Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 47 No. 02
February 2026

A Hybrid Machine Learning–Driven Fraud Risk Assessment Framework for Secure Online Course Transactions

[1]Sandhiya R, [2]Ms. Dhanya Girish

[1]Department of M.Tech Computer Science and Engineering,

Sri Krishna College of Engineering and Technology, Kuniamuthur, Coimbatore, Tamil Nadu - 641008

[2] Assistant professor, Department of Computer Science and Engineering,

Sri Krishna College of Engineering and Technology, Kuniamuthur, Coimbatore, Tamil Nadu - 641008

Abstract

The rapid expansion of online education platforms has led to a significant increase in digital course transactions, along with a rise in chargeback fraud and unauthorized payment disputes. Since online courses are delivered instantly after payment authorization and cannot be revoked, traditional post-transaction fraud detection methods are ineffective and often result in financial losses for education providers. This creates a strong need for proactive fraud prevention mechanisms that assess transaction risk before granting course access.

This project proposes a hybrid machine learning–driven fraud risk assessment framework for secure online course transactions. The system evaluates transactional, behavioral, and contextual features in real time to predict the likelihood of chargeback fraud prior to payment authorization. Logistic Regression is used for interpretable baseline probability estimation, Random Forest captures non-linear behavioral patterns, and Gradient Boosting enhances prediction accuracy by focusing on difficult-to-classify transactions. The outputs of these models are combined to generate a reliable fraud risk score.

Based on the predicted risk level, transactions are classified as low, medium, or high risk and are accordingly approved, verified, or blocked. The proposed framework aims to reduce chargeback incidents, maintain a low false positive rate, and improve transaction security while ensuring a smooth user experience on online education platforms.

Keywords: Online course transactions; chargeback fraud; hybrid machine learning; fraud risk assessment; behavioral analytics; digital payment security

I. INTRODUCTION

The increasing adoption of online education platforms has fundamentally changed the way learning content is delivered and accessed. Digital courses are now offered through web and mobile applications, enabling learners to enroll and begin studying immediately after completing online payments. While this model improves accessibility and scalability, it also places greater emphasis on the security and reliability of transaction processing systems. As online education continues to expand globally, protecting digital course transactions from fraudulent activities has become a critical concern for education service providers.

One of the major challenges faced by online course platforms is chargeback fraud, where completed transactions are later disputed through payment service providers. In the case of digital learning products, course access is typically granted immediately after payment authorization and cannot be withdrawn once delivered. As a result, post-transaction fraud detection mechanisms are largely ineffective and often lead to direct financial losses, penalty fees, and reputational damage. These limitations highlight the need for fraud

prevention strategies that operate before payment authorization rather than after disputes occur.

Many existing fraud detection systems rely on static rule-based checks or single-model classification techniques. Such approaches are limited in their ability to capture complex fraud patterns that involve behavioral anomalies, device inconsistencies, or coordinated misuse across multiple accounts. In addition, reactive fraud management strategies fail to adapt to evolving attack methods, reducing their effectiveness in dynamic online environments. Although transactional and behavioral data are readily available during the payment process, they are often underutilized in real-time decision making.

Recent advances in machine learning have demonstrated strong potential for improving fraud detection accuracy by identifying hidden patterns within large-scale transaction data. However, much of the existing research focuses on traditional e-commerce or financial transactions and does not adequately address the unique characteristics of online education platforms. In particular, limited attention has been given to proactive, pre-transaction fraud risk assessment for digital course transactions.

To address these challenges, a novel hybrid machine learning–based fraud risk evaluation framework is proposed for online course transactions. Our method combines different supervised learning models and processes transaction, behavioral, and contextual characteristics on the fly in order to estimate fraud risk prior to payment authorization. By helping to inform risk-aware transaction decisions, the framework seeks to minimize chargeback incidents and low false positive rates, along with elevating transaction security while keeping user experience smooth.

II. LITERATURE REVIEW

Fraud detection in electronic payment systems have been drawing increasing attention because of the explosive expansion of mobile payments, e-commerce platforms and online services. Early machine learning studies aimed at dealing with the intrinsic class imbalance in fraud datasets. Botchey et al. [1] suggested that sampling methods who are widely used to tackle imbalanced data distribution even achieved the better accuracy in fraud prediction likelihood makes significant contribution on the enhancement of fraud prediction performance in mobile money transactions. Similar findings were observed in more general e-commerce settings that both fraud detection and prevention actions should be taken simultaneously to balance the financial risk with the operational impact [2].

Tree-based and ensemble learning approaches have been extensively studied to learn complex fraud patterns. Hajek et al. [3], is a XGBoost-based method for the detection of mobile payment fraud, where it was demonstrated that the application of boosting performs better than traditional classifiers. Afriyie et al. [6] also showed that optimized machine learning methods enhance the detection performance of credit card transactions, since they manage to properly model non-linearity. Vanini et al. [7] furthered these views by bringing into focus the shift from anomaly detection to full-fledged fraud risk management frameworks.

Chargeback fraud has been treated as a separate issue because it only surfaces at a later date and results in fiscal loss. Günther [4] used artifi cial neural networks to uncover chargeback patterns in transactional records, and Wei et al. [8] analyzed deep learning models in chargeback fraud detection for online gaming sites, verifying the importance of behavioral and temporal features. These findings underscore the shortcomings of post-transaction fraud prevention, especially for digital items where content delivery is non-retractable.

Recent literature emphasizes the importance of adaptive fraud detection systems which are able to manage new types of fraud. Agrahari and Singh [5] surveyed concept drift detection methods, and concluded that static fraud models deteriorate in evolving contexts. On the other hand, graph based methods have been presented for tackling orchestrated and structured fraud. Yu et al. [9] proposed a community-based fraud detection network for e-commerce websites, and Tian et al. [10] used adaptive graph neural networks to model the relationships between fraudulent entities.

Behavioral biometrics has additionally become a critical signal for fraud prevention. Ray-Dowling et al. [11] investigated static mobile behavioural biometrics and identified that these are indeed effective in improving authentication, as well as fraud detection, without escalating user friction. Moreover, some recent studies have also developed explainable models to enhance transparency and trust in fraud detection. Chagahi et al. [12] introduced attention-based explainable models for transparent fraud decisions.

Recent ensemble-based methods have also shown good performance on payment fraud detection. Zeng et al. [13] developed an ensemble learning approach by combining neural networks for e-commerce payment fraud detection and Tayebi and El Kafhali [14] used XGBoost together with Bayesian optimization to improve the detection accuracy. Compagnino et al. [15] did a complete survey on machine learning approaches for fraud detection and discussed that domain specific, proactive and risk-aware systems are required.

However, research remains scant on the proactive fraud risk evaluation in relation to online course transactions. Existing research focus on general application in e-commerce, or in post-transaction fraud detection is not enough for online education platforms, where the access right is delivered promptly after the money transfer has been confirmed. This gap inspires a model of hybrid, behavior-conscious fraud risk assessment for online course transactions.

III. PROBLEM STATEMENT

Many Web-based models of course delivery are made possible with instantaneous digital course transactions where content is accessible after payment authorization. This model, in turn, opens platforms to significant chargeback fraud and unauthorized disputes – as the moment digital content is made available there is no way back. The existing fraud containment mechanisms are, for the most part, reactive ones that catch fraudulent transactions only after a charge back has been filed against them at a revenue lost and cost of penalty fees and limited recoup. And such mechanisms are frequently enforced through static deterministic rules-based checks, or independent-model classifiers which do not adapt to emerging fraud patterns and complex user behavior. Despite that online courses platforms have rich transactional and behavior data, few methods exist to assess pre-transaction fraud risk in order to assist real-time authorization decisions.

IV. RESEARCH GAP

While the literature on fraud detection in e-commerce, mobile payments, and financial transaction systems is vast (see Cerveira et al. [2] and references therein), there are serious gaps when these techniques are adopted to online course platforms. The majority of previous work adopts after-transaction techniques for general payment or fraud detection, which are able to detect there is something wrong only consumers report that the transaction as unauthorized. These methods do not work for digital educational products which give access to the course as soon as payment authorization occurs and cannot be undone.

Moreover, most of the current models are driven by static rule-based approaches or single-model classifiers that have a limited exposure to dynamic and changing fraud patterns. Such systems do not generally account for anomalous behavior, inconsistencies in device and location reports, or orchestrated abuse among a number of user accounts. Furthermore, the existing technique(s) do not often combine transactional, behavioral and contextual risk indicators using a single methodology which operates in real time.

Another important missing point is in reality-related decisions based on fraud predictions that allow the insertion of 7 Transaction. Most of the previous works consider fraud detection as a standard binary classification task and do not provide adaptive responses, like selective verification or controlled transaction blocking. Consequently, such systems result in a higher number of false alerts, or high risk transactions are not alerted.

These constraints demonstrate the requirement of a preventive pre-transaction fraud risk assessment scheme tailored for online course transactions. The task of addressing these gaps is facilitated by the use of a hybrid machine learning capabilities that

uses various models along with Behavior Analysis and risk based Decision Logic to enhance Fraud Detection Accuracy, all while improving user experience.

V. SYSTEM ARCHITECTURE

The proposed Hybrid Machine Learning–Based Fraud Risk Assessment System is developed as a modular layered architecture to facilitate on-the-fly, pre-transaction fraud assessment for online course platforms. Specially, data capture and feature processing, intelligent risk prediction decision control and online learning are collaborated closely in the structure of the architecture to achieve scalability, reliability as for practical deployment.

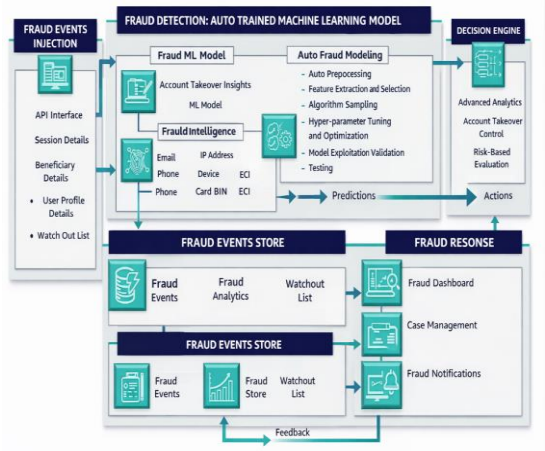

Fig.1. System Architecture of the Hybrid Machine Learning-Driven Fraud Detection

At the back end, User Interaction Layer controls user registration, authentication, course selection and payment opening. Transaction-specific and session-level details, like transaction amount, device characteristics, location data and behavior history, are collected discreetly during checkout. This tier also offers administrative interfaces for tracking transactions, fraud warnings, and system performance statistics.

The **Data Processing and Feature Engineering Layer** does process and cleans raw input data, normalize them and transform them into a structured feature representation for machine learning inference. The transactional aspects cover account age, payment amount and frequency of the transaction and behavioral include number of devices used, location changes and velocity of transaction. Contextual cues from payment gateways are also included. This preprocessing step guarantees the consistence of data and maintains reliable analysis afterwards.

The **Intelligence Layer** constitutes the main analytical part of the system, where a combined ensemble of Multi-Label supervised machine learning models is implemented. Logistic Regression is used to generate human-interpretable baseline estimates of the probability of fraud. Random Forest models describe complex non-linear relationships between transactional and behavioural attributes, while Gradient Boosting incrementally improves on predictive accuracy by concentrating upon misclassified transactions. The scores of these models are then aggregated through the stack ensembling method to form a final fraud risk score for every transaction.

The **Decision and Control Layer** converts "the predicted fraud risk score" to actionable decisions. Transactions are classified into low-risk, medium-risk and high-risk categories by means of predetermined risk threshold rules and rule-assisted logic. Orders with a low risk are authorized automatically, those with medium risk will be referred to an additional verification stage and orders at high risk will be blocked before the payment authorization. This risk-based approach to decision making maximizes fraud detection and minimizes disruption for genuine customers.

Last but not least, the **Feedback and Learning Layer** logs the results of transactions, with good payments/verifications along with confirmed chargebacks. This data is kept in a centralized service and occasionally utilized to retrain the machine learning models.

VI. METHODOLOGY

The Hybrid Machine Learning–Based Fraud Risk Assessment System presented in this paper consists of a well-defined approach for assessing the risk of fraud occurring in transactions opened by customers to buy online courses. The framework incorporates real-time data observation, characteristic pre-processing, hybrid machine learning prediction, predicting strategy based on risk aversion, and the feedback

mechanism to achieve successful large-scale fraud prevention.

1. Data Collection and Feature Identification

The process starts when a user clicks to purchase an online course within the platform. The collected data concerns the transaction, i.e., purchase amount, method of payment and time stamp as well as user---related and contextual information such as account age, type of device used to conduct the transaction, IP addresses of user's machine(s), physical location(s) of a user (s), the time span between two transactions --- a session duration --- and speed at which occurs a transaction. Payment gateway risk signals, and previous dispute history are also considered where available. These multiple data sources provide a 360 degree view of the transactions and potential signs of fraud.

2. Data Preprocessing and Feature Engineering

Raw transaction data is usually noisy, missing, and inconsistent. Thus, the data is cleaned and normalized through preprocessing before inference of the model. The missing values are addressed with methods of statistical imputation, whereas, the categorical variables, like type of device and mode of payment, are coded into numbers. The continuous features are normalized so that there is equal scaling of inputs. Deviations of the normal user behavior are reflected in behavioral features derived, which include the frequency of device changes, a change in the location, and the abnormal velocity of transaction.

3. Hybrid Machine Learning–Based Fraud Prediction

At its heart is a hybrid ensemble learning model which uses multiple supervised machine learning models to estimate the risk of fraud. For the latter a Logistic Regression model is used for pure probabilistic reasons. Random Forest is used here to model nonlinear relationships and complex interactions of transactional and behavioral attributes. Gradient Boosting even improved it by concentrating on examples in the misclassified case during iterative learning. Finally, the outputs of the individual models are combined to generate a unified fraud probability score.

4. Behavioral and Contextual Risk Enhancement

For enhanced prediction confidence, fraud probability score is enriched with the behavioral anomaly indicators and fraudulent network signals. Any irregularities like above, such as sudden increase in device usage, login from an unfamiliar location or strange transaction frequency drive the account being flagged for compromise or grievance. Characteristics that are shared such as IP addresses, device id and payment ids are checked to spot for any indirect links with fraudster transactions happened in the past.

5. Risk Stratification and Decision Logic

After that, the final fraud probability score is mapped into pre-determined risk boundaries for transactions' classifications; low-risk, medium-risk and high-risk. A rule-based decision engine applies the rules of VoC entities while making recommendations to handle outlier cases, which machine learning predictions cannot accurately capture. Low-risk transactions are accepted with no delayed confirmation, medium-risk transactions prompt for extra authorization, while high-risk transactions are refused before payment. This risk-based approach sacrifices security for usability.

6. Transaction Logging and Feedback Learning

A central fraud event store 312 stores all transaction decisions and actions. The captured training dataset, fraud instances and verified results are re-used occasionally to refresh the training process. Models are refreshed systematically to accommodate new fraud patterns and changing user behavior. This constant feedback loop ensures that the system remains performant.

VII. ALGORITHMS

In this section, the algorithms applied in the proposed Hybrid Machine Learning-Based Fraud Risk Assessment System and the mathematical basis are provided. It is also targeted at estimating a credible degree of fraud risk related to any online course transaction that utilizes complementary learning models and decision logic.

1. Logistic Regression for Baseline Fraud Probability

The Logistic Regression estimates the probability of a fraud of each transaction that the baseline is. It

gives results that are interpretable and their exploitation as a stable reference model.

Let, $x = (x_1, x_2, \ldots, x_n)$

The fraud probability is computed as:

$P\_LR = 1 / (1 + e^{\wedge} (- (\beta_o + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n)))$

Where β_o is the bias term and $\beta_1 \ldots \beta_n$ are the learnt model coefficients.

This model represents linear correlation of transaction characteristics and fraud risk.

2. Random Forest for Non-Linear Fraud Pattern Detection

Random Forest is an ensemble learning algorithm that consists in building a number of decision trees at training time and averaging them out at test time.

This probability of fraud for T decision trees is then:

$P\_RF = (1/T) \times \Sigma h\_t(x)$

for t = 1 to T

Here, h_t(x) is the prediction of t-th tree.

Random Forest is also particularly important in modeling the complicated non-linear interactions in behavioral features, such as transaction tempo or switching device and departing location.

3. Gradient Boosting for Error Minimization

Gradient Boosting enhances predictive accuracy by minimizing the loss function in an iterative way, and compensating errors made by previous models. The improved form of the prediction function is given by handler.

$FM(x) = \Sigma \gamma m fm(x)$

for m = 1 to M

Where, f m (x) is the m-th weak learner and γ is the weight of the m-th weak learner.

The likelihood of the fraud is computed by the sigmoid function:

$P\_GB = 1 / (1 + e^{\wedge} (-F\_M(x)))$

The algorithm increases its accuracy by paying attention to transactions which are hard to categorize.

4. Hybrid Ensemble Fraud Probability

Results of the three machine learning models are then synthesized to create an ultimate hybrid score of probability of fraud.

$P\_Hybrid = w_1 P\_LR + w_2 P\_RF + w_3 P\_GB$

Subject to: $w_1 + w_2 + w_3 = 1$

The ensemble methodology eliminates bias and enhances strength through the strength of single models.

5. Behavioral Anomaly Scoring

Behavioral anomaly detection: This is the method used to detect abnormal behavior of users and measures the deviation of the behavior against past norms.

The anomaly score is calculated as:

$A(x) = \Sigma | (x\_i - \mu\_i)/\sigma\_i |$

for i = 1 to k

Where μ_i and σ_i are the mean and standard deviation of the behavioral feature x_i.

A score of higher anomalies shows suspicious behavior.

6. Shared Fraud Network Risk Score

A shared risk score is calculated in order to identify coordinated fraud using shared entities, e.g. IP addresses, devices, or payment identifiers.

$R\_net = (1/N) \times \Sigma F\_j$ for j = 1 to N

Where F_j indicates whether a linked entity has a known fraud history.

7: Risk Thresholding and Decision Logic

The final fraud risk score is computed as:

$R\_final = \alpha P\_Hybrid + \beta A(x) + \gamma R\_net$

Subject to: $\alpha + \beta + \gamma = 1$

Based on R_final, transactions are classified as:

- If $R\_final < \theta_1$ → **Approve Transaction**
- If $\theta_1 \leq R\_final < \theta_2$ → **Additional Verification**
- If $R\_final \geq \theta_2$ → **Block Transaction**

This decision logic enables risk-aware transaction handling.

VIII. EXPERIMENTAL RESULTS

The performance of the Hybrid Machine Learning–Based Fraud Risk Assessment System was verified through the simulated online course transaction datasets. The evaluation emphasizes on the KPIs of real-world FPF systems, namely: fraud detection accuracy, false positive rate, chargeback reduction and overall fitness score.

1. Fraud Detection Accuracy

The detection accuracy is a measure of how well the system is able to detect between fraudulent and non-fraudulent transactions.

Observed Accuracy: 88.5%

Such high accuracy suggests that the hybrid ensemble model could capture linear and non-linear patterns of fraud in fraudulent online course transactions. The fusion of Logistic Regression, Random Forest, and Gradient Boosting not only enhances the stability of the prediction but also reduces misclassification compared with individual model.

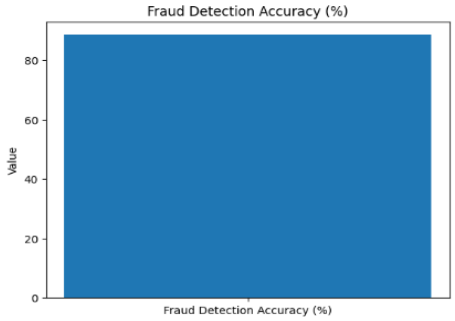

Fig.2. Fraud detection accuracy of the proposed hybrid machine learning framework.

2. False Positive Rate

The false positive (FP) rate is the percentage of good transactions that are classified as fraudulent.

Observed False Positive Rate: 0.8%

For online education applications, a considerably low false positive rate is necessary so as not to disrupt user experience or undermine the trust of the platform due to too many transaction blocks. Results show that the risk-aware decision engine and the ensemble learning mechanism can greatly decrease unnecessary transaction rejections.

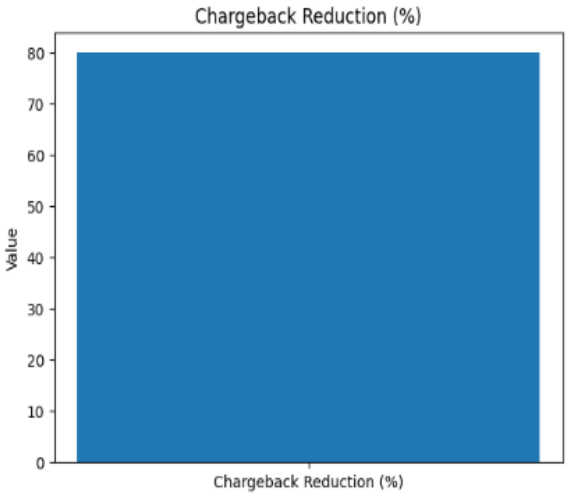

Fig.3. False positive rate of the proposed fraud risk assessment system.

3. Chargeback Reduction

To determine the performance of the system in mitigating post-transaction disputes and loss, chargeback reduction should be measured.

Observed Chargeback Reduction: approximately 80%

A significant decrease in chargebacks proves that pre-transaction fraud risk analysis is effective. The system protects against or disputes high-risk transactions prior to payment authorization to reduce exposure for unrecoverable loss due to digital delivery.

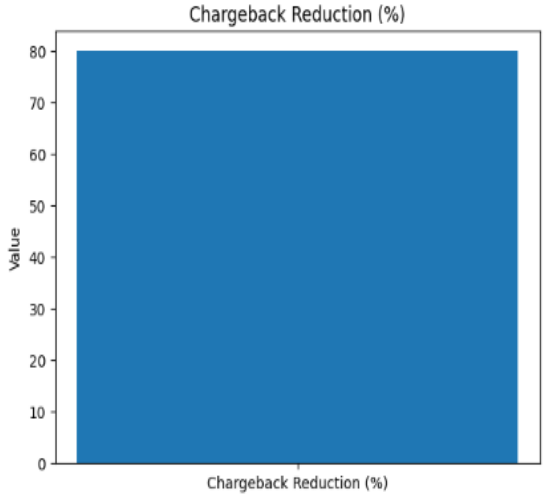

Fig.4. Chargeback reduction achieved by the proposed fraud risk assessment system.

4. Overall Fitness Score

The overall fitness score is a composite measure, which indicates combined benefits (enhanced fraud detection, reduced false positives and prevented chargebacks).

Monitored Overall Fitness Score: 83.3.

This rating indicates that the system is functioning at an equal balance implying that the suggested system can reinforce security without negatively affecting the efficiency of transactions or the user-friendliness.

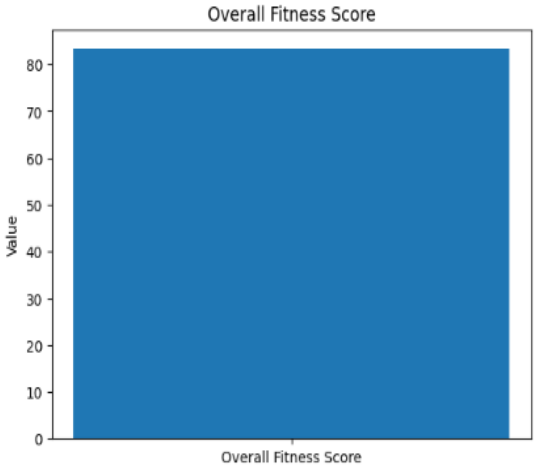

Fig.5. Overall fitness score of the proposed hybrid fraud detection framework.

IX. CONCLUSION

The present study introduced a hybrid machine learning-based fraud risk assessment model that seeks to boost security of online course transactions with proactive, pre-transactional fraud identification. Combining Logistic Regression, Random Forest and Gradient Boosting models allowed to effectively capture both linear and complex non-linear fraud trends existing in digital learning environments. The addition of behavior anomaly analysis and fraud indicator sharing enhances the system's capability to detect higher risk transactions which may not be detectable by transactional features alone.

Experimental test results prove the efficiency of the proposed framework with a high fraud identification accuracy and a very small false positive rate, significantly reducing chargeback cases. The risk-aware decision mechanism facilitates dynamically handling transactions by letting qualifying low-risk transactions go through smoothly and ordering potentially high-risk cases to be verified or rejected if necessary. The modular structure and data\-driven learning scheme enables the framework scalable and flexible to various fraud patterns over time. In general, the proposed method offers a simple and practical tool for enhancing security of transactions and trust in online education systems.

REFERENCES

[1] F. E. Botchey, Z.Qin, K. Hughes- Lartey, and K. E.(Fantom), \Predicting fraud in mobile money transactions using machine learning: The impact of sampling techniques on imbalanced datasets," Informatica, vol. 45, nº. 7, pags." 45–56", 2021, doi: 10.31449/inf. v45i7. 3179.

[2] V. F. Rodrigues, "Fraud detection and preventation on e-commerce," Electronic Commerce Research and Applications, vol. 53, Article 101207, 2022; doi: 10.1016/j.elerap. 2022. 101207.

[3] P. Hájek, M. Henriques, and R. Pereira, "An XGBoost-based framework for fraud detection in mobile payment systems," Appl Soft Comput, vol. 126, Article 109274, 2022.

[4] T. Günther, "Detecting chargebacks in transaction data with artificial neural networks," M.Sc. Thesis, University of Leipzig, Leipzig (Germany), 2022.

[5] S. Agrahari and A. K. Singh, "A literature review on concept drift detection in data stream mining," Journal of King Saud University – Computer and Information Sciences, vol. 34, no. 10, Part B, pp. 9523–9540, 2022., doi: 10.1016/ j.jksuci. 2021. 11. 006.

[6] J. K. Afriyie, R. O. D. Osei, and E. A. Baah, "An optimized machine learning technique for fraud detection and prediction in credit card transactions," Decision Analytics Journal, vol.\ 4(1), pp.\ 11–28, 2017. 6, Art No 100163, 2023, doi:10.1016/j.dajour. 2023. 100163.

[7] P. Vanini, S. Rossi, E. Zvizdic and T. Domenig, "Online payment fraud: From anomaly detection to risk management," Financial Innovation, vol. 9, 66 Article (Published on

the World Wide Web) doi: 10.1186/s40854-023-00470-w [PDF] [Full Source] Journal of Analytical Science and Technology.

[8] Y.-C. Wei, Y.-X. Lai, and M.-E. Wu, "Performance comparison of deep learning models for chargeback fraud detection in online games," Cluster Computing. 26, no. 2, pp. 927–943, 2023, doi: 10.1007/s10586-022-03674-4.

[9] J. Yu, Y. Sun, J. Tang, and H. Liu, "Groupbased fraud detection network over e-commerce platforms," in The 29th ACM International Conference on Information and Knowledge Management (KDD), 2023, pp. 2786–2796.

[10] Y. Tian, K. Daxvongkham, and T.-K. Kim, "Transaction fraud detection using adaptive graph neural network," in IEEE Access, vol. 11, pp. 74231–74243, 2023.

[11] A. Ray-Dowling, D. Hou, and S. Schuckers, "Stationary mobile behavioral biometrics: a survey," Computers & Security, vol. 128, Article 103184, 2023, doi: 10.1016/j.cose. 2023. 103184.

[12] M. H. Chagahi, Subject T Ali, and A. Shamsi, "Explainable artificial intelligence for fraud detection: An attention-based approach," Expert Systems with Applications, vol. 242, Article 122761, 2024.

[13] Q. Zeng, Y. Li, H. Zhang and X. Chen, "NNEnsLeG: An ensemble learning approach for e-commerce payment fraud detection," Knowl. 298, Article 110064, 2025.

[14] M. Tayebi and S. El Kafhali, "A novel system with XGBoost and Bayesian optimization for credit card fraud detection," Cyber Security and Applications, vol. 3, Article 100093, 2025, doi: 10.1016/j.csa. 2025. 100093.

[15] A. A. Compagnino, Merosina L. Polignano G., and G. Semeraro, "An introduction to machine learning methods for fraud detection," Applied Sciences,. 15, n o 21, Article 11787, 2025, doi: 10.3390/app152111787.