A Comprehensive Survey of Fraud Detection for E-Commerce

Abstract— **Fraud detection is essential for safeguarding the security, trust, and integrity of e-commerce ecosystems. In recent years, research has yielded substantial progress through a spectrum of techniques, from traditional machine learning and deep neural networks to graph-based models and hybrid frameworks that combine multiple paradigms. This paper provides a comprehensive review of prevalent e-commerce fraud types and critically evaluates the latest fraud detection methodologies, including real-world deployment scenarios. We address key challenges such as data imbalance, lack of model interpretability, adversarial vulnerabilities, and cross-cultural variations in fraud behavior.**

Unlike prior surveys, this work incorporates cutting-edge developments since 2022 and introduces a unified taxonomy that maps fraud categories to the most effective detection strategies. It also highlights the role of explainable AI, privacy-preserving techniques, and adaptive learning in combating evolving threats. This review serves as a valuable entry point for new researchers and a roadmap for practitioners seeking robust, interpretable, and scalable fraud detection solutions. We conclude by identifying research gaps and proposing a prioritized agenda to guide future innovation in secure and resilient e-commerce systems.

.

Keywords: E-commerce, Machine Learning, Fraud Detection, Deep Learning, Future Directions

I. Introduction

E-commerce platforms connect numerous customers with stores, factories, and independent merchants, providing users with a broad range of goods and services on a daily basis. Such platforms offer a convenient way for users to browse, purchase, and provide feedback on various products and services without time and location restrictions. They have thus attracted an enormous audience, generating significant socio-economic impacts. Frauds inevitably occur in e-commerce due to dishonesty, lack of privacy, and intent to steal money. Frauds cover any form of criminal or illegal deception employed to obtain financial or personal advantages. For instance, a user may use another user's credit card credentials to perform illegal transactions on the other user's behalf. Fraudulent behaviors often manifest through abnormal characteristics reflected by inconsistencies with prior fund operation rules or other normal behaviors [1], such as abnormalities in transaction amount, transaction time, transaction accounts, IP addresses, or personal credit rating.

E-commerce frauds can cause financial losses, sabotage business reputation, and eventually reduce the customer base and generate an adverse societal impact [2]. As an example, malicious marketing could mislead buyers into purchasing substandard products by providing false information, leading to hundreds of millions of dollars in losses globally. As such, it threatens the e-commerce ecosystem by popularizing deceptive advertisements and fostering unfair competition in online advertising [3]. The situation has deteriorated in recent years, due to unscrupulous promotions, especially in industrialized nations, where electronic payment methods replace traditional physical payment methods (e.g., cash) to dominate in everyday transactions [4]. Recent studies [5]–[8] show millions of dollars in financial losses in e-banking due to frauds. A recent study by Juniper Research [9], a leading authority on fintech and payment markets, projects that the global cost of e-commerce fraud will increase from $44.3 billion in 2024 to $107 billion by 2029—an increase of 141%. Among existing frauds, card-related frauds—in contrast to non-card related frauds—account for the majority of direct financial losses, largely due to the wide adoption of credit cards and their inherent feature of supporting electronic payments. Take Mastercard cards for example, an additional 30 million new cards were issued by

mid-2023 when compared to the same period in the previous year in the U.S. alone [10], underscoring the potential scale of impact of card-related frauds.

Many efforts have been made to prevent or detect fraud, given the significance of combating fraud in e-commerce. Since it is challenging and often unfeasible to prevent fraud, existing studies have focused on fraud detection, which aims to identify and report potential frauds as soon as they take place in an e-commerce system [11]. Until now, there have been fruitful results in this line of research, covering various techniques ranging from traditional statistical and rule-based methods to machine learning, deep learning, to more recent graph-based and hybrid methods [1], [12]–[14]. On the one hand, given these fruitful achievements, it calls for a timely review of existing efforts to gain a holistic understanding of recent advances to promote applications. On the other hand, despite the achievements made by existing studies, current research still faces critical challenges, represented by ever-growing scale of e-commerce and evolving fraudulent activities. While the growth in the e-commerce sector (in terms of users and online transactions) poses an expanding venue to fraudsters for vulnerability exploitation, fraudulent activities never stop evolving toward being more sophisticated, requiring novel or more advanced techniques to detect. All the above challenge constantly quest to summarize the successes and pitfalls of existing efforts and identify the key obstacles in current research to derive insights into promising future research directions. Leveraging such insights, researchers and practitioners could better navigating their research efforts by making informed decisions.

Motivated by the above, this survey first presents an overview of the common fraud types and detection techniques in e-commerce, followed by critically evaluating existing techniques through comparative discussions. Finally, we discuss the open challenges and points out some promising future directions.

Compared with existing surveys [15]–[19], this survey significantly distinguishes itself in two aspects: 1) it comprehensively covering various fraud types and detection methods in the field of e-commerce, as opposed to focusing on specific detection methods; 2) it includes the latest advances in this field of research since 2022, which are not covered by existing surveys.

To ensure a comprehensive and systematic review, we adopted a structured literature search strategy. Academic databases including IEEE Xplore, ACM Digital Library, SpringerLink, ScienceDirect, and Google Scholar were queried using keywords such as "e-commerce fraud detection," "machine learning fraud," "deep learning fraud," "graph neural networks for fraud," and "hybrid fraud detection models." We focused on peer-reviewed articles published between 2018 and 2024, with particular emphasis on works appearing after 2022 to ensure coverage of recent advancements. Inclusion criteria required that studies present original research with clear methodological descriptions, utilize e-commerce or closely related financial transaction data, and report empirical results. Exclusion criteria eliminated papers focused solely on traditional banking fraud, insurance fraud without e-commerce context, or theoretical discussions lacking validation. From an initial pool of over 400 articles, approximately 170 were selected based on relevance, recency, and methodological rigor.

In a nutshell, we make the following contributions in this survey:

- We conduct a comprehensive investigation into the fraud detection problem, identifying a number of key fraud types in e-commerce.
- We extensively review existing fraud detection techniques across the identified fraud types andcritically evaluate their strengths and limitations.
- We discuss open challenges for fraud detection in e-commerce, pinpointing recent trends and promising directions for future research.

II. Fraud Types in E-commerce

As e-commerce continues to gain popularity, fraudsters are continually devising new methods to exploit financial systems and services, including credit card transactions, telecommunications, health insurance, vehicle insurance, and online auction platforms. E-commerce fraud can be broadly categorized into two major types: **Card-Present Fraud (CPF)** and **Card-Not-Present**

Fraud (CNPF). **CPF** [20] involves unauthorized transactions in physical retail environments through stolen or cloned cards, and although EMV chip technology has reduced its occurrence, it still persists where card cloning and skimming techniques remain prevalent. In contrast, **CNPF** [21] presents a more significant challenge in online transactions, where cybercriminals use stolen card details for unauthorized purchases without requiring the physical card. This form of fraud thrives on weaknesses in online payment verification processes, often leveraging phishing, data breaches, and malware to acquire sensitive payment credentials. We investigate and summarize the major types of fraud in e-commerce in the following subsections, while Table I illustrates each fraud type's classification as either card-present or card-not-present.

A. Skimming and Counterfeit Card Fraud

Skimming refers to the unauthorized capture of cardholder data from a legitimate transaction, typically through compromised ATMs or point-of-sale (POS) terminals. Fraudsters install skimming devices to extract card information from magnetic stripes, which is subsequently used to create counterfeit cards that mimic genuine payment instruments. Despite advancements in EMV chip technology, counterfeit fraud remains a substantial concern, especially in regions where magnetic stripe transactions persist. The pre-play attack, for instance, exploits weaknesses in random number generation during EMV transactions, enabling attackers to clone chip-based transactions before authentication is completed [22].

An empirical study analyzing underground markets for stolen card data highlights that 97% of compromised payment credentials were derived from magnetic stripe cloning, underscoring the continued reliance on skimming-based fraud [20]. Furthermore, financial institutions in South Asia report that counterfeit card fraud remains a critical issue due to delayed adoption of EMV technology and inadequate risk management practices [23].

In the United States, the migration to EMV cards in 2015 aimed to mitigate counterfeit card fraud. However, the absence of a mandatory PIN requirement has allowed fraud rates to remain elevated compared to other developed countries [24]. Additionally, studies show that in-person fraud rates in the U.S. continue to exceed those in Australia, France, and the UK, largely due to weaker card verification methods.

B. Lost or Stolen Card Fraud

Fraud involving lost or stolen cards occurs when unauthorized individuals gain possession of a legitimate card and misuse it for fraudulent transactions. Unlike counterfeit fraud, which relies on card cloning, lost/stolen fraud is often linked to weak authentication mechanisms. In countries where chip-and-PIN authentication is enforced, such as France and the UK, lost/stolen card fraud rates are significantly lower [24]. Conversely, in the United States, where many transactions still rely on signatures or contactless verification, fraudsters exploit the absence of PIN requirements, increasing the likelihood of successful fraudulent transactions.

Criminal forums and dark web marketplaces facilitate the resale of physically stolen payment cards, often alongside complementary identity data. This enables unauthorized transactions before cardholders report the loss. A longitudinal analysis of a carding marketplace found that fraudsters frequently bundle lost/stolen card details with other personal identifiers to bypass rudimentary security checks [20].

To combat skimming, counterfeit, and lost/stolen card fraud, financial institutions and regulatory bodies have implemented various fraud detection and prevention strategies. Advanced AI-based fraud monitoring systems, including machine learning models such as random forest and neural networks, enhance the detection of anomalous transaction patterns [25]. Additionally, industry-wide adoption of stronger authentication mechanisms, including biometric verification and real-time behavioral analysis, could further reduce the risks associated with physical card fraud.

While EMV migration has reduced counterfeit card fraud in many regions, inconsistent implementation—such as the lack of PIN verification in the U.S.—continues to expose vulnerabilities. Strengthening cross-border regulatory frameworks, increasing customer awareness, and accelerating the phase-out of magnetic stripe transactions re-

TABLE I
E-COMMERCE FRAUD CATEGORIES AND TYPES

Category	Fraud Type
Card-Present Fraud (CPF)	Skimming and Counterfeit Card Fraud Lost or Stolen Card Fraud
Card-Not-Present Fraud (CNPF)	Payment Fraud Friendly Fraud Clean Fraud Affiliated Fraud Account Fraud Triangulation Fraud

main critical measures to mitigate these threats effectively [23], [24].

C. Payment Fraud

Payment fraud in e-commerce, commonly known as identity theft, is a critical issue where cybercriminals utilize various methods of identity fraud to carry out illicit transactions [12]. This type of fraud typically involves the unauthorized use of credit and debit card information obtained through illicit means such as phishing or system hacking. As e-commerce expands globally, the vulnerability to such fraud escalates. According to Juniper Research, global losses to online payment fraud are projected to exceed $362 billion between 2023 and 2028 [26]. Mastercard also reports a continued rise in fraud targeting digital commerce, driven by increasing attack sophistication and mobile wallet adoption [27].

Traditionally, detection efforts have focused on supervised machine learning approaches. Recent studies, such as those by [28], have proposed and evaluated various models including Logistic Regression (LR), Decision Trees (DT), K-Nearest Neighbors (KNN), and Random Forest (RF) on real credit card transaction datasets. These models play a crucial role in detecting fraudulent transactions, with RF and LR demonstrating superior performance in some scenarios. Enhancements such as under-sampling and feature reduction using principal component analysis have been suggested to improve model outcomes. This research underscores the importance of feature selection and sampling techniques in the digital economy and the era of Industry 4.0, characterized by advanced automation and data exchange in manufacturing and service industries, including e-commerce. A comprehensive analysis of machine learning and artificial intelligence algorithms in credit card fraud detection is further elaborated in [29].

More recently, fraud in digital payment ecosystems has shifted toward mobile wallets and cryptocurrency platforms. The increased popularity of services like Apple Pay, Google Pay, and Samsung Pay has introduced new vectors for exploitation. Recent research highlights the vulnerability of these platforms to SIM-swapping, phishing attacks, and mobile malware that can bypass biometric authentication and tokenization mechanisms [30].

To combat these threats, researchers have increasingly applied machine learning to detect behavioral anomalies in real-time transaction data. For example, Muqattash *et al.* [31] demonstrated that logistic regression, decision trees, and support vector machines could effectively detect fraudulent mobile payment patterns by analyzing temporal and behavioral features. Siddiqui*et al.* [32] further emphasize the role of artificial intelligence in enhancing e-payment security through adaptive models that respond to evolving fraud strategies.

Cryptocurrency-based payment systems also face growing security concerns. Cryptocurrency wallets, particularly mobile and non-custodial ones, are vulnerable to malware, phishing, key theft, and replay attacks. Nowroozi *et al.* [33] classify wallets by type and highlight the attack vectors unique to each, noting that mobile wallets are especially susceptible to operating system vulnerabilities and third-party integration flaws. As the use of digital assets expands in e-commerce, fraud detection must integrate blockchain-specific indicators and cross-layered data streams into its analytical models.

D. Friendly Fraud

Friendly fraud can occur intentionally or inadvertently when customers make an online purchase and initiate a chargeback with the issuing bank after receiving the purchased goods or services. This type of fraud is particularly insidious as it exploits the chargeback mechanism designed to protect consumers from unauthorized transactions, making it challenging for merchants to differentiate between legitimate and fraudulent claims. The underlying premise is that a customer, whether genuine or not, pays for a service or commodity and then contends that it was never provided or was damaged upon delivery [34]. Al *et al.* [13] explore the traits and preferences of telecom fraud users using desensitized signaling data, voice message specifics, APP traffic data and billing accounting data. Several studies [13], [34] underscore the complexity of addressing friendly fraud in e-commerce and emphasize the necessity for advanced analytical techniques, such as machine learning, to effectively identify and mitigate such fraudulent activities.

E. Clean Fraud

Clean fraud refers to fraudulent operations that mimic real transactions. As banned fraud accounts typically do not raise flags or reject transactions, this type of fraud is becoming a growing concern for businesses [35]. To identify fraudulent banking transactions, Narsimha *et al.* [36] introduce a cybersecurity defensive mechanism by integrating AI and ML techniques with the current Feedzai security paradigm. Furthermore, this work introduces Feedzai's Open ML fraud detection software tool and a well-known ML & AI model utilizing the random forest algorithm. These tools enhance the intelligent framework for addressing financial fraud detection by adding automatic fraud recognition capabilities.

F. Affiliated Fraud

Malicious actors can manipulate traffic and sign-ups through affiliate fraud, creating the illusion that a merchant is gaining customer attention when, in reality, this is not the case [37]. Such fraud is increasingly common. Individuals often impersonate an organization's spokesperson, creating hazardous situations to gather details about the victim. Dadhich *et al.* [38] explore three types of cybercrime: card payment fraud, mobile payment fraud, and telephone fraud, discussing how systems can identify and prevent these crimes. Additionally, many traditional methods and tools utilize information about credit cards and scams to prevent related fraud. Boubker *et al.* [39] present state-of-the-art data mining strategies to address this issue. The study provides a solid foundation for further research, offering a clear understanding of credit card theft, various methods to counter such fraud, and a summary of the major challenges encountered.

G. Account Fraud

Common fraud detection applications can be account-related, payment-related, and transaction-related. For example, two prevalent types of account fraud are new account fraud and account takeover fraud. In *new account fraud*, fictitious identities are used to open new accounts. Detecting these frauds involves leveraging patterns from various devices and session indicators to identify false identities. *Account takeover fraud* occurs when a hacker utilizes another person's active account to make purchases. To prevent this, users' sessions, devices, and behavioral biometrics can be computed and scored to validate an account. Additionally, tracking user behavior patterns through user journey analysis can aid in identifying account takeovers before they result in any financial impact. Payment fraud refers to any fraudulent or unlawful transaction conducted by a cybercriminal [40]. In this case, the offender defrauds the victim of money, valuables, interests, or private information. This category also encompasses fraud involving stolen goods, unauthorized transactions, and fake refund requests.

H. Triangulation Fraud

This type of fraud involves cybercriminals creating fake or imitation websites to lure customers with discounted goods. These fraudulent websites may appear in advertisements or be sent to a user's email address as part of a phishing attempt [37]. Alharbi *et al.* [41] develop a deep learning (DL)-based strategy for addressing text data problems

using the Kaggle dataset. A novel text2IMG conversion method generates small images, which are then input into a CNN architecture with class weights to address class imbalance issues. DL and ML techniques validate the system's effectiveness and robustness. Coarse-KNN uses the CNN's deep features, achieving an accuracy of 99.87 percent.

Recently, John *et al.* [42] offer practitioners a guide for successfully integrating the development of ML/DL models into business environments to surpass other methods. Moreover, they describe checkpoints for concluding the business case and demonstrate how their framework assists businesses in resolving the problems they have identified.

I. Unified Conceptual Taxonomy

To provide a structured overview of the landscape of e-commerce fraud detection, we introduce a unified conceptual taxonomy that links the principal fraud types to the most effective detection strategies and model types (see Figure 1). This taxonomy is informed by recent literature and empirical results reviewed in Sections II and III. For each fraud type, we map the most relevant detection techniques and the preferred model types commonly used in practice. This framework offers a holistic, at-a-glance guide for researchers and practitioners to align analytic approaches with specific threat types, thereby supporting both the design and evaluation of e-commerce fraud detection systems.

To further clarify the conceptual landscape of e-commerce fraud detection, we present a unified taxonomy that bridges the gap between specific fraud types and the detection strategies most commonly applied to address them (see Figure 2). The framework visually organizes e-commerce fraud into two primary categories—card-present and card-not-present fraud—detailing representative fraud types for each. These are mapped to detection strategies and model types, including traditional statistical, machine learning, deep learning, graph-based, and hybrid approaches, based on recent advances and empirical studies surveyed in Sections II and III. This schematic overview provides a structured understanding for both practitioners and researchers, supporting effective alignment between threat types and analytic solutions.

III. Fraud Detection Techniques

Supervised learning is a common method for e-commerce fraud detection, using labeled transactions (marked as fraudulent or non-fraudulent) to train predictive models [43]–[45]. Algorithms such as Decision Trees (DT), Random Forest (RF), Support Vector Machines (SVM), and Neural Networks (NN) are frequently used to identify known fraud patterns. Although these models excel at recognizing established fraudulent behaviors, they demand substantial, high-quality labeled datasets, which can be difficult to secure due to class imbalance. Moreover, accurately labeling fraudulent transactions can be labor-intensive and prone to errors.

In contrast, unsupervised learning does not rely on labeled data [44]–[46]. Instead, it identifies anomalies or unusual patterns without prior knowledge of fraudulent behavior. Clustering algorithms (e.g., K-Means, DBSCAN), autoencoders, and Isolation Forests (IF) detect potential fraud by pinpointing deviations from normal spending trends [47]. While unsupervised methods are effective at uncovering new and emerging fraud techniques, the absence of explicit fraud labels can lead to a higher incidence of false positives.

To address the inherent challenges of supervised and unsupervised strategies, semi-supervised learning and hybrid models have garnered increasing attention. These approaches leverage a small set of labeled fraudulent transactions alongside self-learning processes, which refine anomaly detection while mitigating false alarms [44], [45].

A. Benchmark Datasets for E-Commerce Fraud Detection

The progress and reproducibility of e-commerce fraud detection research are largely dependent on the availability of robust, well-established benchmark datasets. Before discussing specific detection techniques, it is essential to understand the characteristics of the datasets most widely used in this domain, as these underpin model evaluation and fair comparison across studies.

Table II provides an overview of the most frequently cited public datasets, including a brief description of each dataset, representative studies in

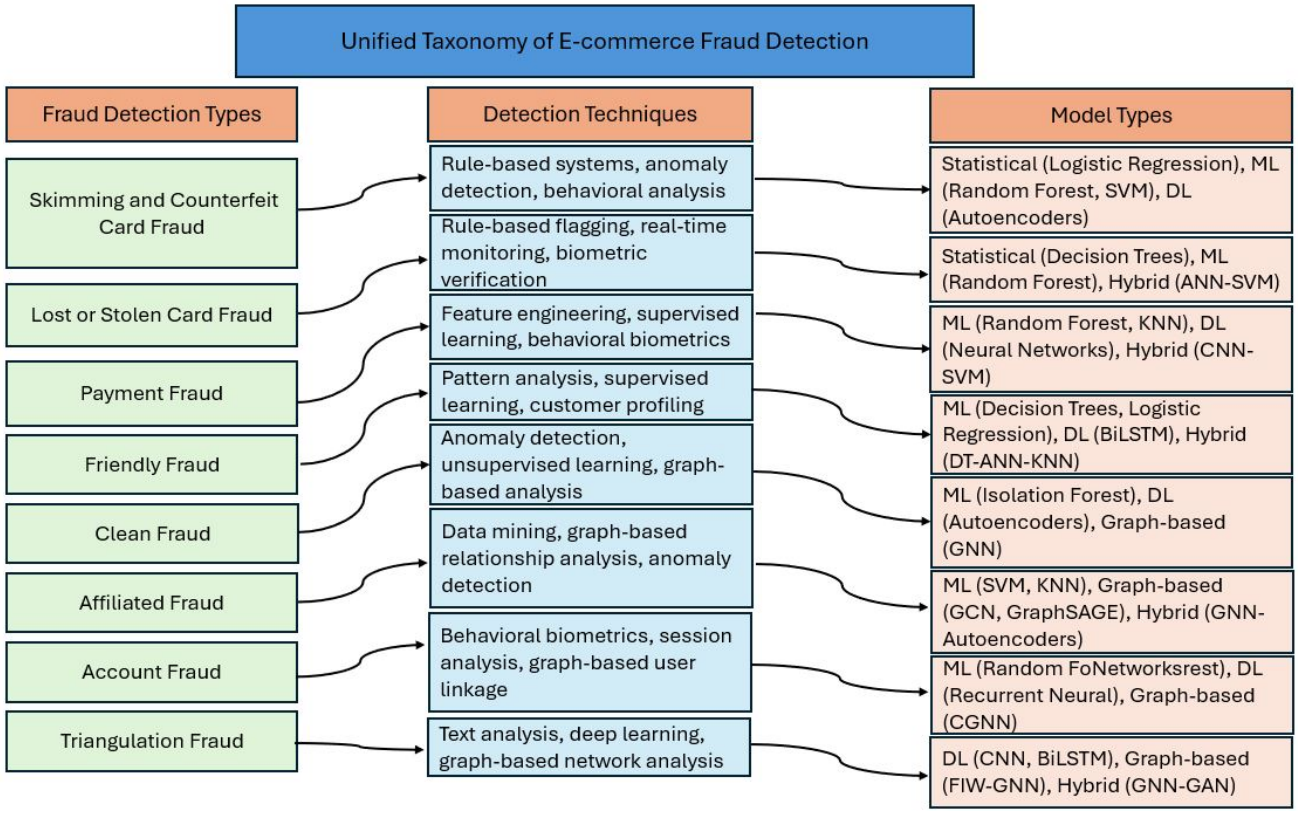

Fig. 1. Unified Taxonomy of E-commerce Fraud Detection. This framework conceptually links the principal types of e-commerce fraud to their most effective detection techniques and corresponding model types, as evidenced in the literature. Each fraud type is mapped to preferred analytic approaches (center and the model types most often used for detection, providing a structured understanding of how specific threats are addressed in practice.

which they have been applied, and their respective domains of application.

Note: When developing or evaluating new fraud detection models, these benchmark datasets are recommended for ensuring reproducibility and facilitating fair comparison, as supported by recent research.

With this context established, the following subsections review and compare the principal categories of fraud detection techniques, analyzing how they have been applied to these datasets and highlighting their respective strengths and limitations.

B. Traditional Statistical and Rule-Based Approaches

Fraudulent transactions often display patterns or anomalies that differentiate them from legitimate transactions, which makes rule-based systems an intuitive approach for fraud detection. These systems use predefined rules, such as flagging transactions that exceed certain amounts or originate from unusual locations. Rule-based systems are highly interpretable and can adapt to new fraud tactics quickly. However, they require regular updates to remain effective, as they may struggle to keep pace with evolving fraud patterns without frequent refinement [53]. The work by Gianini et al. [54] proposes a Game Theory-based method to manage a pool of rules for near real-time fraud detection, where transactions triggering any rule are flagged for review. Using the Shapley Value, the approach assesses each rule's contribution to the pool's overall performance such as precision or recall, rather than evaluating rules in isolation. This highly interpretable system adapts to overlapping or redundant rules, outperforming traditional individual ranking methods by prioritizing collective effectiveness. While computationally intensive, it suits periodic updates, ensuring resilience against

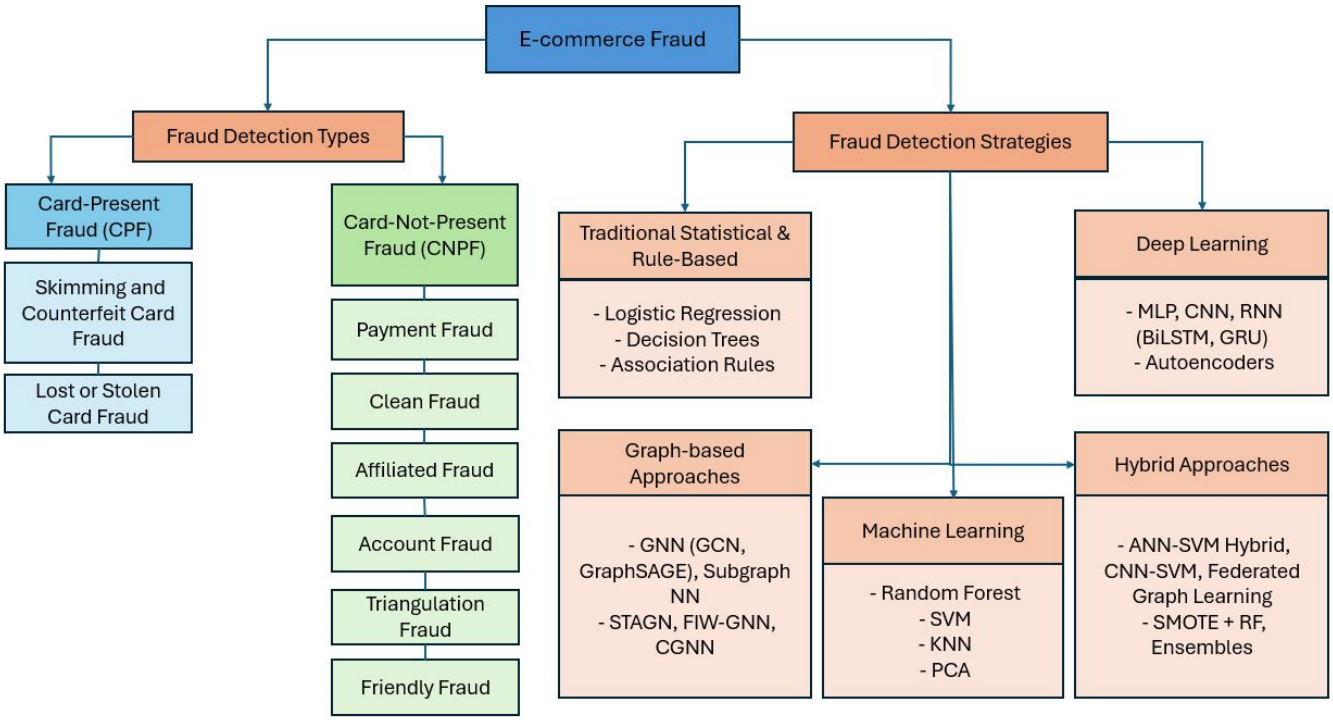

Fig. 2. Unified Conceptual Framework for E-commerce Fraud Detection. This framework presents a high-level taxonomy that organizes e-commerce fraud into card-present and card-not-present categories, and systematically links each category to the main detection strategies reviewed in the literature. The right panel summarizes the principal detection approaches—including traditional statistical, machine learning, deep learning, graph-based, and hybrid methods—alongside typical models used in state-of-the-art fraud detection systems.

TABLE II

OVERVIEW OF COMMONLY USED DATASETS FOR E-COMMERCE FRAUD DETECTION

Dataset	Description	Used In Studies	Domain
IEEE-CIS Fraud Detection	An anonymized dataset including transactional and identity information. Widely used for Kaggle competitions and fraud detection models.	Alkhatib et al. [4], Najadat et al. [12], Tang et al. [48]	Credit card / E-commerce
European Credit Card Dataset	Contains 284,807 transactions over 2 days with 492 frauds. Highly imbalanced.	Dornadula et al. [34], Berhane et al. [49]	Credit card
TaoBao Dataset	E-commerce platform data with labeled fraudulent behaviors. Includes user actions, product interactions, and timestamps.	Song et al. [15], Guo et al. [50]	E-commerce transactions
Amazon Dataset	Reviews and metadata from Amazon. Used for opinion fraud and review-based fraud detection.	Song et al. [15], Li et al. [51]	E-commerce reviews
YelpChi	Subset of Yelp dataset for review fraud detection. Includes spam indicators and user metadata.	Song et al. [15], Li et al. [51]	Opinion fraud
Sparkov Dataset	Simulated transactional data for credit card fraud research. Contains features relevant to fraud classification.	Yan et al. [52]	Credit card fraud

evolving fraud tactics. Moreover, the study by Baumann [55] enhances an existing rule-based fraud detection system for motor insurance by integrating Classification Based on Association Rule Mining (CARM) to uncover dependencies between rules, such as combining ”weekend accident” with

"single witness." A genetic algorithm optimizes rule weights, improving interpretability and precision. Though requiring regular refinement to adapt to new fraud patterns, this approach reduces false positives significantly, offering a scalable way to keep pace with evolving threats in rule-based frameworks.

In contrast, statistical methods employ mathematical models to detect anomalies within transaction data. Techniques such as logistic regression, decision trees, and clustering are commonly used. Logistic regression estimates the probability of fraud based on various transaction attributes, while decision trees classify transactions by applying specific rules derived from data. Clustering techniques, on the other hand, identify transactions that do not fit into established patterns, flagging them as potentially fraudulent [56].

The synergy between rule-based and statistical methods has been a focal point of recent research, particularly in enhancing the adaptability and accuracy of fraud detection systems. For instance, Bekach *et al.* [53] integrated traditional rule-based systems with deep learning techniques, using the CRED algorithm to extract human-understandable rules from deep learning models. This approach was tested on highly imbalanced datasets, demonstrating improved performance and interpretability compared to previous methods.

Moreover, machine learning models have increasingly shown their superiority over traditional statistical methods in fraud detection. Moumeni *et al.* [56] compared logistic regression and principal component analysis (PCA) with more advanced machine learning models, such as multilayer perceptrons (MLPs). Their analysis of transaction data revealed that MLPs, a type of neural network, surpassed traditional methods in classification accuracy.

Finally, the effectiveness of specific algorithms in different contexts has also been a subject of investigation. Valli *et al.* [57] compared logistic regression and random forest algorithms for credit card fraud detection, showing that the random forest algorithm outperformed logistic regression in both accuracy and precision, underscoring its superior effectiveness in this domain.

Note on Fair Comparison: The performance metrics summarized in Table III are drawn from studies using different datasets, data splits, and preprocessing strategies. Direct comparison of results should be interpreted with caution, as these contextual differences may impact reported performance. See Table II for further dataset descriptions.

C. Machine Learning Approaches

Machines generally perform much more effectively than humans, leading to the application of machine-learning algorithms for automatically extracting patterns—in contrast to relying on handcrafted rules—in fraud detection. Until present, various algorithms, represented by Support Vector Machines, Bayesian Classification, k-Nearest Neighbours, Decision Trees and Frequent Pattern Mining, have been applied to build effective fraud detection systems [17], [35], [58]. Figure 3 illustrates the typical pipeline for fraud detection using machine learning approaches. The fundamental pipeline for fraud detection involves data pre-processing and feature extraction, followed by evaluating results using different machine learning algorithms.

A key challenge in fraud detection is the imbalance between normal and fraudulent transactions in datasets. Fang *et al.* [59] address this issue by applying the Synthetic Minority Over-sampling Technique (SMOTE) to mitigate data imbalance in credit card transactions within the e-commerce domain. The focus was on the Light Gradient Boosting Machine (LightGBM) model, comparing its performance to Random Forest and Gradient Boosting Machine algorithms. The results show that LightGBM achieved a total recall rate of 99% on real datasets, demonstrating its effectiveness in combating the imbalance issue and improving fraud detection accuracy.

Dornadula *et al.* [34] analyze past transaction details to extract behavioral patterns and develop a novel fraud detection method for streaming transaction data. In this system, cardholders are categorized into subgroups based on the average dollar amount of their purchases.

Li *et al.* [60] focus specifically on developing a fraud detection model (FDM) for electronic transactions. They address the critical issue of imbalanced data by segregating samples into non-fraud and fraud groups, then training various algorithms

TABLE III

TRADITIONAL STATISTICAL AND RULE-BASED METHODS FOR E-COMMERCE FRAUD DETECTION. *Note: Dataset size, imbalance ratio, and feature information are provided where available. Reported performance metrics are drawn from different studies and may not be directly comparable. See Table II for further dataset details.*

Studies	Technique	Dataset	Size / Imbalance	Feature Types	Results
Bekach *et al.* [53]	CRED algorithm (a Continuous/discrete Rule Extractor via Decision tree induction) MLP	1st dataset: transactional, 2nd dataset: transactional	1st: 151,112 2nd: 140,000 (Imbalance not stated)	Transaction attributes	99.83% Accuracy
Gianini et al. [54]	Game Theory approach using Shapley value for rule contributions	Real-world credit card transaction data	Unspecified (typically highly imbalanced)	Transaction attributes	Reduced operational rules with maintained or improved detection, increased efficiency
Baumann [55]	Classification Based on Association Rule Mining (CARM), Genetic Algorithm	Motor insurance claims	∼140,000; highly imbalanced	Claims, accident context	Reduced false positives by 16.89%, minor increase (0.85%) in false negatives
Moumeni *et al.* [56]	LR PCA Multilayer Perceptrons (MLP)	Transaction data from an American bank	Not stated; (Imbalanced)	Transaction attributes	MLPs outperformed LR and PCA in classification accuracy
Valli *et al.* [57]	LR RF	Financial transaction datasets	Unspecified; (Imbalanced)	Transaction attributes	Random Forest achieved better accuracy and precision than LR

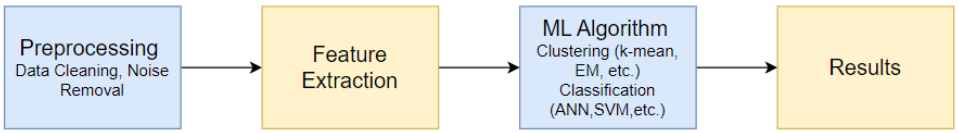

Fig. 3. A Typical Machine Learning Pipeline for Fraud Detection. The third step trains the machine learning model, and the last step makes inferences to finally distinguish frauds from normal data.

to enhance accuracy. Their experiments on a real-world MasterCard dataset reveal that Random Forest outperforms other models, although challenges related to data imbalance persist, with reported accuracies of 0.8236 for Decision Tree, 0.8703 for Logistic Regression, 0.8781 for Random Forest, and 0.8455 for human-judged classification. This study underscores the importance of both algorithm selection and data handling techniques in fraud detection.

Further refining fraud detection approaches, Carta *et al.* [61] propose the Prudential Multiple Consensus model, an ensemble method aimed at improving classification accuracy by addressing common issues such as overfitting and bias in existing state-of-the-art algorithms. This model emphasizes the importance of combining multiple classifiers to enhance the robustness and reliability of fraud detection systems.

While technical research is crucial, it's also essential to understand the broader applications of machine learning in various domains. Pallathadka *et al.* [35] provide a comprehensive survey of how machine learning and AI are applied across e-commerce, corporate management, and financial analysis. This work serves as an essential overview, discussing the role of machine learning in sales expansion, profit maximization, inventory manage-

ment, security, and, notably, fraud detection. By introducing this survey, we gain insight into the diverse applications and challenges of deploying machine learning techniques across industries.

Note on Fair Comparison: Performance metrics summarized in Table IV are reported from studies that may use different datasets, class ratios, feature sets, and experimental setups. Thus, direct comparison of absolute results should be interpreted with caution. See Table II for more detailed dataset descriptions.

D. Deep Learning Approaches

The accuracy and detection rate offered by machine-learning techniques make them the preferred method for fraud detection, and experts continue to work hard on improving detection rates and precision. Moreover, businesses aim to find ways to cut expenses while simultaneously boosting profits, and these studies provide ample opportunities to achieve that goal [50], [62]. With the advent of deep learning models, their implementation in fraud detection has further improved the accuracy of the models. Deep learning models typically provide better accuracy due to their capabilities to analyze features automatically, but come with added complexity. Figure 4 illustrates the process for fraud detection using deep learning approaches (excluding graph-based methods introduced in Section III-E).

Specifically, Alkhatib *et al.* [4] utilize a dataset containing the business dealings of the illustrious Vesta Group to develop a novel model for detecting credit card fraud. In the first step, they performed a proper preprocessing of cleaning the dataset and selected features using a correlation factor. The results of the experiments led to the selection of a deep neural network architecture with a total of seven layers for input, hidden and output. The model outperformed its predecessors across all metrics used for comparison, achieving an area under the ROC curve of 99.1 percent.

Najadat *et al.* [12] utilize the IEEE-CIS Fraud Detection dataset and tested several machine learning and deep learning models, including a BiLSTM-MaxPooling-BiGRU-MaxPooling model based on Bidirectional LSTM and GRU, to determine the legitimacy of online transactions. Undersampling and oversampling using SMOTE were also evaluated as potential solutions for handling highly unbalanced datasets.

Unbalanced datasets can pose a problem for Deep Learning (DL) applications, as a model trained on such datasets may bias predictions toward the class with a greater percentage. Furthermore, effective feature engineering is crucial in building a robust fraud detection DL model. Features that analyze user behavior are particularly essential in this application, aiding in identifying fraud trends [53].

According to Wang *et al.* [63], deep learning can be employed to analyze a user's behavior to predict their current and future actions. This method involves analyzing the user's command-line sequence.

Note on Fair Comparison: The performance results in Table V are reported from different studies, each using different datasets, imbalance ratios, feature types, and experimental protocols. Therefore, direct performance comparison should be interpreted with caution. Please refer to Table II for additional dataset descriptions.

E. Graph-based Approaches

Fraud detection systems require features to train and identify attack patterns. These features necessitate domain knowledge for effective fraud recognition and prevention, highlighting the importance of focusing on the intricacies of the fraud scheme. An attention-based graph network is a credit card fraud detection system that utilizes a focused graph network. Figure 5 illustrates how the model aggregates messages from a node's local graph neighbors (e.g., B, C, and D) to generate an embedding for node A. The figure also shows that messages from these neighbors are based on information aggregated from their respective neighborhoods, forming a hierarchical aggregation structure. This process allows the model to encode information from multi-hop neighborhoods, improving the representation of nodes. At the final layer, initial node attributes contribute to the representation, ensuring both structural and feature-based information are captured. This approach, used in methods like Graph Convolutional Networks (GCNs) and GraphSAGE, helps in tasks such as node classification and link prediction by effectively learning node representations from

TABLE IV

MACHINE LEARNING APPROACHES FOR E-COMMERCE FRAUD DETECTION. *Note: Dataset size, imbalance, and feature information are provided where available. Reported results are drawn from different studies and may not be directly comparable. See Table II for dataset details. DT: Decision Tree, RF: Random Forest, LR: Logistic Regression, SVM: Support Vector Machine.*

Studies	Technique	Dataset	Size / Imbalance	Features	Results
Hasan *et al.* [17]	DT RF LR	MasterCard exchanges dataset	Unspecified (Imbalanced)	Transaction attributes	Accuracy: 84.55% (RF) 71.38% (LR) 78.22% (DT)
Dornadula *et al.* [34]	DT RF LR SVM Isolation Forest Local Outlier Factor	European credit card fraud dataset	total 284,807 0.17% fraud	PCA-applied V1-V28, Time, Amount, Class (Label)	Accuracy: 99% (RF) 97% (DT) 97% (LR)
Pallathadka *et al.* [35]	DT RF Nearest Neighbour SVM	Unspecified	Unspecified	Unspecified	Survey paper No reported results
Wen *et al.* [58]	DT SVM	National banks credit card warehouse	Unspecified	Transaction features	DT outperforms SVM
Fang *et al.* [59]	LightGBM GBM RFt	Credit card transaction dataset	Unspecified; Imbalanced	Transaction features	Recall: 99% (LightGBM)
Li *et al.* [60]	Information Fusion Technology (IFT)-based Fraud Detection Model (FDM) SVM LR	30,000 e-commerce behavior samples	total 30,000 Imbalance not stated	Behavioral and transaction features	IFT-based FDM: Test accuracy = 84.10% Outperformed SVM and LR% accuracy
Carta *et al.* [61]	Prudential Multiple Consensus (PMC)	European credit card fraud dataset	total 284,807 0.17% fraud	PCA-applied V1-V28, Time, Amount, Class (Label)	PMC model AUC: 89%, higher precision than other approaches

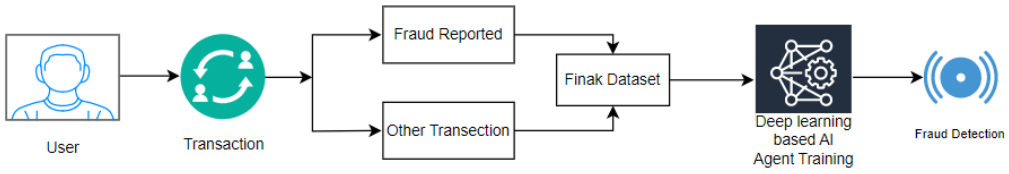

Fig. 4. Deep Learning based Fraud Detection. Deep learning methods notably differ from conventional machine learning in enabling automatic feature extraction within the training process.

graph structures. Below are several research studies related to graph-based models proposed for efficient fraud detection systems.

Cheng *et al.* [65] address the significant issue of card fraud, which incurs substantial costs for cardholders and issuing banks. The proposed solution is a spatial-temporal attention-based graph network (STAGN) for credit card fraud detection. This method involves learning temporal and location-based transaction graph features through a graph neural network and applying spatial-temporal attention along with 3D convolution. The attentional weights are learned end-to-end in conjunction with 3D convolution and detection networks. Empirical studies with domain experts confirm the method's effectiveness in detecting suspicious transactions,

TABLE V

DEEP LEARNING TECHNIQUES FOR E-COMMERCE FRAUD DETECTION. *Note: Dataset size, imbalance, and feature information are provided where available. Reported results are from different studies and may not be directly comparable. See Table II for dataset details. DNN: Deep Neural Network, RNN: Recurrent Neural Network.*

Studies	Technique	Dataset	Size / Imbalance	Features	Results
Alkhatib *et al.* [4]	DNN	IEEE-CIS Fraud Detection (Vesta Group)	∼590,000; 3.5% fraud	434 transaction, identity features	99.1% AUC (ROC curve)
Najadat *et al.* [12]	BiLSTM-MaxPooling-BiGRU-MaxPooling	IEEE-CIS Fraud Detection	∼590,000; 3.5% fraud	434 features	91.37% AUC
Guo *et al.* [50]	Iterative Fast Coordinate Method	Taobao dataset	Unspecified; Imbalance (real-world)	Transaction and user behavior features	Nearly 90% Average Precision
Bekach *et al.* [53]	CRED algorithm MLP	1st: Transactional; 2nd: Transactional	151,112; 140,000; (Imbalance not stated)	Transaction attributes	99.83% Accuracy
Lebichot *et al.* [62]	Baseline DNN Naive DNN Adaptation DNN	Belgian credit card issuer dataset	Unspecified; highly imbalanced	Transaction features	All three methods outperformed the compared baselines
Wang *et al.* [63]	DNN	User clickstream/-command data	Unspecified	User behavioral sequences	98.1% accuracy 3x improvement over baseline

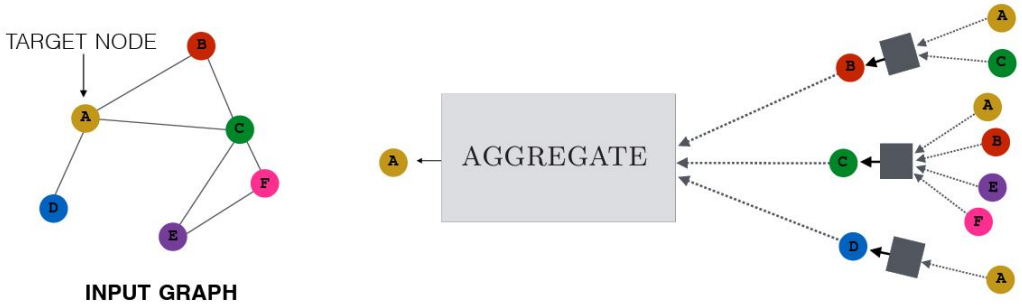

Fig. 5. An Overview of Encoding in Neighborhood Aggregation Methods for Graph-based Approaches [64]. It illustrates how the model aggregates messages from a node's local graph neighbors (e.g., B, C, and D) to generate an embedding for node A, and how messages from these neighbors are based on information aggregated from their respective neighborhoods, forming a hierarchical aggregation structure.

identifying spatial and temporal fraud hotspots, and revealing fraud patterns. Additionally, the method proves effective in other user behavior-based tasks. To tackle big data challenges, STAGN is integrated into the fraud detection system as the predictive model, with detailed implementation insights provided for each module in the system [65].

Yan *et al.* [52] address escalating global economic losses from credit card fraud, highlighting the urgent need for effective and accurate fraud detection. Previous methods, although effective under certain conditions, lack robustness and scalability when applied to real-world credit card transactional datasets with numerous missing values. The proposed solution is a Feature Importance-based Weighted Graph Neural Network (FIW-GNN) [52]. This method constructs a heterogeneous graph tailored for credit card transactional datasets and uses a feature importance-based approach to assign edge weights through a relational graph convolutional network architecture. Evaluation on two benchmark datasets demonstrates that FIW-GNN outperforms state-of-the-art baselines across all selected evaluation metrics.

Zhu *et al.* [66] present CNSGNN, a cost-sensitive graph neural network that addresses two major limitations in graph-based fraud detection: class imbalance and sparse connections between fraud nodes. The model employs a label-aware adaptive neighbor sampler based on reinforcement learning, which selectively aggregates information from similar neighbors. It also introduces a connection-enhancement strategy to create new edges between fraudulent nodes, thereby strengthening intra-class representation. Furthermore, a cost-sensitive matrix optimizer is used to penalize misclassification of minority classes more severely. Experimental results on two e-commerce datasets—MGM and Sichuan—demonstrate that CNSGNN outperforms baseline GNNs in terms of AUC, macro recall, and G-mean, particularly under extreme imbalance conditions.

Li *et al.* [51] propose RAGFormer, a hybrid model that combines graph structure learning with transformer-based semantic encoding. The architecture consists of a semantic encoder using a Transformer to model node attributes, a Relation-Aware GNN to capture topological dependencies, and a self-attention-based fusion module that integrates both feature spaces. The authors highlight that the representations learned by GNNs and Transformers are nearly orthogonal yet complementary, and that combining them leads to more robust fraud detection. The model was evaluated on three datasets: YelpChi, Amazon, and Pay (a real-world credit card transaction dataset). RAGFormer consistently outperformed prior models, achieving state-of-the-art results in AUC, AP, and F1-macro, and demonstrating its strength in capturing both structure and semantics in multi-relational fraud graphs.

Relabeling heterogeneous subgraphs enables the sub-GNN to learn complex knowledge reasoning rules for accurate fraud detection. Publicly available Amazon and Yelp datasets are used for experiments, comparing Sub-GNN with three representative methods: labeled propagation, dense block mining, and GNN-based solutions. Results demonstrate the benefits and potential of the approach. Additionally, a dataset of fraudulent purchases from the popular Chinese e-commerce website Taobao is used to evaluate Sub-GNN. The method achieves a recall rate of over 90% for fraud examples and a precision score exceeding 0.99 [15].

Fraud detection systems in the e-commerce sector face challenges in keeping up with evolving fraud patterns, as they often rely on predefined ”fraudster seeds” set by domain experts. Zhang *et al.* [14] propose eFraudCom, a data processor and fraud detector based on Competitive Graph Neural Networks (CGNN) to address this issue. By independently modeling normal and fraudulent behavior, this system eliminates reliance on predefined fraudulent behavior. It achieves this by using certain normal behaviors as weak supervision information. The data processor constructs a multi-modal network of Taobao shoppers and sellers. To model the distributions of normal and fraudulent behaviors competitively, the detector uses CGNN, which includes dual graph decoders. Additionally, the mutual information regularization term enhances CGNN by improving the embedding subspace's ability to differentiate between normal and fraudulent behaviors [14].

Pourhabibi *et al.* [16] propose a classification framework that systematically explores the implementation of Graph-Based Anomaly Detection (GBAD) techniques, providing researchers with a deeper and more nuanced understanding of the field. This work highlights the need for new empirical research to address existing gaps and offers a roadmap for practitioners to understand the relationship between network characteristics, anomaly types, and the most suitable graph-based approaches for their specific use cases.

Graph-based techniques typically use Machine Learning (ML) or Deep Learning (DL) algorithms, but their implementations differ significantly. These techniques rely on linkages between data instances, which makes the nature of the data distinct from traditional datasets used in DL and ML, often stored in CSV files.

Note on Fair Comparison: The performance metrics summarized in Table VI are reported from different studies using various datasets, size/imbalance, and features, and are not fully comparable. See Table II for detailed dataset descriptions.

TABLE VI

GRAPH-BASED TECHNIQUES FOR E-COMMERCE FRAUD DETECTION. *Note: Dataset size, imbalance, and feature information are provided where available. Results are drawn from different studies and may not be directly comparable. See Table II for dataset details.*

Studies	Technique	Dataset	Size / Imbalance	Features	Results
Zhang *et al.* [14]	Competitive Graph Neural Networks (CGNN)	Two Taobao (multi-modal), two public datasets	Unspecified; Imbalanced	Shopper/seller behavior, transaction, graph structure	99.96% AUC
Song *et al.* [15]	SubGraphs Neural Network (SubGNN)	Amazon, Yelp, Taobao	Amazon/Yelp: Public, Taobao: Not fully public; Imbalance present	Subgraph topology, transaction records	Precision >0.99 recall >90%
Pourhabibi *et al.* [16]	Graph-based anomaly detection methods	39 research papers (review)	Varies (survey)	Multiple (survey)	ROC curves commonly used; survey of results
Li *et al.* [51]	RAGFormer (Transformer+GNN)	YelpChi, Amazon, Pay	YelpChi, Amazon: Public; Pay: Unspecified; Imbalance present	Node attributes, relational structure	SOTA AUC, AP, F1-macro across all datasets
Yan *et al.* [52]	FIW-GNN (heterogeneous GNN)	VESTA and Sparkov	VESTA: ∼590,000; 3.5% fraud; Sparkov: Unspecified	Transactional graph structure, features	Highest F1, AUC, and G-Mean on both datasets
Cheng *et al.* [65]	Spatial-Temporal Attention Graph NN (STAGN)	Major commercial bank	Unspecified; Highly imbalanced	Transaction features, temporal, spatial info	AUC 0.8973
Zhu *et al.* [66]	CNSGNN (cost-sensitive GNN)	MGM, Sichuan e-commerce	Unspecified; Extreme imbalance	Node attributes, edge structure	Outperformed baselines on AUC, macro recall, G-mean

F. Hybrid Approaches

Hybrid models have emerged as a powerful approach in fraud detection, combining multiple machine learning techniques to enhance predictive accuracy and mitigate the limitations of individual classifiers. These models are particularly useful in addressing issues such as data imbalance, false positives, and the adaptability of fraud detection systems across varying datasets. Tiwari *et al.* [67] propose a hybrid fraud detection model that integrates Decision Trees, Artificial Neural Networks (ANN), and K-Nearest Neighbors (KNN) to improve classification accuracy. Their approach assigns a label to each new transaction based on a majority vote from these three classifiers, effectively leveraging the strengths of each technique. Decision Trees provide interpretability and computational efficiency, ANN enhances pattern recognition capabilities, and KNN contributes to high accuracy on smaller datasets. Their experimental results on a European credit card dataset demonstrate an overall accuracy of 95.66%, outperforming individual classifiers such as Decision Trees (94.3%), KNN (96.96%), and ANN (96.55%). This study highlights the robustness of hybrid models in fraud detection, especially in cases where dataset characteristics vary significantly.

In machine learning applications such as fraud detection, imbalanced datasets may lead to suboptimal performance. De Zarza *et al.* [68] examine techniques for optimizing supervised learning algorithms in such situations, with a particular emphasis on resampling methods. Various methods are applied to a four-class spiral dataset, includ-

ing Gaussian Naive Bayes, linear and quadratic discriminant analysis, K-nearest neighbors, support vector machines, decision trees, and multilayer perceptron. Oversampling techniques achieve the best performance in the minority class, with an accuracy of 99.928% and a low number of false negatives. The results demonstrate that resampling strategies significantly improve model performance, especially in the minority class.

Chi *et al.* [69] propose an intrinsically interpretable fraud detection method based on quantitative argumentation. This method combines human-level and data-level knowledge to create an argumentative tree. It extends quantitative argumentation debates (QuAD) frameworks by adding correlation strength between arguments and using Particle Swarm Optimization (PSO) to identify this correlation strength. An empirical study using Ant Financial, Alibaba Group's financial services provider, evaluates the method's performance. The proposed method outperforms the existing DF-QuAD algorithm and is competitive with XGBoost, ANN, SVM, and LR.

Ndama *et al.* [70] introduce a novel hybrid framework that integrates Artificial Neural Networks (ANN) with Support Vector Machines (SVM) to enhance the robustness of credit card fraud detection. Their model employs ANN for feature extraction, followed by SVM for classification, capitalizing on ANN's adaptability in learning complex patterns and SVM's efficiency in maximizing margin separation. Experimental results on a dataset of 284,807 transactions demonstrate that this hybrid model outperforms traditional approaches, achieving a recall of 99.89% with an optimal false negative rate. The study highlights the efficacy of hybrid methodologies in mitigating data imbalance while improving classification accuracy.

Expanding on the integration of machine learning and deep learning, Berhane *et al.* [49] propose a hybrid Convolutional Neural Network-Support Vector Machine (CNN-SVM) model. Their approach replaces the final fully connected layer of CNN with an SVM classifier, optimizing decision boundaries for fraud detection. Evaluated on a real-world credit card dataset, the CNN-SVM model achieves an accuracy of 91.08%, demonstrating its effectiveness in distinguishing fraudulent from non-fraudulent transactions. The authors emphasize the significance of hybridization in fraud detection, particularly in cases where deep learning models struggle with overfitting or fail to generalize across diverse datasets.

Similarly, hybrid ensemble techniques have been explored to address challenges associated with fraud detection. Ndama *et al.* [71] integrate multiple classifiers, including Decision Trees, Gradient Boosting, and Random Forest, with an ANN-SVM framework to enhance predictive performance. Their research highlights that ensemble-based hybrid models significantly reduce false positives while maintaining high recall rates. Moreover, the inclusion of the Synthetic Minority Oversampling Technique (SMOTE) further mitigates class imbalance, reinforcing the importance of data preprocessing in fraud detection methodologies.

Saputra *et al.* [72] also use the SMOTE technique to address class imbalances in e-commerce transaction datasets. They compare the performance of Naive Bayes, Decision Tree, Neural Network, and Random Forest algorithms, finding that Neural Network achieved the highest accuracy at 96%. The application of SMOTE significantly improved the average F1-Score from 67.9% to 94.5% and the average G-Mean from 73.5% to 84.6%, highlighting the critical role of data preprocessing techniques in enhancing machine learning models' performance in fraud detection.

According to the study by Najem *et al.* [11], a comprehensive survey of fraud detection techniques in e-commerce focusing on studies between 2018 and 2020, the research explores data mining, machine learning, and hybrid methods. Among the evaluated methods, Random Forest demonstrated the highest accuracy (99.96%), followed by Artificial Neural Networks (98.69%) and Support Vector Machines (93.96%). The study emphasizes the effectiveness of hybrid approaches, particularly the integration of SMOTE (Synthetic Minority Oversampling Technique) with Random Forest, which significantly improved fraud classification in highly imbalanced datasets.

Tang *et al.* [48] introduce a federated graph learning-based hybrid model, integrating Federated Learning (FL) and Graph Neural Networks (GNN) to address privacy concerns and enhance fraud de-

tection across multiple financial institutions. Their approach constructs transaction graphs using a graph construction algorithm based on weighted feature similarity, mapping local financial data into structured graph representations. Additionally, they propose a graph extension algorithm utilizing convolutional feedforward generative (CFG) networks, which extends transaction graphs across institutions while preserving privacy. Their results demonstrate that this federated approach improves fraud detection performance, achieving 11.87% to 33.9% higher recall and over 3% higher AUC on the IEEE-CIS dataset compared to traditional methods.

Akshay *et al.* [73] introduce a hybrid deep learning model that integrates Graph Neural Networks (GNN) and Autoencoders to enhance fraud detection accuracy. Their model leverages Autoencoders, Generative Adversarial Networks (GANs), XGBoost, Gated Recurrent Units (GRU), and CatBoost to improve detection performance. The proposed approach utilizes community detection algorithms, including Louvain Modularity, Girvan-Newman, and Label Propagation, to enhance graph-based anomaly detection. Experimental results indicate that the three-layer GNN model achieves an accuracy of 79%, outperforming conventional fraud detection models.

In addition, Alarfaj [74] presents a deep learning-based hybrid model that integrates Graph Neural Networks (GNN) and Autoencoders to improve fraud detection in real-time financial transactions. Their model applies Lambda Architecture, combining batch processing for historical data with real-time streaming for immediate fraud detection. A Graph Temporal Attention Network (GTAN) is used to construct dynamic transaction graphs, where nodes represent transactions, accounts, and devices, while edges define transactional dependencies. Evaluated on a dataset of 284,807 transactions from Meezan Bank and UBL, the proposed model achieves an accuracy of 99.75%, effectively identifying fraudulent transactions while maintaining a false positive rate of 0.25%. The study underscores the advantages of graph-based learning and anomaly detection techniques in improving fraud detection accuracy.

Note on Fair Comparison: The performance results in Table VII are reported from studies using various datasets, sample sizes, imbalance ratios, and features, and should not be interpreted as direct comparisons. Please refer to Table II for detailed dataset descriptions.

G. Real-World Applications and Case Studies

While theoretical models and detection strategies form the foundation of e-commerce fraud prevention, real-world implementations illustrate both the challenges and practical benefits of deploying such systems at scale. Major platforms like Amazon, PayPal, and Alibaba have adopted diverse strategies integrating machine learning and cloud-based systems to combat fraudulent transactions.

Amazon utilizes cloud infrastructure such as AWS SageMaker and Amazon Fraud Detector to deploy real-time fraud detection workflows. These systems are integrated with services like AWS Lambda and DynamoDB to enable low-latency responses and high scalability [75], [76].

PayPal has implemented real-time data processing platforms (e.g., Aerospike and Intel Optane) to reduce fraudulent transaction misses by a factor of 30 while reducing infrastructure overhead. Behavioral forensics are also used for tracking fraud vectors, highlighting the importance of digital evidence [77], [78].

Alibaba's Alipay has adopted advanced machine learning frameworks such as Dual Importance-aware Factorization Machines (DIFM) and noisy-label correction models. These frameworks have proven effective in high-volume, label-ambiguous environments common to e-commerce fraud scenarios [79], [80].

These implementations reflect the complexity and dynamism of fraud prevention in global e-commerce. The combined use of ML automation, real-time systems, and human-in-the-loop validation underpins modern fraud detection architectures in leading platforms.

H. Comparison and Discussion

As electronic payment methods, particularly credit cards, have become increasingly popular, they have also become prime targets for criminals seeking to intercept and exploit private data. This growing threat underscores the urgent need for effective credit card fraud detection systems,

TABLE VII

HYBRID MODEL TECHNIQUES FOR E-COMMERCE FRAUD DETECTION. *Note: Dataset size, imbalance, and feature information are provided where available. Results are drawn from different studies and may not be directly comparable. See Table II for dataset details.*

Studies	Technique	Dataset	Size / Imbalance	Features	Results
Najem *et al.* [11]	DT RF KNN NN Fuzzy Logic SVM Naive Bayes	Various (credit card/e-commerce, 2018–2020)	Most <1% fraud, e.g., European dataset: 284,807/492 (highly imbalanced)	Transaction/ behavioral features	RF was the most frequently used method (32 studies reviewed), with the highest reported accuracy of 99.96% among ensemble methods
Tang *et al.* [48]	Federated Learning GNN CFG networks	IEEE-CIS dataset	∼590,000; 3.5% fraud	Graph-based transaction features	Recall improved by 11.87–33.9%, AUC >3% over baselines
Berhane *et al.* [49]	CNN-SVM hybrid	European credit card dataset	284,807; 0.17% fraud	PCA-applied V1-V28, Time, Amount, Class (Label)	Accuracy: 91.08%, better than standalone models
Tiwari *et al.* [67]	DT ANN KNN with voting	European credit card dataset	284,807; 0.17% fraud	PCA-applied V1-V28, Time, Amount, Class (Label)	DT: 94.3% ANN: 96.55% KNN: 96.96%,
De Zarza *et al.* [68]	GNB LDA QDA KNN SVM DT MLP	Credit card transaction dataset	Unspecified; Imbalanced	Transaction features	MLP + Oversampling: 99.93% accuracy; MLP: 99.94% accuracy (original imbalance)
Chi *et al.* [69]	O-QuAD XGBoost ANN SVM LR	Ant Financial (Alibaba)	Unspecified; Imbalance	Transaction and knowledge-based features	O-QuAD: 95.2% recall XGBoost: 94.5% precision
Ndama *et al.* [70]	ANN for feature extraction, SVM for classification	European credit card dataset	284,807; 0.17% fraud	PCA-applied V1-V28, Time, Amount, Class (Label)	Recall: 99.89%, optimal FNR
Ndama *et al.* [71]	DT, GB, RF + ANN-SVM, SMOTE	European credit card dataset	284,807; 0.17% fraud	PCA-applied V1-V28, Time, Amount, Class (Label)	Reduced false positives, high recall, SMOTE mitigates imbalance
Saputra *et al.* [72]	DT NB RF NN SMOTE	E-commerce fraud dataset	Unspecified; Imbalanced	Transaction features	NN: 96% RF/NB: 95% SMOTE improved F1 from 67.9% to 94.5%
Akshay *et al.* [73]	GNN, Autoencoder, GAN, XGBoost, GRU, CatBoost	Credit card transaction dataset	Unspecified; Imbalanced	Graph/sequence features, community detection	3-layer GNN: 79% accuracy, outperforming baselines
Alarfaj [74]	GNN, Autoencoder, GTAN, Lambda Architecture	Meezan Bank & UBL: 284,807	284,807; Imbalance present	Transaction, account, device, temporal graph features	Autoencoder: High accuracy 99.75%, detected all frauds in test set with 5 false positives in 2000 samples

which has driven extensive research in this area. Initial approaches to fraud detection relied heavily on rule-based systems and traditional statistical methods. While these methods provide a strong foundation, they also have notable limitations, particularly in adapting to new and sophisticated fraud patterns. Consequently, the growing importance of safeguarding these transactions has motivated the development of fraud detection systems that span from traditional rule-based approaches to advanced machine learning models like decision trees, logistic regression, and neural networks. While many of these models analyze aggregate data, such as a user's transaction history over time, certain types of fraud can only be detected by closely examining individual actions rather than relying solely on aggregated information [63].

Rule-based systems have long been a staple in fraud detection. These systems operate by applying predefined criteria to flag potentially fraudulent transactions, such as those exceeding specific thresholds or originating from unusual locations. The key advantage of rule-based systems lies in their interpretability and the ease with which they can be adjusted to new fraud tactics. However, they require regular updates to remain effective against the constantly evolving landscape of fraud patterns [53].

In conjunction with rule-based systems, statistical methods such as logistic regression and decision trees have been employed to assess fraud risk quantitatively. These methods analyze multiple transaction attributes and classify transactions based on patterns identified from historical data [56]. However, their reliance on past data can limit their effectiveness in detecting new types of fraud, as they may struggle to keep pace with emerging fraud tactics that do not fit established patterns.

To address the limitations of traditional approaches, machine learning (ML) techniques have emerged as a powerful tool in the fight against fraud. ML models, particularly neural networks and deep learning algorithms, have demonstrated superior accuracy compared to traditional statistical methods. These models excel at processing large datasets and can adapt more effectively to new and evolving fraud patterns [56]. The integration of ML techniques with rule-based systems has further enhanced fraud detection by combining the interpretability of rules with the adaptability and robustness of ML models [53].

Despite the advancements made possible by ML, challenges remain. One major issue is the constant emergence of new and more sophisticated fraud tactics, such as triangulation fraud, which necessitates ongoing adaptation and refinement of detection methods [81]. As fraudsters develop new methods to evade detection, the effectiveness of ML models can diminish over time, requiring frequent updates, retraining, and assessments to maintain their efficiency.

In response to these challenges, new techniques such as graph-based methods have been explored. These methods are particularly useful for analyzing relationships within large volumes of data, which is essential for detecting complex fraud schemes. However, the sheer scale of data that organizations typically manage can lead to performance issues with graph-based models [14], [15]. Despite these challenges, research indicates that graph-based techniques are evolving to better meet the demands of big data environments, offering promise for future fraud detection efforts [1], [16], [82].

Building on these developments, hybrid models represent a significant advancement in fraud detection by effectively integrating multiple machine learning techniques to optimize predictive performance. These hybrid approaches overcome limitations inherent to individual methods, addressing challenges such as data imbalance, high false-positive rates, and the dynamic nature of evolving fraud tactics. For instance, Tiwari et al. [67] combined Decision Trees, Artificial Neural Networks (ANN), and K-Nearest Neighbors (KNN), capitalizing on each classifier's unique strengths to achieve remarkable accuracy. Similarly, hybrid frameworks integrating ANN with Support Vector Machines (SVM) have demonstrated superior performance, significantly reducing false negatives and enhancing overall accuracy [49], [70]. Furthermore, federated graph learning strategies, which merge Graph Neural Networks (GNN) with Federated Learning (FL), effectively address privacy concerns while simultaneously improving detec-

tion accuracy in collaborative environments spanning multiple institutions [48].

Another critical factor in modern fraud detection is the quality and timeliness of training data. As fraud tactics evolve, the data used to train detection models can quickly become outdated, reducing the effectiveness of even the most advanced ML algorithms. Creating features that accurately represent new fraud types requires significant time and expertise. Additionally, while decision trees are valued for their transparency in explaining fraud detection decisions, they may not be sufficient for detecting more complex fraud schemes. This necessitates the use of deep learning algorithms for more challenging detection tasks [81].

In conclusion, while traditional rule-based and statistical methods have laid the groundwork for fraud detection, the evolution of fraud tactics has necessitated the integration of more advanced machine learning techniques. Emerging methods like graph-based analysis show promise but also face challenges related to performance and data management. As fraud tactics continue to evolve, ongoing research, frequent model updates, and the development of specialized detection features are crucial to maintaining the effectiveness of fraud detection systems. An important limitation in current e-commerce fraud detection systems is their insufficient capacity to detect cross-channel fraud, where attackers exploit inconsistencies or blind spots between different customer touchpoints (e.g., website, mobile app, call center, or in-store platforms). Traditional models often operate in silos, failing to correlate fraudulent signals that span across multiple channels or devices. Recent studies emphasize the importance of integrating behavioral, transactional, and contextual data streams to support real-time, cross-channel anomaly detection [83]–[85]. These insights highlight the urgent need for unified fraud detection platforms capable of detecting multi-channel attack vectors and reducing fragmentation across systems.

To facilitate practical decision-making, we summarize the main strengths and weaknesses of each representative technique in Table VIII. Techniques are evaluated by their reported accuracy, scalability to large datasets, interpretability (ability to explain model decisions), and real-world applicability based on the literature.

IV. Geographic and Cultural Context

While e-commerce fraud detection has been widely studied in Western contexts, recent research highlights important geographic and cultural differences in fraud patterns, technology adoption, and regulatory response:

- **Asia-Pacific:** Fraud detection techniques must adapt to the rapid digitization and unique transaction ecosystems found in Asia-Pacific countries. For example, studies using real transaction data from Turkey and Indonesia show that local online behaviors and payment preferences strongly affect the design and success of fraud detection models. Machine learning approaches, such as decision trees, random forests, and neural networks, are increasingly applied to large datasets from these regions to detect local patterns of fraud [86], [87].
- **Africa and Developing Economies:** In Africa, the adoption of digital payments and e-commerce is rapidly growing, but so are fraud risks due to gaps in regulatory enforcement and lower digital literacy. In Nigeria, for example, fraud detection models based on logistic regression have been implemented to address local challenges at point-of-sale, with government initiatives driving broader technology adoption [88]. Research also emphasizes the need for adaptable detection models trained on local transaction patterns and regulatory frameworks [89].
- **Cross-Regional and Global Perspectives:** A major challenge for global e-commerce platforms is transferring fraud detection knowledge across different regions. New methods in continual and transfer learning, such as cross-regional continual learning and heterogeneous trade graph models, are now being used to adapt to regional transaction semantics without retraining from scratch, helping global companies mitigate fraud in diverse markets [90], [91].

These recent advances demonstrate that effective e-commerce fraud detection must consider regional variations in consumer behavior, transaction

TABLE VIII

COMPARATIVE SUMMARY OF E-COMMERCE FRAUD DETECTION TECHNIQUES BY CATEGORY. *Techniques are compared on accuracy, scalability, interpretability, real-world applicability, cost, and data requirements, synthesized from the cited literature.*

Category	**Technique**	**Accuracy**	**Scalability**	**Interpretability**	**Real-World Applicability**	**Cost**	**Data Requirements**
Rule-Based	CRED (Bekach et al. [53])	99.83%	High	High (explicit rules)	Widely used	Low	Low (hand-crafted rules)
Statistical/ML	Logistic Regression (Valli et al. [57])	Moderate	High	High	Yes	Low	Moderate (labeled features)
Statistical/ML	Random Forest (Hasan et al. [17])	84.55%	High	Medium	Yes	Low–Moderate	Moderate (labeled data)
Statistical/ML	LightGBM (Fang et al. [59])	Recall: 99%	High	Low	Yes	Moderate	High (large labeled data)
Deep Learning	BiLSTM–GRU (Najadat et al. [12])	AUC: 91.37%	Medium	Low	Yes	High	High (large labeled sequences)
Deep Learning	DNN (Alkhatib et al. [4])	ROC AUC: 99.1%	Medium	Low	Yes	High	High (large labeled data)
Graph-Based	Sub-GNN (Song et al. [15])	Prec >0.99, Rec >90%	Medium	Medium (rule learning)	Yes	High	Very High (graph-structured data)
Graph-Based	CGNN (Zhang et al. [14])	AUC: 99.96%	High	Medium	Yes	High	Very High (graph, behavior, multi-modal)
Graph-Based	CNSGNN (Zhu et al. [66])	High AUC, Macro Recall	High	Medium (label-aware)	Yes	High	Very High (graph, labeled)
Graph-Based	RAGFormer (Li et al. [51])	SOTA AUC, F1-macro	Medium	Medium–High (fusion attention)	Yes	High	Very High (graph+attribute)
Hybrid	CNN–SVM (Berhane et al. [49])	91.08%	Medium	Low	Yes	High	High (diverse, labeled)
Hybrid	ANN–SVM (Ndama et al. [70])	Recall: 99.89%	Medium	Low	Yes	High	High (diverse, labeled)
Hybrid	Federated GNN (Tang et al. [48])	AUC ↑3%, Recall ↑33.9%	Medium	Medium	Yes	High	Very High (distributed, privacy-sensitive)

technologies, and regulatory environments. Future research should further prioritize building models that can be efficiently transferred and adapted across geographic contexts.

A. Gaps in the Existing Literature

Despite significant advances, notable gaps remain in the research and deployment of e-commerce fraud detection systems. The following summarizes the principal limitations for each major technique or fraud type, as reported in recent literature:

- **Rule-Based Systems**: These methods often rely on hand-crafted rules derived from historical fraud patterns. Literature highlights that rule-based systems are not robust to novel or adaptive fraud strategies and require frequent manual updates [53], [55]. They are rarely

evaluated on truly large or evolving datasets, limiting their adaptability in real-world scenarios.

- **Traditional Statistical and Machine Learning Models**: Models such as logistic regression and random forest require large, high-quality labeled datasets, which are often unavailable due to privacy and class imbalance [56], [57]. Many studies lack robustness testing across diverse e-commerce environments and seldom address concept drift or real-time processing needs. Feature engineering remains largely manual, and models may underperform on new fraud types.
- **Deep Learning Approaches**: While deep learning models (such as DNNs and BiLSTM-GRU) achieve high accuracy, they typically demand massive labeled datasets, extensive computational resources, and are seldom deployed in real-time production environments [4], [12]. Many published models have only been tested on static, retrospective datasets, with little discussion of robustness to adversarial attacks or evolving fraud tactics. Model interpretability also remains a major barrier to adoption.
- **Graph-Based Methods**: The literature shows that graph neural networks excel at capturing relational fraud but require complex, often proprietary, graph-structured data that is rarely shared publicly [14], [15], [51]. There is limited benchmarking across large, dynamic, or cross-platform datasets, and few studies report real-time scalability. Current research also lacks standard protocols for evaluating robustness or generalization to unseen fraud schemes.
- **Hybrid and Ensemble Models**: Hybrid approaches combining ML, DL, or GNNs show strong performance but suffer from compounded data and resource requirements [49], [48], [70], [67]. Few works rigorously test these models in live, production-scale e-commerce systems or address their maintainability over time. Real-time fraud detection using hybrid models remains largely experimental.
- **Fraud Type Coverage**: Most research focuses on credit card or transaction fraud; emerging fraud types such as triangulation fraud, synthetic identities, or cross-platform collusion are underexplored [1], [81]. Datasets representing these complex fraud patterns are scarce, and few benchmarks address the unique challenges they pose.

Addressing these gaps requires: (1) public release of large, representative, and regularly updated datasets; (2) rigorous robustness and adversarial testing; (3) development of interpretable, scalable, and real-time capable models; and (4) expanded coverage of novel and complex fraud types.

B. Practical Implementation in Real-World E-Commerce Platforms

Deploying fraud detection techniques in production e-commerce environments involves multiple technical challenges and considerations:

- **Integration Architecture:**
  - *Microservices & REST APIs:* Most fraud detection models are implemented as RESTful microservices, which can be invoked by the main transaction processing workflow (e.g., via Python Flask, FastAPI, or Java Spring Boot).
  - *Batch vs. Real-time Scoring:* Lightweight models (rule-based, tree-based) can be deployed directly in transaction flows for real-time scoring. Heavier models (deep learning, GNNs) may run asynchronously or as batch jobs, with high-risk scores flagged for immediate review.
  - *Cloud and On-Premises:* Cloud platforms (AWS SageMaker, GCP AI Platform, Azure ML) enable scalable deployment and auto-scaling, while on-premises deployments may require GPU/TPU servers for deep learning inference.
- **Latency and Throughput:**
  - *Low-Latency Inference:* For fraud detection in e-commerce, response times must be in the low milliseconds to avoid user friction. Rule-based and tree-based models (Random Forest, LightGBM) can be optimized with ONNX or XGBoost serving for sub-10ms inference.

  - *Accelerated Deep Models:* Use TensorRT, ONNX Runtime, or TorchServe for deploying DNNs; employ model quantization and pruning to reduce inference time.
  - *Edge/On-Device Inference:* For ultra-fast or privacy-sensitive use cases, lightweight models can run on the e-commerce server or even in-app, but most complex models remain server/cloud-side.
- **Scalability:**
  - *Horizontal Scaling:* Use Kubernetes or Docker Swarm to deploy multiple replicas of the scoring service for high transaction volumes.
  - *Big Data Pipelines:* For offline analysis and model retraining, tools such as Apache Spark, Kafka Streams, or Airflow are used to process billions of records at scale.
  - *Graph Data Handling:* Graph-based models may require distributed graph databases (e.g., Neo4j Aura, Amazon Neptune, TigerGraph) and graph partitioning for scalable GNN inference [14], [51].
- **Continuous Learning and Monitoring:**
  - *Concept Drift Handling:* Use model monitoring tools (e.g., Evidently AI, Amazon SageMaker Model Monitor) to detect drift in fraud patterns and trigger retraining.
  - *Automated Retraining:* Implement CI/CD pipelines (e.g., with GitHub Actions, Jenkins) to retrain models as new labeled data arrives.
  - *Explainability:* Integrate SHAP or LIME for tree/ML models and GNNExplainer for graph models to support regulatory compliance and analyst trust [53].
- **Security and Privacy:**
  - *Secure APIs:* Use authentication (OAuth2, JWT) and encryption (TLS) for all fraud detection endpoints.
  - *Federated Learning:* For cross-institutional fraud detection, apply federated learning frameworks (e.g., TensorFlow Federated) to train models without sharing sensitive transaction data [48].
  - *Data Anonymization:* Apply tokenization or masking to sensitive features during both model training and inference.

In summary, while rule-based and ML models can be integrated with minimal engineering effort, deep learning, hybrid, and graph-based techniques require advanced infrastructure for real-time, scalable, and secure deployment in high-volume e-commerce environments.

V. Open Challenges

The challenge of fraud detection is non-trivial, as it involves adopting an adversarial stance against fraudsters. Existing systems in the e-commerce industry may experience performance decay due to their limited adaptability to changes in fraud patterns [14]. The core challenges of fraud detection systems are outlined below, along with their potential solutions. However, despite the availability of these solutions, limitations persist that necessitate further research and innovation.

While many technical advances in e-commerce fraud detection have demonstrated strong experimental results, real-world implementation remains complex. Each of the challenges discussed below has significant implications for operational deployment. These include integration into legacy systems, maintaining real-time responsiveness at scale, ensuring compliance with regulations, and managing resource constraints in smaller organizations. By exploring these challenges through both a technical and practical lens, we aim to highlight critical gaps between research innovation and deployment feasibility.

A. Ethical and Regulatory Considerations

Given the sensitive nature of fraud detection involving financial and personal data, addressing ethical and regulatory considerations is essential for real-world deployment [92], [93]. Privacy concerns stem from the extensive collection and processing of user information, which elevates the risk of data breaches and unauthorized access if not protected by robust safeguards [93], [94]. Recent research emphasizes the need for privacy-by-design architectures, continuous data governance,

and proactive oversight to ensure that fraud detection systems maintain trust and comply with data protection frameworks such as the GDPR and the EU AI Act [92], [94], [95].

Algorithmic bias presents another major challenge. Fraud detection models trained on historical or imbalanced data may inadvertently perpetuate discrimination, resulting in unfair or opaque outcomes [96]–[98]. Addressing bias requires the adoption of fairness-aware model development, routine algorithmic audits, and increased transparency in decision-making to ensure that fraud prevention systems are both equitable and accountable.

Compliance with evolving regulatory frameworks, including GDPR, the EU AI Act, and similar policies, is now a fundamental requirement for the responsible deployment of AI-powered fraud detection [92], [95]. Effective governance calls for clear accountability structures, regular audits, and transparent reporting, which are essential for meeting legal and ethical standards and fostering public trust. As regulatory landscapes evolve, organizations must prioritize comprehensive and adaptive governance strategies to ensure fair, ethical, and legal use of AI in fraud prevention [92], [95].

B. Organization Centric Schema

Tax *et al.* [99] recently proposed an organization-centric schema for fraud detection systems emphasizing machine learning's potential to enhance both accuracy and efficiency. However, a key challenge lies in the manual involvement required during the system's initial setup. Although subsequent actions are handled by automated detectors early inaccuracies stemming from this manual phase can lead to erroneous decisions potentially degrading customer experience. While machine learning approaches show promise, their effectiveness hinges on overcoming these foundational limitations.

One proposed solution is the adoption of automated onboarding processes that integrate seamlessly with fraud detection systems. By minimizing manual input automation can streamline the setup phase, reducing inefficiencies and the risk of human error. Furthermore, as Tax *et al.* [99] suggest, continuous monitoring and adaptation through machine learning can refine the system over time, enabling it to better detect evolving fraud patterns. This dynamic evolution could mitigate the customer experience issues caused by incorrect decisions, offering a more robust and scalable framework.

Despite these advances, existing solutions exhibit notable drawbacks that necessitate further research. Automated onboarding while reducing manual effort, often lacks the flexibility to adapt to highly variable organizational contexts or rapidly shifting fraud tactics limiting its scalability in dynamic e-commerce environments. Similarly, continuous monitoring via machine learning relies heavily on the quality and diversity of training data biased or incomplete datasets can perpetuate errors undermining detection accuracy. These limitations are particularly pronounced in real-time applications, where delays in adapting to new fraud patterns or false positives can erode user trust and satisfaction. Moreover, the computational overhead of constant model retraining may strain resources posing challenges for smaller organizations.

To address these gaps researchers must focus on enhancing the scalability and adaptability of fraud detection systems. For instance, developing more context-aware onboarding algorithms could better accommodate diverse operational needs, while advanced data augmentation techniques might improve model robustness against emerging threats. Additionally, optimizing resource efficient learning methods could broaden accessibility. The persistent risk of incorrect decisions further underscores the need to refine automation processes ensuring minimal impact on customer experience. Future work should prioritize these areas optimizing onboarding flexibility, improving adaptability to novel fraud patterns, and reducing error rates to advance the state-of-the-art in e-commerce fraud detection.

C. Imbalanced Dataset

The challenge posed by imbalanced datasets in e-commerce fraud detection arises from the substantial abundance of legitimate transactions compared to the limited number of fraudulent ones. This imbalance adversely impacts the performance of machine learning models employed for fraud detection [68], [100]. Consequently, the model's ability to identify and classify fraudulent

transactions is compromised, leading to potential oversight of true threats.

This class imbalance often leads to misleadingly high accuracy while masking poor recall on the minority (fraudulent) class. Since fraudulent transactions can represent less than 0.5% of total activity, models trained on raw data tend to overwhelmingly predict the majority class, resulting in high false negative rates. In fraud detection, these false negatives are particularly costly because they represent undetected threats. Therefore, evaluation metrics such as F1-score, AUC-PR, and recall for the minority class are preferred over overall accuracy [101], [102].

To address the issue of imbalanced datasets, researchers have explored various techniques, including oversampling the minority class, undersampling the majority class, cost-sensitive learning, and advanced methods like synthetic data generation [103]–[107]. Hybrid models have also been proposed, such as MLPs and ensemble learning architectures optimized with GAN-augmented sampling [108]–[110]. Recent innovations like DOS-GNN apply graph-based over-sampling and dual feature aggregation to improve detection of rare events in highly imbalanced environments [111].

These approaches have demonstrated notable improvements in fraud detection performance, especially in recall and AUC-PR. However, research continues to show that there is no universal solution; model generalizability and real-time deployment remain open challenges [102], [112]. Future work must also focus on developing benchmark datasets and domain-specific evaluation strategies that fairly reflect model performance on minority classes.

By investing in continued innovation on imbalanced learning, the research community can contribute to building more robust and accurate fraud detection systems, ultimately strengthening financial safety and consumer trust in e-commerce environments.

D. Model interpretability

In e-commerce fraud detection, a key challenge is developing machine learning models that are both highly effective and interpretable [113]. Advanced models, such as deep neural networks often rely on complex decision making processes that obscure the reasoning behind their predictions [114]. This lack of interpretability reduces transparency, making it difficult for stakeholders to understand why a transaction is flagged as fraudulent. As a result, trust in fraud detection systems may be weakened, and decision-making processes can become more complex.

Various approaches have been proposed to address this challenge. Jhangiani et al. [115] advocate for supervised machine learning models, such as Random Forest and ensemble techniques, prioritizing predictive accuracy and robustness in fraud detection. In contrast, Tax et al. [99] emphasize the need for interpretability and transparency, highlighting their role in supporting organizational decision-making and investigator trust. Bhowmik et al. [116] introduce a Deep Belief Network-based approach aimed at balancing effectiveness with interpretability. These perspectives align with broader scholarly discussions [117]–[119], which underscore that interpretability is essential not only for model acceptance but also for ensuring compliance with societal and ethical norms—key factors in fostering stakeholder trust and the successful deployment of e-commerce fraud detection systems.

Despite advancements in fraud detection, existing solutions have notable limitations that underscore the need for further research. Jhangiani et al. [115] primarily focus on optimizing predictive accuracy in fraud detection using supervised machine learning models but do not explicitly address interpretability. In contrast, Tax et al. [99] highlight the importance of interpretability and transparency, particularly in supporting fraud investigators and enhancing organizational decision-making. While they explore Explainable AI methods, such as SHAP and rule-based explanations, they acknowledge the trade-offs between interpretability and predictive performance, especially in complex e-commerce fraud scenarios. Similarly, Bhowmik et al. [116] propose a Deep Belief Network-based approach that, despite its promise, struggles to fully explain intricate feature interactions, limiting its interpretability. Moreover, these approaches often require significant trade-offs in computational efficiency, scalability, or adaptability to rapidly evolving fraud patterns, reducing their practicality in dynamic e-commerce environments. General

interpretability techniques, such as those discussed by Ribeiro et al. [118], typically provide post-hoc explanations that may not accurately reflect the model's internal logic, further complicating trust and validation.

These shortcomings highlight the urgent need for continued research on model interpretability in e-commerce fraud detection. Interpretable models are crucial for ensuring transparency and accountability, particularly in sensitive domains where automated decisions affect both customers and businesses. Providing clear rationales behind predictions fosters stakeholder trust, facilitates system evaluation, and encourages broader adoption. Additionally, interpretability offers valuable insights into fraud patterns, enabling organizations to develop targeted countermeasures. As regulatory frameworks tighten, explainable models also support compliance by providing justifiable decisions. Future research should focus on mitigating current trade-offs—designing models that preserve high accuracy without compromising transparency, enhancing real-time adaptability, and ensuring explanations are both precise and actionable. Advancing these areas is essential to improving the effectiveness, trustworthiness, and regulatory alignment of fraud detection systems in the evolving e-commerce landscape.

E. Adversarial Attacks

E-commerce fraud detection systems face growing vulnerability to adversarial attacks, which compromise their ability to accurately and efficiently identify fraudulent behavior. These attacks exploit system weaknesses by injecting manipulated data or deploying sophisticated adversarial examples that target machine learning models directly [120]. When successful, such assaults degrade detection performance, leading to significant financial losses, reputational harm, and diminished trust in business operations.

To counter these threats, researchers have proposed multiple strategies. Data preprocessing techniques, such as anomaly and outlier detection, filter problematic inputs before they affect model training [121]. Advanced machine learning models that detect complex patterns beyond traditional rule-based systems bolster resilience against evasion tactics [122]. Feature engineering refines detection accuracy by improving the extraction of pertinent data, while training models with adversarial examples strengthens resistance to manipulation [123]. Real-time monitoring and anomaly detection enable swift identification of suspicious activities, and adaptive strategies ensure systems evolve with emerging fraud patterns [86], [124].

Targeted defenses against adversarial attacks further enhance robustness. Techniques like adversarial training, input sanitization, robust optimization, and defensive distillation proactively equip models to recognize and resist manipulative inputs [120], [123]. Hybrid approaches combining these methods create layered defenses, offering greater protection against diverse and evolving threats.

Despite these efforts, existing solutions exhibit significant limitations that necessitate further research. Data preprocessing techniques like outlier detection [121] often struggle to distinguish between legitimate anomalies and adversarial inputs, risking false positives or missed threats. Machine learning models designed for pattern recognition [122] may fail to keep pace with rapidly evolving attack strategies, while feature engineering requires extensive domain expertise and may not scale across diverse e-commerce contexts. Adversarial training [123] is resource-intensive, typically tailored to known attack vectors, and may lack effectiveness against novel, unforeseen threats. Real-time monitoring and adaptive approaches [86], [124], though effective in theory can incur high computational costs and latency, limiting their practicality for large-scale systems. Techniques designed explicitly for adversarial defense, like robust optimization or defensive distillation, may introduce significant computational overhead, hindering scalability in high volume transaction environments.

These shortcomings highlight the urgent need for continued research into adversarial attacks on fraud detection systems. Enhanced solutions must overcome trade-offs between accuracy, scalability, and resilience, ensuring systems adapt effectively to sophisticated, unseen threats without compromising performance. Developing more efficient preprocessing methods that reliably isolate adversarial inputs, designing scalable feature

engineering frameworks, and creating lightweight adaptive defense models are critical steps forward. Additionally, integrating model interpretability within adversarial robustness strategies will facilitate transparency, enabling proactive identification and mitigation of potential vulnerabilities. Addressing these gaps is essential to safeguarding e-commerce systems against the escalating sophistication of adversarial attacks, protecting businesses and consumers in an ever-shifting threat landscape.

F. Dynamic Adaptation to Emerging Fraud Techniques

The ever-evolving nature of e-commerce fraud demands detection systems that can swiftly adapt to new tactics devised by fraudsters. Traditional models, and even some advanced machine learning approaches, often depend on historical data patterns, rendering them ill-equipped to anticipate or detect novel fraud techniques [1], [113]. This reliance on outdated indicators highlights the urgent need for systems capable of real-time or near-real-time adaptation to counter emerging threats effectively.

Researchers have proposed several solutions to tackle this challenge. Continual learning models and adaptive algorithms, which incrementally update their knowledge base without requiring full retraining, offer a promising avenue for keeping pace with evolving fraud patterns [125]. Similarly, anomaly detection systems that focus on deviations from normal transaction behavior rather than solely relying on past fraud signatures provide an alternative means to identify previously unseen schemes.

Despite these advancements, current solutions exhibit significant limitations that necessitate further research. Continual learning models [125], while flexible, often suffer from catastrophic forgetting, where updates to accommodate new patterns degrade performance on previously learned fraud types. This trade-off compromises long term reliability, especially in diverse e-commerce settings. Adaptive algorithms, though efficient in theory, require substantial computational resources to process incoming data continuously, limiting their scalability for platforms handling massive transaction volumes. Anomaly detection systems, meanwhile, face challenges with high false positive rates, as distinguishing benign outliers from fraudulent deviations remains imprecise without robust contextual understanding. Moreover, these approaches often lack cross domain adaptability, struggling to generalize insights across varied e-commerce platforms or industries due to differences in transaction norms and fraud behaviors.

These shortcomings emphasize the need for continued research to enhance the adaptability of fraud detection systems. Future efforts should prioritize developing algorithms that mitigate forgetting in continual learning, ensuring consistent performance across both old and new fraud patterns. Lightweight adaptive models that scale efficiently with high transaction throughput are also critical, as is refining anomaly detection to reduce false positives through improved context awareness. Additionally, creating comprehensive, up-to-date datasets that capture emerging fraud instances will bolster model training, while cross domain frameworks could enable knowledge transfer across platforms and sectors. Addressing these gaps—balancing adaptability, scalability, and accuracy is essential to equip e-commerce fraud detection systems to handle the dynamic, unpredictable nature of modern fraud threats.

VI. Future Directions

A. Explainable AI

Advancements in Explainable AI (XAI) techniques [126], [127] are pivotal for enhancing transparency, trust, and operational efficacy in e-commerce fraud detection. As detection models become more complex and data-driven, XAI approaches, including SHAP, LIME, rule-based explanations, and model-agnostic frameworks—— play a vital role in demystifying model predictions and providing actionable, understandable rationales for flagged transactions [127]–[129]. Integrating XAI into fraud detection supports regulatory compliance, improves risk mitigation, facilitates collaboration between analysts and AI, and ultimately enhances decision-making and customer trust in digital commerce [126], [128].

Despite its benefits, implementing XAI in real-world fraud detection poses ongoing challenges, such as the need for scalable and domain-adaptive interpretability, balancing explanation fidelity with model performance, and ensuring explanations are

meaningful for human experts [127], [129]. Continued research is focused on refining interpretability techniques tailored to the unique complexities of fraud detection, with the goal of making explanations more accurate, actionable, and auditable. As XAI methods advance, they are expected to strengthen the resilience and accountability of fraud detection systems, supporting responsible, trustworthy, and transparent AI use in high-stakes financial contexts [126], [128].

B. Use of IDS in Fraud Detection

Intrusion Detection Systems (IDS) can enhance security measures and monitoring capabilities in fraud detection, significantly contributing to e-commerce fraud prevention. IDS are instrumental in identifying suspicious activities and patterns within network traffic, enabling real-time detection of fraudulent transactions and activities. Integrating IDS into fraud detection systems offers benefits such as heightened detection accuracy [11], [12], [17], [18], [35], early identification of fraudulent behavior, and improved response capabilities to mitigate potential risks. By continuously monitoring network traffic and analyzing data for anomalies or suspicious patterns, IDS enhances fraud detection processes, swiftly identifying potential threats and enabling prompt response actions. Utilizing IDS in fraud detection is crucial for addressing the challenges of e-commerce fraud, providing real-time monitoring, detection, and response capabilities essential for safeguarding against evolving fraudulent tactics and protecting businesses and consumers from financial losses and reputational damage.

IDS can also address the challenge of dynamic adaptation to emerging fraud techniques. By identifying real-time suspicious activities and patterns within network traffic, IDS enhances security measures and monitoring capabilities. This capability is essential for detecting and responding to new and evolving fraud tactics. Continuous monitoring and analysis of data for anomalies or suspicious patterns enable IDS to swiftly identify potential threats and facilitate prompt response actions. Thus, integrating IDS in fraud detection helps tackle the challenge of dynamic adaptation, offering real-time monitoring, detection, and response capabilities vital for protecting against evolving fraudulent tactics and mitigating financial losses and reputational damage.

C. AI-Powered Automation

AI-powered automation significantly enhances fraud detection in e-commerce and financial services by delivering robust, real-time monitoring and rapid response to suspicious transactions. Advanced systems employ a combination of machine learning, deep learning, and robotic process automation (RPA) to process high-volume transactional data, uncover complex fraud patterns, and streamline detection workflows [130]–[133]. By automating the analysis and classification of transactional behavior, these systems offer substantial improvements in detection accuracy, early identification of fraudulent activity, and operational efficiency. The integration of AI and RPA also minimizes manual intervention, enabling organizations to swiftly mitigate potential threats and reduce financial losses.

A key advantage of AI-powered automation lies in its adaptability to emerging and evolving fraud tactics. Self-learning models continuously analyze new data, enabling automated fraud detection systems to update their strategies in real time and respond to increasingly sophisticated attacks [130], [132], [133]. This continuous adaptation ensures that organizations can proactively counteract novel fraud methods, safeguard customer assets, and protect their reputations in dynamic digital ecosystems. Ongoing research addresses remaining challenges, such as model interpretability, integration with legacy systems, and privacy, to further strengthen the reliability and transparency of automated fraud prevention for the future [131], [132].

D. Multi-Factor Authentication Expansion

Multi-Factor Authentication (MFA) has emerged as a critical advancement in e-commerce fraud prevention, offering layered protection against unauthorized access and credential-based attacks. MFA strengthens identity verification by requiring users to present two or more factors, typically a combination of knowledge (e.g., password), possession (e.g., OTP), and inherence (e.g., biometrics). Innovations in this area now incorporate advanced biometric systems, such as facial recognition, fingerprint scanning, and

liveness detection, which significantly improve the reliability of user validation [134]. These techniques are particularly effective in mitigating impersonation and identity theft, two of the most common fraud methods in e-commerce transactions [135].

Recent research has explored the integration of adaptive machine learning models into MFA frameworks, enabling systems to dynamically assess user behavior and adjust authentication challenges based on contextual risk [136]. Blockchain-based MFA has also been proposed to securely manage secondary authentication data and reduce exposure to man-in-the-middle attacks [137]. Moreover, hybrid security architectures that combine traditional MFA with real-time fraud detection systems using supervised learning algorithms have shown promising results, both in detecting suspicious activities and in reducing false positives [138], [139]. As fraud tactics continue to evolve, the expansion of MFA to include intelligent, context-aware authentication mechanisms will be essential to enhance both system resilience and user trust in digital commerce.

E. Cross-Channel Fraud Prevention Strategies

To address the growing complexity of fraud tactics, future research must prioritize the development of integrated fraud detection systems that span multiple e-commerce channels. Cross-channel fraud occurs when attackers exploit gaps across user access points, such as desktop sites, mobile apps, and offline interfaces, to bypass detection. Effective prevention strategies include consolidating user behavioral data, transaction logs, and session metadata into unified machine learning pipelines. AI-driven frameworks that combine real-time analysis and cross-channel behavioral profiling have shown promise in detecting multi-vector threats [83]–[85]. Future systems must also balance security with user experience by employing adaptive thresholds and seamless identity verification across platforms.

F. Federated Learning for Privacy-Preserving

Federated learning is emerging as a highly promising approach for privacy-preserving fraud detection, enabling collaborative model training across multiple institutions without the need to exchange sensitive customer data. By leveraging decentralized data and secure aggregation protocols, federated learning can help institutions detect fraudulent activity while minimizing the risk of privacy breaches [140]–[142]. Recent research demonstrates that federated learning frameworks can match or surpass the accuracy of traditional centralized approaches, especially when enhanced with techniques such as differential privacy, secure multi-party computation, and blockchain integration [141], [143], [144].

Future work will need to address challenges such as heterogeneous data distributions, communication overhead, scalability, and robust defense against adversarial attacks in federated environments [145]–[147]. As these issues are addressed, federated learning is expected to play a central role in enabling real-time, privacy-conscious fraud detection in complex financial and e-commerce ecosystems.

G. Integration of Behavioural Biometrics for Continuous Authentication

The integration of behavioral biometrics, such as keystroke dynamics, mouse movement trajectories, and contextual user behaviors, represents a promising direction for continuous authentication in e-commerce and financial fraud detection. Recent research demonstrates that fusing multimodal behavioural data enables fraud detection systems to continuously and unobtrusively authenticate users, leveraging individual interaction patterns that are extremely difficult to replicate or steal [148], [149]. By combining inputs from mouse dynamics, keystrokes, and user context, advanced models can significantly enhance fraud detection accuracy and provide rapid responses to suspicious activity, even in complex real-world environments [150], [151].

These approaches have been shown to operate with low latency and minimal user friction, making them practical for large-scale deployment in online platforms [148], [149]. As fraudsters develop increasingly sophisticated evasion tactics, the integration of behavioural biometrics with machine learning will offer an adaptive and effective security layer for continuous authentication and anomaly detection, ultimately strengthening defences against account takeover and unauthorized activity [150], [151].

H. Continual Learning

Continual learning [125] emerges as a highly promising approach for the future of e-commerce fraud detection, offering numerous advantages. It facilitates adaptability to the dynamic landscape of fraudulent activities, ensuring that fraud detection systems remain effective against evolving tactics employed by fraudsters [125]. Continual learning also enables efficient resource utilization by leveraging existing knowledge and incorporating new data seamlessly, enhancing the scalability and cost-effectiveness of fraud detection processes. Additionally, it improves model robustness by allowing systems to continuously refine and update their predictive capabilities in response to changing fraud patterns and emerging threats. Embracing continual learning techniques can revolutionize e-commerce fraud detection, ensuring heightened effectiveness, efficiency, and resilience in combating fraudulent activities.

Continual learning provides a promising solution to the challenge of evolving attack patterns and the need for adaptability in fraud detection systems [125]. Continuously updating models with new data and insights enables fraud detection systems to stay ahead of emerging threats, effectively mitigating risks posed by evolving attack strategies. Further research should focus on developing algorithms and frameworks that efficiently incorporate new knowledge while minimizing the risk of catastrophic forgetting, ensuring that fraud detection systems remain robust and effective in dynamic e-commerce environments.

I. Adversarial Robustness

Emphasizing adversarial robustness in e-commerce fraud detection is crucial for safeguarding financial assets, upholding customer trust, complying with regulations, and ensuring sustained system effectiveness [99], [120]. Addressing the complexities of adversarial attacks, expanding defense mechanisms, and enhancing interpretability and transparency are essential. Integrating adversarial mechanisms will strengthen systems against evolving fraud tactics, improving their resilience in identifying and mitigating fraudulent activities. Key benefits include enhanced accuracy, reliability, transparency, and compliance, which help reduce financial losses and advance the field of fraud detection.

Adversarial robustness targets the challenge of adversarial attacks, where attackers manipulate input data to deceive machine learning models, leading to incorrect classifications. By incorporating adversarial mechanisms, fraud detection systems become more resilient against these evolving tactics, enhancing their effectiveness in identifying and mitigating fraudulent activities. This approach results in improved accuracy, reliability, transparency, and compliance, thereby reducing financial losses and advancing fraud detection.

J. Deep Ensemble Learning

Deep Ensemble Learning [152] offers the potential for enhancing the accuracy and resilience of e-commerce fraud detection systems by training multiple deep learning models and aggregating their predictions. However, challenges such as algorithmic complexity, interpretability, data availability, and integration with operational systems must be addressed to fully leverage its benefits. Efficient algorithms and architectures need to be developed to manage the computational demands of ensemble learning while improving interpretability to help stakeholders understand the contributions of individual models. Additionally, acquiring representative datasets and ensuring seamless integration with existing fraud detection systems are essential for practical deployment. Addressing these challenges will contribute to developing more effective fraud detection systems, improving accuracy, adaptability, and robustness in combating fraudulent activities in e-commerce.

Deep Ensemble Learning also holds promise for addressing the challenge of organization-centric schemas in fraud detection, which introduces complexities related to manual involvement and potential impacts on customer experience. By training multiple models and aggregating their predictions, deep ensemble learning techniques can enhance the accuracy and reliability of fraud detection systems. Further research should focus on developing efficient algorithms and architectures for deep ensemble learning, acquiring diverse and representative datasets, and ensuring seamless integration with existing systems. These efforts are crucial for

the scalability and effectiveness of fraud detection mechanisms in varied e-commerce environments.

K. Zero-Shot Learning (ZSL)

Zero-shot learning [19] has significant potential to revolutionize e-commerce fraud detection. Realizing these benefits and applying them practically requires ongoing research and development. Overcoming challenges associated with zero-shot learning is essential to fully harness its potential in fraud detection scenarios. Addressing these challenges can significantly enhance the adaptability, scalability, and generalization capabilities of fraud detection systems. Zero-shot learning enables systems to generalize to unseen fraud patterns and adapt to evolving tactics used by fraudsters, helping to stay ahead of emerging threats in the dynamic e-commerce landscape.

Additionally, zero-shot learning improves scalability by efficiently utilizing limited labeled data and leveraging auxiliary information to guide the learning process. This capability allows organizations to deploy fraud detection solutions across diverse e-commerce platforms and industries without extensive data labeling efforts. Furthermore, zero-shot learning facilitates transfer learning across domains, enhancing the generalization capabilities of fraud detection systems. Models trained in one e-commerce domain can effectively detect fraud in another domain, thus improving overall performance and effectiveness.

L. Few-Shot Learning

Few-shot learning techniques [153] hold significant promise for enhancing fraud detection systems in e-commerce. These techniques address challenges such as limited labeled data, adaptability to emerging fraud patterns, transfer learning across diverse domains, and model robustness. By advancing few-shot learning, fraud detection capabilities can be greatly enhanced. Few-shot learning enables models to effectively use limited labeled data and adapt to evolving fraud patterns with minimal supervision. Additionally, it facilitates knowledge transfer across different e-commerce domains, ensuring scalability and versatility. Few-shot learning also improves model robustness, potentially promoting interpretability and transparency in decision-making processes. Integrating few-shot learning into e-commerce fraud detection systems can lead to more accurate, efficient, and adaptable mechanisms, ultimately improving the detection of fraudulent activities.

Zero-shot learning and few-shot learning techniques offer promising avenues for addressing the imbalanced dataset challenge in fraud detection systems. Few-shot learning leverages limited labeled data effectively and adapts to new and evolving fraud patterns with minimal supervision. Zero-shot learning generalizes to unseen fraud patterns without explicit training data. By employing both approaches, organizations can enhance the scalability, adaptability, and effectiveness of fraud detection systems in handling imbalanced datasets and detecting emerging threats. Further research in these areas is crucial for advancing e-commerce fraud detection and ensuring the security of online transactions.

In the future, the domain of detecting fraud in e-commerce is set to undergo evolution and innovation. Continuous studies, technological advancements, and the development of methods will influence how detection and prevention are approached. Keeping systems up to date, promoting communication, and working together will ensure that fraud detection remains accurate, responsive, and robust. By staying alert, encouraging teamwork, and being adaptable, it can successfully tackle internet and online e-commerce fraud while safeguarding and protecting the trustworthiness of e-commerce transactions [154].

M. Large Language Models

Large Language Models (LLMs), such as GPT-4 and domain-specific variants, are rapidly advancing anomaly detection for fraud in e-commerce and finance. By interpreting complex, unstructured, and semistructured data, such as transaction notes, customer communications, and claim narratives, LLMs can identify subtle, novel, and context-dependent fraudulent behaviors that traditional models often miss. Recent research shows that integrating LLMs with advanced data analytics improves both the accuracy and speed of fraud detection, enabling the identification of complex fraud schemes by analyzing diverse textual and numerical patterns [92], [155]–[160]. LLM-based systems also support the generation of synthetic

fraud data, provide narrative explanations for suspicious activity, and benchmark detection performance across multiple real-world scenarios.

However, practical challenges remain for LLM-based frameworks, including concerns around interpretability, data privacy, and computational demands [92], [160]. Ongoing research emphasizes the need for integrating LLMs with explainable AI and federated learning to achieve robust, privacy-preserving, and transparent fraud detection in dynamic, high-risk e-commerce environments. As these challenges are addressed, LLM adoption is expected to further enhance the sensitivity, adaptability, and trustworthiness of future fraud detection systems.

N. Multimodal Detection Techniques

Multimodal fraud detection uses a combination of structured and unstructured data sources, such as transaction records, user behavior patterns, device metadata, social relationships, and even images or text, to build more robust and adaptive detection systems [92], [161]–[163]. By integrating these diverse modalities, multimodal approaches can capture complex and evolving fraud schemes, including collusion rings and synthetic identities, that are often missed by unimodal systems. Recent research in high-impact journals demonstrates the effectiveness of advanced architectures, such as contrastive multimodal dialogue networks and fine-grained attention-based deep learning, in improving fraud detection accuracy, reducing false positives, and supporting real-time analysis in large-scale financial and e-commerce environments [161]–[163].

Despite their promise, multimodal models face challenges in data integration, scalability, and interpretability [92], [161]. Ongoing research focuses on improving feature fusion, developing efficient hybrid models, and designing explainable frameworks to support human analysts in interpreting complex model decisions. As e-commerce ecosystems continue to generate richer and more varied data, advancements in multimodal fraud detection are expected to drive the next wave of innovation in robust, scalable, and adaptive fraud prevention systems.

O. Prioritized Research Agenda and Solution Strategies

While all the future directions outlined above are critical, the most **urgent priorities for advancing e-commerce fraud detection** are: **(1) Adversarial robustness and continual learning**, (2) handling data scarcity via few-shot and zero-shot learning, (3) scalable and interpretable AI, and (4) privacy-preserving/federated methods. *Adversarial robustness* is especially crucial, as adversarial attacks pose direct operational threats and undermine system reliability [99], [120]. Close behind is *continual learning* [125], which addresses the need for adaptive models able to keep pace with evolving fraud tactics.

To visually summarize and clarify these priorities, Figure 6 presents a conceptual framework for the research agenda. The figure organizes the most urgent future directions in a tiered structure, associating each area with actionable solution strategies. This graphical overview highlights the practical steps and methodological advances required to address each major challenge in e-commerce fraud detection.

Recommended solution strategies include: adversarial training, defensive distillation, and anomaly detection pipelines for adversarial robustness [120]; lightweight continual/adaptive learning frameworks to minimize catastrophic forgetting [125]; advanced data augmentation, transfer, few-shot, and zero-shot learning to address imbalanced and novel fraud patterns [19], [153]; integrating SHAP, LIME, and rule extraction to balance accuracy with explainability [126], [127]; and federated learning protocols, differential privacy, and secure aggregation to ensure privacy-preserving collaboration across organizations [140], [141].

In summary, research efforts should first prioritize building robust, adaptive, and privacy-conscious fraud detection systems, while ensuring interpretability and scalability. These advances will enable real-world deployment of trustworthy AI, able to withstand sophisticated threats and adapt to the rapidly changing landscape of e-commerce fraud.

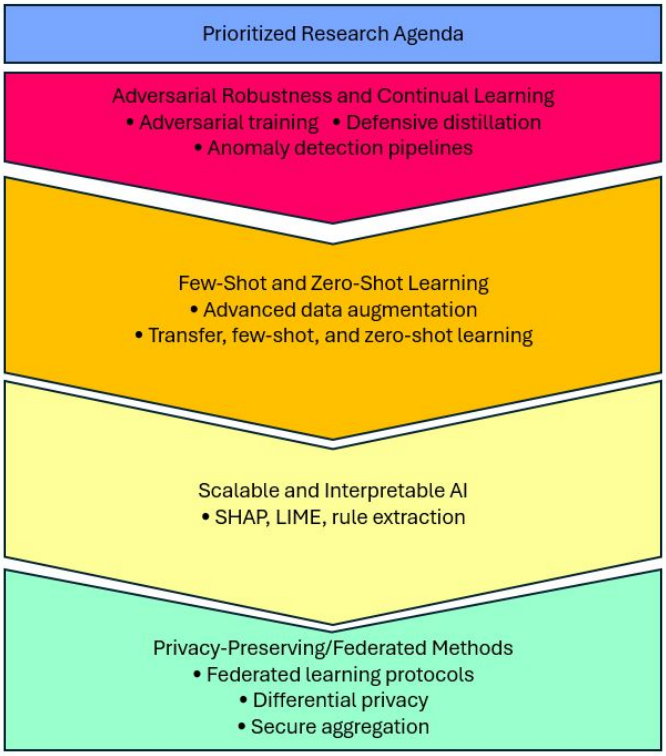

Fig. 6. Prioritized research agenda and recommended solution strategies for advancing e-commerce fraud detection. The figure presents a tiered visual framework, emphasizing the urgency of adversarial robustness and continual learning, solutions for data scarcity (few-shot/zero-shot learning), scalable and interpretable AI, and privacy-preserving/federated methods, each with actionable technical strategies.

VII. Conclusions

The rapid growth of e-commerce and accompanying fraudulent activities has underscored the critical need for effective fraud detection mechanisms. Traditional methods often fall short in cost, time, and accuracy, while the emergence of Machine Learning (ML) techniques, especially deep learning techniques, has revolutionized fraud detection by offering cost-effective, fast, and accurate solutions. This survey explores state-of-the-art fraud detection systems in e-commerce, conducting an extensive review of fraud detection techniques, focusing on recent advancements in ML-based, graph-based, deep learning techniques and hybrid model techniques. It identifies several challenges for the current research, including organization-centric schema, imbalanced datasets, model interpretability, adversarial attacks, and dynamic adaptation to emerging fraud techniques. Addressing these challenges is crucial for developing robust fraud detection systems in e-commerce. Additionally, this survey highlights some promising future directions, such as Explainable AI (XAI), Intrusion Detection Systems (IDS), continual learning, adversarial robustness, deep ensemble learning, Zero-Shot Learning (ZSL), and Few-Shot Learning (FSL), which warrant exploration to improve the resilience and effectiveness of fraud detection systems in e-commerce. By comprehensively investigating state-of-the-art e-commerce fraud detection, this study contributes significantly to the existing body of knowledge and literature on e-commerce fraud detection. It offers valuable insights for future research endeavors and facilitates the advancement of fraud detection approaches and strategies in the e-commerce field.

Author contributions This work was a collaborative effort by the authors, each of whom contributed substantially to its conception, writing, and revision. SA played a pivotal role in data acquisition and delivering the first draft.

Funding This work was partially supported by the Australian Research Council (grant numbers DP220103717, LE220100078).

Data availability All the data are available within the manuscript.

DECLARATIONS

Conflict of interest The authors declare no competing interests.

REFERENCES

[1] H. Zhou, G. Sun, S. Fu, L. Wang, J. Hu, and Y. Gao, "Internet financial fraud detection based on a distributed big data approach with node2vec," *IEEE Access*, vol. 9, pp. 43 378–43 386, 2021.

[2] B. Omair and A. Alturki, "A systematic literature review of fraud detection metrics in business processes," *IEEE Access*, vol. 8, pp. 26 893–26 903, 2020.

[3] H. Weng, Z. Li, S. Ji, C. Chu, H. Lu, T. Du, and Q. He, "Online e-commerce fraud: a large-scale detection and analysis," in *2018 IEEE 34th International Conference on Data Engineering (ICDE)*. IEEE, 2018, pp. 1435–1440.

[4] K. I. Alkhatib, A. I. Al-Aiad, M. H. Almahmoud, and O. N. Elayan, "Credit card fraud detection based on deep neural network approach," in *2021 12th International Conference on Information and Communication Systems (ICICS)*. IEEE, 2021, pp. 153–156.

[5] S. A. M. Shaddad, "Effects of cybercrime in e-banking systems," Master's thesis, Eastern Mediterranean University (EMU)-Doğu Akdeniz Üniversitesi (DAÜ), 2023.

[6] P. Vanini, S. Rossi, E. Zvizdic, and T. Domenig, "Online payment fraud: from anomaly detection to risk management," *Financial Innovation*, vol. 9, no. 1, p. 66, 2023.

[7] I. D. Mienye and N. Jere, "Deep learning for credit card fraud detection: A review of algorithms, challenges, and solutions," *IEEE Access*, 2024.

[8] L. Hernandez Aros, L. X. Bustamante Molano, F. Gutierrez-Portela, J. J. Moreno Hernandez, and M. S. Rodríguez Barrero, "Financial fraud detection through the application of machine learning techniques: a literature review," *Humanities and Social Sciences Communications*, vol. 11, no. 1, pp. 1–22, 2024.

[9] J. Research, "ecommerce fraud to exceed $107 billion in 2029." 2024. [Online]. Available: https://www.juniperresearch.com/press/pressreleasesecommerce-fraud-to-exceed-107bn-in-2029

[10] R. de Best, "Quarterly number of mastercard issued credit cards in u.s. and worldwide 2006-2024," Nov 2024. [Online]. Available: https://www.statista.com/statistics/618137/number-of-mastercard-credit-cards-worldwide-by-region/

[11] S. M. Najem and S. M. Kadeem, "A survey on fraud detection techniques in e-commerce," *Tech-Knowledge*, vol. 1, no. 1, pp. 33–47, 2021.

[12] H. Najadat, O. Altiti, A. A. Aqouleh, and M. Younes, "Credit card fraud detection based on machine and deep learning," in *2020 11th International Conference on Information and Communication Systems (ICICS)*. IEEE, 2020, pp. 204–208.

[13] W. A. Al-Khater, S. Al-Maadeed, A. A. Ahmed, A. S. Sadiq, and M. K. Khan, "Comprehensive review of cybercrime detection techniques," *IEEE Access*, vol. 8, pp. 137 293–137 311, 2020.

[14] G. Zhang, Z. Li, J. Huang, J. Wu, C. Zhou, J. Yang, and J. Gao, "efraudcom: An e-commerce fraud detection system via competitive graph neural networks," *ACM Transactions on Information Systems (TOIS)*, vol. 40, no. 3, pp. 1–29, 2022.

[15] J. Song, X. Qu, Z. Hu, Z. Li, J. Gao, and J. Zhang, "A subgraph-based knowledge reasoning method for collective fraud detection in e-commerce," *Neurocomputing*, vol. 461, pp. 587–597, 2021.

[16] T. Pourhabibi, K.-L. Ong, B. H. Kam, and Y. L. Boo, "Fraud detection: A systematic literature review of graph-based anomaly detection approaches," *Decision Support Systems*, vol. 133, p. 113303, 2020.

[17] F. Hasan, S. K. Mondal, M. R. Kabir, M. A. Al Mamun, N. S. Rahman, and M. S. Hossen, "E-commerce merchant fraud detection using machine learning approach," in *2022 7th International Conference on Communication and Electronics Systems (ICCES)*. IEEE, 2022, pp. 1123–1127.

[18] K. S. Lim, L. H. Lee, and Y.-W. Sim, "A review of machine learning algorithms for fraud detection in credit card transaction," *International Journal of Computer Science & Network Security*, vol. 21, no. 9, pp. 31–40, 2021.

[19] R. Chalapathy and S. Chawla, "Deep learning for anomaly detection: A survey," *arXiv preprint arXiv:1901.03407*, 2019.

[20] M. Aliapoulios, C. Ballard, R. Bhalerao, T. Lauinger, and D. McCoy, "Swiped: Analyzing ground-truth data of a marketplace for stolen debit and credit cards," in *30th USENIX Security Symposium (USENIX Security 21)*, 2021, pp. 4151–4168.

[21] S. C. Alliance, "Card-not-present fraud: A primer on trends and authentication processes," 2014.

[22] M. Bond, O. Choudary, S. J. Murdoch, S. Skorobogatov, and R. Anderson, "Chip and skim: cloning emv cards with the pre-play attack," in *2014 IEEE Symposium on Security and Privacy*. IEEE, 2014, pp. 49–64.

[23] A. Ijaz, "Risk management of counterfeit cards fraud: An empirical study of challenges among south asian financial institutions," Master's thesis, Nord universitet, 2018.

[24] F. Hayashi, "Payment card fraud rates in the united states relative to other countries after migrating to chip cards," *Economic Review*, vol. 104, no. 4, pp. 23–40, 2019.

[25] Y. Jain, N. Tiwari, S. Dubey, and S. Jain, "A comparative analysis of various credit card fraud detection techniques," *International Journal of Recent Technology and Engineering*, vol. 7, no. 5, pp. 402–407, 2019.

[26] Juniper Research, "Online payment fraud losses to exceed $362 billion globally by 2028," 2023, https://www.juniperresearch.com/press/losses-online-payment-fraud-exceed-362-billion.

[27] Mastercard, "Ecommerce fraud trends and statistics merchants need to know in 2024," 2024, https://b2b.mastercard.com/news-and-insights/blog/ecommerce-fraud-trends-and-statistics-merchants-need-to-know-in-2024/.

[28] V. Chang, A. Di Stefano, Z. Sun, G. Fortino *et al.*, "Digital payment fraud detection methods in digital ages and industry 4.0," *Computers and Electrical Engineering*, vol. 100, p. 107734, 2022.

[29] M. Sánchez-Aguayo, L. Urquiza-Aguiar, and J. Estrada-Jiménez, "Fraud detection using the fraud triangle theory and data mining techniques: A literature review," *Computers*, vol. 10, no. 10, p. 121, 2021.

[30] S. Alshebli and C. Y. Yeun, "Examining the security landscape of mobile payment systems," in *2024 2nd International Conference on Cyber Resilience (ICCR)*. IEEE, 2024, pp. 1–5.

[31] R. Muqattash and F. Kharbat, "Detecting mobile payment fraud: Leveraging machine learning for rapid analysis," in *2023 Tenth International Conference on Social Networks Analysis, Management and Security (SNAMS)*. IEEE, 2023, pp. 1–5.

[32] M. K. Siddiqui and K. K. Goyal, "A study the use of e-payment systems based on artificial intelligence," *Artificial Intelligence and Communication Technologies*, pp. 1063–1076, 2023.

[33] E. Nowroozi, S. Seyedshoari, Y. Mekdad, E. Savaş, and M. Conti, "Cryptocurrency wallets: Assessment and security," in *Blockchain for Cybersecurity in Cyber-Physical Systems*. Springer, 2022, pp. 1–19.

[34] V. N. Dornadula and S. Geetha, "Credit card fraud detection using machine learning algorithms," *Procedia computer science*, vol. 165, pp. 631–641, 2019.

[35] H. Pallathadka, E. H. Ramirez-Asis, T. P. Loli-Poma, K. Kaliyaperumal, R. J. M. Ventayen, and M. Naved, "Applications of artificial intelligence in business management, e-commerce and finance," *Materials Today: Proceedings*, vol. 80, pp. 2610–2613, 2023.

[36] B. Narsimha, C. V. Raghavendran, P. Rajyalakshmi, G. K. Reddy, M. Bhargavi, and P. Naresh, "Cyber defense in the age of artificial intelligence and machine learning for financial fraud detection application," *IJEER*, vol. 10, no. 2, pp. 87–92, 2022.

[37] K. G. Al-Hashedi and P. Magalingam, "Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019," *Computer Science Review*, vol. 40, p. 100402, 2021.

[38] M. Dadhich, K. K. Hiran, S. S. Rao, R. Sharma, and R. Meena, "Study of combating technology induced fraud assault (tifa) and possible solutions: The way forward," in *International Conference on Emerging Technologies in Computer Engineering*. Springer, 2022, pp. 715–723.

[39] M. B. Boubker, S. Ouahabi, K. Elguemmat, and A. Eddaoui, "A comprehensive study on credit card fraud prevention and detection," in *2021 Fifth International Conference On Intelligent Computing in Data Sciences (ICDS)*. IEEE, 2021, pp. 1–8.

[40] C. Phua, V. Lee, K. Smith, and R. Gayler, "A comprehensive survey of data mining-based fraud detection research," *arXiv preprint arXiv:1009.6119*, 2010.

[41] A. Alharbi, M. Alshammari, O. D. Okon, A. Alabrah, H. T. Rauf, H. Alyami, and T. Meraj, "A novel text2img mechanism of credit card fraud detection: A deep learning approach," *Electronics*, vol. 11, no. 5, p. 756, 2022.

[42] M. M. John, H. H. Olsson, and J. Bosch, "Towards an ai-driven business development framework: A multi-case study," *Journal of Software: Evolution and Process*, p. e2432, 2022.

[43] D. H. Reddy, "An analysis of the supervised learning approach for online fraud detection," *Computational Intelligence and Machine Learning*, 2022.

[44] A. Cherif, A. Badhib, H. Ammar, S. Alshehri, M. Kalkatawi, and A. Imine, "Credit card fraud detection in the era of disruptive technologies: A systematic review," *Journal of King Saud University-Computer and Information Sciences*, vol. 35, no. 1, pp. 145–174, 2023.

[45] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial fraud: a review of anomaly detection techniques and recent advances," *Expert systems With applications*, vol. 193, p. 116429, 2022.

[46] N. Husnaningtyas and T. Dewayanto, "Financial fraud detection and machine learning algorithm (unsupervised learning): Systematic literature review," *Jurnal Riset Akuntansi Dan Bisnis Airlangga*, 2023.

[47] S. Beigi and M. Amin Naseri, "Credit card fraud detection using data mining and statistical methods," *Journal of AI and Data Mining*, vol. 8, no. 2, pp. 149–160, 2020.

[48] Y. Tang and Y. Liang, "Credit card fraud detection based on federated graph learning," *Expert Systems with Applications*, vol. 256, p. 124979, 2024.

[49] T. Berhane, T. Melese, A. Walelign, and A. Mohammed, "A hybrid convolutional neural network and support vector machine-based credit card fraud detection model," *Mathematical Problems in Engineering*, vol. 2023, no. 1, p. 8134627, 2023.

[50] Q. Guo, Z. Li, B. An, P. Hui, J. Huang, L. Zhang, and M. Zhao, "Securing the deep fraud detector in large-scale e-commerce platform via adversarial machine learning approach," in *The World Wide Web Conference*, 2019, pp. 616–626.

[51] H. Li, S. Jiang, L. Zhang, S. Du, G. Ye, and H. Chai, "Ragformer: Learning semantic attributes and topological structure for fraud detection," *arXiv preprint arXiv:2402.17472*, 2024.

[52] K. Yan, J. Gao, and D. Matsypura, "Fiw-gnn: A heterogeneous graph-based learning model for credit card fraud detection," in *2023 IEEE 10th International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, 2023, pp. 1–10.

[53] B. Youssef, F. Bouchra, and O. Brahim, "Rules extraction and deep learning for e-commerce fraud detection," in *2020 6th IEEE Congress on Information Science and Technology (CiSt)*. IEEE, 2021, pp. 145–150.

[54] G. Gianini, L. G. Fossi, C. Mio, O. Caelen, L. Brunie, and E. Damiani, "Managing a pool of rules for credit card fraud detection by a game theory based approach," *Future Generation Computer Systems*, vol. 102, pp. 549–561, 2020.

[55] M. Baumann, "Improving a rule-based fraud detection system with classification based on association rule mining," in *INFORMATIK 2021*. Gesellschaft für Informatik, Bonn, 2021, pp. 1121–1134.

[56] L. Moumeni, M. Saber, I. Slimani, I. Elfarissi, and Z. Bougroun, "Machine learning for credit card fraud detection," in *WITS 2020: Proceedings of the 6th International Conference on Wireless Technologies, Embedded, and Intelligent Systems*. Springer, 2022, pp. 211–221.

[57] L. N. Valli, N. Sujatha, and D. Divya, "A novel approach for credit card fraud detection using lr method-comparative studies," *Eduvest-Journal of Universal Studies*, vol. 2, no. 12, pp. 2611–2614, 2022.

[58] H. Wen and F. Huang, "Personal loan fraud detection based on hybrid supervised and unsupervised learning," in *2020 5th IEEE international conference on big data analytics (ICBDA)*. IEEE, 2020, pp. 339–343.

[59] Y. Fang, Y. Zhang, and C. Huang, "Credit card fraud detection based on machine learning." *Computers, Materials & Continua*, vol. 61, no. 1, 2019.

[60] J. Li, "E-commerce fraud detection model by computer artificial intelligence data mining," *Computational Intelligence and Neuroscience*, 2022.

[61] S. Carta, G. Fenu, D. R. Recupero, and R. Saia, "Fraud detection for e-commerce transactions by employing a prudential multiple consensus model," *Journal of Information Security and Applications*, vol. 46, pp. 13–22, 2019.

[62] B. Lebichot, Y.-A. L. Borgne, L. He-Guelton, F. Oblé, and G. Bontempi, "Deep-learning domain adaptation techniques for credit cards fraud detection," in *INNS Big Data and Deep Learning conference*. Springer, 2019, pp. 78–88.

[63] S. Wang, C. Liu, X. Gao, H. Qu, and W. Xu, "Session-based fraud detection in online e-commerce transactions using recurrent neural networks," in *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 2017, pp. 241–252.

[64] W. L. Hamilton, R. Ying, and J. Leskovec, "Representation learning on graphs: Methods and applications," *arXiv preprint arXiv:1709.05584*, 2017.

[65] D. Cheng, X. Wang, Y. Zhang, and L. Zhang, "Graph neural network for fraud detection via spatial-temporal attention," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 8, pp. 3800–3813, 2022.

[66] H. Zhu, W. Zhong, Z. Huang, and Z. Wang, "Research and application of e-commerce marketing fraud detection method based on graph neural network," in *Proceedings of the 2024 2nd International Conference on Artificial Intelligence, Systems and Network Security*, 2024, pp. 20–27.

[67] P. Tiwari, S. Mehta, N. Sakhuja, I. Gupta, and A. K. Singh, "Hybrid method in identifying the fraud detection in the credit card," in *Evolutionary Computing and Mobile Sustainable Networks: Proceedings of ICECMSN 2020*. Springer, 2021, pp. 27–35.

[68] I. de Zarzà, J. de Curtò, and C. T. Calafate, "Optimizing neural networks for imbalanced data," *Electronics*, vol. 12, no. 12, p. 2674, 2023.

[69] H. Chi, Y. Lu, B. Liao, L. Xu, and Y. Liu, "An optimized quantitative argumentation debate model for fraud detection in e-commerce transactions," *IEEE Intelligent Systems*, vol. 36, no. 2, pp. 52–63, 2021.

[70] O. Ndama, I. Bensassi, and E.-N. El Mokhtar, "Innovative credit card fraud detection: A hybrid model combining artificial neural networks and support vector machines," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, p. 2674, 09 2024.

[71] O. Ndama, I. Bensassi, and E. M. En-Naimi, "Integrating artificial neural networks and support vector machines machine learning algorithms for advanced credit card fraud detection," in *Modern Artificial Intelligence and Data Science 2024: Tools, Techniques and Systems*. Switzerland: Springer, 2024, pp. 453–461.

[72] A. Saputra *et al.*, "Fraud detection using machine learning in e-commerce," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 9, 2019.

[73] J. S. Akshay, T. Vinusha, R. S. Bianca, C. S. Krishna, and G. Radhika, "Enhancing credit card fraud detection with deep learning and graph neural networks," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, 2024, pp. 1–6.

[74] F. Khaled Alarfaj and S. Shahzadi, "Enhancing fraud detection in banking with deep learning: Graph neural networks and autoencoders for real-time credit card fraud prevention," *IEEE Access*, vol. 13, pp. 20 633–20 646, 2025.

[75] V. Wadne, S. Bhandari, S. Deokar, A. Sonavale, and G. Shinde, "Detect credit card fraud in automated systems using machine learning with aws sagemaker," in *Multidisciplinary Approaches for Sustainable Development*. CRC Press, 2024, pp. 91–97.

[76] N. S. Chougule, C. J. Awati, and R. Deshmukh, "Using aws sagemaker to deploy ml credit card fraud detection model," in *2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)*. IEEE, 2024, pp. 150–156.

[77] "Paypal solves fraud challenges with aerospike® and intel® optane™ persistent memory," Intel Corporation, Tech. Rep., August 2023, accessed: 2025-06-12. [Online]. Available: https://www.intel.com/content/dam/www/central-libraries/us/en/documents/2023-08/aerospike-paypal-case-studyl.pdf

[78] L. Standare, D. Hayes, N.-A. Le-Khac, and K.-K. R. Choo, "Forensic investigation of paypal accounts," *Cyber and Digital Forensic Investigations: A Law Enforcement Practitioner's Perspective*, pp. 141–174, 2020.

[79] D. Xi, B. Song, F. Zhuang, Y. Zhu, S. Chen, T. Zhang, Y. Qi, and Q. He, "Modeling the field value variations and field interactions simultaneously for fraud detection," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 17, 2021, pp. 14 957–14 965.

[80] C. Zhang, Q. Wang, T. Liu, X. Lu, J. Hong, B. Han, and C. Gong, "Fraud detection under multi-sourced extremely noisy annotations," in *Proceedings of the 30th ACM international conference on information & knowledge management*, 2021, pp. 2497–2506.

[81] A. Mutemi and F. 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.

[82] Z. Li, P. Hui, P. Zhang, J. Huang, B. Wang, L. Tian, J. Zhang, J. Gao, and X. Tang, "What happens behind the scene? towards fraud community detection in e-commerce from online to offline," in *Companion Proceedings of the Web Conference 2021*, 2021, pp. 105–113.

[83] U. Dolu and E. Sefer, "A novel gbt-based approach for cross-channel fraud detection on real-world banking transactions," in *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, 2022, pp. 73–84.

[84] L. Magomaeva, "Instruments of retail banking business: Using information cross-channels to fight against fraud and falsification," *Vestnik of Astrakhan State Technical University. Series: Economics*, 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:169236300

[85] T. Siqin, L. Yang, S.-H. Chung, and X. Wen, "Cross-channel influences in mobile-app-website e-commerce supply chains: When to weaken the influence?" *Transportation Research Part E: Logistics and Transportation Review*, vol. 182, p. 103408, 2024.

[86] M. Golyeri, S. Celik, F. Bozyigit, and D. Kılınç, "Fraud detection on e-commerce transactions using machine learning techniques," *Artificial Intelligence Theory and Applications*, vol. 3, no. 1, pp. 45–50, 2023.

[87] N. Suardiman, S. Dhanny, D. Tjahjadi, B. Permana, K. Ukar *et al.*, "E-commerce fraud detection using generated data from banksim using machine learning approach: A pilot study," in *2024 18th International Conference on Ubiquitous Information Management and Communication (IMCOM)*. IEEE, 2024, pp. 1–4.

[88] O. F. Alabi and A. A. David, "Framework for detection of

fraud at point of sale on electronic commerce sites using logistic regression," *EAI Endorsed Transactions on Scalable Information Systems*, 2022.

[89] U. Chilaka, G. Chukwudebe, and A. Bashiru, "A review of credit card fraud detection techniques in electronic finance and banking," *Conic Res. Eng. J*, vol. 3, pp. 456–467, 2019.

[90] Y. Li, X. Yang, Q. Gao, H. Wang, J. Zhang, and T. Li, "Cross-regional fraud detection via continual learning with knowledge transfer," *IEEE Transactions on Knowledge and Data Engineering*, 2024.

[91] Y. Li, Y. Yang, X. Yang, Q. Gao, and F. Zhou, "Forgetting prevention for cross-regional fraud detection with heterogeneous trade graph," *arXiv preprint arXiv:2204.10085*, 2022.

[92] E. Dritsas and M. Trigka, "Machine learning in e-commerce: Trends, applications, and future challenges," *IEEE Access*, 2025.

[93] G. O. Mbah, "Data privacy in the era of ai: Navigating regulatory landscapes for global businesses," *International Journal of Science and Research Archive*, 2024.

[94] H. Ijaiya, "Harnessing ai for data privacy: Examining risks, opportunities and strategic future directions," *International Journal of Science and Research Archive*, 2024.

[95] V. O. Eghaghe, O. S. Osundare, C. P.-M. Ewim, and I. C. Okeke, "Navigating the ethical and governance challenges of ai deployment in aml practices within the financial industry," *International Journal of Scholarly Research and Reviews*, vol. 5, no. 2, 2024.

[96] C. Bura, S. Kamatala, and P. K. Myakala, "Ethical challenges in data science: Navigating the complex landscape of responsibility and fairness," *International Journal of Current Science Research and Review*, 2025.

[97] J. K. Bahangulu and L. Owusu-Berko, "Algorithmic bias, data ethics, and governance: Ensuring fairness, transparency, and compliance in ai-powered business analytics applications," *World J Adv Res Rev*, pp. 1746–1763, 2025.

[98] E. E. Agu, A. O. Abhulimen, A. N. Obiki-Osafiele, O. S. Osundare, I. A. Adeniran, and C. P. Efunniyi, "Discussing ethical considerations and solutions for ensuring fairness in ai-driven financial services," *International Journal of Frontier Research in Science*, vol. 3, no. 2, pp. 001–009, 2024.

[99] N. Tax, K. J. de Vries, M. de Jong, N. Dosoula, B. van den Akker, J. Smith, O. Thuong, and L. Bernardi, "Machine learning for fraud detection in e-commerce: A research agenda," in *Deployable Machine Learning for Security Defense: Second International Workshop, MLHat 2021, Virtual Event, August 15, 2021, Proceedings 2*. Springer, 2021, pp. 30–54.

[100] F. Thabtah, S. Hammoud, F. Kamalov, and A. Gonsalves, "Data imbalance in classification: Experimental evaluation," *Inf. Sci.*, vol. 513, pp. 429–441, 2020.

[101] R. George and B. Roy, "Handling class imbalance in fraud detection using machine learning techniques," in *Lecture Notes in Electrical Engineering*, 2021.

[102] W. Priatna, H. Purnomo, I. Sembiring, and T. Wellem, "Integrating class imbalance solutions into fraud detection systems: A systematic literature review," in *2024 2nd International Conference on Technology Innovation and Its Applications (ICTIIA)*. IEEE, 2024, pp. 1–6.

[103] S. Alharbi, A. Alorini, K. Alahmadi, H. Alhosaini, Y. Zhu, and X. Wang, "Exploring oversampling techniques for fraud detection with imbalanced classes," in *11th International Conference on Informatics, Electronics & Vision*, 2023.

[104] A. Dal Pozzolo, O. Caelen, Y.-A. Le Borgne, S. Waterschoot, and G. Bontempi, "Learned lessons in credit card fraud detection from a practitioner perspective," *Expert systems with applications*, vol. 41, no. 10, pp. 4915–4928, 2014.

[105] R. Qaddoura and M. M. Biltawi, "Improving fraud detection in an imbalanced class distribution using different oversampling techniques," in *2022 International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI)*. IEEE, 2022, pp. 1–5.

[106] C. Charitou, S. Dragicevic, and A. d. Garcez, "Synthetic data generation for fraud detection using gans," *arXiv preprint arXiv:2109.12546*, 2021.

[107] E. Strelcenia and S. Prakoonwit, "A survey on gan techniques for data augmentation to address the imbalanced data issues in credit card fraud detection," *Machine Learning and Knowledge Extraction*, vol. 5, no. 1, pp. 304–329, 2023.

[108] ——, "A new gan-based data augmentation method for handling class imbalance in credit card fraud detection," in *2023 10th International Conference on Signal Processing and Integrated Networks (SPIN)*. IEEE, 2023, pp. 627–634.

[109] Z. Li, M. Huang, G. Liu, and C. Jiang, "A hybrid method with dynamic weighted entropy for handling the problem of class imbalance with overlap in credit card fraud detection," *Expert Systems with Applications*, vol. 175, p. 114750, 2021.

[110] W. Priatna, H. Purnomo, A. Iriani, I. Sembiring, and T. Wellem, "Optimizing multilayer perceptron with cost-sensitive learning for addressing class imbalance in credit card fraud detection," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 2024.

[111] S. Jing, L. Chen, Q. Li, and D. Wu, "Dos-gnn: Dual-feature aggregations with over-sampling for class-imbalanced fraud detection on graphs," in *2024 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2024, pp. 1–8.

[112] M. Isangediok and K. Gajamannage, "Fraud detection using optimized machine learning tools under imbalance classes," in *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022, pp. 4275–4284.

[113] X. Zhang, F. Guo, T. Chen, L. Pan, G. Beliakov, and J. Wu, "A brief survey of machine learning and deep learning techniques for e-commerce research," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 18, no. 4, pp. 2188–2216, 2023.

[114] W. Samek, G. Montavon, S. Lapuschkin, C. J. Anders, and K.-R. Müller, "Explaining deep neural networks and beyond: A review of methods and applications," *Proceedings of the IEEE*, vol. 109, no. 3, pp. 247–278, 2021.

[115] R. Jhangiani, D. Bein, and A. Verma, "Machine learning pipeline for fraud detection and prevention in e-commerce transactions," in *2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*. IEEE, 2019, pp. 0135–0140.

[116] A. Bhowmik, M. Sannigrahi, D. Chowdhury, A. D. Dwivedi, and R. R. Mukkamala, "Dbnex: Deep belief network and explainable ai based financial fraud detection," in *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022, pp. 3033–3042.

[117] Z. C. Lipton, "The mythos of model interpretability: In machine learning, the concept of interpretability is both important and slippery." *Queue*, vol. 16, no. 3, pp. 31–57, 2018.

[118] M. T. Ribeiro, S. Singh, and C. Guestrin, "" why should i trust you?" explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international*

conference on knowledge discovery and data mining, 2016, pp. 1135–1144.

[119] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data & Society*, vol. 3, no. 2, p. 2053951716679679, 2016.

[120] D. Lunghi, A. Simitsis, O. Caelen, and G. Bontempi, "Adversarial learning in real-world fraud detection: Challenges and perspectives," in *Proceedings of the Second ACM Data Economy Workshop*, 2023, pp. 27–33.

[121] A. Boukerche, L. Zheng, and O. Alfandi, "Outlier detection: Methods, models, and classification," *ACM Comput. Surv.*, vol. 53, no. 3, jun 2020. [Online]. Available: https://doi.org/10.1145/3381028

[122] M. Ozkan-Ozay, E. Akin, Ö. Aslan, S. Kosunalp, T. Iliev, I. Stoyanov, and I. Beloev, "A comprehensive survey: Evaluating the efficiency of artificial intelligence and machine learning techniques on cyber security solutions," *IEEE Access*, 2024.

[123] R. R. Popat and J. Chaudhary, "A survey on credit card fraud detection using machine learning," in *2018 2nd international conference on trends in electronics and informatics (ICOEI)*. IEEE, 2018, pp. 1120–1125.

[124] P. Gupta, "Leveraging machine learning and artificial intelligence for fraud prevention," *SSRG International Journal of Computer Science and Engineering*, vol. 10, no. 5, pp. 47–52, 2023.

[125] B. Lebichot, W. Siblini, G. Paldino, Y.-A. Le Borgne, F. Oblé, and G. Bontempi, "Assessment of catastrophic forgetting in continual credit card fraud detection," *Expert Systems with Applications*, p. 123445, 2024.

[126] J. Qureshi, "Ai-powered cloud-based e-commerce: Driving digital business transformation initiatives," 2024.

[127] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins *et al.*, "Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai," *Information fusion*, vol. 58, pp. 82–115, 2020.

[128] J. Černevičienė and A. Kabašinskas, "Explainable artificial intelligence (xai) in finance: a systematic literature review," *Artificial Intelligence Review*, vol. 57, no. 8, p. 216, 2024.

[129] P. M. Varatharajoo, N. H. Zakaria, J. A. Bakar, and M. Mahmuddin, "Explainable artificial intelligence (xai) model for online fraud detection: A critical review in malaysia's digital economy," in *2024 7th International Conference on Internet Applications, Protocols, and Services (NETAPPS)*. IEEE, 2024, pp. 1–8.

[130] R. Sharma, D. Nalawade, P. Negi, R. Dhabliya, S. Bhattacharya, and V. Khetani, "Aipowered automation of fraud detection in financial services," in *Proceedings of the 5th International Conference on Information Management Machine Intelligence*, 2023.

[131] K. Venigandla and N. Vemuri, "Rpa and ai-driven predictive analytics in banking for fraud detection," *Tuijin Jishu/Journal of Propulsion Technology*, 2022.

[132] A. Dalsaniya, O. Id, K. Patel, and P. R. Swaminarayan, "Challenges and opportunities: Implementing rpa and ai in fraud detection in the banking sector," *World Journal of Advanced Research and Reviews*, 2025.

[133] F. Johora, R. Hasan, S. F. Farabi, M. Z. Alam, M. I. Sarkar, and M. A. A. Mahmud, "Ai advances: Enhancing banking security with fraud detection," in *2024 First International Conference on Technological Innovations and Advance Computing (TIACOMP)*, 2024, pp. 289–294.

[134] R. S. PV *et al.*, "Multi-level authentication: Combining face, palm, and liveness detection for improved security," *Journal of Innovative Image Processing*, vol. 5, no. 2, pp. 181–191, 2023.

[135] M. Ojewale and P. Yomsi, "Multi-factor authentication and fingerprint-based debit card system," *Journal of Informatics and Data Mining*, vol. 5, no. 2, pp. 19–28, 2019.

[136] A. Cherif, S. Alshehri, M. Kalkatawi, and A. Imine, "Towards an intelligent adaptive security framework for preventing and detecting credit card fraud," in *2022 IEEE/ACS 19th International Conference on Computer Systems and Applications (AICCSA)*. IEEE, 2022, pp. 1–8.

[137] S. Mercan, M. Cebe, K. Akkaya, and J. Zuluaga, "Blockchain-based two-factor authentication for credit card validation," in *International Workshop on Data Privacy Management*. Springer, 2021, pp. 319–327.

[138] R. Premsai, "Evaluating the impact of multi factor authentication on cybersecurity effectiveness," *International Scientific Journal of Engineering and Management*, 2024. [Online]. Available: https://api.semanticscholar.org/CorpusID:274622763

[139] A. M. Aburbeian and M. Fernández-Veiga, "Secure internet financial transactions: A framework integrating multi-factor authentication and machine learning," *AI*, vol. 5, no. 1, pp. 177–194, 2024.

[140] K. D'souza, S. Puthusseri, and A. G. Samuel, "Scalable federated learning for privacy-preserving credit card fraud detection," in *2023 IEEE International Carnahan Conference on Security Technology (ICCST)*. IEEE, 2023, pp. 1–6.

[141] A. Abadi, B. Doyle, F. Gini, K. Guinamard, S. K. Murakonda, J. Liddell, P. Mellor, S. J. Murdoch, M. Naseri, H. Page *et al.*, "Starlit: Privacy-preserving federated learning to enhance financial fraud detection," *arXiv preprint arXiv:2401.10765*, 2024.

[142] T. Awosika, R. M. Shukla, and B. Pranggono, "Transparency and privacy: the role of explainable ai and federated learning in financial fraud detection," *IEEE Access*, 2024.

[143] A. A. Ahmed and O. Alabi, "Secure and scalable blockchain-based federated learning for cryptocurrency fraud detection: A systematic review," *IEEE Access*, 2024.

[144] T. Baabdullah, A. Alzahrani, D. B. Rawat, and C. Liu, "Efficiency of federated learning and blockchain in preserving privacy and enhancing the performance of credit card fraud detection (ccfd) systems," *Future Internet*, vol. 16, no. 6, p. 196, 2024.

[145] R. Sun, "A comprehensive investigation of fraud detection behavior in federated learning," in *ITM Web of Conferences*, vol. 70. EDP Sciences, 2025, p. 03030.

[146] T. El Hallal and Y. El Mourabit, "Federated learning for credit card fraud detection: Key fundamentals and emerging trends," in *2024 International Conference on Circuit, Systems and Communication (ICCSC)*. IEEE, 2024, pp. 1–6.

[147] H. Zheng, "Federated learning-based credit card fraud detection: A comparative analysis of advanced machine learning models," in *ITM Web of Conferences*, 2025.

[148] J. Li, Q. Yi, M. K. Lim, S. Yi, P. Zhu, and X. Huang, "Mbbfauth: Multimodal behavioral biometrics fusion for continuous authentication on non-portable devices," *IEEE Transactions on Information Forensics and Security*, 2024.

[149] G. Ciaramella, G. Iadarola, F. Martinelli, F. Mercaldo, and A. Santone, "Continuous and silent user authentication

through mouse dynamics and explainable deep learning: A proposal," in *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022, pp. 6628–6630.

[150] Z. Zhang, H. Yin, S. X. Rao, X. Yan, Z. Wang, W. Liang, Y. Zhao, Y. Shan, R. Zhang, Y. Lin *et al.*, "Identifying e-commerce fraud through user behavior data: Observations and insights," *Data Science and Engineering*, pp. 1–16, 2025.

[151] S. R. Adapa, "Enhancing credit card fraud detection: A novel approach with random forest and behavioral biometrics," *International Journal for Research in Applied Science and Engineering Technology*, 2024.

[152] I. D. Mienye and Y. Sun, "A deep learning ensemble with data resampling for credit card fraud detection," *IEEE Access*, vol. 11, pp. 30 628–30 638, 2023.

[153] Y. Song, T. Wang, P. Cai, S. K. Mondal, and J. P. Sahoo, "A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities," *ACM Comput. Surv.*, vol. 55, no. 13s, jul 2023. [Online]. Available: https://doi-org.ezproxy.lib.uts.edu.au/10.1145/3582688

[154] K. B. Soni, M. Chopade, and R. Vaghela, "Credit card fraud detection using machine learning approach," *Appl. Inf. Syst. Manag*, vol. 4, no. 2, pp. 71–76, 2021.

[155] S. Korkanti, "Enhancing financial fraud detection using llms and advanced data analytics," in *2024 2nd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS)*. IEEE, 2024, pp. 1328–1334.

[156] X. Cao, S. Li, V. Katsikis, A. Khan, H. He, Z. Liu, L. Zhang, and C. Peng, "Empowering financial futures: Large language models in the modern financial landscape," *EAI Endorsed Transactions on AI and Robotics*, 2024.

[157] J. Nicholls, A. Kuppa, and N.-A. Le-Khac, "Large language model xai approach for illicit activity investigation in bitcoin," *Neural Computing and Applications*, pp. 1–13, 2024.

[158] L. Jiang, "Detecting scams using large language models," *arXiv preprint arXiv:2402.03147*, 2024.

[159] J. Chakraborty, W. Xia, A. Majumder, D. Ma, W. Chaabene, and N. Janvekar, "Detoxbench: Benchmarking large language models for multitask fraud & abuse detection," *arXiv preprint arXiv:2409.06072*, 2024.

[160] J. Su, C. Jiang, X. Jin, Y. Qiao, T. Xiao, H. Ma, R. Wei, Z. Jing, J. Xu, and J. Lin, "Large language models for forecasting and anomaly detection: A systematic literature review," *arXiv preprint arXiv:2402.10350*, 2024.

[161] Q. Lu, W. Du, S. Yang, W. Xu, and J. L. Zhao, "Can earnings conference calls tell more lies? a contrastive multimodal dialogue network for advanced financial statement fraud detection," *Decision Support Systems*, vol. 189, p. 114381, 2025.

[162] G. Wang, J. Ma, and G. Chen, "Attentive statement fraud detection: Distinguishing multimodal financial data with fine-grained attention," *Decision Support Systems*, vol. 167, p. 113913, 2023.

[163] J. Kaikaus, J. L. Hobson, and R. J. Brunner, "Truth or fiction: Multimodal learning applied to earnings calls," in *2022 IEEE International Conference on Big Data (Big Data)*. IEEE, 2022, pp. 3607–3612.