International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V73I6P114

Volume 73 Issue 6, 157-172, June 2025

Original Article

E-commerce Security: An Overview of AI Driven Threat / Anomaly Detection

M. Sakthivanitha[1], S. Silvia Priscila[2], Praveen B.M.[3]

[1]*Institute of Computer Science and Information Science, Srinivas University, Karnataka, India.*
[1]*Department of Computer Applications, Vels Institute of Science, Technology & Advanced Studies, Chennai, Tamil Nadu, India.*
[2]*Department of Computer Science, Bharath Institute of Higher Education and Research, Tamil Nadu, India.*
[3]*Institute of Engineering and Technology, Srinivas University, Karnataka, India.*

[1]*Corresponding Author : sakthivanithamsc@gmail.com*

Received: 16 December 2024 Revised: 19 May 2025 Accepted: 03 June 2025 Published: 28 June 2025

Abstract - *The current marketplaces have changed from physical to online electronic markets (e-commerce) due to substantial advancements in information technology. Governments worldwide, especially those in developing countries, have supported and encouraged smaller and medium-sized businesses to conduct business online. Businesses need to discover a better approach to integrating e-commerce (EC) while maintaining its security as it develops. Resolving the core issue of insufficient security on EC web servers and customers' computer systems is necessary for the rapid expansion of EC. This study aims to conduct a Systematic Literature Review (SLR) to identify the vulnerabilities found in EC platforms. Globally, EC has enhanced consumer convenience and technology. EC has a fraud issue. Platforms and merchants combat fraud to safeguard their clients and companies. One effective technique for spotting odd trends and possible fraud is anomaly detection. The article addresses various methods for putting anomaly detection technology into practice and examines its application in EC fraud detection. Therefore, security concerns and suggested fixes pertaining to EC systems will be discussed in the conclusion. These findings can be used as a starting point for more research on EC security concerns, including suggestions and solutions for a safe EC platform.*

Keywords - *E-Commerce Security, Artificial Intelligence (AI), Threats, Anomaly Detection, Machine Learning (ML), Deep Learning (DL), Fraud Detection.*

1. Introduction

The most common use of the Internet is to make purchases, which allows businesses all over the world to advertise their products online. Amazon is an online marketplace as well. Fast and simple communication is made possible via the Internet, as used by Facebook, Instagram, WhatsApp, and other social media platforms. As a result, a privacy policy that protects the integrity of communication and its users must be followed [1]. The rules, procedures, and safeguards to create safe and secure online transactions are called EC security. The dimensions of EC security include:

- Privacy: Using firewalls, encryption, and other safeguards to protect consumer data from unauthorized internal and external threats.
- Authentication: Verifying users using distinct credentials, including usernames, login IDs, passwords, or email addresses. Brands can employ multi-factor authentication to lower authentication fraud because criminals can impersonate users with weak passwords and login IDs.
- Non-repudiation: Because of cryptographic evidence, like digital signatures, neither the company nor the client can dispute a transaction that took place.
- Integrity: Ensuring accurate and unaffected recording of shared consumer data. Customers' confidence will be damaged if information is intentionally or inaccurately manipulated.
- Confidentiality: Defence mechanisms against the unapproved release of data
- Availability: protection against the deletion or delay of data.

The continued rise of EC and its integration into many parts of daily life has increased the cost of advanced security measures. The advancement of AI and ML technology has enabled EC businesses to improve their security procedures and protocols for safeguarding sensitive company and customer data. AI and ML will revolutionize EC security positively during the next ten years [2].

AI increases EC security by detecting and preventing fraudulent transactions. ML models detect anomalous

tendencies, such as abnormally high-value purchases or abrupt changes in billing information, which may indicate fraud. Protecting their financial and personal data protects the organization from losses and wins over customers. The EC is significant for the following reasons.

Securing customer information: In essence, EC security gathers and saves all private data from the client, including bank account information, credit card numbers, address information, and personal information. Identity theft, monetary losses, and personal injury may result from not protecting this data. Because not protecting your consumer data can also result in a loss of your company's reputation.

Eliminating monetary loss: Insecure EC platforms are highly susceptible to cyber-attacks; numerous malevolent actions and cyberattacks are carried out daily to compromise the website. It may surprise you to learn that some 30,000 websites are hacked daily. Customers and the company may suffer financial losses as a result of any breach.

Maintaining a positive reputation: Reputation has become essential for EC businesses since any harm might have catastrophic repercussions. A single data leak could negatively impact the consumer's confidence in the company you represent. Since the majority of consumers first favour websites they trust, any breach from that website will result in a loss of trust and, ultimately, income.

Advantage over competitors: Popular brands are a common source of rivalry for EC businesses. They can, therefore, gain an advantage over their competitors and other businesses by showcasing a strong commitment to security. Smaller businesses can gain a competitive edge by concentrating on their security procedures and providing clients with a secure shopping environment. EC fraud is prevented and detected using ML and AI through data analysis, pattern recognition, and trend adaptation. By increasing the precision and effectiveness of fraud detection, these technologies lessen the need for review by humans and rule-based systems.

ML and AI are the key advancements that can help add more intelligence, adaptive technology, and efficiency to fraud detection and prevention in EC. One key application is anomaly detection. ML algorithms can determine anomalies within transactional data that are out of the ordinary from patterns typically seen and assign an anomaly flag for additional scrutiny. AI can also assist with risk scoring by synthesizing numerous data points - such as transaction history, behaviour, device information, and geolocation - to determine the confidence level for fraud. This can allow EC to place enhanced levels of authentication through multiple-factor authentications or require human review for high-risk transactions. Predictive analytics allow EC companies to identify fraud from past trends, allowing them to make proactive decisions regarding exhibiting fraud activities. Company behaviour analytics do the same with behaviours - such as noticing unusual activity could indicate a prior account takeover - by taking multiple behavioural readings measuring the established behaviour to alert on anything that deviates. The last benefit is having real-time monitoring, where EC sites can identify issues immediately, analyze them, and take mitigation action while limiting the windows of opportunity for the attacker. The other component allows detection and prevention methods with ML and AI to reduce their effectiveness, reduce false positives, and offer a better customer experience; organizations sacrifice no proper security for an improved customer experience to make "better security decisions." Unlike conventional means, ML and AI interventional methods are very collaborative with traditionally validated and fully efficient methods.

1.1. E-Commerce Security Threats

1.1.1. Phishing

Phishing attacks occur when hackers attempt to gain a victim's personal information by impersonating a business over the phone, text, or email. They will try to trick the user into revealing important information, including account numbers, passwords, and other private data. Phishing attacks, which typically include social engineering and messaging tools, have risen in popularity as a means of obtaining private user data such as passwords, bank transaction details, social security numbers, and so on. Setting up sophisticated phishing detection mechanisms on computers is critical for safeguarding users from such threats [3]. The cornerstone of phishing attacks is to direct the user to a website that resembles legitimate. These sites look to be from a reputable business or web application. Nowadays, a phishing website can be created utilizing various complex technologies to appear legitimate, and non-technical people can quickly set up a phishing website. As a result, even inexperienced users may prefer this type of attack because it is simple. Phishing attacks can be carried out by intruders who target a specific user group or many people.

1.1.2. Structured Query Language (SQL) Injections

Hackers can obtain critical information in your company's database by manipulating everyday coding language with rogue commands. SQL injection attacks allow hackers to get beyond a website's authentication measures, gaining access to data and perhaps jeopardizing the system. A security flaw known as SQL injection (SQLI) happens when unauthorized SQL code is added to online applications to corrupt databases. This can seriously affect an organization, including data breaches, server outages, and data loss [4].

1.1.3. Malware

Malware is malicious software, any code or program created to compromise systems, steal data, or interfere with regular operations. Malware code's nature includes a range of traits and actions that specify its intent and capabilities.

Hackers can use malware to spy on people, obtain backdoor access to networks, and steal client data. Adware, trojan horses, and ransomware are just a few of the various types of malware. The hacker's sole objective is to infect a victim's PC with malware. Direct attacks are difficult or impossible since some firewalls secure most computers. As a result, hackers try to fool the computer into executing the harmful code. Using papers or executable files is the most popular method for doing this.

1.1.4. Brute Force Attacks

Brute force is a technique for account hacking that involves guessing passwords, as shown in Figure 1. It is among the most often used techniques for breaking website account passwords. The Brute-Force programme creates and verifies password variations or lists. Hackers can guess passwords until they find a combination that works using botnets and specialized software. Once they have access to your client database, they can profitably sell it using customer names, bank account information, and other data. The best defences against brute-force attacks include captcha challenges, two-factor verification, and complicated password requirements [5].

1.1.5. Distributed Denial of Service (DDoS)

An anonymous IP address can flood and spam a website with traffic to compromise it. Regular users will no longer be able to access your website once it has been infiltrated by malicious traffic. A DDoS attack is more difficult to defend against since it originates from several sources, whereas a DoS attack only comes from one [6].

1.1.6. Social Engineering

Social engineering (SE) is a broad term used to describe various malicious actions performed through human relationships. It manipulates users' thoughts to trick them into revealing personal information or committing security mistakes. SE attacks involve one or more steps. Before the attack, a criminal investigates to gather information about the target, such as possible entry points and weak security systems. The attacker then tries to connect with the victim, gradually building trust and preparing for further security breaches, such as sharing personal information or gaining unauthorized access to important resources.

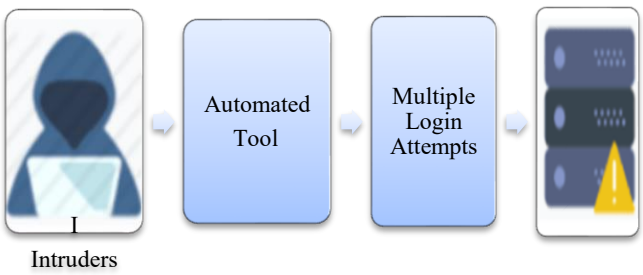

Fig. 1 Brute force attacks

1.1.7. Bots

A bot is an application meant to do specific functions. Humans do not need to launch bots; they can function autonomously and follow instructions. Many bots are designed to accomplish duties that humans would typically complete, such as repetitive tasks, but they do so much faster than humans. Hackers can execute DDoS assaults using a botnet, a network of bots that send bogus requests to a web server to overload it and disrupt its functionality [7]. Certain attackers create customized bots that can scrape a website for pricing and inventory information. To reduce revenue and earnings, these hackers, typically business rivals, can subsequently utilize the data to alter or lower the prices on their websites.

1.1.8. Financial Frauds (FF)

Nowadays, fraud is a big problem for many people. There are numerous methods to define FF. From one perspective, FF exploits regulatory flaws in financial products to generate illicit profits. Any unfair practices in the financial market that prioritize one's interests over others is commonly known as FF. Depending on the nature of the financial products, there are various forms of fraud, such as securities fraud, insurance fraud, bank card fraud, bill fraud, deposit fraud, and loan fraud. Depending on its origin, fraud can be divided into two: internal and external. Figure 3 presents the fraud detection process in EC.

- Attempting to obtain illegitimate benefits by conducting unlawful operations or intrusions through the banking transaction system. This category includes identity theft, bank card fraud, and other internal bank infractions, corporate process fraud, or fraudulent activities that take advantage of weaknesses in corporate procedures to make money, is typical.
- Conveying fraudulent information about credit guarantees or deceptive promises. This category includes the majority of financial and investment fraud. Scammers frequently use High-profit investments as bait to get investors to provide financial fraud information.
- Creating information imbalance fraudulently and hiding crucial information. Fraud in the securities market includes a lot of insider trading, purposefully concealing hazards when promoting derivative products, and manipulating the securities market in many ways to profit from arbitrage [8].

1.2. ML/DL Techniques in E-Commerce Security

ML is a subfield of AI that includes methods that may gain insight from and forecast data. ML can improve overall security posture, automate actions, and increase threat detection in cyber security. ML continuously scans the network's activity for irregularities to identify risks. Large volumes of data are processed by ML engines very instantly in order to identify important events. These methods enable

identifying policy infractions, unknown malware, and insider risks. The framework in Figure 2 can detect ransomware that was never detected before trying to infect endpoints. It uses the characteristics and actions of known malware to identify new dangerous files and activities. ML can protect productivity by spotting risks and hazards in cloud platforms and apps, spotting anomalies based on location, and evaluating questionable cloud app login behaviour.

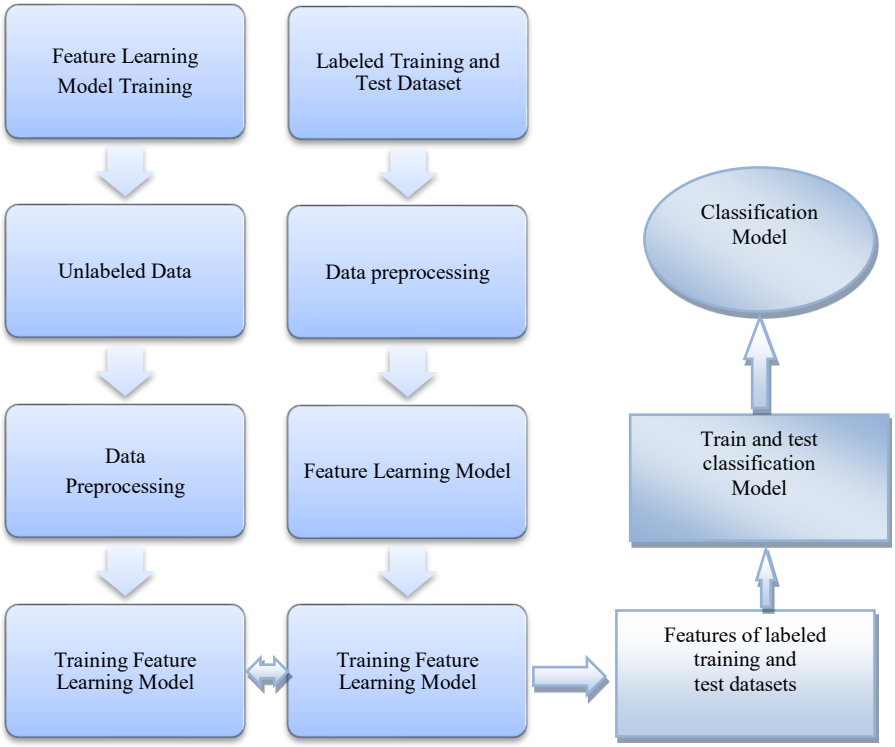

Fig. 2 Fraud detection process for financial fraud

Analyzing traffic data in encrypted form in standard network telemetry allows ML to identify malware in encrypted communication. ML algorithms identify harmful patterns to uncover risks concealed by encryption instead of decrypting. The following are some crucial tactics that online retailers might use:

1.2.1. Detection of Anomalies

One of the primary applications of ML in cyber security is this. ML systems can identify deviations that indicate fraudulent activity by defining a norm of regular user behaviour. For example, the system may flag users for more investigation if they abruptly try to purchase from a different geographic location or if their transaction frequency noticeably increases. ML and predictive algorithms keep up with new threats and continuously enhance fraud detection. The general procedures for AI-based threat/anomaly identification are depicted in Figure 3. The following are the essential steps in the AI threat/anomaly detection process:

Data collection and preparation involves gathering data from several sources and preparing it for analysis, which may involve data normalization, cleaning, and segmentation to ensure the AI model.

1.2.2. Feature Selection

Identifying the most crucial data points to let the AI model distinguish between typical and abnormal patterns. Identification of the dataset's informative qualities, or "features," comes next once the data has been gathered. The selected traits are indicative of behavioural tendencies. The AI accelerates and increases the accuracy of the anomaly identification process by focusing on these traits.

Training the model. Employing historical data to train the AI model allows it to learn the typical activity within the dataset. The anomaly identification process is based on developing a dynamic baseline or "typical" behaviour model. This usually means using ML or statistical methods for the selected features. The optimal strategy, whether density, aggregation, or another method, depends on the organization's objectives and system requirements. Identifying anomalies or hazards. Following training, the model may identify anomalies by examining and contrasting new data with the learned patterns.

1.2.3. Feedback Loop

Implementing recommendations to enhance the algorithm's performance and responsiveness further.

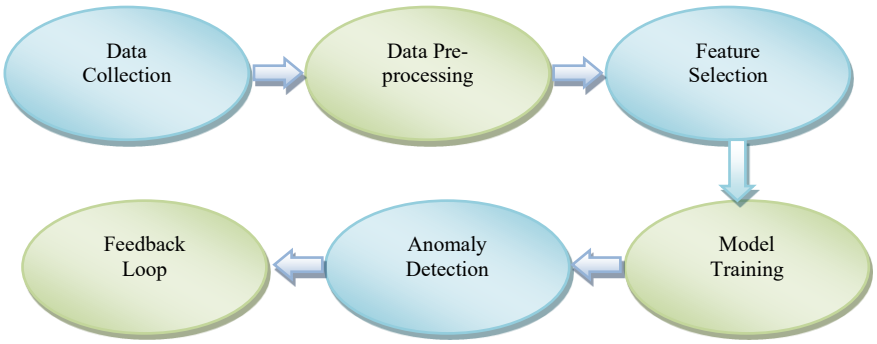

Fig. 3 AI-based threat/anomaly detection processes

1.2.4. Predictive Analytics

It is the process of using ML to forecast possible cyber threats by examining past data. Predictive models can predict the possibility of future incidents by looking at patterns from past attacks. EC platforms can fortify their defences prior to attacks.

1.2.5. Fraud Detection

Real-time transaction data analysis by ML algorithms can spot fraudulent activity. These systems can identify irregularities and flag questionable transactions for manual review or automated blocking by analyzing a number of variables, including transaction value, user activity, and device reputation.

1.2.6. Threat Intelligence

Gathering and evaluating data from several sources to spot new risks is part of using ML for threat intelligence. EC companies can stay ahead of attacks and put preventive measures in place by examining threat trends and Indications of Compromise (IOCs). Online "bad neighbourhoods" can be predicted by ML, which helps stop users from visiting dangerous websites. ML automatically detects attack infrastructures set up for emerging and existing threats by analyzing Internet behavior.

1.2.7. Automated Response Systems

Automated response systems that promptly address hazards can be made possible by ML. For instance, the system can instantly isolate impacted accounts or ban the suspicious IP address when an anomaly is found, significantly speeding up response times.

Despite the growing dependency on EC, many platforms are still susceptible to various forms of cyber threats, especially fraud. It remains a relevant issue that the security of EC web servers and end-user systems is inadequate, as they tend to be the target of malicious attacks that hold sensitive identity and financial information. Although technology continues to advance, massive strides remain in systematically identifying and classifying vulnerabilities in an EC environment. Further, although fraud detection has become paramount in creators and vendors, many methods are straightforward rule-based models that hardly consider the complexities of fraudulent behaviour, making it hard to understand if the rule still applies to new fraudulent behaviours. Anomaly detection has shown promise by detecting anomalies away from standard transaction patterns to improve EC fraud detection mechanisms. However, while there is much literature on the application of anomaly detection on EC fraud detection systems, it has not yet constructed a clear view of how vulnerability anomaly detection has been applied in the field of EC for FD.

To address this need, this study conducts an SLR to identify and classify the vulnerabilities found in EC platforms. It examines the use of anomaly detection in minimizing fraud. The SLR helps make sense of the available literature by pulling together existing research findings to provide a complete picture of the challenges faced by EC platforms, proposed solutions and emerging technologies. The findings can be used as a baseline for future research and, ultimately, create more robust and secure EC systems.

2. Literature Review

AI solutions in the finance sector are transforming how organizations engage with risk management, fraud mitigation, and customized investment services as they make your systems safer and enhance your value propositions. In the EC, AI and ML provide value to the user experience through intelligent chatbots, dynamic pricing strategies, and customized recommendations. Concerning big data and IoT, Rane et al., 2024, address how organizations can align AI with big data interpretation and IoT to develop more innovative businesses and solutions through data [9]. The authors examine potential ethical dilemmas, including algorithmic integrity, bias, data privacy, related concerns, and regulatory policy. The assessment pointed to meaningful topics for future research, such as the potential for AI in supporting organizations that promote sustainability and the need for Explainable AI (XAI) solutions to increase transparency and confidence.

The study employs an SLR methodology and searched and selected relevant studies into EC security and anomaly detection, including primarily studies on developing

countries, and studied English peer-reviewed journal articles, conference proceedings, and review articles, published for the first time between 2013-2024. Studies were selected if they focused on EC security, fraud detection, anomaly detection, and/or application of ML and DL techniques in an EC context, especially if they outlined applications and implications in developing countries. Studies that were not peer-reviewed, not focused on EC, duplicates, and those not available in full-text were excluded. The literature was identified through search engines and/or databases to acquire an exhaustive literature review. The specific search engines/databases were the IEEE Xplore, ACM Digital Library, Science Direct, SpringerLink, Scopus, Web of Science, etc. Search queries utilized 'Boolean' operators and relevant keyword combinations, e.g., "EC security", "anomaly detection", "fraud detection", "machine learning", and "developing countries", with appropriate searches/filters/modifiers. The selection process was robust to ensure that this systematic literature review used relevant peer-reviewed and high-quality literature.

The study by Ma et al., 2024, aims to identify efficient technological security solutions that will improve the sustainable growth of business models, safeguard the rights of users and data, and strengthen the security of EC platforms [10]. The authors explore the creation of the Big Data-enabled EC Business Model in response to identifying the challenges with EC platforms and positively increasing the awareness of those who make decisions about marketing approaches.

Reddy et al., 2024, review various predictive models and how they are applied in a real-world context, representing substantial enhancements to the performance of operations indices such as customer satisfaction, inventory turnover, and order delivery [11]. It also highlights how real-time data streams and ML algorithms can predict consumer behaviour, exploit Supply Chain (SC) capabilities, and provide an enhanced delivery model. Joshi et al., 2024 explore the AI methodologies to improve the multiple facets of EC, including but not limited to consuming interactions, personalizing, prediction systems, fraud detection, inventory management, and supply chain operations.

By leveraging AI technology, such as ML, computational language-based modellers, computer vision, and predictive analytics, EC organizations can advance their operational execution and improve their decision-making loop while offering their customers tailored experiences [12]. Gorantla et al., 2024, developed a framework for improving security and data storage in EC applications through hybrid clouds [13]. Hybrid, cloud-based implementations include the latest community-driven cloud security models to assure data safeguarding while providing the security, multi-functionality, capacity, and economies of scale available to traditional cloud computing infrastructures. Hybrid cloud deployments offer greater availability, better data processing capabilities, and enhanced storage performance than traditional on-premises equivalents.

2.1. ML Techniques

Many studies have published their results about their phishing attack awareness, detection, and prevention findings, yet there is no comprehensive and appropriate way of prevention right now; therefore, ML is crucial to safeguard against phishing-related cybercrime.

Nalla et al., 2024 investigate AI methods that can measure consumer behaviours, preferences, and purchasing behaviours using approaches like ML, NLP, and NNs. It also discussed how AI could inform trends in the marketplace, improve SC management, and increase customer satisfaction. The research paper highlighted the challenges and potential of using AI in large data environments, such as the need for scalable systems and data privacy challenges [14].

Hasan et al., 2024 examine the challenges and issues of detecting fraudulent online payments and EC transactions. AI and data analytics solutions detect, deter, and regulate online fraud in banking and EC transactions [15]. Lastly, Rahman et al., 2024 present a novel price-scraping attack scenario and outline the steps to execute the invented attack scenario. Additionally, the study uses time and byte entropy analysis to discuss security mechanisms [16].

The reason for the investigation by Li et al., 2024 is to distinguish EC extortion, utilize Big Data Mining (BDM) to address the financial dangers of EC businesses, examine more proficient arrangements utilizing Information Fusion Technology (IFT), and create an EC fraud detection based on IFT, AI, and information mining (DM). Meanwhile, EC extortion concerns are considered when utilising BDM innovation, the LR Model (LRM), SVM, and the recommended IFT-based FDM [17]. Karim et al. (2023) use the popular ML models, which combine LR, SVC, and decision tree (LR+SVC+DT) with soft and hard voting, to accurately and successfully defend against phishing attacks using the dataset repository's URL-based dataset [1].

Bell et al. (2020) showed how ML methods could resolve categorisation problems by leveraging URL properties. Functional attributes for training were selected based on a working phishing detection algorithm. PHISHTANK is an online tool being engineered to detect phishing attacks. PHISHTANK employs multiple characteristics to assess whether a site is safe or malicious. They develop a URL structure to detect a phishing attack with the URL [18]. Reddy et al. (2024) propose a big data platform to assist online retailers in dealing with some of the issues associated with EC. Fraud is a global issue that can affect both individuals and companies. AI and ML have proved crucial in the fight against fraud in today's tech-driven world [19].

Tung et al. present a new real-time SQLI attack detection method. The system is optimized by using ML techniques to practice and enhance its ability to identify and defend against SQLI attacks accurately. Machine learning algorithms include CNN, LR, NB, SVM, and RF. The system encompasses project pre-development, data preprocessing, feature selection, ML model selection, training, testing, deployment, and evaluation. Incorporating the system in the web application server's backend would enhance the online application's safety and security features [4].

Golyeri et al., 2023 use a new dataset that includes different features regarding online shopping behaviours on Boyner Group's EC site and application. Four fundamental ML algorithms, DT, LR, RF, and XGB, are used to detect fraud in EC transactions. This work is progressive since it experiments with various ML classifiers and uses features that differ from existing literature and methods [20].

Dong et al., 2021 propose a technique for identifying fraud in the e-market in real-time, the SVM-based Fraud Detection Framework (SVM-FDF). The FD framework is used to improve a product's exposure and reputation when it is presented to the e-market, replacing a limited marketing plan that aims to draw in new clients [21].

Potla et al.,2024 look at the potential of real-time ML models to change fraud detection. Where static, rule-based systems infrequently learn from the past, a real-time ML model learns continuously from big data sets and flagged behaviour in real time. Real-time FD increases the likelihood that legitimate client transactions will not be delayed, and as FD speeds up, so will the number of false positives generated by real-time ML.ML has been able to look through vast quantities of transaction data across the industry in a more timely manner than it was ever able to before-supervised ML models for classification tasks, including RF and Gradient Boosting Machines. Unsupervised learning methods, which are used for anomaly detection, include Autoencoders [22].

Rajeshwari et al., 2024 examine how using quantum networks to detect fraud in EC transactions may cause a significant change in online security. Quantum networks, which are based on quantum physics concepts, provide the greatest level of data security [23]. Companies that support online payment collect an enormous amount of data on customer transactions and use ML, ML algorithms to discern fraud from legitimate behavior. ML methodologies were used to identify online payment fraud in EC transactions to enhance fraud detection abilities.

The dataset consisted of structured data at the transaction level to examine and identify patterns that separate non-fraud from fraudulent actions. Modelling approaches such as Light Gradient Boosting Machines (LGBMs) typically utilize feature engineering when identifying patterns, such as aggregating user-level metrics such as mean and standard deviation. The fraud detection problem is very complicated due to the difference in the data of fraud and non-fraud. The metrics used to examine the performance of the model include but are not limited to, accuracy and an F1 score. The current process implements a refinement of the LGBM and XGBoost models that employ Bayesian optimization strategies.

Wang et al. 2024, suggest a novel method that combines Multi-Layer Perceptron (MLP) with the ensemble model. In particular, this method uses the advantages of the first-layer feature extractors RF, GBDT, and XGBoost. The remaining training samples are fed into the second-layer MLP after samples incorrectly identified in the first layer are eliminated. This method reduces the issue of over-reliance on samples, improves the model's capacity for generalization, and fortifies its global modelling capabilities to a certain degree. The Gitee dataset demonstrates that this two-layer model is more accurate than traditional ensemble models and can extract valuable information from the data [24].

2.2. Deep Learning Techniques

DL algorithms are in high demand because of their tolerable resilience, high recognition rate, and excellent implementation. A DL-based quantitative financial fraud detection system for EC large data is proposed by Liu et al. in 2021 [8]. In addition to limiting feature extraction to the space-time volume of the dense trajectory, the encoders will be used to detect the behaviour initially, hence reducing computing complexity. Second, feature fusion will take place utilising weighted correlation techniques to give the features classification capabilities once the neural network model has produced behavioural visual word representations for the features. Finally, financial fraud will be assessed and categorised using sparse reconstruction mistakes. Artificial Neural Networks (ANNs) are disrupting EC because of improvements to personalized experiences, operational efficiencies, and security. ANNs can provide advanced approaches to deal with large amounts of data, recognize complex patterns, and generalize these findings. This means they can thrive in fast-paced, data-rich, trend-responsive EC platforms. ANNs are effective when applied to large amounts of data and employ data and information processing methods modelled after the behaviour of biological neural networks. Their ability to be trained using vast quantities of high-velocity transaction data accounts for their appeal in credit card fraud detection [25].

Sahingoz et al. created a DL-based phishing detection system employing five distinct algorithms: ANN, CNN, Recurrent Neural Networks (RNN), Bidirectional RNN (Bi-RNN) and Attention Networks (AN). The system primarily aimed to classify web pages as quickly as possible using URLs. A larger dataset of labelled URLs with over five million records was collected and published to test the system [3]. Nafees et al., 2024 propose a threat detection framework

for large EC communication settings, which is ML-based. More specifically, DL in ML is incorporated. The proposed threat detection framework uses LSTM Deep Neural Network (DNN) and DL methods [26]. A type of cybercrime known as phishing entails obtaining login information from online platforms such as online marketplaces, online banking, online businesses, EC, and online educational institutions. Phishers send unwanted emails to people and make fake websites that look like the real ones. Phishers exploit a user's login credentials when they visit a phoney website through spam. Researchers have created strong tools, including antivirus software, whitelists, and blacklists, to detect phishing websites. Attackers constantly devise new strategies to circumvent cyber defences by exploiting network and human weaknesses. A DL-based data-driven system for detecting fake websites is presented by Saha et al. in 2020. More specifically, an MLP, a feed-forward NN, is used to anticipate phishing web pages [27]. Retailers use shrewd strategies like product reviews and ratings to build a strong brand and increase sales and profits. This may indirectly affect the purchase made by a customer unaware of the retailer's actions. A study by Chatrath et al., 2022 established a handling strategy for repurposed product photos on EC websites based on user sensitivity [28]. Product Image-based Vulnerability Detection (PIVD) uses three phases of CNN to detect dishonest traders, allowing customers to acquire products more confidently and with fewer damages-table 1 lists ML and DL models in the recent literature.

Table 1. ML and DL models for anomaly detection

Author /Year	Methods	Purpose	Score /Evaluation
Karim et al., 2023[1]	DT, LR, RF, NB, GBM, KNN, SVC and LSD	Phishing URL's	Precision, accuracy, recall, F1-score, and specificity,
Tung et al., 2024[4]	LR, NB, SVM, RF	SQLI Attacks	Accuracy
Rane et al., 2024[9]	ML, IoT, XAI	Intelligent chatbots, dynamic pricing and tailored recommendations	-
Golyeri et al., 2023[19]	DT, LR, RF, XGB	FD	Accuracy
Dong et al., 2021[20]	SVM-FDF	Real-time FD	Accuracy, Recall, Precision, Error Rate
Potla et al., 2024[21]	RF and Gradient Boosting	FD	Accuracy
Rajeshwari et al., 2024[22]	Quantum Networks, LGBM and XGBoost models, SMOTE	Detect fraudulent online payments in transactions made via EC.	Accuracy, F1-Score
Wang et al., 2024[23]	Ensemble model- RF, GBDT, and XGBoost, MLP	Analysis of anomalous data in online orders	Accuracy
Sahingoz et al., 2024[3]	ANN, RNN, CNN, Bi-RNN, AN	Phishing detection system	Accuracy 98.74%
Nafees et al., 2024[26]	LSTM, DNN	Detection of threats	Accuracy
Bensaoud et al., 2024[29]	Deep transfer Learning	Malware detection	
Liu et al., 2020[8]	Stacked Denoising Encoder	Fraud detection	Precision, recall, AUC-ROC
Saha et al., 2020[27]	MLP	Phishing attacks	Accuracy
Chatrath et al., 2022[28]	CNN	Handling strategy for product images that have been repeated based on user vulnerabilities	F1-Score

The studies reviewed represent a growing and changing field of study in EC security, where the different ML and DL methodologies are widely used to address a range of threats (e.g., relevant fraud, phishing, malware, and SQL injection attacks), and there seem to be various variations of their ML and DL applications.

The resurgence of FD is seen, whereby research was primarily focused on tackling fraud via hybrid and ensemble models to improve anomaly-based detection and deal with the challenge of imbalanced data. Phishing and SQL injection threats are also comparatively established and comparable to fraud, as models using CNNs and traditional classification models were identified. Most studies use evaluation metrics associated with accuracy and precision, recall, F1-score and AUC-ROC, focusing on studies balancing the detection performances whilst attempting to control false positives. Importantly, while many models demonstrated high performance (e.g., a phishing detection accuracy of 98.74%), several models depended on synthetic or proprietary datasets, limiting benchmarking and reproducibility efforts.

In addition, more advanced techniques, such as quantum networks and explainable AI (XAI), are beginning to emerge, although they merit further validation. The primary gap identified is in the implementation and integration with live

EC activity, the model(s) explainability and regulatory compliance. Overall, while there appear to be innovative increases in the field, there is still considerable work to do before establishing these security solutions in real-world EC environments reliably and trustily.

2.3. Statistical Evidence and Case Studies

According to Juniper Research, annual global losses due to online payment fraud are expected to surpass $48 billion annually by 2025, almost doubling from $26 billion in 2020. This unsettling trend is coupled with PwC's 2023 survey findings that 85% of consumers stated they would avoid a company with poor data security, and 67% would 'never go back to that platform' after only one data breach. In addition, IBM's 2023 "Cost of a Data Breach Report" stated that the average data breach cost in retail was $3.28 million, with customer attrition being the most significant contributing factor in addition to regulatory fines. High-profile events such as the eBay data breach, which compromised 145 million users, and British Airways' £20 million GDPR fine for a payment card data breach highlight some of the operational and reputational risks faced by companies who get this wrong. Companies like Amazon still struggle with fraud from compromised seller accounts, totalling losses on buyers and legitimate sellers. These examples show how cybersecurity lapses can risk huge costs, loss of trust, and potentially irreparable damage to the business's credibility. For EC businesses, and especially EC businesses in a fast-changing and digitizing world, a business must ensure robust security not just for compliance and data breach mitigation but for customer confidence and competitive survivability. Some of the notable case studies are as follows.

In 2025, UK retailer Marks & Spencer (M&S) suffered a significant cyberattack perpetrated by the Scattered Spider hacking group. The cyan pressure point was breached because it used third-party contractor systems, resulting in the suspension of M&S's online activities. It was estimated that this incident cost M&S up to £300 million (US$400 million), leading to a decline in M&S's market value of £1 billion. M&S also breached customers' personal data, names, and contact information. This case alarmingly identified vulnerabilities related to reliance and dependencies linked to third parties, stressing significant cumulative financial and reputational risks of breaches [30].

In 2023, personal Genomics Company 23andMe experienced a breach of personal data that affected approximately 6.9 million users' sensitive personal and genetic data. 23andMe was breached by a credential-stuffing attack from hackers who used a series of previous data loss data leaks that involved testing user passwords. The 23andMe breach raised substantial privacy issues and generated multiple class action lawsuits and regulatory investigations, as well as evidence of the necessity for a functioning user authentication strategy for customer accountability for sensitive data [31]. In 2024, the known retail brand Hot Topic suffered a data breach presenting personal data for over 54 million customer accounts. Names, emails, physical addresses, phone number purchase history, and partially encoded credit card numbers were all compromised. The malware was identified as an information stealer, and the attackers were able to breach Hot Topic's 3rd party data unification service, which also compromised the Hot Topic customer's login credentials. This case demonstrates the threat of third parties and securing elements of the supply chain [32].

These case studies provide perspectives on the scaling danger of cyber threats that companies face in the EC environment and the sensitivity of customer trust and business continuity. These examples highlight the need for a robust understanding of cybersecurity threats, such as being proactive with regular security audits, familiarising employees with cyber risks, using strong authentication practices, and monitoring your third-party ecosystem. The more we do to mitigate risk, the better our chance to protect customer information and our company brand.

3. Materials and Methods

ML algorithms are trained to process enormous datasets and automatically find recurrent patterns or clusters among important variables and data points. An ML system may have detected an anomaly when encountering data that does not match any preexisting patterns. Figure 4 shows the ML models in anomaly detection in ecommerce.

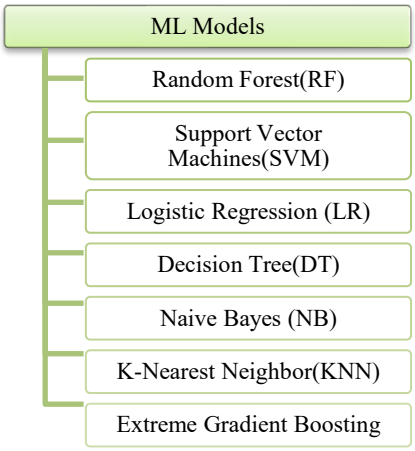

Fig. 4 ML models for anomaly detection

3.1. Random Forest (RF)

The RF classifier exhibits far more desirable and logical findings than the other classifiers, outperforming them in identifying various types of attacks. The RF prediction is based on the majority vote for the result of the aggregated forecasts of many DTs. The dataset will be poured into sub-trees after the decision tree has been built and the optimal set of variables has been determined. However, it is not always simple to determine the ideal combination of variables. One of

its major advantages is that this classifier requires less time. In addition, it classifies enormous volumes of data with a lower percentage error. The RF Classifier was chosen because of its capacity to reliably identify complex patterns and effectively address the multicollinearity issue [33-35].

3.2. Support Vector Machines

SVM is a supervised ML technique used for classification and regression problems. They are extensively employed in several domains, including NLP, image analysis, and pattern identification. A hyperplane is a 1-D plane that splits an n-dimensional plane in half. It is a point in 1D. A hyperplane is a line in a 2D setting. It is a plane in 3D. It is impossible to depict a hyperplane in more than three dimensions simultaneously.

In contrast to earlier classification techniques, SVMs identify the decision border that maximises the distance between each class and the closest data points. The decision boundary produced by SVMs, also known as maximum margin classifiers, is the maximum margin hyperplane. Samples nearer the hyperplane are called support vectors. It establishes the orientation and position of the hyperplane. SVM uses these support vectors to maximize the margin of the classifier. Support vectors can be added or subtracted to change the position of the hyperplane [36].

3.3. Logistic Regression

Data Envelopment Analysis (DEA) is used to verify the link between assessment and stimuli to comprehend consumer perception, lower hazards associated with online shopping, and preserve online security. It creates a stimuli-organism response model and employs regression analysis to investigate the connections between consumer behaviour, risk perception, and unfavourable online shopping assessments. Regression analysis is employed when the dependent variable is dichotomous or binary (i.e., it can only take two values). Modelling the likelihood of a specific result based on one or more predictor variables is the goal of logistic regression [37, 38].

3.4. Decision Trees

Fraud detection can be expressed as a knowledge extraction work with imbalanced classes or a classification problem using case-based reasoning in EC. A popular ML approach for tasks involving regression and classification is the DT. This approach to supervised learning creates a model of decisions and their potential outcomes that resemble a tree. The nodes in the tree structure represent the decision or conclusion, while the potential outcomes are represented by the edges [39-41].

3.5. Naïve Bayes Classifier (NB)

Modelling the distribution of inputs within a class or category is the goal of the generative learning algorithm family that includes NB. Unlike discriminative classifiers like LR, it cannot identify the most important features for class separation. It is extensively utilised in recommendation systems, spam filtering, and text classification. The algorithm may be able to produce predictions quickly and accurately with this method, which is predicated on the idea that the properties of the input data are conditionally independent given the class [42].

3.6. K-Nearest Neighbour (KNN)

An algorithm for supervised ML is the KNN. Despite its ability to conduct regression, KNN is primarily employed for data point classification. The KNN Algorithm categorizes new data points according to their similarity to those in that category. The KNN algorithms translate the input data to the target data by grabbing the freedom to choose any functional form from the training data. Since KNN makes no particular assumptions about the mapping function, it falls within the category of nonparametric algorithms [43].

3.7. Extreme Gradient Boosting (XGB)

A tree-based ensemble technique called XGB generates a final prediction by aggregating the predictions of several DTs. The technique can represent non-linear relationships between variables and works particularly well with high-dimensional data [23, 24].

3.8. Convolutional Neural Networks (CNN)

CNNs are an effective tool for detecting anomalies in EC, especially for image-based, sequence, and multimodal data. CNNs are particularly effective in finding tiny patterns, making them appropriate for detecting anomalies in product photos, user activity, or transaction records. Multimodal detection combines transaction characteristics and visual representations (such as transaction amount histograms) for anomaly detection. CNNs use spatial features extracted from these visualizations to identify outliers [45].

3.9. Recurrent Neural Networks (RNN)

An important DL model explicitly designed for handling sequential data is an RNN. An RNN may capture the relationship between surrounding input values since it has recurrent connections between successive hidden states (ht) and accepts a sequence of discrete time values (xt) as input. An RNN can handle variable-length input data by utilizing all of the model's components share parameters, such as wx, wy, wz, and. Bidirectional RNNs are a common variant of RNNs that add two hidden state layers with hidden cells coupled in two opposite directions. This allows the model to identify information from a sequence that is both past and future [46].

3.10. Long Short Term Memory (LSTM)

LSTM networks are an extremely successful solution for EC security and anomaly detection because of their capacity to examine sequential data and capture temporal connections. EC platforms create massive amounts of time-stamped data, such as user activity, transactions, and navigation patterns,

which makes LSTM an effective tool for detecting abnormalities. A unique type of RNN model called LSTM solves the long-term dependency issue in RNNs by imposing a gate-controlled structure in the recurrent module. LSTM networks use a mathematical model that manages cell state (C_t) and hidden state (h_t) to record long-term dependencies and prevent vanishing gradients. The LSTM model's major components are the input gate, forget gate, output gate, and cell state update mechanism. LSTM cell state is a memory track for long-term dependencies and handles vanishing and exploding gradient issues better than vanilla RNNs [47]. Figure 5 presents the DL models for anomaly detection.

3.11. Auto Encoders (AEs)

Neural networks, or AEs, are very helpful in EC for data reduction, recommendation systems, and anomaly detection. They are designed to learn the input data's compressed form, or encoding, and then use that encoding to reconstruct the input. AEs are a class of NNs that learn a reduced-dimensional representation from input data, which are composed of an encoder that will learn to present input information to a reduced-dimensional representation and a decoder that will learn to present this reduced-dimensional representation back to the original input information.

By organizing the learning problem this way, the encoder network learns a practical "compression" function that allows the input information to be represented as a definable lower-dimensional representation. This allows the decoder to represent the input data using a reconstruction from a lower-dimensional representation. The model is trained by minimizing the reconstruction error (or the MSE), which is the difference between the input to the AE and the output produced by the decoder after reconstruction.

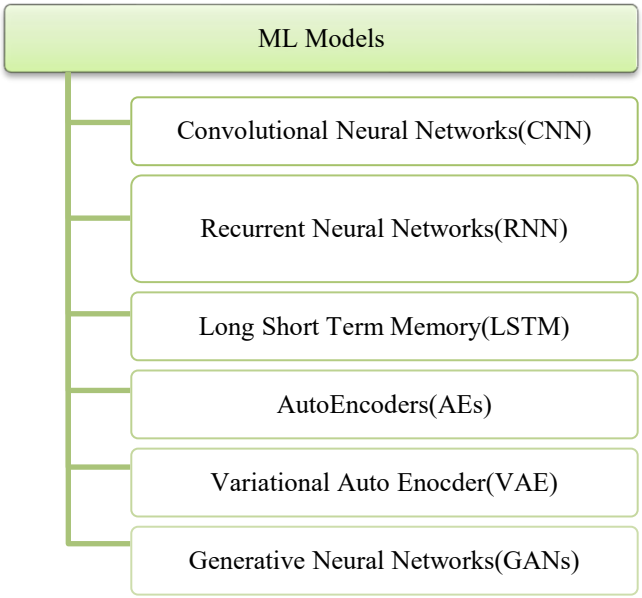

Fig. 5 DL models for anomaly detection

3.12. Variational Auto Encoders (VAE)

The VAE is an extension of the AE. The VAE adds a variational inference component to the learning problem in addition to an encoder and decoder network. Instead of mapping from a fixed constraint input data to a vector (point estimate), a VAE learns a mapping from input data to a distribution. We try to recover what we started with by sampling a latent code of this distribution. According to standard Bayesian terminology, the posterior is the distribution of the latent code given our data, the likelihood is the distribution of the data given the latent code, and the prior is the distribution of the latent code [48].

3.13. Generative Adversarial Networks

Neural networks termed Generative-Adversarial Networks (GAN) are intended to build a generative model of input data distribution. Traditionally, they consist of two neural networks, usually feed-forward, labelled a Discriminator (D) and a Generator (G). The goal of training both networks simultaneously is to find a representation of the distribution of the input data, X, through a competitive skills game. From random noise of fixed dimension (Z), G learns a mapping to sample X similar to samples from the input distribution. At the same time, D gains the ability to distinguish between real samples taken from the source data (X) and fakes (X) constructed by G. The parameters of D are adjusted at each epoch of training to maximize its effectiveness in distinguishing X from X. Meanwhile the parameters of G are adjusted to maximize its effectiveness in creating examples that cannot be distinguished from X by D. As training continues D gets better at distinguishing real and fake samples from X while G becomes better at making samples similar to X [49].

4. Results and Discussions

Anomaly detection has become an important aspect in the fight against EC fraud, enabling systems to monitor for highly dubious behavior that diverges from expected behavior that may suggest malicious activities. While the ability to detect novel threats constantly developing is apparent, employing anomaly detection can be fraught with various technical and implementation challenges. EC environments can be highly complicated and extensive, consisting of diverse customers with highly variable buying behaviours, high transaction volumes, and constant changes in fraud models, which propose several challenges.

These challenges may include processing, in real-time, massive, and heterogeneous data sources, identifying genuine anomalies and abnormal variations, and managing the high level of false positives often produced by the detection system, which may overwhelm analysts and erode the level of trust toward the detection system, and finally having limited labelled data, concept drift with respect to changes in the EC retail fraud, and constantly changing fraud schemes making previously supervised learning categorizations obsolete. For

these reasons, implementing anomaly detection systems for EC environments that are reliable, adaptable, and highly scalable is still ongoing. Advancements must be made in developing algorithms and implementing systems.

4.1. Strengths and Limitations of Current Methods

Anomaly detection methods, including those based on identifying transaction patterns for fraud detection, have been implemented successfully in some EC situations. In general, anomaly detection methods also have much potential, for instance, in identifying previously unknown fraud patterns for possible behavioural changes because they learn spending patterns from history and also allow for feature-based real-time deviations.

Anomaly detection methods are adaptable and scalable, and detecting previously undetected fraud behaviours will likely benefit changing high-volume EC transactions. While there are some benefits, none of the possibilities of the complete potential of these methods are limiting. One of the main challenges has been data imbalance; that is, the ratio of legitimate transactions to fraud transactions is disproportionate (many-one), which makes it very difficult for models to use sustainable features to learn fraud. Additionally, if this problem exists, data imbalance would likely lead to many false positive results, leading to user frustration and, if not, a fracture of the business process.

Another challenge derives from concept drift, or the evolution of fraud over time, while it is clear that models built on dated, static data will be less accurate. This is one of several reasons analysts often find and use less-explainable methods when considering flagged transactions. However, even when reported progress is made in a research context, there is still a big gap in the real world.

The reason that techniques will not be deployed on live EC systems revolves around wondering how to mitigate issues like a lack of well-labelled training and test data sets, computation time, integration, complexity with legacy systems, and data privacy. As a result, mitigating some of these obstacles is an important next step to enhance anomaly detection's credibility and operational usefulness in real-world fraud prevention systems.

4.2. Strategies for Improving EC Security

Educating and raising awareness among employees: Human error is among the leading causes of cybersecurity events. Comprehensive staff training programs will ensure staff members understand their obligations regarding an organization's safety online, the best practices regarding their cybersecurity, and how to recognize if a phishing attempt has targeted them.

Incident response plan: The key to reducing the impact of a cyber catastrophe is having an incident response plan. An incident response plan outlines what the organization should do in case of a security breach. Implementing a prompt/action-coordinated response is necessary to minimize damage to their business and protect customer information.

Frequent security audits and vulnerability assessments: Conducting frequent security audits and assessments will reveal and mend potential vulnerabilities in the EC infrastructure. Organizations may strengthen their security measures by acting proactively while avoiding new vulnerabilities. Employee education and awareness: Human error is one of the biggest causes of cybersecurity issues. Providing employees with thorough training programs ensures that they understand their role in protecting online safety, know the best practices for cybersecurity and can recognize phishing attempts.

Incident response strategy: Reducing the harm caused by a cyber incident requires a well-documented incident response plan. An incident response plan helps to minimize harm and safeguard client information by outlining what should happen in the case of a security breach and facilitating a prompt, coordinated response.

4.3. Trends and Research Gaps Identified

In recent literature published between 2022 and 2025, the scope of investigation into where ML and Artificial Intelligence (AI) can benefit EC fraud detection has rapidly expanded. The most significant observation in the literature is the move from rule-based systems to automated, adaptive, and real-time detection using anomaly detection, DL, and ensemble methods. Recent work also highlights researchers exploring the incorporation of AEs, LSTM networks, and graph-based methods for ML fraud detection, where learner and fraud patterns may be more complex, contextualized, and evolving. While any move to ML is substantial to rule-based systems, an interesting trend in using ML for FD is the application of behavioural analytics, such as user behaviour profiling and device fingerprinting, to inform practical risk analysis. It is also seen that hybrid approaches - combining supervised and unsupervised learning - aim to increase performance and modelling under data imbalance conditions. Finally, we have also noted the emergence of real-time monitoring frameworks that demonstrate time-to-market incentives for FD strategies to be of practical interest and deployable.

While considerable progress has been made, there are still a number of research gaps. The biggest gap is the small number of widespread real-world EC fraud datasets, affecting the results' reproducibility and benchmarks for detection models. Most studies still use synthetic or highly curated datasets, limiting the conclusions' applicability. The next significant gap is concept drift; few studies examine how models will specifically adjust to fast environmental changes, even though adaptive learning is crucial for changing fraud

strategies. The explainability idea is frequently overlooked; for instance, most DL-based models are considered black boxes and are typically rejected due to the business's requirement for auditability and transparency. Finally, while some methods show accuracy in controlled environments, there has been little to no exploration of implementation phases in real-world deployment settings, such as integration with legacy systems, scalability, etc. Addressing these gaps is important for the field and the development of FD systems, keeping the demands of the EC industry in mind that are valid in reality and robust in application.

4.4. Challenges in Anomaly Detection

Anomaly detection systems are developed to identify changes in normal behavior that represent potential security incidents, malfunctioning systems or performance issues. However, it is not easy to detect a large-scale anomaly from an anomaly detection system for the following reasons:

- Massive data: One of the biggest challenges is the volume and variety of data produced in networks. In order to classify an anomaly as benign or threatening, advanced algorithms and credible computing power must be utilized to process the data and distinguish between anomalies meaningfully.
- Complex normal behavior: Normal behavior can change considerably over time; hence, anomaly detection systems must be trained and continue learning.
- Fraudulent events: False positives waste time and resources and might also cause IT staff to become desensitized to real threats, and false positives might generate unnecessary alerts.
- Execution and optimization: A balance between sensitivity and specificity must be struck to implement anomaly detection effectively. If the system is overly sensitive, it might produce too many false positives; if it is too specific, objective abnormalities might escape unnoticed.
- Dynamic distribution of data: User behaviour, product inventory, and pricing patterns on EC platforms continually change, resulting in non-stationary data. Models may require periodic retraining to adjust to new trends.
- Lack of Labeled Data: Anomalies are rarely recognized since identifying them requires significant effort and knowledge. Supervised learning approaches are challenging to implement.

5. Discussions

With the global transition to digital markets continuing to accelerate, EC has changed the game of business conduct and the consumer experience. Especially in developing countries, governments are encouraging More Small and Medium Enterprises (SMEs) to transition to digital platforms for business operations. However, while commerce has moved into the digital sphere and increased its accessibility for the consumer base, there are also substantial security implications, including vulnerabilities in web servers, devices at the consumer level and transactional systems. Moreover, fraud is a cause for concern, encouraging firms to implement sufficient trade application security mechanisms. The implications of this study through an SLR noted the common vulnerabilities associated with EC applications and discussed how anomaly detection technologies could assist in fraud. The implications for developing countries with limited resources and technical know-how are particularly pertinent.

These businesses should have begun utilizing cheaper and scalable options such as cloud security services, open-source anomaly detection systems, or secure payment gateways. In particular, one of the more proactive options would be security auditing, implementing basic but effective measures (i.e. multi-factor authentication), training staff, and educating customers. In addition, creating a mobile-first security model and working with local governments or global cybersecurity initiatives can also help build resilience. In this case, as long as developing countries prioritize practical and contextual solutions, they can help businesses adhere to regulation requirements, create trust, and provide a means of sustainable growth...in the digital economy.

5.1. Future AI Trends and Policy Recommendations

The future of AI concerning the security of EC platforms will be characterized by a greater dependence on self-learning AI that will continuously and accurately respond to new fraud patterns in real-time, new FD methods through multimodal FD leveraging text, transaction data, behavioural data, and biometric data in a single algorithm, and greater utilization of federated learning to create AI models based on data from multiple decentralized sources while preserving each user's privacy. Using XAI can provide better clarity and trust about how complex models make decisions; this is important for compliance and auditing reasons.

Regarding the policy recommendations for EC security, the government/relevant authorities need to develop relevant policies to encourage or promote the development of appropriate means of governing AI to secure EC platforms. Standardizing security protocols and guidelines is important as this provides universal guidelines that will be meaningful in using AI for FD and help ensure security. This provides consistency and interoperability across EC platforms. Privacy and data governance should be recognized as a top priority, with enforceable data protection laws outlining how AI systems identify and use personal data because models utilize sensitive behavioural and financial data. The law should also require transparency and accountability, outlining that AI processes and decisions are explainable, particularly where AI is used to flag or deny a customer transaction. Third-party risks must also be managed - for example, an EC platform should not only be responsible for its acts and omissions. However, it should be held accountable for the security

protocols and standards held by vendors and other integrated services. Public-private partnerships between technology companies, academic research, and governments should be encouraged to share threat intelligence, develop secure AI tools, and open datasets for research and benchmarking. Lastly, there needs to be ongoing assessments and compliance audits to understand how AI models perform long-term and how they need to change or adapt over time to avoid impaired performance (for example, concept drift) or adversarial security vulnerabilities. Taken together, these policy directions will form the basis for a trustworthy consumer environment, less fraud, and safer and more resilient digital commerce

6. Conclusion

EC security is essential for shielding businesses and consumers from online purchasing and marketing risks and hazards. However, EC platforms may ensure compliance, build trust, and safeguard their brand by using robust security measures. ML and DL models have become essential tools for identifying anomalies in EC because of their capacity to process vast volumes of complex data, adjust to shifting trends, and identify minute inconsistencies. Their tool has proven helpful in FD, bot activity tracking, and customer behavior analysis.

However, properly exploiting these models necessitates solving major obstacles and tailoring solutions to the dynamic nature of EC platforms. Hybrid AI systems that integrate many ML models are bolstered by real-time data processing capabilities and are improved by blockchain technology for increased security and transparency, which are the way for FD's future. Developments in XAI transfer learning and semi-supervised techniques are expected to improve the effectiveness of ML/DL for detecting anomalies in EC. Firms can balance accuracy, efficiency, and ethical issues by integrating these strategies with strong data engineering processes.

References

[1] Abdul Karim et al., "Phishing Detection System through Hybrid Machine Learning Based on URL," *IEEE Access*, vol. 11, pp. 36805-36822, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] B. Girimurugan, "AI and Machine Learning in E-Commerce Security: Emerging Trends and Practices," *Strategies for E-Commerce Data Security: Cloud, Blockchain, AI, and Machine Learning*, pp. 29-53, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Ozgur Koray Sahingoz, Ebubekir BUBEr, and Emin Kugu, "DEPHIDES: Deep Learning Based Phishing Detection System," *IEEE Access*, vol. 12, pp. 8052-8070, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Tung, Tean Thong, "*Detection of SQL Injection Attack Using Machine Learning*," Bachelor's Thesis, Faculty of Information and Communication Technology (Kampar Campus), Universiti Tunku Abdul Rahman, 2024. [Publisher Link]

[5] Ruchi Rai, Ankur Rohilla, and Abhishek Rai, "Understanding Cybersecurity Threats in E-Commerce," *Strategies for E-Commerce Data Security: Cloud, Blockchain, AI, and Machine Learning*, pp. 501-522, 2024. [Publisher Link]

[6] Dipok Deb, "*Application and Analysis of Machine Learning and Deep Learning Algorithms in Detection of DDoS Cyberattacks*," Master's Thesis, The University of Texas Rio Grande Valley, 2024. [Google Scholar] [Publisher Link]

[7] Masoud Hemmatpour, Changgang Zheng, and Noa Zilberman, "E-Commerce Bot Traffic: In-Network Impact, Detection, and Mitigation," *2024 27th Conference on Innovation in Clouds, Internet and Networks (ICIN)*, Paris, France, pp. 179-185, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Jian Liu, Xin Gu, and Chao Shang, "Quantitative Detection of Financial Fraud Based on Deep Learning with Combination of E-Commerce Big Data," *Complexity*, vol. 2020, no. 1, pp. 19-11, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] Nitin Rane, Saurabh Choudhary, and Jayesh Rane, "Artificial Intelligence and Machine Learning in Business Intelligence, Finance, and E-commerce: A Review," *SSRN*, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Xiuli Ma, and Zehao Wang, "Computer Security Technology in E-Commerce Platform Business Model Construction," *Heliyon*, vol. 10, no. 7, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] Surendranadha Reddy Byrapu Reddy et al., "Effective Fraud Detection in e-Commerce: Leveraging Machine Learning and Big Data Analytics," *Measurement: Sensors*, vol. 33, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Meet Ashokkumar Joshi, "Artificial Intelligence in E-commerce: a Comprehensive Analysis," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 3, no. 3, pp. 1309-1314, 2024. [Google Scholar] [Publisher Link]

[13] Venkata Ashok K. Gorantla et al., "Utilizing Hybrid Cloud Strategies to Enhance Data Storage and Security in E-Commerce Applications," *2024 2nd International Conference on Disruptive Technologies (ICDT)*, Greater Noida, India, pp. 494-499, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[14] Lakshmi Nivas Nalla, and Vijay Mallik Reddy, "AI-Driven Big Data Analytics for Enhanced Customer Journeys: A New Paradigm in E-Commerce," *International Journal of Advanced Engineering Technologies and Innovations*, vol. 2, no. 1, pp. 719-740, 2024. [Google Scholar] [Publisher Link]

[15] Iqbal Hasan, and Sam Rizvi, "AI-Driven Fraud Detection and Mitigation in e-Commerce Transactions," *Proceedings of Data Analytics and Management*, Singapore, pp. 403-414, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Rizwan Ur Rahman, and Deepak Singh Tomar, "Threats of Price Scraping on e-Commerce Websites: Attack Model and its Detection Using Neural Network," *Journal of Computer Virology and Hacking Techniques*, vol. 17, pp. 75-89, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] JiaoLong Li, "E-Commerce Fraud Detection Model by Computer Artificial Intelligence Data Mining," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1-9, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[18] Simon Bell, and Peter Komisarczuk, "An Analysis of Phishing Blacklists: Google Safe Browsing OpenPhish and PhishTank," *ACSW '20: Proceedings of the Australasian Computer Science Week Multiconference*, pp. 1-11, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] Vijay Mallik Reddy, Lakshmi Nivas Nalla, and Manish Vishwanath, "Optimizing E-Commerce Supply Chains through Predictive Big Data Analytics: A Path to Agility and Efficiency," *International Journal of Machine Learning Research in Cybersecurity and Artificial Intelligence*, vol. 15, no. 1, pp. 555-585, 2024. [Google Scholar] [Publisher Link]

[20] Murat Golyeri et al., "Fraud Detection on E-Commerce Transactions Using Machine Learning Techniques," *Artificial Intelligence Theory and Applications*, vol. 3, no. 1 pp. 45-50, 2023. [Google Scholar] [Publisher Link]

[21] Yanjiao Dong et al., "Real-Time Fraud Detection in e-Market Using Machine Learning Algorithms," *Journal of Multiple-Valued Logic & Soft Computing*, vol. 36, pp. 1-3, 2021. [Google Scholar] [Publisher Link]

[22] Ravi Teja Potla, "AI in Fraud Detection: Leveraging Real-Time Machine Learning for Financial Security," *Journal of Artificial Intelligence Research and Applications*, vol. 3, no. 2, pp. 534-549, 2023. [Google Scholar] [Publisher Link]

[23] G. Rajeshwari et al., "Fraud Detection in E-Commerce Transactions Using Machine Learning Techniques and Quantum Networks," *Quantum Networks and Their Applications in AI,* pp. 146-162, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[24] WenLong Wang et al., "Two-Layer Generalization Boosting Model for Anomaly Detection of e-Commerce Orders," *Sixth International Conference on Computer Information Science and Application Technology (CISAT 2023)*, vol. 269, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[25] Abed Mutemi, and Fernando Bacao, "E-Commerce Fraud Detection Based on Machine Learning Techniques: Systematic Literature Review," *Big Data Mining and Analytics*, vol. 7, no. 2, pp. 419-444, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[26] Shabeena Nafees et al., *Intelligence Framework to Identify Malicious Activities for Safety*, Emerging Trends in IoT and Computing Technologies, 21st ed., CRC Press, 2024. [Google Scholar] [Publisher Link]

[27] Ishita Saha et al., "Phishing Attacks Detection Using Deep Learning Approach," *2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, pp. 1180-1185, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[28] Sarvjeet Kaur Chatratha, G.S. Batrab, and Yogesh Chaba, "Handling Consumer Vulnerability in e-Commerce Product Images Using Machine Learning," *Heliyon*, vol. 8, no. 9, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[29] Ahmed Bensaoud, Jugal Kalita, and Mahmoud Bensaoud, "A Survey of Malware Detection Using Deep Learning," *Machine Learning with Applications*, vol. 16, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[30] The Times, Marks & Spencer faces £300m hit from cyberattack. [Online]. Available: https://www.thetimes.com/business-money/companies/article/marks-and-spencer-300m-cost-cyberattack-9gln92jd9?utm_source=chatgpt.com&region=global

[31] Ryan Holthouse, Serena Owens, and Suman Bhunia, "The 23 and Me Data Breach: Analyzing Credential Stuffing Attacks, Security Vulnerabilities, and Mitigation Strategies," *arXiv Preprint*, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[32] SOCRadar, Top 10 Major Cyber Attacks Targeting E-Commerce Industry, 2024. [Online]. Available: https://socradar.io/top-10-cyber-attacks-targeting-e-commerce-industry/?utm_source=chatgpt.com

[33] Ban Mohammed Khammas, "Ransomware Detection Using Random Forest Technique," *ICT Express*, vol. 6, no. 4, pp. 325-331, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[34] Bhupinder Singh, and Christian Kaunert, "Unraveling Financial Fraud with AI and Machine Learning: Screening into Ad Clicks, Credit Card Management, and E-Commerce Transactions," *Strategies for E-Commerce Data Security: Cloud, Blockchain, AI, and Machine Learning*, pp. 406-429, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[35] Mustafa Mohamed Ismail, and Mohd Anul Haq, "Enhancing Enterprise Financial Fraud Detection Using Machine Learning," *Engineering, Technology & Applied Science Research*, vol. 14, no. 4, pp. 14854-14861, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[36] Samuel Augustina Lata Jeyaraj et al., "Machine Learning Algorithms for E-Commerce Security: A Practical Approach," *Strategies for E-Commerce Data Security: Cloud, Blockchain, AI, and Machine Learning*, pp. 361-385, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[37] Feng Lu, "Online Shopping Consumer Perception Analysis and Future Network Security Service Technology Using Logistic Regression Model," *PeerJ Computer Science*, vol. 10, pp. 1-23, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[38] Renuka Sharma, Kiran Mehta, and Poonam Sharma, "Role of Artificial Intelligence and Machine Learning in Fraud Detection and Prevention," *Risks and Challenges of AI-Driven Finance: Bias, Ethics, and Security*, pp. 90-120, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[39] L.A. Anto Gracious et al., "Advancing E-Commerce Security: Strategic Innovations and Future Directions in AI and ML," *Strategic Innovations of AI and ML for E-Commerce Data Security*, pp. 79-106, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[40] Yury Y. Festa, and Ivan A. Vorobyev, "A Hybrid Machine Learning Framework for e-Commerce Fraud Detection," *Model Assisted Statistics and Applications*, vol. 17, no. 1, pp. 41-49, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[41] Mark E. Lokanan, and Vikas Maddhesia, "Supply Chain Fraud Prediction with Machine Learning and Artificial Intelligence," *International Journal of Production Research*, vol. 63, no. 1, pp. 286-313, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[42] T. Maheshwaran et al., "Securing E-Commerce Strategies with Cloud, Blockchain, AI, and ML," *Strategies for E-Commerce Data Security: Cloud, Blockchain, AI, and Machine Learning*, pp. 470-500, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[43] Sami Ahmed Haider et al., "Energy-Efficient Self-Supervised Technique to Identify Abnormal User Over 5G Network for E-Commerce," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 1631-1639, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[44] Yue Lin, "Anomaly Detection Combining Bidirectional Gated Recurrent Unit and Autoencoder in the Context of e-Commerce," *Engineering Research Express*, vol. 6, no. 3, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[45] Oluwambo Tolulope Olowe et al., "Enhancing Cybersecurity through Advanced Fraud and Anomaly Detection Techniques: A Systematic Review," *2024 International Conference on Science, Engineering and Business for Driving Sustainable Development Goals (SEB4SDG)*, Omu-Aran, Nigeria, pp. 1-12, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[46] Alexander I-Chi Lai, and Hung-Chih Yang, "Deep Learning Based Behavior Anomaly Detection within the Context of Electronic Commerce," *2023 IEEE International Conference on Intelligence and Security Informatics (ISI)*, Charlotte, NC, USA, pp. 1-6, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[47] Sara Farag, and Nahla Barakat, "Data and Model Centric Approaches for Card Fraud Detection," *2023 International Conference on Computer and Applications (ICCA)*, Cairo, Egypt, pp. 1-6, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[48] Jia Li, Shaojiang Liu, and Jiajun Zou, "E-Commerce Data Anomaly Detection Method Based on Variational Autoencoder," *2024 3rd International Conference on Artificial Intelligence, Internet of Things and Cloud Computing Technology (AIoTC)*, Wuhan, China, pp. 231-234, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[49] Sachin Dixit, "Advanced Generative AI Models for Fraud Detection and Prevention in FinTech: Leveraging Deep Learning and Adversarial Networks for Real-Time Anomaly Detection in Financial Transactions," *TechRxiv*, 2024. [CrossRef] [Google Scholar] [Publisher Link]