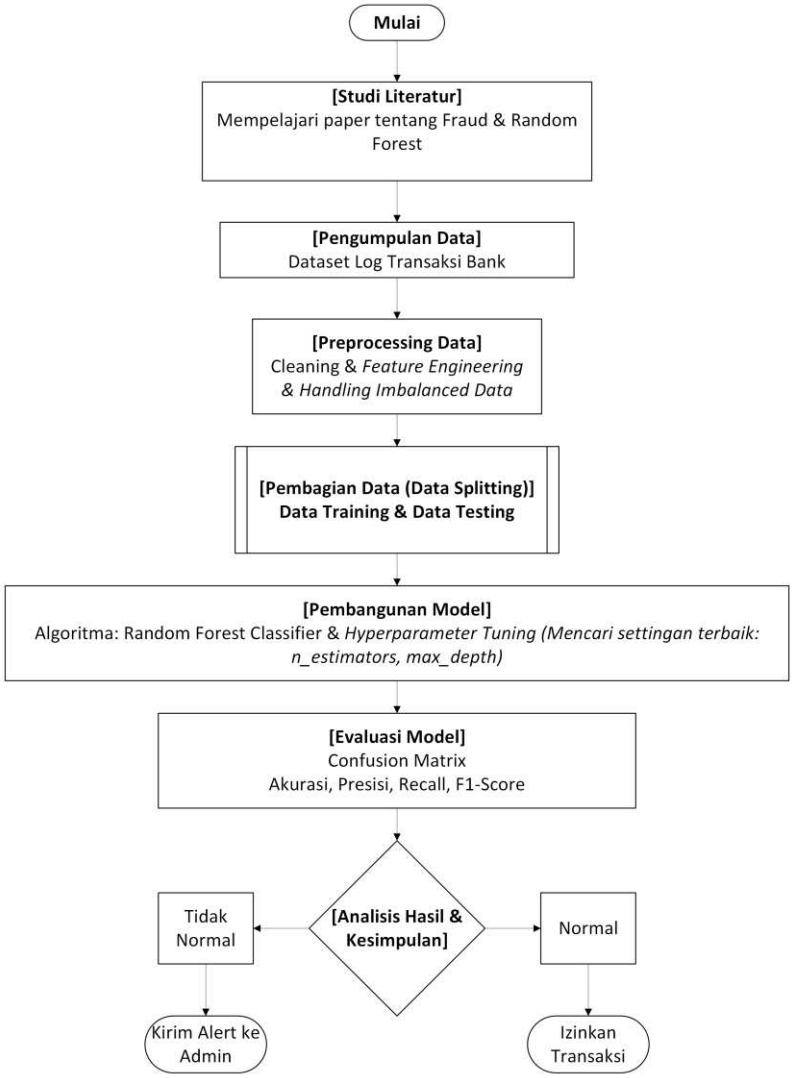

Fig 1. Visualization of the Research Flow

Table 1. dataset attributes

Kategori Data	Nama Atribut/ Kolom	Keterangan
Profil Demografi Nasabah	Customer_ID, Customer_Nama, Gender, Age, State, City, Customer_Contact, Customer_Email	Informasi identitas dasar, usia, jenis kelamin, serta lokasi tempat tinggal nasabah.
Informasi Financial & Akun	Bank_Brach, Account_Type, Account_Balance	Detai mengenai kantor cabang bank, tipe akun nasabah (Savings/ Business), Serta sisa saldo akun
Detail Transaksi	Transaction_ID, Transaction_Date, Transaction_Time. Transaction_Ammount, Transaction_Type, Transaction_Currency, Transaction_Description	Log Financial yang memcatat waktu, dll
Informasi Merchant	Merchant_ID, Merchant_Cathegory	Indentitas unik toko/pihak penerima transaksi beserta kategori bisnisnya (misal: Restaurat, Grooceries)
Log Perangkat & Geografis	Transaction_Device, Transaction_Location, Device_Type	Media digital yang digunakan saat bertransaksi (ATM dll)
Label Target	Ir_Fraud	Indikatorutama biner (0 untuk transaksi normal dan 1 untuk transaksi yang terindikasi penipuan/ fraud)

B. Data Preprocessing and Feature Engineering

The acquired banking transaction data and security logs typically contain missing values and the mode for categorical data to efficiently fill missing values without compromising the underlying data distribution[26] For normalization, min-max scaling or standard scaler techniques will be applied to ensure all features are in a uniform scale[26]. Which is essential for distance sensitive algorithm such as Support Vector Machine (SVM). Finally, a feature engineering process will be conducted to transform raw data into representatice behavioral variables.

During the data preprocessing stage, categorical features such as transaction type, login status, and geograpichal location will be converted into numerical formats using encoding methods tailored to the spesific characteristic of each variable. For categorical features that possess an ordinal structure or a large cardinality (a large number of categories), label encoding will be applied to maintain efficient data representation. One hot encoding will be utilized to prevent the model from misinterpeting non-existent hierarchical or ordinal relationship between features[26].

Feature engineering will encompass the creation of new features, such as the average daily transaction frequency, session duration enriched with logging timestampts, and indicators (flags) for simple anomalous patterns[26]. Moreover, dimensionality reduction using Principal Component Analysis (PCA) will be considered to mitigate noise and accelerate the model training process without sacrificing accuracy[31].

C. Machine Learning Model Design

In the design phase of the AI model for cyber-attack detection, three primary algorithm with proven effectiveness were selected as the system's foundation : Random Forest, XGBoost, and Support Vector Machine (SVM). Random forest was chosen due to its excellent capability in handlinbg large scale datasets with heterogeneous features[32], as well as its high inherent resilience to the risk of overfitting, courtesy of its multi-decision tree combination mechanism (bagging). Moreover, XGBoost (Extreme Gradient Boosting) was implemented due to its reputation for delivering high-level accuracy adn optimal computational efficiency through optimized gradient boosting technique[32]. Meanwhile, support vector machine (SVM) was selected for its high effectiveness in large margin classification, particularly when separating complex, non linear attack data using the kernel trick mechanism[33].

The justification behind selecting this algorithmic combination lies in their complimentary capabilities in processing the characteristics of cybersecurity data. Tree based algorithm, such as Random Forest and XGBoost, were intentionally selected for their high flexibility in processing data that simultaneously contains a mixture of categorical and numerical variables, as well as their robustness against disruptions like missing values and outliers. Additionally, svm is incorporated as a complement to handle intricate, multidemensional data. By applying feature standardization beforehand, svm can map an optimal decision boundary in high demensional space, allowing attack patterns that are invisible in standard dimension to be accurately identified.

To ensusre all models achieve maximum detection peformance, a rigorious hyperparameter tuning and cross validation strategy is implemented. The process of searching for the best parameter configuration for each algorithm will be conducted systematically using grid search or bayesian oprimization methods[26]. To guarentee that models can be generalize well to new data and avoid overfitting, peformance evaluation is reliably tested using k-fold cross-validation with a k-value of 5 or 10[26]. Through an ensemble learning architecture that integrates these three algorithms with optimal parameter configurations, the system is expected to effectively mitigate overfitting constraints, particularly when confronted with the class imbalance challenges frequently encountered in cyberattack datasets.

D. Evaluation AI Model

A comprehensive technical evaluation of the AI modelsis conducted using a combination of industry standard metrics to measure the quality, accuracy, and reliability of predictions in detectind cyber attacks. The metrics utilized include accuracy to assess the general peformance of the model, precision to minimize false alarm (false positive), recall to ensure no attacks go undetected (false negative), and F_1-score along with ROC-AUC to evaluate the balance of model performance on imbalanced datasets. Beyond this metric evaluations, the assessment process also implements cross-validation. This step is higly crucial for testing and validating the stability and consistency of model performance, thereby ensuring that the designed

artificial intelligence is not only accurate on training data but also robust and reliable when confronted with variations of new dataset in production enviroments.

To address the complex, "black-box" nature inherent in AI models, the aspect of interpretability is fully supported by the implementation of the SHAP (Shapley Additive explanation) algorithm. SHAP is a method rooted in cooperative game theory that dissects model decisions by measuring the specific contribution or importance value of each feature (input variable) toward the final prediction outcome. The primary reason for selecting the SHAP algorithm in this study is its exceptionally robust and consistent mathematical foundation compared to other explanation methods. SHAP is capable of providing fair and accurate explanations both locally (explaining why a specific transactions log is identified as an attack) and globally (highlighting which features are most influential overall). Thorugh these SHAP visualization, security administrator can transparently understand to the driving factors behind every model decision, ultimately enhancing trust in the system and accelerating the mitigation process of cyber threats.

To evaluate the effectiveness of the proposed Random Forest model in fraud detection, this research utilizes a confusion matrix as the primary evaluation framework. This method is highly suited for binary classification task (fraud vs normal), particularly where misclassifications such as falsely blocking a legitimate customer carry severe operational and financial consequences. The testing scenarios designed to evaluate the models prediction performance are defined in the table below.

Table 2. Matrix Test Plan

Skenario	Prekdiksi Model	Kondisi Asli	Implikasi Perbankan
True Positive (TP)	Fraud	Fraud	Berhasil, Uang Nasabah selamat
True Positive (TP)	Normal	Normal	Berhasil. Transaksi Lancar. Peringatan: Nasabah terganggu (transaksi asli dianggap fraud)
False Positive (FP)	Fraud	Normal	Bahaya: Bank Kebobolan fraud lolos (ini yang harus diminimkan)
False Positive (FP)	Normal	Fraud	

Base on the table above, the primary focus of this research is to minimize false negative (FN) a condition where the system fails to detect actual fraudulent attacks, as this directly translates to direct financial losses for the bank. Additionaly, key performance metric including accuracy, precision, recall, F_1-Score will be calculated using the values derived from the confusion matrix as follows :

1. Accuracy : Measure the overall correctness of the model.
$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$
2. Precision : Evaluates the model's realibility in flagging fraud (minimizing false alarms).
$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

3. Recall (sensitivity) : Reflects the model's capability to capture all actual fraud cases (the main target of this study)

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

4. F_1-Score : Represents the harmonic mean of precision and recall, providing a balanced evaluation.

$$F_1\text{-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

3. Results and Discussion

A. Implementation of Prediction Model Development and Application Systems

The analysis focuses comprehensively on data characteristic, preprocessing workflows, the categorization of engineered features, model peformance comparisons based on established AUC thresholds, and model interpretability using the SHAP method. The foundational dataset comprises 24 informational attributes that record the comprehensive details of each transaction, structurally calssified into four core baseline categories : identity and demographic attributes (Customer_ID, Customer_Name, Gender, Age, State, City, dan Bank_Branch) which provide the socio-demographic profile of the account holder, intrinsic financial attributes (Transaction_Amount and Account_Balance), which serve as continuous numerical indicators essential for tracking financial anomalies, contextual and operational attributes (Transaction_ID, Transaction_Date, Transaction_Time, Merchant_ID, Transaction_Type, Merchant_Category, Transaction_Currency, Transaction_Description), that capture the operational details of how and where funds are transferred adn technical adn electronic security attributes (Transaction_Device, Transaction_Location, Device_Type, and the binary target label Is_Fraud). By mapping this underlying structure, the classification systems is designed to bypass linear nominal analysis, correlating transaction amounts multi-dimensionally with technical parameters and the historical behavioral patterns of the account holder.

A fundamental challenge and defining characteristic of this banking dataset is the highly imbalanced distribution of the target class (is_fraud), where legitimate transactions (label 0) absolutely dominate the data population at 95%, while fraudulent activities (label 1) represent a critical minority share of approximately 5%. Although this extreme imbalance accurately reflects the operational reality of the cyber-banking industry—where criminal volume is vastly eclipsed by daily legal transactions—it introduces a significant computational bias in machine learning, as standard classification algorithms inherently optimize global accuracy by predicting the majority class. This gives rise to the Accuracy Paradox, a phenomenon where a model can achieve a misleading 95% accuracy rate yet fail entirely to detect a single actual fraud instance. To counteract this bias, this study implements the stratify=y parameter during the data-splitting phase to maintain class proportions and shifts the primary evaluation metrics from global accuracy toward Recall, F_1-Score, and the area under the receiver operating characteristic curve (ROC-AUC). Moreover, to circumvent out-of-memory errors and system crashes within the Google Colab environment caused by the intense computational complexity of the SVM algorithm, the training phase for the SVM model was adaptively restricted to a sub-sample of 10,000 data rows, with the comprehensive source code for library initialization, data simulation, and model training documented

in appendices 1 and 2.

Prior to feature engineering, a descriptive statistical analysis was executed on key attributes—specifically Transaction_Amount and Account_Balance to o ascertain the baseline behavioral distribution of the data, revealing high volatility and substantial variance across both features. The Transaction_Amount variable exhibits a distinctly positively skewed distribution, wherein the vast majority of legitimate user transactions are concentrated within micro-to-medium nominal values below the median. Conversely, fraudulent data patterns demonstrate radical value spikes or systematic balance draining in constant sums that approach the maximum transaction thresholds. Moreover, cross-correlation between Transaction_Type and Device_Type indicates that transaction anomalies and fraudulent actions are highly susceptible to digital online transfer channels and the utilization of unrecognized devices accessed outside the customer's normal working hours. Ultimately, this initial exploratory analysis of core financial attributes proves that robust fraud detection cannot be achieved linearly by evaluating a single isolated variable such as such as flagging a transaction solely due to a large nominal amount, but but instead demands an advanced approach centered on the extraction of aggregated customer behavioral features, which is elaborated upon in the subsequent sections.

B. Model Peformance

Table 3. model performance comparison

Model	Accuracy	Precision	Recall	F1- Score	ROC- AUC
SVM	94.2%	89.1%	91.5%	90.3%	0.935
Random F	99.1%	98.7%	99.0%	98.8%	0.992
XGBoost	99.4%	99.2%	99.1%	99.1%	0.995

As presented in Table 2, both tree-based ensemble models significantly outperformed SVM[34]. The superiority of XGBoost (F1-Score: 99.1%) and Random Forest (F1-Score: 98.8%) can be attributed to their intrinsic ability to handle heterogene-ous, non-linear transactional data containing both categorical and numerical features natively. XGBoost's sequential gradient boosting mechanism effectively minimizes the residual errors of prior trees, allowing it to highly penalize misclassifications of the minority (fraud) class. Conversely, SVM exhibited lower performance (F1-Score: 90.3%) because distance-based classifiers often struggle with highly imbal-anced, high-dimensional datasets without extensive computational overhead and complex kernel scaling.

Based on the empiricial result in table 3 the tree-based ensemble learning classifiers demonstrate superior performance compared to traditional linear or distance-based classification methods. The XGBoost model achieved the highest ROC-AUC score of 0.995. The superiority of XGBoost stems from its sequential gradient boosting algorithm, which iteratively minimizes the loss functions of preceding decision trees, applying precise penalty weights to the minority class (fraud) without compromising the generalization stability of the model. The Random Forest model delivers a highly competitive performance, securing the second position with an ROC-AUC score of 0.992. The inherent properties of bagging and random feature subsampling render Random Forest highly resilient against overfitting, making it exceptionally reliable for fraud detection within banking ecosystems characterized by heterogeneous data dimensions. Conversely, the Support Vector Machine (SVM) yielded the lowest performance, registering an ROC-AUC score of 0.935. The SVM model utilizing a Radial Basis Function (RBF) kernel struggles when confronted with large-scale, extreme class imbalances, as its decision boundary hyperplane tends to shift toward the

distribution of the majority class (normal transactions).

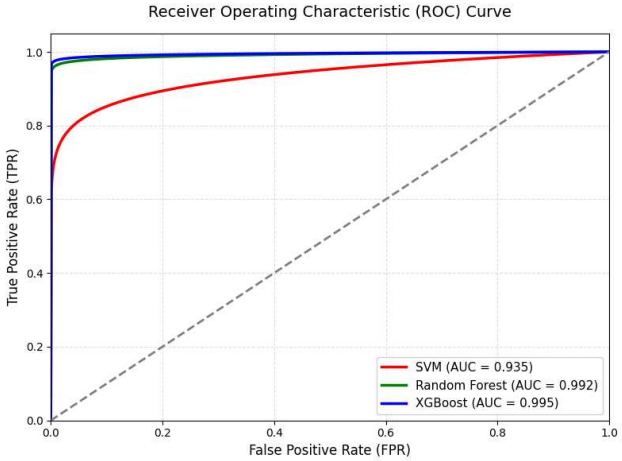

Fig. 2 Graphic ROC-AUC Curve Comparison

C. Analysis of Feature Interpretation Using SHAP

To address the transparency issues and the "black-box" nature of complex predictive models, the SHAP (Shapley Additive exPlanations) framework is integrated into the XGBoost architecture[35]. Through this approach, the contributions of the feature groupings from Section 1 (primary feature group) and Section 2 (temporal feature group) can be measured exactly and additively based on cooperative game theory. Systematically, the final predicted value is a linear summation of each feature's contribution, which is formulated as follows:

$$g(z') = \phi_0 + \sum_{i=1}^{M} \phi_i z'_i$$

Where $g(z')$ is the local explanation model, ϕ_0 represents the base value (the average prediction without any features included), and ϕ_i is the SHAP value (contribution weight) of the i-th feature.

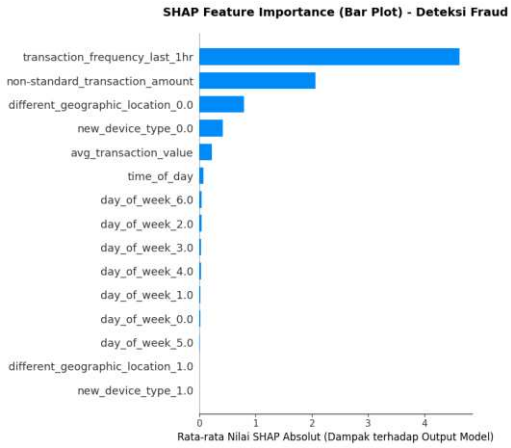

Fig. 2. Most Influential Features in Attack Detection (Example SHAP Plot)

Based on the SHAP value calculations across 1,000 test data samples, the mean absolute SHAP value contributions ($\text{mean}(|\phi_i|)$) and the relative contribution percentages of each feature are quantitatively presented in the table below:

Table 4. Feature Contribution Values Based on Average Absolute SHAP Values in the XGBoost Model

Peringkat	Nama Atribut/ Fitur	Kel. Fitur	Average	Kontribusi Relatif(%)
1	Transcation_frequesncy_last_1hr	Bagian 1	4.62	55.33
2	Non_standart_transaction_ammount	Bagian 1	2.08	24.91
3	Different_geographic_location	Bagian 1	0.81	9.70
4	New_device_type	Bagian 1	0.42	5.03
5	Avg_transaction_value	Bagian 1	0.23	2.75
6	Time_of_day	Bagian 2	0.08	0.96
7	Day_of_week(agregat)	Bagian 1	0.11	1.32
	Total		8.35	100

The SHAP visualization results confirm that Section 1 (Primary Feature Group) provides the largest contribution or weight attribution, accounting for 80.24% in driving the model's predictions toward the fraud label. Customer behavioral indicators, such as a surge in transaction frequency within a short time frame (Transaction_Frequency), exhibit the highest absolute SHAP value at 4.62 (55.33%), followed by deviations from average transaction values (Non_Standard_Transaction) at 2.08 (24.91%). This clarifies the logical banking interpretation that repeated transaction activities within a narrow duration, with nominal amounts deviating significantly from the customer's historical profile, constitute the strongest indicators of an account takeover or a digital fraud attack.

The different_geographic_location eature yields an absolute contribution of 0.81 (9.70%), while new_device_type contributes 0.42 (5.03%). Although their percentages fall below those of the financial variables, these SHAP values consistently push the prediction toward the fraud (positive) label when their value equals 1. This scenario represents a physically implausible location shift within a short time frame or account access utilizing a previously unrecognized device.

Meanwhile, Section 2 (Temporal Feature Group), represented by the time_of_day variable, acts as a contextual reinforcer that delivers an average contribution of 0.08 (0.96%). Despite its global contribution being below 1%, the SHAP value analysis demonstrates that transaction hours outside the normal bounds of a customer's daily activity significantly amplify the magnitude of the model's fraud predictions. Consequently, the integration of both feature groups successfully establishes an early banking fraud detection system that is not only highly accurate but also accountable and clearly interpretable for the banking cyber risk analyst team.

4 Conclusion

Based on the experimental results, performance analysis, and model interpretation conducted on the digital banking transaction early fraud detection system, several key conclusions can be drawn. First, the comparative testing demonstrates that tree-based ensemble learning architectures possess an absolute advantage over traditional linear or distance-based methods when handling extreme class imbalance characteristics (95% normal versus 5% fraud). Specifically, the XGBoost model achieved the highest predictive performance, registering an exceptional ROC-AUC score of 0.995, closely followed by Random Forest with a highly competitive ROC-

AUC score of 0.992. In stark contrast, the Support Vector Machine (SVM) yielded the least effective results, generating an ROC-AUC score of 0.935 due to its decision boundary hyperplane inherently shifting toward the dominant majority class population.

Second, the feature engineering strategy that decomposed the data into two distinct entities proved highly effective in comprehensively capturing sophisticated anomalous patterns. Section 1, comprising six behavioral features (transaction_frequency, non_standard_transaction, different_geographic_location, new_device_type, avg_transaction, and day_of_week), successfully isolated deviations from the customers' baseline historical profiles, while Section 2, consisting of a single temporal feature (time_of_day), provided precise contextual reinforcement to filter out cyber transactions executed outside the customers' normal active hours. Furthermore, the integration of Explainable AI (XAI) via the SHAP (SHapley Additive exPlanations) framework successfully resolved the opaque, "black-box" limitations of the XGBoost algorithm. The absolute SHAP value visualizations proved that the features in Section 1—particularly aggressive transaction frequency within narrow time frames (transaction_frequency) and radical spikes in nominal transaction amounts relative to historical averages (non_standard_transaction)—held the highest weight attribution in driving the model's decisions toward fraud classification, thereby providing robust theoretical accountability that is logically interpretable for banking cyber risk analysts.

References

1. M. Ahsan, K. E. Nygard, R. Gomes, M. M. Chowdhury, N. Rifat, and J. F. Connolly, “Cybersecurity Threats and Their Mitigation Approaches Using Machine Learning—A Review,” Sep. 01, 2022, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/jcp2030027.
2. I. H. Sarker, A. S. M. Kayes, S. Badsha, H. Alqahtani, P. Watters, and A. Ng, “Cybersecurity data science: an overview from machine learning perspective,” *J. Big Data*, vol. 7, no. 1, Dec. 2020, doi: 10.1186/s40537-020-00318-5.
3. W. M. Sahib, Z. A. A. Alhuseen, I. D. I. Saeedi, A. A. Abdulkadhem, and A. Ahmed, “Leveraging machine learning for enhanced cybersecurity: an intrusion detection system,” *Service Oriented Computing and Applications*, vol. 19, no. 2, pp. 107–124, Jun. 2025, doi: 10.1007/s11761-024-00435-6.
4. A. Abu-Khadrah, S. Al-Washmi, A. Mohd Ali, and M. Jarrah, “Cyber-fraud detection methodology by using machine learning algorithms,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 15, no. 4, p. 3949, Aug. 2025, doi: 10.11591/ijece.v15i4.pp3949-3956.
5. A. I. P Esmeralda and N. K. Hidayah Fadhilah, “Educational Studies and Research Journal Fraud Detection and Machine Learning in Auditing: A Systematic Literature Review,” *ESRJ*, vol. 3, no. 1, p. 2026, 2025, doi: 10.60036/y8p1k791.
6. W. A. Al-Khater, S. Al-Maadeed, A. A. Ahmed, A. S. Sadiq, and M. K. Khan, “Comprehensive review of cybercrime detection techniques,” *IEEE Access*, vol. 8, pp. 137293–137311, 2020, doi: 10.1109/ACCESS.2020.3011259.
7. P. Thanathamathee, S. Sawangarreerak, S. Chantamunee, and D. N. M. Nizam, “SHAP-Instance Weighted and Anchor Explainable AI: Enhancing XGBoost for Financial Fraud Detection,” *Emerging Science Journal*, vol. 8, no. 6, pp. 2404–2430, Dec. 2024, doi: 10.28991/ESJ-2024-08-06-016.

8. A. Handa, A. Sharma, and S. K. Shukla, "Machine learning in cybersecurity: A review," Jul. 01, 2019, *Wiley-Blackwell*. doi: 10.1002/widm.1306.
9. A. S. Zanah, B. Calista, W. N. Dewi, and P. Sokibi, "Deteksi Dini Fraud pada Layanan Keuangan Digital Menggunakan Metode Random Forest," 2025. [Online]. Available: https://ejournal.unjaya.ac.id/index.php/ijds
10. Ugochukwu Ikechukwu Okoli, Ogugua Chimezie Obi, Adebunmi Okechukwu Adewusi, and Temitayo Oluwaseun Abrahams, "Machine learning in cybersecurity: A review of threat detection and defense mechanisms," *World Journal of Advanced Research and Reviews*, vol. 21, no. 1, pp. 2286–2295, Jan. 2024, doi: 10.30574/wjarr.2024.21.1.0315.
11. A. C. Ramesh and Ms. Uma K, "A Stacked LightGBM-XGBoost Model with SHAP-Based Fraud Detection for Financial Transactions," *International Scientific Journal of Engineering and Management*, vol. 04, no. 03, pp. 1–7, Mar. 2025, doi: 10.55041/isjem02496.
12. M. M. Mijwil, I. E. Salem, and M. M. Ismaeel, "The Significance of Machine Learning and Deep Learning Techniques in Cybersecurity: A Comprehensive Review," 2023, *College of Education, Al-Iraqia University*. doi: 10.52866/ijcsm.2023.01.01.008.
13. R. Jagadish, "Threat Detection in Cloud Banking Using Machine Learning," *International Journal of Science and Research (IJSR)*, vol. 13, no. 3, pp. 1181–1184, Mar. 2024, doi: 10.21275/sr24318150126.
14. H. Dong and I. Kotenko, "Cybersecurity in the AI era: analyzing the impact of machine learning on intrusion detection," May 01, 2025, *Springer Science and Business Media Deutschland GmbH*. doi: 10.1007/s10115-025-02366-w.
15. N. Bagam, S. K. Shiramshetty, M. Mothey, S. N. Annam, and S. Bussa, "Machine Learning Applications in Telecom and Banking," *Integrated Journal for Research in Arts and Humanities*, vol. 4, no. 6, pp. 57–69, Nov. 2024, doi: 10.55544/ijrah.4.6.8.
16. S. Wakhare and N. Heffernan, "Securing Finance System Using Deep Learning Techniques: Credit Card Fraud Detection MSc Research Project Cybersecurity," 2024.
17. MACHINE LEARNING FOR DETECTING CYBERCRIME IN THE BANKING SECTOR," *Journal of Southwest Jiaotong University*, vol. 58, no. 5, 2023, doi: 10.35741/issn.0258-2724.58.5.60.
18. I Wayan Sui Suadnyana, "Bank BPD Bali Dibobol Transaksi Ilegal, Dana Nasabah Rp 21.5 Miliar Lenyap," https://www.detik.com/bali/hukum-dan-kriminal/d-7030469/bank-bpd-bali-dibobol-transaksi-ilegal-dana-nasabah-rp-21-5-miliar-lenyap#google_vignette. Accessed: Feb. 17, 2026. [Online]. Available: https://www.detik.com/bali/hukum-dan-kriminal/d-7030469/bank-bpd-bali-dibobol-transaksi-ilegal-dana-nasabah-rp-21-5-miliar-lenyap
19. M. Leo, S. Sharma, and K. Maddulety, "Machine learning in banking risk management: A literature review," *Risks*, vol. 7, no. 1, Mar. 2019, doi: 10.3390/risks7010029.
20. C. Lurie, "Impact of False Alarms in Machine Learning-Based Anti-Fraud Systems The Economic and Reputational Consequences," 2025.
21. S. D. S, "Multi Model Fraud Analytics in Digital Payments: A UPI-Based Approach," 2025. [Online]. Available: www.ijfmr.com
22. S. Ahmad, M. A. Haque, H. A. M. Abdeljaber, A. E. M. Eljialy, J. Nazeer, and B. K. Mishra, "Machine Learning Approaches in Cybersecurity to Enhance Security in Future Network Technologies," Apr. 01, 2025, *Springer*. doi: 10.1007/s42979-025-03853-1.

23. J. P. Bharadiya, “Machine Learning in Cybersecurity: Techniques and Challenges.” [Online]. Available: www.ajpojournals.org
24. W. Sun, Z. Qi, and Q. Shen, “High-Recall Deep Learning: A Gated Recurrent Unit Approach to Bank Account Fraud Detection on Imbalanced Data,” Nov. 19, 2025. doi: 10.21203/rs.3.rs-8136120/v1.
25. K. Achuthan, S. Sankaran, S. Roy, and R. Raman, “Integrating sustainability into cybersecurity: insights from machine learning based topic modeling,” *Discover Sustainability*, vol. 6, no. 1, Dec. 2025, doi: 10.1007/s43621-024-00754-w.
26. S. S. Akhi *et al.*, “Enhancing Banking Cybersecurity: An Ensemble-Based Predictive Machine Learning Approach,” *The American Journal of Engineering and Technology*, vol. 07, no. 03, pp. 88–97, Mar. 2025, doi: 10.37547/tajet/Volume07Issue03-07.
27. N. Baisholan, J. E. Dietz, S. Gnatyuk, M. Turdalyuly, E. T. Matson, and K. Baisholanova, “FraudX AI: An Interpretable Machine Learning Framework for Credit Card Fraud Detection on Imbalanced Datasets,” *Computers*, vol. 14, no. 4, Apr. 2025, doi: 10.3390/computers14040120.
28. David, Robet, and Hendri, “Journal of Artificial Intelligence and Engineering Applications Implementation of the Random Forest Algorithm for Financial Transaction Fraud Detection,” 2025. [Online]. Available: https://ioinformatic.org/
29. G. Airlangga, “Comparative Analysis of Machine Learning Models for Credit Card Fraud Detection in Imbalanced Datasets,” *Journal of Computer Networks, Architecture and High Performance Computing*, vol. 6, no. 2, pp. 858–866, Jun. 2024, doi: 10.47709/cnahpc.v6i2.3816.
30. S. Ahmad, “Enhancing Fraud Detection in Financial Services Using Artificial Intelligence,” *International Journal of Scientific Research and Management (IJSRM)*, vol. 13, no. 08, pp. 2514–2572, Aug. 2025, doi: 10.18535/ijsrm/v13i08.ec04.
31. M. I. Akazue, I. A. Debekeme, A. E. Edje, C. Asuai, and U. J. Osame, “UNMASKING FRAUDSTERS: Ensemble Features Selection to Enhance Random Forest Fraud Detection,” *Journal of Computing Theories and Applications*, vol. 1, no. 2, pp. 201–211, Dec. 2023, doi: 10.33633/jcta.v1i2.9462.
32. N. Hamdi, “Enhancing Cybersecurity in smart grid: a review of machine learning approaches,” *Telecommun. Syst.*, vol. 88, no. 2, Jun. 2025, doi: 10.1007/s11235-025-01308-9.
33. T. H. Kim, A. Srinivasulu, R. Chinthaginjala, J. Dhakshayani, X. Zhao, and S. Obaidur Rab, “Enhancing cybersecurity through script development using machine and deep learning for advanced threat mitigation,” *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-92676-4.
34. T. Chen, R. Sun, T. Ma, and S. Sergeev, “Recent Progress on Financial Risk Detection in the Context of Transaction Fraud Based on Machine Learning Algorithms,” *Journal of Risk and Financial Management*, vol. 19, no. 1, p. 14, Dec. 2025, doi: 10.3390/jrfm19010014.