DIGIVENTURE
Journal of Applied Digital Business Research

Vol. 1, No. 3, (2026), pp. 555-567
e-ISSN: 3123-5212
Website: https://journal.unesa.ac.id/digiventure

Comparative Analysis of XGBoost and TabNet Approaches For E-commerce Fraud Detection

Wildan Arsthasya Putra Budi[1] and Riska Dhenabayu[2]

[1,2] Digital Business, Universitas Negeri Surabaya, Indonesia

*) Corresponding Author (Email: wildanarsthasya2202@gmail.com)

ABSTRACT

The rapid growth of e-commerce has increased transaction convenience but has also intensified the risk of fraudulent activities, particularly in highly imbalanced transaction datasets. Effective fraud detection systems are therefore essential to minimize financial losses and maintain consumer trust. This study aims to compare the performance of an ensemble learning approach, Extreme Gradient Boosting (XGBoost), and a deep neural network model designed for tabular data, TabNet, in detecting fraud in e-commerce transactions. The research employs a quantitative comparative method using a publicly available e-commerce fraud detection dataset consisting of 299,695 transaction records, with only 2.21% fraudulent transactions and 97.79% non-fraudulent transactions. Model performance is evaluated using accuracy, precision, recall, and F1-score, supported by confusion matrix and precision-recall curve analysis. The results show that both models achieve high accuracy of approximately 99%; however, accuracy alone is insufficient due to the highly imbalanced nature of the dataset. TabNet demonstrates higher precision (92%), indicating fewer false positives, while XGBoost achieves higher recall (72%) and F1-score (80%), reflecting a better balance between detecting fraudulent transactions and minimizing misclassification. These findings indicate that XGBoost provides more reliable overall performance for e-commerce fraud detection under imbalanced data conditions. This study provides empirical insights into the comparison of ensemble learning and deep learning approaches for fraud detection applications in e-commerce environments..

Keywords: *Machine learning, Fraud, E-commerce*

I. INTRODUCTION

In recent years, the development of e-commerce has grown rapidly. This growth has been influenced by various factors, one of which is the convenience offered to consumers in conducting online transactions. Furthermore, the advancement of internet technologies has stimulated the expansion of e-commerce into mobile commerce (m-commerce), enabling users to shop directly through their mobile devices (Roxana Badircea, Nicoleta Mihaela Doran, Alina Georgiana Manta, & Jenica Popescu, 2021). Global e commerce sales have shown consistent year-over-year growth. In 2024, e-commerce sales reached USD 6.634 trillion, and projections indicate a further increase to approximately USD 7.956 trillion by 2026–2027 (Ethan Cramer-Flood, 2025).

Despite its economic benefits, e-commerce is highly vulnerable to fraudulent activities. Common forms of e-commerce fraud include credit card fraud, chargeback fraud, and triangulation fraud, all of which result in significant financial losses for both consumers and businesses (Haq, 2022; Paige Tester, 2022) Global online fraud is projected to grow at a compound annual growth rate (CAGR) of 18.5%, reaching USD 5.7 trillion by 2025, with phishing and fraudulent platforms accounting for nearly half of all incidents (Dodiya, Dudhat, & Vidani, 2024). From a product category perspective, fraudsters tend to focus on items with high value and strong resale potential. This includes collectible goods, which account for up to 106%, as well as luxury products, reaching approximately 104% (Opena, 2025)

Previously, fraud detection systems were conducted manually using traditional rule-based approaches. However, rule-based models have limitation rule-based systems rely on predefined "if–then" rules derived from expert knowledge. While they are transparent and interpretable, they face major limitations, including poor adaptability to new fraud patterns, rule complexity, and difficulty in scaling as the number of rules increases (Faisal, 1990) As fraud methods continue to evolve, traditional models struggle to adapt to the available transaction data. Moreover, when it comes to fraud, there is usually a significant imbalance between fraudulent and non-fraudulent transactions (Hassan & Kareem, 2025)

Machine learning has become a widely adopted approach for fraud detection due to its ability to learn patterns from large-scale transaction data. However, conventional machine learning models often struggle with issues such as class imbalance and overfitting, which can negatively impact detection performance (Islam, Haque, & Rezaul Karim, 2024) To address these limitations, ensemble learning techniques have been introduced to reduce bias and variance by combining multiple base learners, resulting in improved robustness and accuracy (Mishra, Tripathi, Chaurasia, & Chaurasia, 2025) Among ensemble methods, XGBoost has demonstrated strong performance in fraud detection tasks due to its scalability, efficient handling of sparse data, and robustness to class imbalance (Chen & Carlos, 2016)

Previous research has shown that XGBoost outperforms Logistic Regression and LightGBM in distinguishing between legitimate and fraudulent transactions (Wibowo, Hartono, & Lusiana, 2024) Given the challenge of highly imbalanced fraud datasets, where fraudulent instances are significantly outnumbered by legitimate ones (Syahbani, Firdaus, & Musodo, 2025), XGBoost represents an appropriate choice. As an ensemble learning method, it is well recognized for its capability to model complex patterns and effectively handle class imbalance (Tayebi & Kafhali, 2025)

This research uses deep neural network model, TabNet for fraud detection. TabNet is a novel deep neural network architecture for tabular data that directly processes raw inputs using gradient descent optimization and employs sequential attention to select salient features at each decision step, thereby enabling efficient end-to-end learning and improved interpretability (Pfister & Arik, 2020). Previous research indicates that TabNet effectively handles imbalanced fraud datasets and achieves high classification accuracy. (Pfister & Arik, 2020) Previous research indicates that TabNet effectively handles imbalanced fraud datasets and achieves high classification accuracy (Meng, Lim, Lee, & Lim, 2023) Although XGBoost and TabNet have individually demonstrated strong performance in fraud detection, comparative studies evaluating ensemble-based and deep neural network-based approaches in e-commerce fraud detection remain limited. Most existing research focuses on single-model evaluations rather than direct performance comparisons (Duwadi & Sharma, 2024)

Therefore, this study aims to conduct a comparative analysis of XGBoost and TabNet for detecting fraud in imbalanced e-commerce transaction data, using evaluation metrics such as precision, recall, and F1-score These results will be visualized with confusion matrix, and precision-recall curve. The findings are expected to provide insights into the suitability of ensemble learning and deep learning models for real-world e-commerce fraud detection.

II. LITERATURE REVIEW

Fraud Detection in E-Commerce

The rapid growth of digital transactions has increased the importance of fraud detection systems, as companies are estimated to lose up to 5% of their annual revenue due to fraudulent. activities (Nathaniel Andrew Davis & Sophia Anne Harris, 2024). Digital fraud has evolved into various forms, including identity theft, clean fraud, affiliate fraud, triangulation fraud, and merchant fraud, each involving different mechanisms to exploit transaction systems and user information (Dr Padmalatha N a, 2020)

Traditional fraud detection methods were predominantly based on rule-based systems utilizing predefined “if–then” rules. While these approaches offer transparency and interpretability, they suffer from limited adaptability, increasing complexity, and scalability challenges as fraud patterns evolve (Faisal, 1990). To overcome these limitations, recent studies have increasingly applied machine learning techniques to improve fraud detection performance.

Machine learning models are capable of automatically learning complex patterns from transaction data and have demonstrated strong effectiveness in identifying fraudulent behavior (Jin & Zhang, 2025). Among these models, XGBoost and TabNet have shown promising results. Previous research reported that XGBoost achieved high classification accuracy in fraud detection tasks (Tayebi & Kafhali, 2025), while TabNet also demonstrated strong performance in detecting fraudulent activities due to its capability in handling tabular data (Meng et al., 2023). Overall, the application of machine learning approaches such as XGBoost and TabNet provides improved accuracy and robustness in fraud detection, contributing to the development of more intelligent and data-driven fraud prevention systems in the digital commerce environment.

Ensemble Learning

Ensemble learning is a machine learning technique that combines two or more models to achieve higher predictive performance than a single algorithm. The effectiveness of ensemble learning is influenced by two main factors: the accuracy of the base learners and the diversity among them. Model diversity refers to the variation in predictions produced by individual learners, which enables the ensemble to capture different patterns within the data (Bian, Wang, Yao, Chen, & Member, 2019). Ensemble learning methods can generally be categorized into parallel and sequential approaches. In parallel ensembles, base models are trained independently, and their predictions are aggregated to generate the final output (Mienye & Sun, 2022). Parallel ensembles are further classified into homogeneous and heterogeneous types. Homogeneous ensembles employ the same algorithm across all base learners, such as decision tree–based methods used in XGBoost and CatBoost (Bian et al., 2019)). While this approach simplifies model construction, it may limit diversity due to the use of a single algorithm (Alshdaifat, Al-hassan, & Aloqaily, 2021) In contrast, heterogeneous ensembles integrate multiple algorithms to increase model diversity (Ravichandran, Gavahi, Ponnambalam, Burtea, & Mousavi, 2021)

XGBoost

XGBoost algorithm is a decision tree-based ensemble that use the gradient boosting framework. XGBoost has an improvement of gradient boosting with it's ability to to achieve scalable and highly accurate algorithms for classification and regression (Mienye & Sun, 2022) Compared with gradient boosting that focusing on reducing bias but still having a high chance of overfitting, the XGBoost algorithm contains a regularization that prevent overfitting (Li & Chen, 2020). This is possible because XGBoost employs techniques such as limiting tree depth, adjusting the learning rate, and applying subsampling to reduce the risk of overfitting (Mienye & Sun, 2022). One of the key advantages of the XGBoost algorithm is that it requires minimal feature engineering, such as data normalization or feature scaling, as it can inherently handle such cases. Additionally, it is capable of managing missing values effectively. XGBoost can also generate feature importance scores, which help in understanding input variables and performing feature selection. Moreover, it operates faster than most machine learning algorithms, can process large datasets efficiently, and is less prone to overfitting (Nobre & Neves, 2019)

TabNet

Deep neural networks have demonstrated remarkable capabilities in handling image, text, and audio data (Sun, Chen, Wang, Liu, & Liu, 2016) However, they exhibit certain limitations when dealing with tabular data (Shwartz-Ziv & Armon, 2022) To address this issue, Pfister & Arik (2020) developed a new deep neural network architecture called TabNet.

TabNet was specifically designed for tabular data and can handle raw inputs without requiring heavy preprocessing or feature normalization. It is trained using gradient descent optimization, allowing it to be easily integrated into an end-to-end learning process. Unlike traditional models, TabNet uses a sequential attention mechanism to focus on the most important features at each decision step. This feature selection happens for each individual input, so the model can pay attention to different features depending on the data (Pfister & Arik, 2020)

TabNet performs better or at least as well as traditional tabular learning models in both classification and regression tasks. It also provides two types of interpretability: local interpretability, which explains the importance of features for individual predictions, and global interpretability, which shows how much each feature contributes to the overall model. In addition, TabNet enhances model performance and generalization across various datasets through an unsupervised pre-training technique that predicts missing features. (Pfister & Arik, 2020)

Analysis Variable and Visualization

In fraud detection research, model performance evaluation commonly relies on multiple classification metrics to provide a comprehensive assessment. Tayebi & Kafhali (2025) emphasize the use of four primary metrics accuracy, precision, recall, and F1-score to evaluate machine learning models in fraud detection tasks. These metrics are also adopted in this study to assess the performance of XGBoost and TabNet.

Accuracy measures the proportion of correctly classified transactions among all observations (Zheng, 2015) However, in highly imbalanced datasets such as fraud detection, accuracy alone may not sufficiently reflect model performance. Precision evaluates the proportion of correctly identified fraudulent transactions among all transactions predicted as fraud, while recall measures the model's ability to correctly identify actual fraudulent transactions (Martin Ward Powers, 2011). The F1-score, as the harmonic mean of precision and recall, provides a balanced evaluation by accounting for both false positives and false negatives, making it particularly suitable for fraud detection scenarios where misclassification costs are significant (Sharma, 2023)

III. METHOD

In conducting this research, this study uses quantitative research with a comparative approach. The main objective of this study is to compare the performance of the XGBoost and TabNet algorithms in ecommerce fraud detection. A comparative approach is used to examine the differences between two or more treatments of a variable, or even multiple variables at once. The purpose of this type of quantitative research method is to identify and analyze distinctions between two or more situations, events, activities, or programs (Fadilla, Ketut Ngurah Ardiawan, Eka Sari Karimuddin Abdullah, Jannah Ummul Aiman, & Hasda, 2021) In this comparative analysis, two models will be evaluated based on their performance using metrics such as, precision, recall, F1-score, PR-AUC, and the confusion matrix.

The data used in this research is secondary data introduced by Uygur (2024) named "E-commerce Transactions Fraud Detection" dataset. This dataset consists of 299,695 records of information related to fraud detection in e-commerce transactions, with each record containing features such as transaction_id, user_id, account_age_days, total_transaction_user, avg_amount_user, amount, country, bin_country, channel, merchant_category, promo_used, avs_match, cvv_result, three_ds_flag, transaction_time, shipping_distance_km, and is_fraud. This dataset was selected because it is an imbalanced dataset, making it suitable for evaluating machine learning models in real-world fraud detection scenarios. Based on previous studies by Meng et al. (2023) and Tayebi & Kafhali (2025) both XGBoost and TabNet demonstrate strong performance in handling imbalanced data. Therefore, this dataset is appropriate for comparing the capabilities of the two models in managing imbalance and predicting fraudulent transactions. This study begins with the importation of the e-commerce fraud dataset as the primary data source. After the dataset is successfully loaded, the data undergoes a pre-processing stage to ensure that it is suitable for model development and evaluation. Following the pre-processing phase, the dataset is divided into training and testing sets through a dataset splitting process.

This step is conducted to enable an objective evaluation of the models' performance on unseen data. The training process is then performed using two different approaches. The first approach applies the XGBoost algorithm, representing an ensemble learning method, while the second approach utilizes TabNet, a deep neural network architecture designed for tabular data. Both models are trained independently using the same dataset split to ensure a fair comparison. After training, both models are subjected to a model evaluation stage using predefined performance metrics. The evaluation results are subsequently analyzed in the model performance comparison phase to assess the effectiveness of each model in detecting fraudulent transactions. Finally, the outcomes of the model comparison are presented through data visualization, which facilitates clearer interpretation of the results and supports conclusions regarding the comparative performance of XGBoost and TabNet in e commerce fraud detection.

IV. RESULTS AND DISCUSSION

Table 1 Class Distribution Table

Id	Transaction Type	Total Transaction	Percentage
0	Non-Fraud	293.083	97.79%
1	Fraud	6.612	2.21%

Based on the class distribution analysis presented in Table 1, the dataset consists of 293,083 non-fraudulent transactions (97.79%) and 6,612 fraudulent transactions (2.21%). This indicates that the dataset is highly imbalanced, with fraudulent transactions representing only a small portion of the total data. The class imbalance was intentionally preserved in this study to simulate real-world e-commerce fraud detection conditions, where fraudulent transactions usually occur much less frequently than legitimate transactions. This approach allows the models to be evaluated based on their ability to detect fraud under challenging and realistic data distributions.

Table 2. Comparative Table

Model	Accuracy	Precision	Recall	F1-score
XGBoost	99%	90%	72%	80%
TabNet	99%	92%	64%	75%

Based on the evaluation results presented in Table 2, both XGBoost and TabNet were assessed using four primary performance metrics: accuracy, precision, recall, and F1-score. In the context of e-commerce fraud detection, both models achieved very high accuracy, reaching approximately 99%. However, as highlighted in previous studies, accuracy is not a sufficiently informative metric for highly imbalanced datasets, as it is strongly influenced by the majority class, in this case, non-fraudulent transactions, rather than the model's effectiveness in identifying fraudulent activities (Ghanem et al., 2023). Therefore, further analysis focuses on precision, recall, F1-score, confusion matrix, and PR-AUC based on the testing results.

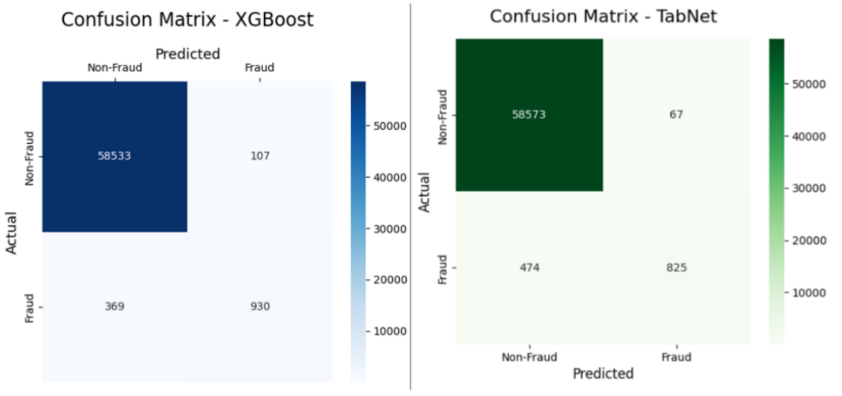

Figure 1 Confusion Matrix

In terms of precision, TabNet outperformed XGBoost by approximately 2%. TabNet correctly classified 92% of transactions predicted as fraudulent, compared to 90% for XGBoost. This indicates that TabNet produces fewer false positives and adopts a more conservative approach when identifying fraud. This result is also reflected in the confusion matrix, where TabNet generated fewer false positives, with only 67 non-fraudulent transactions incorrectly classified as fraudulent. Meanwhile, XGBoost produced 107 false positives. This indicates that TabNet is less likely to raise false alarms and is more cautious when classifying transactions as fraud.

However, this higher precision comes at the expense of recall. TabNet achieved a recall of only 64%, meaning that 36% of actual fraudulent transactions were not detected. Based on the confusion matrix analysis, TabNet correctly identified 825 fraudulent transactions as fraud (True Positive) but failed to detect 470 actual fraudulent transactions (False Negative). In contrast, XGBoost demonstrated stronger recall performance at 72%, correctly identifying 930 fraudulent transactions while producing 369 false negatives. This indicates that XGBoost has a greater ability to capture the minority class and detect a higher proportion of fraudulent transactions.

The F1-score, which represents the harmonic mean of precision and recall, provides a balanced measure of model performance in imbalanced classification tasks (Ghanem et al., 2023). While optimizing precision or recall individually often leads to trade-offs, the F1-score reflects the overall effectiveness of a model in minimizing both false positives and false negatives. Based on the results, XGBoost achieved an F1-score of 80%, outperforming TabNet, which obtained an F1-score of 75%. This indicates that XGBoost maintains a better balance between detecting fraudulent transactions and minimizing incorrect classifications.

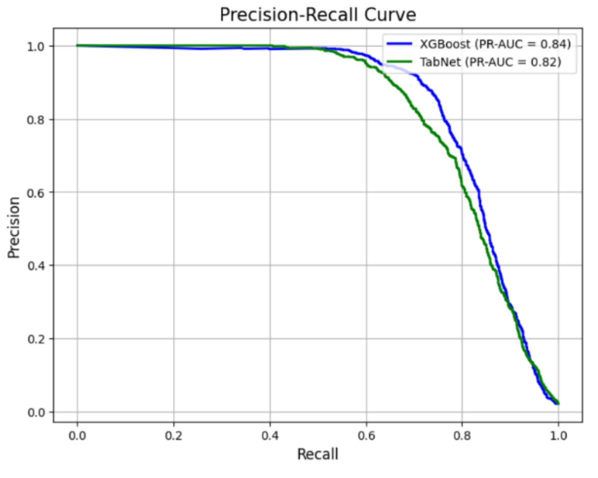

Figure 2 PR-AUC Curve

The PR-AUC evaluation further supports these findings. XGBoost achieved a PR-AUC score of 0.84, while TabNet obtained 0.82. In highly imbalanced fraud detection datasets, PR-AUC provides a more representative evaluation because it focuses on the model's ability to identify the minority class through precision and recall rather than being influenced by the majority class. The higher PR-AUC score indicates that XGBoost demonstrates stronger capability in distinguishing fraudulent transactions while maintaining fewer classification errors.

Overall, the comparison highlights a trade-off between the two models. TabNet demonstrates advantages in precision by reducing false positives, making it more conservative in fraud classification. However, its lower recall indicates that a considerable number of fraudulent transactions remain undetected. In fraud detection systems, missed fraud cases can lead to significant financial losses, making recall a critical performance indicator. XGBoost, on the other hand, provides a more balanced performance by capturing more fraudulent transactions while maintaining an acceptable level of false positives.

From an e-commerce perspective, effective fraud detection is essential to minimize financial losses, reduce chargebacks, protect customer data, and maintain user trust. Undetected fraudulent transactions can result in revenue loss, reputational damage, and decreased customer confidence (Lindemulder & Kosinski, n.d.). Based on the combined evaluation of precision, recall, F1-score, and PR-AUC, XGBoost is identified as the more suitable model for e-commerce fraud detection in this study. While TabNet demonstrates strength in precision, XGBoost provides more consistent and reliable performance in capturing fraudulent transactions, making it more appropriate for real-world e-commerce applications operating under highly imbalanced data conditions.

V. CONCLUSION AND RECOMMENDATION

Based on the evaluation results, both XGBoost and TabNet achieved high accuracy of approximately 99%. However, due to the highly imbalanced nature of the dataset, accuracy alone is not sufficient to represent model performance. Therefore, this study focused on precision, recall, F1-score, confusion matrix, and PR-AUC to evaluate fraud detection capability.

The results show that TabNet achieved slightly higher precision (92%) compared to XGBoost (90%), indicating that TabNet produced fewer false positives. However, XGBoost demonstrated better recall (72%) and F1-score (80%) than TabNet (64% recall and 75% F1-score), allowing it to detect a higher proportion of fraudulent transactions. Based on these results, XGBoost provides a more balanced performance between detecting fraud cases and minimizing incorrect classifications.

From an e-commerce perspective, XGBoost shows greater potential for implementation in fraud detection systems by improving the identification of suspicious transactions, reducing potential financial losses, and enhancing transaction security. Future studies may further explore real-world transaction datasets, advanced model optimization, and the development of real-time fraud detection systems to improve practical applicability.

REFERENCES

Alshdaifat, E., Al-hassan, M., & Aloqaily, A. (2021). Effective heterogeneous ensemble classification: An alternative approach for selecting base classifiers. *ICT Express*, *7*(3), 342–349. https://doi.org/10.1016/j.icte.2020.11.005

Bian, Y., Wang, Y., Yao, Y., Chen, H., & Member, S. (2019). Ensemble Pruning based on Objection Maximization with a General Distributed Framework. In *IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS*.

Chen, T., & Carlos, G. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ?16) (p*. 785). New York: ACM.

Dodiya, G., Dudhat, D., & Vidani, J. (2024). To Study the Impact of Online Fraud and Scams. *International Journal of Law, Human Rights and Constitutional Studies*, 24–36.

Dr Padmalatha N a. (2020). E-Commerce Frauds and the role of fraud Detection Tools in managing the risks associated with the frauds. *International Journal of Advanced Science and Technology*, *29*(4s), 38–46. Retrieved from https://www.researchgate.net/publication/348297729

Duwadi, N., & Sharma, A. (2024). *Comparative Study of Machine Learning Techniques for Insurance Fraud Detection . Aviation Electronics, Information Technology, Telecommunications, Electricals, and Controls (AVITEC), 131*.

Ethan Cramer-Flood. (2025, April). Worldwide Retail Ecommerce Forecast 2025. Retrieved October 22, 2025, from Website website:

https://www.emarketer.com/content/worldwide-retail-ecommerce-forecast-2025

Fadilla, Z., Ketut Ngurah Ardiawan, M., Eka Sari Karimuddin Abdullah, M., Jannah Ummul Aiman, M., & Hasda, S. (2021). *METODOLOGI PENELITIAN KUANTITATIF* (N. Saputra, Ed.). Yayasan Penerbit Muhammad Zaini. Retrieved from http://penerbitzaini.com

Faisal, K. (1990). Overcoming Limitations of Rule-Based Systems: An Example of a Hybrid Deterministic Parser. *SpringerProceedings of the 6th Austrian Conference on Artificial Intelligence (KONNAI)*, 48–49. Berlin: Springer.

Ghanem, M., Ghaith, A. K., El-Hajj, V. G., Bhandarkar, A., de Giorgio, A., Elmi-Terander, A., & Bydon, M. (2023, December 1). Limitations in Evaluating Machine Learning Models for Imbalanced Binary Outcome Classification in Spine Surgery: A Systematic Review. *Brain Sciences*, Vol. 13. Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/brainsci13121723

Haq, I. u. (2022). RISK OF FRAUD AND SUSTAINABILITY OF E-COMMERCE. *Journal of Research in Economics and Finance Management*, 27–37.

Hassan, Y. A., & Kareem, O. S. (2025). Credit Card Fraud Detection: A Comparative Study of Machine Learning and Deep Learning Methods. *Engineering and TechnologyJournal*, *5165*.

Islam, S., Haque, M. M., & Rezaul Karim, A. M. (2024). A rule-based machine learning model for financial fraud. *International Journal of Electrical and Computer Engineering*, *760*.

Jin, J., & Zhang, Y. (2025). The analysis of fraud detection in financial market under machine learning. *Scientific Reports*, *15*(1). https://doi.org/10.1038/s41598-025-15783-2

Li, Y., & Chen, W. (2020). A Comparative Performance Assessment of Ensemble Learning for Credit Scoring. *Mathematics*, *8*(10). https://doi.org/10.3390/math8101756

Martin Ward Powers, D. (2011). EVALUATION: FROM PRECISION, RECALL AND F-MEASURE TO ROC, INFORMEDNESS, MARKEDNESS & CORRELATION. *Article in Journal of Machine Learning Technologies*, *2*(1), 37–63. https://doi.org/10.9735/2229-3981

Meng, C. C., Lim, K. M., Lee, C. P., & Lim, J. Y. (2023). Credit Card Fraud Detection using TabNet. *2023 11th International Conference on Information and Communication Technology (ICoICT)*. Melaka, Malaysia: IEEE.

Mienye, I. D., & Sun, Y. (2022). A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects. *IEEE Access*, Vol. 10, pp. 99129–99149. Institute of

Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2022.3207287

Mishra, D., Tripathi, S. M., Chaurasia, A., & Chaurasia, P. K. (2025). A Review on Ensemble Learning Methods: Machine Learning. *International Journal of Research Publication and Reviews*, *3796*.

Nathaniel Andrew Davis, & Sophia Anne Harris. (2024). Leveraging Machine Learning Models for Real-Time Fraud Detection in Financial Transactions. *International Journal of Computer Technology and Science*, *1*(1), 01–06. https://doi.org/10.62951/ijcts.v1i1.56

Nobre, J., & Neves, R. F. (2019). Combining Principal Component Analysis, Discrete Wavelet Transform and XGBoost to trade in the financial markets. *Expert Systems with Applications*, *125*, 181–194. https://doi.org/10.1016/J.ESWA.2019.01.083

Opena, A. (2025, September 22). 70+ eCommerce Fraud Statistics [2025]: Trends, Data, & Facts. Retrieved December 19, 2025, from Website website: https://cropink.com/ecommerce-fraud-statistics

Paige Tester. (2022, December 5). 7 Types of E-Commerce Fraud & How to Prevent Them. Retrieved October 22, 2025, from Website website: https://datadome.co/learning-center/7-types-of-ecommerce-fraud-how-to-prevent-them/#types

Pfister, T., & Arik, S. O. (2020). TabNet: Attentive Interpretable Tabular Learning. *The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21) (p*. 1): Association for the Advancement of Artificial Intelligence (AAAI.

Ravichandran, T., Gavahi, K., Ponnambalam, K., Burtea, V., & Mousavi, J. S. (2021). Ensemble-based machine learning approach for improved leak detection in water mains. *Journal of Hydroinformatics*, *23*(2), 307–323. https://doi.org/10.2166/HYDRO.2021.093

Roxana Badircea, Nicoleta Mihaela Doran, Alina Georgiana Manta, & Jenica Popescu. (2021). E-Commerce and the Factors Affecting Its Development in the Age of Digital Teology: Empirical Evidence at EU?27 Level. *MDPI*, *1*.

Sharma, N. (2023, June 10). Understanding and Applying F1 Score: AI Evaluation Essentials with Hands-On Coding Example. Retrieved November 6, 2025, from Website website: https://arize.com/blog-course/f1-score/

Shwartz-Ziv, R., & Armon, A. (2022). Tabular data: Deep learning is not all you need. *Information Fusion*, *81*, 84–90. https://doi.org/https://doi.org/10.1016/j.inffus.2021.11.011

Sun, S., Chen, W., Wang, L., Liu, X., & Liu, T.-Y. (2016). *On the Depth of Deep Neural Networks: A Theoretical View*. Retrieved from www.aaai.org

Syahbani, A. M., Firdaus, W., & Musodo, K. A. (2025). A Comparative Study of Data Mining Algorithms for Fraud Detection in Financial Transactions. *Sinkron : Jurnal Dan Penelitian Teknik Informatika*, *814*.

Tayebi, M., & Kafhali, S. E. (2025). A novel approach based on XGBoost Classifier and Bayesian Optimization for Credit Card Fraud Detection. *Cyber Security and Application*, *2*.

Uygur, U. (2024, January 1). E-Commerce Fraud Detection Dataset. https://doi.org/https://doi.org/10.34740/kaggle/dsv/13601574

Wibowo, J. S., Hartono, B., & Lusiana, V. (2024). Online Payment Fraud Detection Optimization with XG Boost. *Journal of Software Engineering and Simulation*, *41*.

Zheng, A. (2015). *Evaluating Machine Learning Models: A Beginner's Guide to Key Concepts and Pitfalls*. Sebastopol.