UNIABUJA Journal of Engineering and Technology

https://ujet.uniabuja.edu.ng/

ISSN: 2714-3236 (Online); 2714-3228 (Print)

Volume 2, Issue 1, 2025; 196-204

Rule-Based Approach to e-Commerce Fraud Detection

Bisallah H. IBRAHIM, Habiba U. SALIHU and Yusuf A. ALESHINLOYE

Department of Computer Sciences, Faculty of Science, University of Abuja, Abuja, Nigeria

*Corresponding Author: habee.salih@gmail.com

Abstract

With the rapid development of information technologies, e-commerce now becomes prevalent worldwide. However, huge number of transactions in e-commerce raises the potential for new problems namely fraud in e-commerce transactions in which the most difficult aspect of the problem is how to detect it. Many solutions are based on machine learning approach which may be expensive due to time and resources needed for training of dataset any time there is change in business policy. In this work, therefore, a system is proposed to detect fraud in e-commerce transaction using analytical technique built on top of rule-based engine. First, at the analytical stage, feature selection and feature engineering were carried out to define the appropriate data features that determines the accuracy of fraud detection in e-commerce transaction. Secondly, a rule-based engine is constructed that defines the policy for every feature and set their threshold to detect deviation from normal policy and then predict the status of the transaction. The constructed rule engine contains various rules starting from single to combination and aggregation of two or more rules. The purpose of the rule is to compute an integer score based on each feature of the transaction and then aggregate the score as more rules are applied. A non-fraud transaction eventually has zero (0) score otherwise the score greater than zero (0) signifies fraud and magnitude of the fraud is determined by the size of the score. Thirdly, strategy is developed for modification of rules without affecting the system performance. The strategy involved creating interfaces for rule update. Finally, a graphical user interface application is developed to demonstrate all the proposed objectives and experiment is carried out using online sourced e-commerce dataset which was only used for evaluation. The system was fine-tuned and experiment performed in three phases. The result of the experiment shows that the proposed system achieved high performance with average accuracy of 93.3% while precision, recall and F1-score are 96.3%, 95.5% and 95.9% respectively. This implies that true positive (TP) and true negative (TN) is high while false positive (FP) and false negative (FN) is very low.

Keywords: *Ecommerce fraud, data analytics, rule-based modelling, machine learning models.*

1.0 Introduction

It is hard to underestimate the role of e-commerce in a world where most communications happen on the web (Shafiyah et. al, 2013) and our virtual environment is full of advertisements with attractive products and services to buy (Išoraitė, 2018). Meanwhile, it is obvious that many criminals are trying to take advantage of it, using scams and malware to compromise users' data.

Fraud in E-commerce transaction occurs when a fraudster goes to an online store and makes an unauthorized transaction using the compromised details of a stolen or fake credit card leaving the merchant without legal payment for the goods; thus, the store will have to charge money back to the compromised customer. This scenario differs from real-life credit card fraud (Rodrigues et.al, 2022) because there may not be card physically involved, and the victim does not always have to have some type of interaction with the fraudster in real life in order to be compromised.

Having said that, this chapter lays a background to the research that will be carried out by describing the topic as well as defining the scope of the work. There is no doubt that e-commerce fraud detection and prevention still remain an open problem as the world continues to go deep down into virtuality and e-commerce scope is expanding. Although electronic commerce (e-commerce) is not something new, it is an evolution of traditional business practices to take advantage of the new technologies of the internet age. In fact, International Chamber of Commerce (2019) noted that global trade is increasingly digitalized such that buyers, sellers, and intermediaries now rely on technologies that enable commerce at a speed, scale, and efficiency unimaginable just a few decades ago.

Over the last few years, e-commerce has become an indispensable part of the global retail framework. According to Statista Research (2020), country with highest retail e-commerce Compound Annual Growth Rate (CAGR) currently is Turkey, retail e-commerce sales worldwide is amount to 3.5tr USD and by 2023 e-commerce share of total global retail sales will be 22%. Also, it noted that 2019 Business-to-Customer (B2C) index value for e-commerce in Nigeria stood at 53.2 points, the fourth highest in Africa, and currently the number of online shoppers in Nigeria is approximately 76.6 million with the most valuable e-commerce

sector being travel and accommodation while the fastest growing e-commerce sector is food and personal care. Therefore, one can say that Nigeria's digital landscape is flourishing. The country has one of the biggest internet economies in Africa. With the continent's largest population and one of the youngest worldwide, Nigeria presents a vast digital audience.

Along with the growth of e-commerce sector, the count of e-commerce related frauds are also increasing in every year since 1993. As per a report in 2013, 5.65 cents are lost due to frauds out of every $100 in e-commerce turnover (Shini, 2018). Fraud has never been a new thing, although the trend for E-commerce fraud rises as the number of cash-free transactions increase. It is especially obvious now, when the world is moving away from in-store purchases. Due to the COVID-19 quarantine, people have to make more purchases online to stay safe or because the products they need are unavailable in closed local shops.

Problem Statement:- As the use of e-commerce continues to expand, so does the complexity of fraudulent activities. Existing fraud detection methods are limited by the need for large data volumes and computational resources, as well as their inability to adapt to new fraud tactics in real-time. There is a need for a more efficient, adaptable fraud detection system capable of addressing these challenges.

The work aims to develop a model for detecting ecommerce fraud using Data analytics and rule-based approach addressing the limitations of existing models such as Islam et'als (2024) model.The objectives include:

a. To develop a scalable and adaptive rule-based fraud detection system.
b. To identify the most significant features in e-commerce transaction data that contribute to fraud detection.
c. To evaluate the system's performance against traditional models, using accuracy, precision, recall, and F1-score as metrics. Our findings demonstrate improvements in accuracy,false positive rate validating the effectiveness of the proposed approach.

1.1 Review of Related Works

1.2 E-Commerce Fraud Detection Techniques

Fraud detection has been studied for various applications, but e-commerce presents unique challenges due to the complexity of online transactions. The traditional approach has focused on using machine learning algorithms, including decision trees, support vector machines (SVM), and neural networks. However, these systems often require large datasets and significant computational resources.

1.3 Machine Learning-Based Approaches

Machine learning techniques, particularly Random Forest and Support Vector Machines, have gained popularity for fraud detection. These models are capable of learning from transaction data and detecting anomalies that traditional rule-based systems cannot identify. For instance, Random Forest models offer high accuracy and robustness but require significant data preprocessing (Ayo et al., 2017).

1.4 Rule-Based Systems

Rule-based systems use predefined conditions to classify transactions as fraudulent or legitimate. These systems are simple to implement but struggle to adapt to evolving fraud tactics. Recent work has explored how dynamic rule modifications can improve adaptability and maintain the performance of rule-based systems over time (Rodrigues et al., 2022).

Combining rule-based systems with machine learning techniques has proven to be an effective way to address the limitations of both approaches. Hybrid models leverage the strengths of predefined rules for quick classification and the adaptability of machine learning to detect complex fraud patterns (Shafiyah et al., 2013).

Khatri (2024) proposed a hybrid framework combining rule-based and anomaly-detection techniques for payment fraud detection. The study demonstrated improved accuracy, adaptability, and scalability by leveraging domain knowledge and expert-defined rules alongside anomaly detection techniques. However, the study lacks extensive real-world deployment and evaluation.

Khanum et al. (2024) study compared machine learning (ML) and rule-based systems (RBS) for fraud detection, concluding that ML achieved higher accuracy and F1 scores but required significant computational resources. RBS was faster but less accurate. Hybrid models were suggested as a balance between performance and efficiency. However, the study did not delve deeply into interpretability or broader applications.

Mutemi & Bacao (2024) conducted a systematic literature review on e-commerce fraud detection using ML techniques, highlighting gaps in real-world dataset availability and insufficient focus on marketplace-

specific challenges. Despite identifying trends and ML techniques, the study lacked experimental validation.

Islam et al. (2024) introduced a rule-based ML model for financial fraud detection that achieved 99% accuracy and precision on benchmark datasets. The study outperformed traditional ML models but did not explore scalability or real-time application.

Alsubari et al. (2023) explored detecting fake reviews in e-commerce using a CNN-BiLSTM hybrid model. This approach achieved better performance than Random Forests but did not address generalizability across other domains.

Yuanyuan Tang (2023) presented a fraud detection model using reinforced deep learning and the ABC algorithm, achieving high accuracy. However, computational expense and lack of testing on broader datasets were noted as limitations.

Ahmed et al. (2021) proposed a semantic rule-based model for financial fraud detection using ontology-based reasoning and an alert system. While efficient at generating alerts, the system was not tested in highly dynamic fraud scenarios.

Youssef et al. (2021) combined deep learning with rule extraction techniques to enhance interpretability in e-commerce fraud detection. The approach improved performance but required significant computational power for training.

Gayam (2020) focused on anomaly detection, transaction monitoring, and risk mitigation using AI. While comprehensive, the study struggled with adapting to evolving fraud tactics and lacked implementation of real-time updates.

Padmalatha (2020) highlighted effective fraud detection tools and trends in managing e-commerce risks. The study provided a comprehensive overview of tools tailored to business-specific needs, leveraging ML and big data technologies. However, it did not delve deeply into specific fraud detection methodologies for prominent fraud types.

Adi Saputra & Suharjito (2019) applied the SMOTE method to improve classification metrics on imbalanced datasets, achieving better F1-scores and G-Mean across various machine learning models. However, the approach was limited in addressing interpretability issues and computational costs.

Pradheepan Raghavan et al. (2019) demonstrated that combining SVMs and CNNs performed well on large datasets and ensemble models were effective for smaller datasets. However, the study struggled with unsupervised learning and frequently required retraining to adapt to dynamic fraud patterns.

Hangjun et al. (2019): Achieved improved precision, recall, and accuracy in real-time fraud detection using Spark-based big data analytics. Despite its success, the model was not tested in industrial big data environments like cloud systems and lacked robustness verification for broader applications (Hangjun et al., 2019).

Dahee & Kyungho (2018): Surveyed and implemented ML-based methods for financial fraud detection in IoT environments. While ML outperformed ANN methods in efficiency, the processing time for real-time detection posed significant challenges when combining ML and deep ANN processes.

Suryanarayana et al. (2018): Developed logistic regression-based models for credit card fraud detection, achieving high accuracy and assisting fraud investigators. Performance could have been enhanced through ensemble approaches combining multiple data mining techniques.

Shini (2018) proposed an SVM-based framework for detecting fraudulent sellers in online marketplaces. This approach utilized historical marketplace data and social media analytics but faced a cold start problem, making it ineffective for evaluating new sellers.

Evandro et al. (2014) utilized computational intelligence techniques like neural networks and Bayesian networks for fraud detection, achieving up to 43.66% economic efficiency. However, the unbalanced dataset, with minor fraud classes representing less than 1% of data, significantly affected performance.

Shaji & Panchal (2017) introduced Adaptive Neuro-Fuzzy Inference Systems for fraud detection in e-commerce transactions, which demonstrated adaptability to newer instances of fraud. However, real-time fraud detection and scalability remained challenging.

Roldán-García et al. (2017) proposed an ontology-driven approach to resolve semantic conflicts in anti-fraud rule repositories, effectively reducing errors in fraud detection. However, the scalability of this approach across diverse domains was not adequately addressed.

Milo et al. (2018) introduced the RUDOLF system for refining fraud detection rules, demonstrating effective rule refinement for detecting credit card fraud and network attacks. However, the reliance on domain expertise limited its scalability and adaptability to broader fraud scenarios.

This work presents a system for detecting fraud in e-commerce transactions while assisting domain experts in defining and refining rules in dynamic environments. Addressing challenge of rule adaptation, the system combines analytical and rule-based approaches, working interactively with experts until satisfactory rules are achieved. Experiments using real-world datasets confirm the system's effectiveness

and efficiency for rule refinement. the system can process transactions in batches, optimizing performance for large datasets. This optimization enhances usability, allowing for efficient analysis of large volumes of data without memory constraints.

2.0 Methodology

Data analytics and rule-based approach were used in the implementation of this system. The steps taken to implement this are:

I. **Data collection:** This is the first step in implementation of the solution. It involves collecting ecommerce transaction data set from open source data repository. The dataset that will be used for the proposed system's development to detect fraud on e-commerce transaction is "merged_dataset" which includes the following columns: customerEmail, customerPhone, customerDevice, customerIPAddress, customerBillingAddress, No_Transactions, No_Orders, No_Payments, Fraud_statusAddress.The dataset provides valuable information about ecommerce transaction records and their associated attributes. Each entry includes: -customerEmail, customerPhone, customerDevice customerIPAddress,customerBillingAddress, No_Transactions, No_Orders, No_Payments, Fraud_status .

II. **Data Preprocessing and Feature Selection:** In the data preprocessing phase, the dataset is prepared and transformed to ensure its quality, consistency, and compatibility with the subsequent analysis. Incorrect data entry was manually filtered out of the data gathered to allow the models to train using correct and authentic data.

III. **Data Cleaning:** The data gathered contained some inaccurate entries which were inconsequential to the research. Data cleaning was done by manually going through the data and filtering out the incorrect entries in order to remove any form of discrepancies.

IV. **Feature Selection:** The relevant features or attributes for the analysis are selected from the dataset. This step involves evaluating the importance and relevance of each attribute and choosing the most informative ones for the analysis.

2.1 Develop a Heuristic Rule-based Model

A rule-based model is built, which is a model based on expert knowledge and understanding of the problem. This model is not based on any existing maths/statistical model rather it is based on predefined rules. We use heuristic because it's easy to interpret and rules can be easily updated to capture new or misclassified fraud. The rules are in the form of IF-THEN (conditional statements) statements. We created a set of rules and conditions based on known patterns of fraudulent behavior and applied them to the dataset to flag suspicious transactions when these conditions are met or otherwise. Some of the rules;

a. If a user makes a purchase and pays a large amount from multiple newly created accounts
b. If a user makes multiple transactions using different payment cards.
c. If the phone number of a user does not conform to a certain country code.
d. If someone from the same IP address is creating multiple accounts and paying for items with credit cards.

2.2 Improving the Heuristic Rule-based Model

To improve the performance of the system, the following techniques are adopted;

1. We assigned a weight to each rule by assigning a score of 1 or 0 i.e. 1 if fraudulent and 0 for non-fraud.
2. We set a threshold for each rule between 1 and 0 to determine the intensity of the fraud case
3. Blacklisting of violated features in case of future occurrence. If a parameter for a particular user is already blacklisted, the rule assigns a score of 1 otherwise 0.
4. Rule combination and Aggregate rule score to determine the final weight of violation.

3.0. Experiments and Evaluations

In this work, the result of the proposed system will be evaluated in terms of Recall, Precision, Accuracy, and F1-Score, which are defined below respectively.

True Positive (Tp):

Means the number of correctly predicted fraud transactions among all the true fraud transactions,

False Positive (Fp):

This means the number of normal transactions which are incorrectly predicted as fraud transactions,

True Negative (Tn):

Means the number of correctly predicted normal transactions among all the true normal transactions,

False Negative (Fn):

This means the number of fraudulent transactions that are incorrectly predicted as normal transactions.

The Precision:

is the synonym for the positive predictive value. It tells us how many of the correctly predicted cases turned out to be positive.

Precision = [TP÷(TP+FP)] ……. 1

Recall:

Is a synonym for the true positive rate and the accuracy indicates the general fraud detection performance. It also tells us how many of the actual positive cases we were able to predict correctly.

Recall= [TP÷(TP+FN)]……. 2

f1-Score:

is the harmonic mean of precision and recall (Saputra & Suharjito, 2019; Zhou et.al, 2019).

$$\boldsymbol{F1 - Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad 3$$

Accuracy:

This is important for evaluating the classification model.it is also the fraction of prediction that the model got right.

Accuracy= {(TP +TN) ÷ (TP + TN+ FP + FN)} …….. 4

3.1 System Design/The Proposed Fraud Detector

The proposed techniques are implemented as application software. The ideal method of using the application is to integrate it with an e-commerce system such that it checks every transaction before completion and decides whether the transaction is fraudulent or not. The program consists of major four engines namely the database layer, the rule layer, the analytical, and the interface layer.

Database layer: This layer consists of a database engine that stores the persistent data in the system. This is implemented with SQLite and it is to be packed as part of the application. The most important data stored by the database is the blacklist feature. This is the feature perceived to be malicious or abnormal based on the rules.

Rule Layer: This is the major part that decides whether a transaction is fraud or not. The rule engine is also responsible for scoring a transaction. Based on the research methodology, the rule is of two types; single and combined rules. The complete code is attached to the appendix.

a. **Single rule:** This set of rules checks each feature such as IP address, physical address, mobile number, device information, etc. Also, for most of the single rules, we first check whether the feature exists in the blacklist and if it does a score of 1 is allotted otherwise 0 score is allotted. The rule then further checks if the feature is above the threshold if it is, the score is incremented otherwise the score is returned.
b. **Combined rule**: this rule is formed from a combination and aggregation of two or more single rules based on the research mythology. The score is also determined from the existing scores from the combined party of a single rule.

Analytical layer: this layer is responsible for data pre-processing as discussed in the methodology behind the scenes. The purpose of the layer is to work on the dataset and ensure it conforms to the structure required by the rule engine. Some of the tasks in this layer include the conversion of data from one structure to another such as List, Dictionary, and

The User Interface layer: this is an interactive layer accessible to the user. It is designed with various widgets. The first task for the user is to select the file on the local system that contains the transaction. The transaction file is a comma-separated value (CSV) file that may contain single or multiple records. However, the interface provides a transaction data view, analysis information view, and blacklist data view. It also provides an interface for experts to modify rules as discussed in the methodology.

3.2 System Evaluation and Result Analysis

The system allows a single transaction data file to be uploaded for analysis as well as batch analysis of large data files. Once the data file is uploaded, the system performs the pre-processing and displays it on the transaction data view. When analyze button is triggered, the system carries out the analysis of the data and applies the rules and blacklist check to generate a score that determines its fraudulent status. The score is an integer ranging from 0 and above depending on various issues discovered with the transaction data. A zero (0) score means that the transaction is not fraud otherwise the level of suspicion is then determined by how big the value of the score is. Also, the system is able to generate a report for every suspicious transaction and explains which feature is actually causing the issue. Thereby allowing the expert to blacklist the feature for future reference.

Furthermore, the system provides an interface for rule updates and it allows the experts to adjust the rule to

cater for new events that might affect the decision of the system. Let us suppose, for instance, we receive information about a fraudster operating from a particular location or some of the information is revealed to the security personnel. The rule engine can be updated to use geolocation information to map a set of location addresses that are around the geolocation that can be blacklisted. This is one of many advantages of rule modification and updating.

3.2.1 Result Analysis

The performance of the proposed system was evaluated using precision, recall, accuracy, and F1-score which have already been discussed. The test was done in three phases. At each phase, the dataset contains 100 records. For the first phase, 10 fraud transactions were injected. The second phase contains 14 frauds while the third phase contains 20 fraudulent transactions.

The performance of any classification or detective system greatly depends on the true positive (TP), true negative (TN), false positive (FN), and false negative (FN). All must be accurate or close to accurate. TP and TN should be high while FP and FN should be low. The system was fine-tuned to achieve high performance according to the tables presented for the three phases of test both FN and FP are kept under 5% each while all the TP are above 70%.

The Summary Table of the evaluation is presented also which shows that the average **accuracy** is 96.9% while **precision, recall**, and **F1-score** are 96.3%, 95.5%, and 95.9% respectively. This is not bad for a system like this as things can still be improved as more rules are being encoded.

Table 1: Confusion Matrix of 1st evaluation

Confusion Matrix		**Predicted Classes**	
		Fraud	**Non-Fraud**
Actual Classes	**Fraud**	83 (TP)	3 (FN)
	Non-Fraud	4 (FP)	10 (TN)

The evaluation matrix for the first phase of evaluation has high TP(83)

Table 2: Confusion Matrix of 2nd evaluation

Confusion Matrix		**Predicted Classes**	
		Fraud	**Non-Fraud**
Actual Classes	**Fraud**	80 (TP)	4 (FN)
	Non-Fraud	2 (FP)	14 (TN)

Table 2: Displays results from the 2nd phase process, ranking variables by their predictive importance. Key contributors include transaction time, amount, and device fingerprint.

Table 3: Confusion Matrix of 3rd evaluation

Confusion Matrix		**Predicted Classes**	
		Fraud	**Non-Fraud**
Actual Classes	**Fraud**	73 (TP)	4 (FN)
	Non-Fraud	3 (FP)	20 (TN)

Table 3. Compares the performance metrics of different algorithms (e.g., SVM, Logistic Regression, Random Forest) tested during model development, emphasizing the proposed model's superiority.

Table 4: Summary Table of the Evaluation

Parameter	**1st Evaluation**	**2nd Evaluation**	**3rd Evaluation**	**Average**
Precision	0.9540	0.9756	0.9605	0.9634
Recall	0.9651	0.9524	0.9481	0.9552
Accuracy	0.9550	0.9550	0.9690	0.9693
F1-Score	0.9595	0.9639	0.9542	0.9592

Received: 05-02-2025 / Accepted: 21-02-2025 /Published: 31-03-2025

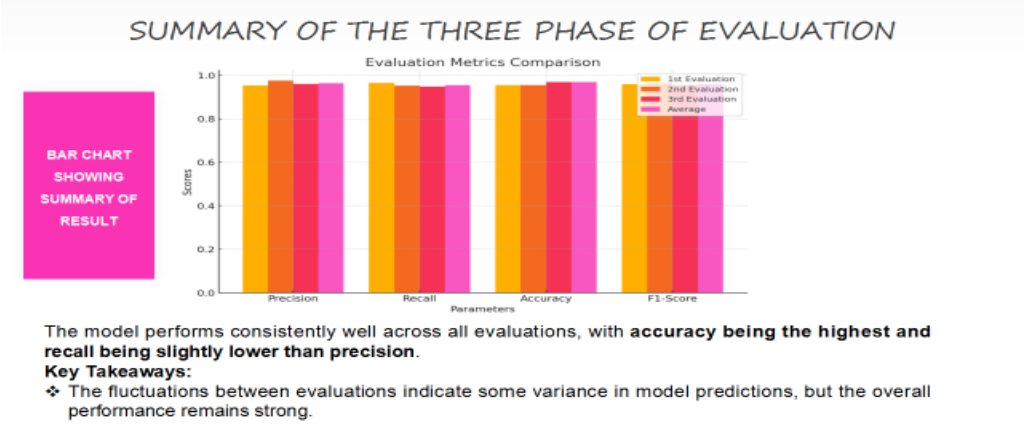

Figure 1: Summary of the three phases Evaluations

Model Comparison

Metric	Existing Model (Islam et al., 2024)	Proposed Model (This Study)
Accuracy	96%	**96.5%**
Precision	95.5%	**96.9%**
Recall	94%	**95.5%**
F1-Score	95.5%	**95.9%**
Efficiency	High	**Very High**

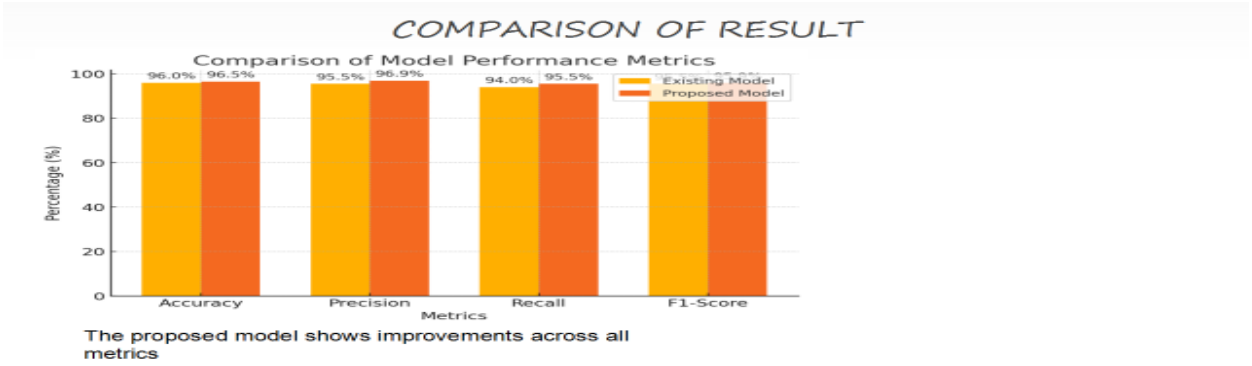

Figure 2: Comparison of Result

3.4 Discussion of Results

Experimental evaluation demonstrated the model's effectiveness in real-world scenarios. The system achieved an accuracy of 96.5%, a precision of 96.9%, a recall of 95.5%, and an F1-score of 95.9%. These metrics highlight the model's ability to minimize false positives and false negatives, ensuring reliable fraud detection with minimal disruption to legitimate transactions. Compared to the benchmark model proposed by Islam et al. (2024), the system consistently outperformed across all evaluation metrics, confirming its efficiency.

Key Differences:

1. **Accuracy**: The proposed model's approach improves prediction accuracy by integrating domain-specific rules in the model.
2. **Adaptability**: Unlike the static nature of Islam et al.'s model, the proposed model incorporates dynamic rule updates and feedback loops, enhancing adaptability to new fraud patterns.
3. **Efficiency**: The use of rule-based approach improves computational efficiency, reducing false positives while maintaining high precision and recall.

References

ACI Worldwide. (2020). Global eCommerce retail sales up 209 percent in April, ACI Worldwide research reveals [Blog post]. Business Wire. https://www.businesswire.com/news/home/20200511005666/en/

Adrian, B. (2015). Emerging markets queries in finance and business: Detecting and preventing fraud with data analytics. Procedia Economics and Finance, 32, 1827–1836.

Ahmed, M., Ansar, K., Muckley, C. B., Khan, A., Anjum, A., & Talha, M. (2021). A semantic rule-based digital fraud detection. PeerJ Computer Science, 7, e649.

Alexander, S. L., & Arvid, L. (2019). An overview of deep learning in medical imaging focusing on MRI. Zeitschrift für Medizinische Physik, 29(2), 102–127. https://doi.org/10.1016/j.zemedi.2018.11.002

Alsubari, S. N., Deshmukh, S. N., Aldhyani, T. H., Al Nefaie, A. H., & Alrasheedi, M. (2023). Rule-based classifiers for identifying fake reviews in e-commerce: A deep learning system. In Fuzzy, Rough and Intuitionistic Fuzzy Set Approaches for Data Handling: Theory and Applications (pp. 257-276). Singapore: Springer Nature Singapore.

Andrea, D. P., Giacomo, B., Olivier, C., Cesare, A., & Gianluca, B. (2017). Credit card fraud detection: A realistic modeling and a novel learning strategy. IEEE Transactions on Neural Networks and Learning Systems. https://doi.org/10.1109/TNNLS.2017.2736643

August, M. (2017). Multi-faceted evolution of mobile payment strategy, authentication, and technology. Federal Reserve Bank of Boston. https://www.bostonfed.org/publications/mobile-payments-industryworkgroup/multi-faceted-evolution-of-mobile-payment-strategy-authentication-and-technology.aspx

Ayo, C. K., Adewoye, J. O., & Oni, A. A. (2007). Business-to-consumer e-commerce in Nigeria: Prospects and challenges. African Journal of Business Management, 5, 5109.

Belás, J., Korauš, M., Kombo, F., & Korauš, A. (2016). Electronic banking security and customer satisfaction in commercial banks. Journal of Security and Sustainability Issues, 5, 411–422. https://doi.org/10.9770/jssi.2016.5.3(9)

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

Brownlee, J. (2014). Discover feature engineering: How to engineer features and how to get good at it. Machine Learning Mastery. https://machinelearningmastery.com/discover-feature-engineering-how-to-engineer-features-and-how-to-get-good-at-it/

del Mar Roldán-García, M., García-Nieto, J., & Aldana-Montes, J. F. (2017). Enhancing semantic consistency in anti-fraud rule-based expert systems. Expert Systems with Applications, 90, 332-343.

Dahee, C., & Kyungho, L. (2018). An artificial intelligence approach to financial fraud detection under IoT environment: A survey and implementation. Hindawi Security and Communication Networks, 2018, Article ID 5483472. https://doi.org/10.1155/2018/5483472

Elluru, P. K. (2021, March 11). Step-by-step process of feature engineering for machine learning algorithms in data science. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2021/03/step-by-step-process-of-feature-engineering-for-machine-learning-algorithms-in-data-science/

Ehikioya, S. A., & Guillemot, E. A. (2020). Critical assessment of the design issues in e-commerce systems development. Engineering Reports, 2, e12155. https://doi.org/10.1002/eng2.12155

Gayam, S. R. (2020). AI-Driven Fraud Detection in E-Commerce: Advanced Techniques for Anomaly Detection, Transaction Monitoring, and Risk Mitigation. Distributed Learning and Broad Applications in Scientific Research, 6, 124-151.

Gee, S. (2015). Fraud and fraud detection: A data analytics approach. Wiley Corporate F&A Series.

Heaton, J. (2016). An empirical analysis of feature engineering for predictive modeling. SoutheastCon 2016, 1–6. https://doi.org/10.1109/SECON.2016.7506650

Heaton, J. (2020). An empirical analysis of feature engineering for predictive modeling. arXiv. https://arxiv.org/abs/1701.07852v2

Islam, S., Haque, M. M., & Karim, A. N. M. R. (2024). A rule-based machine learning model for financial fraud detection. International Journal of Electrical & Computer Engineering (2088-8708), 14(1).

Išoraitė, M. (2018). Electronic commerce: Theory and practice. Integrated Journal of Business and Economics, 2, 73. https://doi.org/10.33019/ijbe.v2i2.78

Johannes, S. K., & Rajasvaran, L. (2020). Auto-insurance fraud detection: A behavioral feature engineering approach. Journal of Critical Reviews, 7(3), 125–129. https://doi.org/10.31838/jcr.07.03.23

Khanum, A., Chaitra, K. S., Singh, B., & Gomathi, C. (2024, January). Fraud Detection in Financial Transactions: A Machine Learning Approach vs. Rule-Based Systems. In 2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE) (pp. 1-5). IEEE.

Khatri, M. R. (2024). COMBINING THE STRENGTHS OF RULE-BASED AND ANOMALY DETECTION TECHNIQUES FOR ROBUST AND COMPREHENSIVE PAYMENT FRAUD DETECTION.

International Journal of Applied Machine Learning and Computational Intelligence, 14(4), 11-20.

Kolkman, D. (2020). The usefulness of algorithmic models in policy making. Government Information Quarterly, 37(3). https://doi.org/10.1016/j.giq.2020.101488

Li, X., Zhang, M., Liu, Y., Ma, S., Jin, Y., & Ru, L. (2014). Search engine clicks spam detection based on bipartite graph propagation. Proceedings of the 7th ACM International Conference on Web Search and Data Mining, 93–102. https://doi.org/10.1145/2556195.2556214

Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on MRI. Zeitschrift für Medizinische Physik, 29(2), 102–127. https://doi.org/10.1016/j.zemedi.2018.11.002

Milo, T., Novgorodov, S., & Tan, W. (2016). Rudolf: Interactive rule refinement system for fraud detection. Proceedings of the VLDB Endowment, 9(13), 1465–1468. https://doi.org/10.14778/3151251.3151274

Mutemi, A., & Bacao, F. (2024). E-Commerce Fraud Detection Based on Machine Learning Techniques: Systematic Literature Review. Big Data Mining and Analytics, 7(2), 419-444.

Najem, S. M., & Kadeem, S. M. (2021). A survey on fraud detection techniques in e-commerce. Tech-Knowledge Journal, 1(1), 33–47.

Nithya, C., & Saravanan, V. (2018). A survey of feature extraction and feature engineering in data mining. IOSR Journal of Engineering, 57(1), 83–87.

Rodrigues, V. C., Policarpo, L. M., da Silveira, E., Righi, R., da Costa, C. A., Barbosa, J. L., Antunes, R. S., Scorsatto, R., & Arcot, T. (2022). Fraud detection and prevention in e-commerce: A systematic literature review. Electronic Commerce Research and Applications, 56, 101207. https://doi.org/10.1016/j.elerap.2022.101207

Saputra, A., & Suharjito, A. (2019). Fraud detection using machine learning in e-commerce. International Journal of Advanced Computer Science and Applications, 10(9), 333–339. https://doi.org/10.14569/IJACSA.2019.0100943

Shaji, J., & Panchal, D. (2017, April). Improved fraud detection in e-commerce transactions. In 2017 2nd International Conference on Communication Systems, Computing and IT Applications (CSCITA) (pp. 121-126). IEEE.

Shafiyah, N., Shaker, H., Alsaqour, O., & Uddin, M. (2013). Review on electronic commerce. Middle East Journal of Scientific Research, 18, 1357–1365. https://doi.org/10.5829/idosi.mejsr.2013.18.9.12421

Shahid, A., Keshav, K., & Jenifur, M. (2016). A review paper on e-commerce. TIMS 2016-International Conference, Gwalior. https://www.researchgate.net/publication/304703920_A_Review_Paper_on_E-Commerce

Shini, R. (2018). Detection of fraudulent sellers in online marketplaces using support vector machine approach. International Journal of Engineering Trends and Technology, 57(1), 48–53.

Shpyrko, V., & Koval, B. (2019). Fraud detection models and payment transactions analysis using machine learning. SHS Web of Conferences, 65, 02002. https://doi.org/10.1051/shsconf/20196502002

Sharmila, S., & Suvasini, P. (2018). Detection of automobile insurance fraud using feature selection and data mining techniques. International Journal of Rough Sets and Data Analysis, 5, 1–20. https://doi.org/10.4018/IJRSDA.2018070101

Suryanarayana, S. V., Balaji, G. N., & Rao, G. V. (2018). Machine learning approaches for credit card fraud detection. Int. J. Eng. Technol, 7(2), 917-920.

Vlasselaer, V., Eliassi-Rad, T., Akoglu, L., Snoeck, M., & Baesens, B. (2016). GOTCHA! Network-based fraud detection for social security fraud. Management Science, 63. https://doi.org/10.1287/mnsc.2016.2489

Youssef, B., Bouchra, F., & Brahim, O. (2021, June). Rules Extraction and Deep Learning for e-Commerce Fraud Detection. In 2020 6th IEEE Congress on Information Science and Technology (CiSt) (pp. 145-150). IEEE.