

Machine Learning-Based Fraud Detection in Banking Transactions

Yiliu Hu^{1*}, Jiongyu Chen²

¹School of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing, China

²School of Applied Mathematics, Xi'an Jiaotong-Liverpool University, Suzhou, China

**Corresponding Author. Email: 1842471705@qq.com*

Abstract. With the development of globalization, the bank transactions has great changed. This study systematically examines the significant role of FinTech in the global banking industry, integrating academic research results from 2015 to 2025. The study analyzed the change of traditional banking models by innovative technologies such as blockchain, artificial intelligence (AI), peer-to-peer (P2P) lending, and mobile payments, and explored the opportunities and challenges brought by these innovations. The study evaluated the solution of the banking industry to the impact of financial technology, analyzed the relevant regulatory environment, and identified future research directions, particularly in emerging markets. The findings indicate that financial technology plays a significant role in improving bank operational efficiency, increase financial inclusiveness, while emphasizing the necessity of promoting ethical technology integration.

Keywords: FinTech, Fraud detection, Digital transformation, Artificial intelligence, Compliance Supervision

1. Introduction

The global financial services industry is undergoing a transformative era influenced by FinTech. In 2016 alone, the investment in FinTech companies exceeded 24.7 billion US dollars [1]. This technological revolution has changed the traditional banking operation model and enhanced transaction efficiency and personalized services. According to the [2], FinTech refers to innovations driven by technology in the financial services sector, potentially bringing new business models, application scenarios, business processes or products, fundamentally altering the structure and delivery methods of financial services. Academic discussions on the impact of FinTech on the banking industry mainly focus on three core issues: how technological innovation reconfigures traditional banking business processes, what kind of competitive situation has emerged between traditional banks and FinTech companies, and how the regulatory framework needs to be adjusted to adapt to this digital transformation. This literature review aims to answer these questions, covering aspects such as technological changes, strategic adaptations, regulatory landscapes, and future research agendas. The research analysis is based on formal literature, industry reports.

2. Literature review

2.1. Technological innovation and disruptive mechanisms

2.1.1. Blockchain and Decentralized finance (DeFi)

The distributed ledger technology (DLT) of blockchain has changed the transparency and security of bank transactions. Unlike traditional systems, blockchain supports peer-to-peer transactions with immutable records, reducing the settlement time from 3-5 days to nearly real-time processing [3]. The smart contracts encoded in the blockchain, as self-executing agreements, automate compliance processes such as "Know Your Customer" (KYC) checks, reducing operational costs for banks by 30-50% [4]. Decentralized finance (DeFi) platforms built on blockchain challenge the roles of traditional financial intermediaries. The decentralized exchange (DEX) of Ethereum supports peer-to-peer transactions without the involvement of banks, while stablecoins such as USDC enable smooth global payments. However, scalability remains a key bottleneck: the transaction processing speed of public blockchains such as Bitcoin is only 7 transactions per second (TPS).

2.2. Product and risk management innovation

Product and Service Innovation Fintech has driven the development of new financial products, including robo-advisors, peer-to-peer (P2P) lending platforms, and mobile payment solutions. Indonesian banks now offer payment services based on Near Field Communication (NFC) and QR codes, which has improved financial services in underdeveloped areas. For instance, mobile wallets like OVO and Dana have enabled millions of unbanked individuals to participate in digital transactions. P2P lending platforms such as Investree and Modalku have facilitated financing for small and medium-sized enterprises, addressing key gaps in traditional banking services.

Advanced Risk Management and Security Due to concerns over data security, risk management systems have been strengthened. Banks now use AI-driven behavioral analysis and real-time threat detection to identify fraudulent activities. Cryptography and blockchain have also enhanced transaction security, such as the use of distributed ledger technology by Bank Rakyat Indonesia (BRI) to secure cross-border payments. However, small banks often lack advanced security infrastructure and are vulnerable to cyber attacks and data breaches.

2.3. Regulatory uncertainty

The fintech industry in Indonesia is confronted with regulatory gaps that impede sustainable growth [5]. argue that the incomplete legal framework for fintech services has created numerous risks, particularly in the areas of P2P lending and digital payments. The absence of standardized regulations leads to inconsistent supervision, with some banks taking advantage of regulatory arbitrage while others struggle to adapt [6]. suggest that regulatory authorities adopt a method similar to the UK's regulatory sandbox to foster innovation while safeguarding consumers.

2.4. Electronic banking fraud transactions

Fraud activities in electronic banking are dynamic and complex [7]. pointed out that such frauds are "difficult to detect because the fraud activities are dynamic and distributed across different customer files", and occur in irregular data sets. Common methods include phishing or vishing (deceptive

emails or phone calls to obtain sensitive information) and malware attacks (such as Zeus Trojan horses and other malicious software used to invade accounts) [8].

Security Measures and Shared Responsibility Banks adopt a mixed fraud detection method, combining internal transaction data with third-party analysis to identify and detect [9]. Preventive measures include identity verification, key encryption, and anti-spyware tools. However, as responsibility shifts to customers, banks must clearly communicate protection measures, such as installing antivirus software or avoiding clicking on suspicious links [10]. The shared responsibility policy is used to define boundaries, but ambiguity still exists. For example, if a customer ignores warnings about known phishing attacks, the responsibility may shift to the customer. This requires banks to proactively communicate risks to avoid legal disputes and maintain trust [11].

3. Methodology

This paper uses logistic regression, XGBoost, random forest, and neural network models for modeling and fraud prediction.

3.1. Resource of data

This paper utilizes the publicly available "Bank Transaction Fraud Detection" dataset from Kaggle, published by Sagar Maru in February 2025. The dataset contains 200,000 complete bank transaction records from LOL Bank Pvt. Ltd., integrating multi-system data sources including 24 well-structured features with no null values or outliers.

The binary target variable `Is_Fraud` is analyzed in relation to three main feature categories: Customer Information (`Customer_ID` as unique identifier, Demographics including Gender and Age, Location data with State, City and Bank_Branch, and Account details specifying `Account_Type` as Business/Savings/Checking), Transaction Attributes (`Transaction_ID` for identification, Temporal elements of `Transaction_Date` and `Transaction_Time`, Financial components including `Transaction_Amount` and `Account_Balance`, along with Contextual information such as `Transaction_Type`, `Merchant_Category` and `Device_Type`), and Supplementary Features (`Transaction_Location`, `Transaction_Currency` which is exclusively INR in this dataset, and `Customer_Contact/Email` details), forming a comprehensive framework for fraud detection analysis.

3.2. Data features analysis

The comparison between fraudulent and non-fraudulent samples (see Figure 1) reveals a significant class imbalance, with fraudulent samples being substantially outnumbered by non-fraudulent ones. Such data imbalance can severely impact model performance, necessitating appropriate preprocessing techniques to mitigate bias and enhance predictive accuracy.

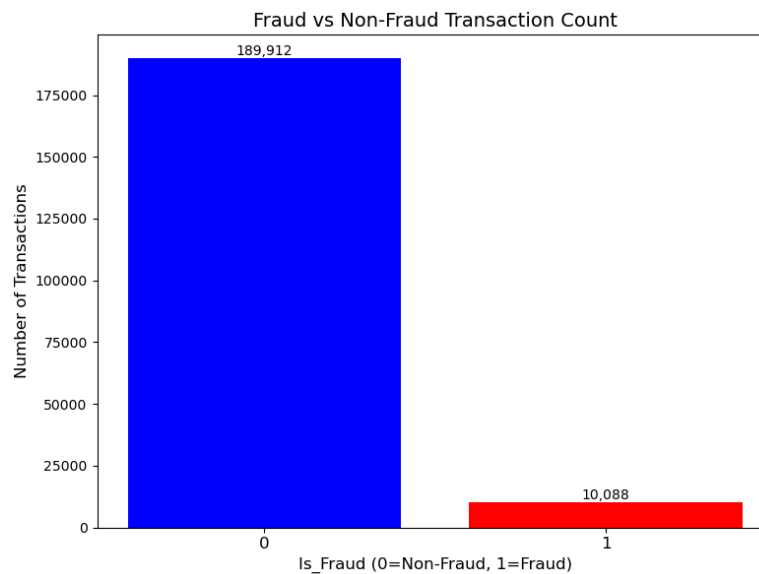


Figure 1. Comparison between fraudulent and non-fraudulent transactions

3.3. Data preprocessing & feature engineering

3.3.1. Data cleaning

The raw dataset underwent rigorous preprocessing to ensure data quality and enhance model performance. Confidential identifiers (e.g., customer names, transaction IDs) were removed to preserve privacy. Temporal features were processed to capture transaction patterns: timestamps were sorted chronologically, and cyclical time variables (e.g., hour of day) were encoded using trigonometric transformations to maintain periodicity.

3.3.2. Feature engineering

To enhance dataset quality, this paper engineered the following features to capture potential fraud-related patterns:

(1) High-risk time window flag

Transactions occurring between 00:00-04:00 were flagged as high-risk based on empirical evidence that fraudulent activities are more prevalent during these hours, when legitimate users are typically inactive.

(2) Transaction amount ratio

The ratio of transaction amount to account balance was calculated and log-transformed to reduce skewness. This metric helps identify suspicious transactions where the amount is disproportionately large relative to the account balance - a potential indicator of fraudulent transfers.

(3) Spatiotemporal features

Rolling fraud rates by geographical region (state/city) were computed to detect localized fraud clusters, as fraudulent transactions often exhibit spatial concentration patterns.

(4) Interaction terms

Multiplicative interaction features (e.g., transaction amount $\times$ hour) were constructed to capture temporal risk patterns, particularly "large transactions during specific time windows" which may indicate coordinated fraudulent activity.

3.4. Model selection

This study employs four models for comparative analysis: Logistic Regression, XGBoost [12], Random Forest [13], and Neural Network [14]. The rationale for selecting these models is as follows:

Logistic Regression: As a classic linear classifier, it offers high computational efficiency and strong interpretability, making it suitable for preliminary analysis. Its probabilistic outputs can directly serve as risk scores for operational use.

XGBoost: This gradient-boosted tree model excels at capturing nonlinear patterns in structured data while mitigating overfitting through regularization. It demonstrates superior performance in fraud detection, particularly for imbalanced datasets.

Random Forest: By aggregating multiple decision trees, it reduces variance and maintains robustness against noisy data and feature redundancy. Its built-in feature importance evaluation helps identify critical fraud-related attributes.

Neural Network: Neural networks can autonomously capture such non-linear combinations through hierarchical representations, surpassing manual feature engineering.

A comprehensive comparison of these models' performance will systematically evaluate their efficacy for bank fraud detection and inform practical deployment strategies.

3.5. Evaluation metrics

In bank transaction fraud detection, where fraud cases are extremely rare, conventional accuracy becomes misleading. Thus, this study adopts the following targeted metrics:

3.5.1. Precision

The indicator evaluates the accuracy of fraud detection.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

TP (True Positive) represents the number of samples correctly predicted as positive, and FP (False Positive) represents the number of samples incorrectly predicted as positive. High precision ensures alarm credibility, reducing operational costs for manual review teams.

3.5.2. Recall

The indicator measures the model's performance in detecting real fraudulent transactions.

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

FN (False Negative) is the number of samples which are incorrectly predicted as negative. High recall means the model has ability to capture more positive samples. Maximizing recall is critical as

a missed fraud (FN) incurs direct financial losses, whereas false alarms (FP) only entail review costs.

3.5.3. F1-score

This indicator represents the harmonic mean of precision and recall, combining the importance of both [15].

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (3)$$

It provides a unified optimization goal when bank needs to balance risk control costs (precision) and fraud coverage (recall).

3.5.4. AUC-ROC

AUC-ROC *quantifies the model's overall discriminative power [16].* This indicator evaluates true positive rates (TPR) and false positive rates (FPR) across decision thresholds, *Values closer to 1 indicate optimal trade-offs between fraud detection and false alarm suppression, which is vital to guide threshold tuning in banking workflows.*

4. Model training

4.1. Logistic Regression model

4.1.1. Dataset splitting

The relevant features for modeling were selected from the dataset. The dataset was then partitioned into training set and testing set by the `train_test_split` function, with the test set comprising 20% of the total data. The `stratify=y` parameter was applied to ensure that the ratio of positive (fraudulent) to negative (legitimate) samples in both the training set and test set remained consistent with the original dataset distribution.

4.1.2. Class imbalance handling

To address the severe class imbalance (5% fraud prevalence), this study implemented manual oversampling to boost fraudulent samples to 10% of the training set (`target_ratio=0.1`). The required additional samples n_{new} were calculated via:

$$n_{new} = \left\lceil \frac{target_ratio \times n_{total} - n_{fraud}}{1 - target_ratio} \right\rceil \quad (4)$$

and generated through random sampling with replacement from existing fraud cases. This conservative approach, unlike SMOTE, preserves data authenticity by avoiding synthetic sample generation - particularly crucial for high-dimensional financial data where artificial samples may

distort feature relationships. The selected 10% balance optimally enhances fraud pattern recognition while preventing overfitting.

4.1.3. Results

This model achieved an AUC-ROC score of 0.6315, demonstrating moderate discriminative capability. As shown in Table 1, while the model exhibited high precision (0.96) for legitimate transactions (class 0), its performance on fraud detection (class 1) revealed critical trade-offs: despite moderate recall (0.52), the low precision (0.07) indicates substantial false positives among predicted fraud cases. The overall accuracy of 0.63 reflects the dataset's inherent class imbalance, where the model's tendency to favor majority-class predictions artificially inflates this metric.

Table 1. The Classification report of Logistic Regression model

	Precision	Recall	F1-score	Support
0	0.96	0.63	0.76	9486
1	0.07	0.52	0.13	514
Accuracy			0.63	10000

4.2. XGBoost model

4.2.1. Dataset splitting

Consistent with the data splitting methodology outlined in Section 4.1.1, this model maintained the 80:20 train-test ratio with stratified sampling.

4.2.2. Class imbalance handling

In order to address the immanent class imbalance in fraud detection datasets (where fraudulent transactions typically represent <5% of cases), we implement a weighted classification approach through XGBoost's `scale_pos_weight` parameter. This critical parameter is calculated as:

$$scale_{pose_weight} = \frac{n_{normal}}{n_{fraud}} \quad (5)$$

n_{normal} and n_{fraud} refer to the volume of genuine and fraudulent transactions in the training samples. This approach assigns greater importance to correctly identifying the minority fraud class by adjusting misclassification penalties."

4.2.3. Model evaluation

The XGBoost model achieved an AUC-ROC score of 0.6006, demonstrating limited discriminative capability in fraud detection. As shown in Table 2, the model reliably identifies legitimate transactions (precision=0.96, recall=0.71), but shows limited fraud detection capability. While detecting 40% of fraud cases (recall=0.40), its low precision (0.07) generates excessive false alarms (93% false positive rate).

Table 2. The classification report of XGBoost model

	Precision	Recall	F1-score	Support
0	0.96	0.71	0.81	9486
1	0.07	0.40	0.12	514
Accuracy			0.69	10000

4.3. Random Forest model

4.3.1. Dataset splitting

Consistent with the data splitting methodology outlined in Section 4.1.1, this model maintained the 80:20 train-test ratio with stratified sampling.

4.3.2. Class imbalance handling

The Random Forest classifier addresses class imbalance through its built-in `class_weight='balanced'` parameter, which can automatically adjusts weights inversely proportional to class frequencies. This weighting strategy assigns higher importance to minority class (fraudulent transaction) samples during both individual tree construction and aggregate voting process.

4.3.3. Model evaluation

This model achieved an AUC-ROC score of 0.6215, demonstrating moderate discriminative capability for fraud detection. As shown in Table 3, the model achieves high precision (0.96) but moderate recall (0.62) for legitimate transactions, while fraud detection shows balanced recall (0.52) but critically low precision (0.07), resulting in a 93% false positive rate.

Table 3. The classification report of Random Forest model

	Precision	Recall	F1-score	Support
0	0.96	0.62	0.75	9486
1	0.07	0.52	0.12	514
Accuracy			0.62	10000

4.4. Neural Network model

4.4.1. Dataset splitting

Consistent with the data splitting methodology outlined in Section 4.1.1, this model maintained the 80:20 train-test ratio with stratified sampling.

4.4.2. Class imbalance handling

To address severe class imbalance, the model implements a composite approach combining data-level and algorithm-level interventions. At the data level, manual oversampling strategically increases fraud sample representation to 30% of the training set (fraud-to-legitimate ratio 3:7)

through randomized duplication of minority-class instances. At the algorithm level, a customized weighted loss function applies differential penalties, assigning a $3\times$ greater weight (`weight_for_1` = 3.0) to misclassified fraud cases compared to legitimate transactions. The model further refines its decision boundary through precision-recall guided threshold optimization, dynamically selecting an operational threshold of 0.28 (versus the default 0.5) to prioritize recall - a critical requirement in fraud detection systems.

4.4.3. Model evaluation

This model achieved an AUC-ROC score of 0.5084 which suggests near-random classification performance for fraud cases. As shown in Table 4, this model demonstrates strong performance for legitimate transactions (Class 0), achieving high precision (0.95), recall (0.95), and F1-score (0.95). However, fraud detection (Class 1) remains challenging, with precision, recall, and F1-score all at 0.07, indicating poor discriminative capability for minority-class samples.

Table 4. The classification report of Neural Network model

	Precision	Recall	F1-score	Support
0	0.95	0.95	0.95	9486
1	0.07	0.07	0.07	514
Accuracy			0.91	10000

5. Results and comparison

To comprehensively evaluate the effectiveness of different approaches, the study compare the performance of four machine learning models on the fraud detection task. As summarized in Table 5, each model demonstrates distinct strengths and limitations across key evaluation metrics.

Table 5. Comparative performance of fraud detection models

	Non-Fraud F1	Fraud F1	Non-Fraud Recall	Fraud Recall	AUC-ROC
Logistic Regression	0.76	0.13	0.63	0.52	0.6315
XGBoost	0.81	0.12	0.71	0.31	0.6006
Random Forest	0.75	0.12	0.62	0.10	0.6215
Neural Network	0.95	0.07	0.95	0.07	0.5084

Based on the results, it can be seen that the models perform well in predicting non-fraudulent transactions, but their detection levels for fraudulent transactions need to be improved.

In terms of the Non-Fraud F1 score, the Neural Network model performs the best, reaching 0.95. This indicates that the model achieves a high balance between precision and recall when identifying non-fraudulent transactions. The Non-Fraud F1 score of the XGBoost model is 0.81, placing it at an intermediate level. The Non-Fraud F1 scores of Logistic Regression and Random Forest are 0.76 and 0.75, respectively, lagging behind those of the Neural Network and XGBoost models.

From the perspective of the Non-Fraud Recall metric, the Neural Network model has the highest recall rate of 0.95, which means that it can effectively identify most non-fraudulent transactions. In contrast, the Non-Fraud Recall values of Logistic Regression, XGBoost, and Random Forest are 0.63, 0.71, and 0.62, respectively, which are relatively lower.

In terms of the Fraud F1 score, all models perform poorly and require further improvement. The Fraud F1 scores of Logistic Regression, XGBoost, and Random Forest are in the range of 0.12 - 0.13, while the Fraud F1 score of the Neural Network model is even lower, at only 0.07. This suggests that in bank transaction data, the identification of fraud transaction samples is challenging, and all models face difficulties in balancing precision and recall.

For the Fraud Recall metric, the Logistic Regression model performs the best, reaching 0.52. This means that 52% of the transactions judged as fraud by the model are indeed fraudulent transactions. However, the Fraud Recall values of XGBoost, Random Forest and Neural Network are 0.31, 0.10 and 0.07, respectively, all at low levels, showing that these models have significant room for improvement in detecting fraudulent transactions.

Regarding the AUC-ROC metric, the Logistic Regression model achieves a value of 0.6315, performing the best. The AUC-ROC of Random Forest is 0.6215, close to that of Logistic Regression, indicating a relatively good performance. The AUC-ROC of XGBoost is 0.6006, performing less well. The near-random AUC-ROC (0.5084) of the neural network suggests that deep learning models may fail to generalize without sufficient fraud samples. Despite this, it excels in non-fraudulent transaction recall. Overall, the AUC-ROC values of all models are not high, further illustrating the complexity of bank transaction fraud detection tasks.

Based on a comprehensive analysis of the above metrics, the Neural Network model has significant advantages in identifying non-fraudulent transactions, especially in terms of recall, and can effectively avoid misjudging non-fraudulent transactions as fraud. However, all models have certain limitations in terms of the Fraud F1 score and AUC-ROC. This may be caused by the relatively small number of fraud transaction samples in bank transaction data and the complex and diverse fraud patterns, which makes it difficult for models to accurately capture fraud characteristics. Future research could consider incorporating more data features, adopting more advanced model fusion strategies, or conducting more in-depth feature engineering to improve the performance of bank transaction fraud detection models.

6. Conclusions and recommendations

6.1. Conclusions

This study addresses the issue of bank transaction fraud detection by constructing four models—logistic regression, XGBoost, random forest and neural network—and conducting a comprehensive performance evaluation. Through comprehensive evaluation of model performance across key metrics - including class-specific F1-scores (for both legitimate and fraudulent transactions), recall rates, and AUC-ROC values, the following conclusions are drawn:

6.1.1. Performance in non-fraudulent transaction identification

All models demonstrate certain capabilities in identifying non-fraudulent transactions, with the neural network model standing out as the most prominent. It achieves high values in both the F1-score and recall for non-fraudulent transactions, showing that the model effectively balances precision and recall, accurately identifying the majority of non-fraudulent transactions and avoiding misclassifying legitimate transactions as fraudulent. This helps maintain normal banking operations and customer satisfaction.

6.1.2. Limitations in fraudulent transaction detection

All models perform poorly in detecting fraudulent transactions. The F1-scores for fraudulent transactions are generally low, suggesting significant challenges in balancing precision and recall, making it difficult to accurately identify fraudulent transactions. The overall low AUC-ROC values across all models further underscore the complexity of bank transaction fraud detection tasks, where the scarcity of fraudulent samples and the diversity of fraud patterns pose substantial difficulties for model identification.

6.1.3. Different models have distinct advantages

Each model excels in different aspects, and no single model can be definitively declared the best. The logistic regression model performs best in terms of the AUC-ROC metric, fraud recall and F1-score, indicating its relative advantage in comprehensively distinguishing between fraudulent and non-fraudulent transactions. The XGBoost model exhibits intermediate performance across all metrics, demonstrating stable performance. The neural network model shows clear superiority in identifying non-fraudulent transactions. Therefore, in practical applications, the choice of model should be tailored to specific requirements.

6.2. Recommendations

Drawing upon empirical analysis, this study proposes several strategic enhancements to optimize fraud detection efficacy in banking transactions:

6.2.1. Select models based on practical application scenarios

When choosing a fraud detection model, banks should thoroughly consider real-world application scenarios and business requirements, weighing the cost implications of false positives and false negatives. If the business has low tolerance for false positives—i.e., misclassifying legitimate transactions as fraudulent incurs high costs (e.g., deteriorating customer experience, triggering complaints)—models with superior performance in identifying non-fraudulent transactions, such as neural networks or XGBoost, should be prioritized. In scenarios prioritizing maximal fraud detection where minimizing false negatives is critical (accepting higher false positives to reduce missed fraud cases), high-recall models like logistic regression become preferable despite potential impacts on legitimate customer experience.

6.2.2. Adopt model ensemble strategies

Given that different models exhibit distinct strengths and weaknesses, a practical approach is to employ ensemble methods that combine predictions from multiple models. For instance, the outputs of the four models used in this research can be aggregated through weighted averaging, voting mechanisms, or other fusion techniques. This strategy leverages the complementary advantages of individual models, thereby improving overall predictive performance.

6.2.3. Optimize data and feature engineering

The scarcity of fraudulent transaction samples in bank transaction data is a critical factor limiting model performance. To address this, banks should explore advanced techniques to augment the

representation of minority-class samples, such as sophisticated oversampling methods (e.g., SMOTE [17], ADASYN) or the collection of additional historical fraud data. Furthermore, in-depth feature engineering—extracting latent patterns and refining discriminative features—can significantly improve a model’s ability to detect fraudulent transactions.

6.2.4. Implement continuous monitoring and model updates

With the continuous advancement of fraudulent techniques and ongoing transformations in financial regulatory frameworks, fraud detection systems require corresponding upgrades to maintain effectiveness. Banks should establish robust monitoring mechanisms to periodically evaluate model performance. If performance degradation is detected, models should be promptly retrained or optimized to address emerging fraud patterns and business needs. Additionally, integrating expert knowledge and manual review processes can provide secondary validation of model predictions, further increasing the accuracy and reliability of fraud detection.

These recommendations aim to provide actionable insights for banks seeking to enhance their fraud detection systems while balancing operational efficiency and risk mitigation.

References

- [1] KPMG. (2016). The pulse of fintech. London: KPMG. Retrieved from <https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2017/02/pulse-of-fintech-q4-2016.pdf>
- [2] Financial Stability Board. (2017). Financial stability implications from FinTech: Supervisory and regulatory issues that merit authorities’ attention. Retrieved from <http://www.fsb.org/wp-content/uploads/R270617.pdf>
- [3] Peters, G. W., & Panayi, E. (2016). Understanding modern banking ledgers through blockchain technologies: Future of transaction processing and smart contracts on the internet of money. In *Banking Beyond Banks and Money* (pp. 239–278). Springer.
- [4] Basel Committee on Banking Supervision (BCBS). (2017). Sound practices: Implications of fintech developments for banks and bank supervisors. Basel, Switzerland: Bank for International Settlements. Retrieved from <https://www.bis.org/bcbs/publ/d415.htm>
- [5] Abubakar, L., & Handayani, T. (2022). Strengthening regulations: Efforts to accelerate banking's digital transformation in the digital economy era. *Legal Issues*, 51(3), 259–270. <https://doi.org/10.14710/mmh.51.3.2022.259-270>
- [6] Qothrunnada, NA, Iswanto, J., S, DF, Hendratri, BG, & Subekan, S. (2023). Digital Transformation of Sharia Financial Institutions : Opportunities and Implementation in the Industrial Era 4.0. *Indonesian Journal of Humanities and Social Sciences*, 4(3), 741–756. <https://doi.org/10.33367/ijhass.v4i3.4585>
- [7] Carminati, M., Toberto, C., Maggi, F., Epifani, I., & Zanero, S. (2015). Bank Sealer: A decision support system for online banking fraud analysis and investigation. *Computers and Security*, 53, 175–186. <https://doi.org/10.1016/j.cose.2015.04.002>
- [8] Jansen, J., & Leukfeldt, R. (2015). How people help fraudsters steal their money: An analysis of 600 online banking fraud cases. Workshop on Socio-Technical Aspects in Security and Trust (STAST). Electronic ISSN: 2325-1697.
- [9] Bhasin, M. L. (2015). Data mining: A competitive tool in the banking and retail industries. *The Chartered Accountant*, October, 588–594.
- [10] Van der Meulen, N. S. (2013). You’ve been warned: Consumer liability in internet banking fraud. *Computer and Security Review*, 32, 713–718.
- [11] Andrews, L., & Boyle, M. V. (2008). Consumers’ accounts of perceived risk online and the influence of communication sources. *Qualitative Market Research: An International Journal*, 11(1), 59–75. <https://doi.org/10.1108/13522750810845559>
- [12] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). San Francisco, CA, USA. <https://doi.org/10.1145/2939672.2939785>
- [13] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- [14] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>

- [15] Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63.
- [16] Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861–874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- [17] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357. <https://doi.org/10.1613/jair.953>