

# Artificial Intelligence in Combating Synthetic Identity Fraud: A Comparative Case Study of Amazon and Shopify E-Commerce

ADEPEJU DEBORAH BELLO<sup>1</sup>, OLUWASEYI BABATUNDE OGUNTOLA<sup>2</sup>, JOHN ACHIDOK<sup>3</sup>,  
AYODEJI TEMITOPE AJIBADE<sup>4</sup>, OLUWATOSIN OMOTORIOGUN<sup>5</sup>, FLORENCE OLABISI  
OGUNLEYE<sup>6</sup>, OLUWADAMILOLA AYOOLA<sup>7</sup>

<sup>1</sup> CyberFraud R&D Analyst, Barclays UK,

<sup>2</sup> Senior Auditor, Cybersecurity & Technology Infrastructure, Citi Bank, US,

<sup>3</sup> Masters Candidate, University of North Dakota,

<sup>4</sup> Fraud Analyst, Barclays UK,

<sup>5</sup> KYC Onboarding Manager, Barclays UK

<sup>6</sup> KYC Quality Control Analyst, Barclays UK,

<sup>7</sup> Doctoral Candidate, University of Fairfax, US

**Abstract-** Synthetic identity fraud is a rapidly growing threat, accounting for over \$12 billion in losses (≈25% of global identity fraud) in 2024 and is projected to rise further. This study evaluates how AI-driven methods detect and prevent synthetic identity fraud on e-commerce platforms. Using a qualitative multiple-case approach (Amazon, Shopify). This study reviews secondary sources including technical documentation, industry reports, and academic literature. The findings of the study show that advanced AI techniques such as anomaly detection, behavioural analytics, and hybrid supervised models have proven to substantially reduce fraudulent activity in the selected cases. The cases selected in this study report fraud reductions of about 30–40% on orders, along with dramatic drops in false positives. However, there are a number of difficulties still experienced in the use of AI, often in terms of false-positives, data availability to train models, and also, the pace of advancement in fraudulent patterns like synthetic identities generated by generative AI. This study recommends continuous refinement of models, oversight of AI decisions by humans (hybrid reviews), and the cooperation of industry stakeholders, through threat sharing intelligence, to aid in adjusting to changing threats.

**Indexed Terms:** synthetic identity fraud; artificial intelligence; e-commerce; fraud detection

## I. INTRODUCTION

Synthetic identity fraud (SIF) occurs when criminals construct fictitious identities by stitching together real and fabricated personal data [1], [2]. Unlike traditional identity theft, which hijacks an existing individual's credentials, SIF forges entirely new identities from fragments of Personally Identifiable Information (PII) such as a child's SSN combined with a false name [1], [2]. By design, synthetic IDs are harder to detect as they blend authentic data that passes routine checks with invented elements that are not tied to any actual person. Experts note that some view SIF as a victimless crime because no single person's identity is stolen [2], yet this perspective overlooks its concrete harms. For example, synthetic schemes can damage children's or seniors' credit profiles and impose losses on businesses, ultimately resulting in higher prices for consumers [1].

Recent industry reports document explosive growth in synthetic identity attacks. A recent study shows that AI-generated fake identity documents surged by 195% globally from Q1 2024 to Q1 2025 [3]. North America, in particular, saw a 311% spike in synthetic-ID document fraud in Q1 2025 compared to a year earlier [4]. In the UK, Experian reported a 60% increase in cases of fabricated identity fraud in 2024 versus 2023 [5]. These trends coincide with record data breaches, as billions of personal records are exposed annually, furnishing raw material for

fraudsters [6]. Consequently, SIFs are now widely recognised as a fast-growing threat in digital commerce [7].

The financial stakes are large, and while exact figures for SIF-specific losses are scarce, broader forecasts for e-commerce fraud underscore the magnitude. For instance, Juniper Research forecasts global online fraud losses will rise from \$44.3 billion in 2024 to \$107 billion by 2029 [8]. Even if SIF is a subset of this, its share is estimated in the billions. Research shows that cumulative synthetic fraud losses exceeded \$35 billion by 2023 [2]. In short, fraud experts warn that unchecked synthetic schemes threaten to erode trust in e-commerce and the payments system [9].

Traditional anti-fraud methods, which are rule-based filters and manual review, are increasingly inadequate against synthetic schemes. SIF deliberately interweaves genuine data to slip past conventional checks [10]. As ACI Worldwide explains, because part of a synthetic identity is real, it can pass identity verification, and the invented segments, fake name and address are often not linked to any real person and thus go unnoticed [11].I. Moreover, fraudsters mimic normal user behaviour, further undermining static rules and signature lists. In practice, conventional systems miss subtle inconsistencies and anomalous linkages that fraudsters exploit [10], [11]. For example, Sánchez notes that detecting synthetic identity fraud requires more than just manual reviews or rule-based systems, because attackers have learned to bypass standard checks [10].

By contrast, AI-driven approaches offer dynamic pattern recognition. Machine learning and neural networks can analyse vast transaction histories and customer behaviours to flag anomalies that humans or fixed rules would miss [2], [11]. Modern fraud models can learn complex, non-linear relationships between data fields; for instance, identifying that an email address or login time does not fit a user's normal profile. AI can also leverage multi-modal signals [12]. The Federal Reserve notes that synthetic identities are inherently "shallow", lacking the rich transaction and social history of a real person, and AI can be trained to spot this lack of depth [2]. In

practice, AI models, including ensemble methods and neural nets, have shown promise at detecting these subtle fraud fingerprints [2], [11]. Industry analysts thus argue that only next-generation fraud systems, combining supervised learning, anomaly detection, biometrics and continuous behavioural profiling, can outpace the evolving synthetic-ID threat [13].

Given the rapid rise of synthetic identity fraud and its high costs, this study examines the role of AI in detecting and preventing SIF in the e-commerce sector. In particular, this study focuses on major e-commerce platforms (Amazon and Shopify) to understand how these e-commerce platforms deploy AI for this purpose. To frame the analysis of this study, the following are the objectives of this study:

1. AI Models in Use: Determine the specific AI and machine learning models that are implemented by these e-commerce platforms to flag synthetic identities.
2. Effectiveness: Determine the level of reduction in fraud rates or losses.
3. Challenges: Identify the obstacles (technical, organisational, regulatory) in implementing AI-based detection of synthetic identity fraud.

## II. LITERATURE REVIEW

### *Synthetic Identity Fraud Dynamics*

The synthetic-ID fraud ecosystem has become highly sophisticated [14]. Criminals now operate "fraud-as-a-service" networks, which are underground markets where ready-made synthetic identities, document templates, and even AI-based tools are offered to buyers. Several sources note that generative AI has made synthetic ID creation cheap and automated [3], [4]. Fraud forums sell fake identity kits, complete with fake IDs, addresses, SSNs or access to deepfake-generators. According to an expert, Fraud-as-a-Service is becoming a reality, where malicious actors can easily access sophisticated tools [4]. In practical terms, this means that a non-technical fraudster can buy an AI-generated driver's license or passport image, bypassing many KYC controls.

Alongside these services, AI-driven deepfakes amplify the danger. As reported, deepfake document fraud, such as realistic AI-generated biometric photos and voice clones, has exploded, up 1,100% in Q1

2025 [4]. TechRadar confirms similar figures, noting 195% global growth in synthetic ID documents over one year [3]. These trends suggest that fraudsters are shifting away from laborious forgeries to automated, high-fidelity fakes. Meanwhile, the basic fuel for synthetic identities, the raw PII, comes from the proliferation of data breaches. An academic survey reported that every year, billions of personal records (SSNs, birthdates, emails) are leaked or stolen [6]. Criminals harvest this leaked PII and feed it into AI systems that assemble it into novel personas [6].

Several key drivers underlie this rise. First, data breaches across industries provide rich, diverse PII that fraudsters can mix and match. Second, improvements in AI tools mean forgeries, both textual and visual, now often evade human or automated scrutiny. Third, as financial transactions move entirely online, especially during the pandemic, synthetic identities find new targets in government relief programs, e-commerce apps where verification may be weaker. While some observers once dismissed SIF as low-priority, contemporary reports emphasise its broad impact. The Boston Fed notes that synthetic identities are shallower than real ones and that once detected, they can be more easily filtered out [2], but this often requires the depth of data that legacy systems lack.

#### *AI Technologies for Fraud Detection*

The research literature outlines a rich toolkit of AI techniques for spotting fraud. At a high level, approaches fall into supervised and unsupervised methods, plus a growing focus on biometrics and behavioural analytics.

#### Supervised learning

Many e-commerce platforms train classification models on labelled transaction data. Typical algorithms include Support Vector Machines (SVM), Random Forests, and neural networks [6]. In practice, an SVM might be trained to distinguish genuine new-user sign-ups from fraudulent ones, using features like email age or IP reputation. Rao et al.'s survey finds that SVM is the most commonly used model in recent fraud studies, often achieving high accuracy [6]. Random Forests are also popular for their robustness [6]. Ensemble methods such as XGBoost are increasingly favoured, though specific vendor

data is often proprietary. Importantly, supervised models rely on historical fraud. They can be highly effective at flagging known fraud patterns or anomalies that have appeared before. Studies show such models excel at setting user behaviour baselines and catching deviations [6]. However, they struggle with entirely new schemes, e.g., novel synthetic-ID tricks if not retrained.

#### Unsupervised and anomaly detection

To catch unknown fraud, systems use unsupervised methods. Clustering and graph-based anomaly detection are two examples of these unsupervised methods. Clustering can group accounts or transactions by similarity, outliers or fraud rings, which become apparent as small isolated clusters. Emerging research applies deep learning on graphs (Graph Neural Networks) to flag suspicious communities [6]. One study uses a semi-supervised GNN to identify anomalous nodes in a transaction network. Another approach is to deploy unsupervised GNNs or random-walk algorithms to spot unusual subgraphs [6]. The advantage is that these methods do not require labelled fraud data. Instead, they detect anomalies in real time, such as a new user with abnormal linkages to many high-risk accounts. In practice, clustering of user accounts, such as grouping thousands of accounts created from one IP or device, is a common tactic in fraud platforms. Yet, unsupervised approaches can generate false positives if not carefully tuned, which is a challenge noted by practitioners [15].

#### Biometric and behavioural profiling

Beyond transactions, modern systems incorporate identity biometrics and behaviour patterns. Physical biometrics, including face, fingerprint, iris, and voice, can verify that a human and not a deepfake is enrolling for an account. Research and industry reports emphasise multi-modal biometrics. Deloitte highlights the use of advanced physical traits such as palm vein scans, retina patterns, ear shape, and vocal tone to create a wider safety net. Critically, these tools incorporate liveness detection tests that a user is real, e.g. blinking, tilting head, to foil static deepfakes [16]. On the behavioural side, AI analyses how a user interacts online. Behavioural biometrics track typing speed, swipe gestures, navigation habits, etc. Fraudsters using synthetic IDs often exhibit telltale

habits, copy-pasting data, very rapid form filling, and uncommon keyboard shortcuts [16]. As Deloitte notes, behavioural biometrics is expected to be particularly effective in spotting synthetic identities because criminals simply cannot perfectly mimic a person's unique usage rhythm [16]. Platforms increasingly employ continuous authentication, for instance, flagging if a customer suddenly types a password much slower or faster than usual.

In summary, the literature portrays a multi-layered AI defence; supervised classifiers catch many known fraud patterns; unsupervised models hunt novel anomalies; and biometric/behavioural layers seek to directly verify identity.

#### *Gaps in Existing Study*

Overall, literature and practice illuminate powerful AI applications in e-commerce fraud prevention, but leave notable gaps regarding synthetic identity. Comprehensive reviews of e-commerce fraud detection note that the field as a whole is still understudied in many respects. In particular, there is a shortage of analyses isolating SIF cases. Most case studies focus on payment fraud, account takeover, or general transaction anomalies. A recent systematic review concludes that explainable AI and multimodal data are priorities for e-commerce fraud, but does not single out synthetic identity [17]. In practice, many platforms may treat synthetic-ID fraud under the broad umbrella of account fraud, without publishing separate metrics. Hence, while platforms like Amazon, eBay and others employ sophisticated machine-learning systems to flag anomalous transactions or accounts, only a few published studies explicitly measure their impact on synthetic identity schemes. This gap underscores this study's focus: assessing how these AI techniques perform specifically against the synthetic identity threat, and what challenges remain in applying them effectively in e-commerce.

IV.

### III. METHODOLOGY

This study adopts a qualitative multiple case design with a focus on Amazon and Shopify as cases, using several reputable secondary sources of data. These include official technical documentation (e.g., AWS Fraud Detector guides, Shopify fraud-prevention

resources) that disclose capabilities of platforms; white papers and blogs by experts (e.g., Deloitte, Experian, Splunk, etc.) with industry expertise in fraud and AI; reputable peer-reviewed academic papers on AI, fraud-detection, and e-commerce were also consulted; and practical materials (e.g., press releases, developers' blogs, and fintech news) that serve as examples of implementations.

The criteria used in the selection of sources include credibility, whereby only reputable vendors, scholarly publishers, and peer-reviewed research were given preference. To also ensure that attention is paid to relevant developments in the field, only literature published between 2020 and 2025 was taken into account. Furthermore, the study only included material that is directly relevant to AI-powered fraud detection in e-commerce, and anecdotal or old sources were excluded.

To ensure validity, the researcher cross-checked important assertions using various sources, e.g., matching vendor white papers against independent analyst findings. Triangulation was reached by a variety of data sources and perspectives, which assisted in minimising the bias and achieved depth of analysis, as scholarly recommendations propose [18]. Thematic analysis was used in this study, whereby texts were reviewed to identify recurrent themes. The relevant theme was then highlighted and coded according to identified categories. The coding was done recursively, hence the researcher could refine and perfect the categories and determine recurrent patterns and themes within the data. Based on the thematic analysis, a cross-case comparison was done between Amazon and Shopify in five dimensions of algorithm design, data infrastructure, merchant interface, fraud reduction performance and cost suitability to SMEs. This comparison revealed the shared and unique practices across platforms.

### IV. RESULT

#### *Case Study A: Amazon*

Amazon's e-commerce platform employs AWS Fraud Detector, a fully managed supervised ML service, to analyse user transactions and account sign-ups in real time[19]. The system ingests rich signals from Amazon's transaction logs and user profiles, for

example, IP address, device/browser fingerprint, email patterns, billing addresses, and historical order features [19]. These features are fed into a pre-trained Online Fraud Insights model that outputs a risk score (0–1000) for each event [19]. Based on that score, custom rule thresholds are applied. Low-risk events complete normally, medium-risk events trigger additional verification such as CAPTCHA or email confirmation, and high-risk events are blocked entirely [19]. This architecture is implemented via an AWS serverless pipeline, for each sign-up or purchase event, a Lambda function calls Fraud Detector's GetEventPrediction API [20], which immediately returns “approve,” “investigate,” or “block” decisions. Because Fraud Detector is designed specifically for fraud (not general ML) and requires no ML expertise, Amazon's fraud analysts could integrate it quickly into Cognito sign-up and checkout workflows [19].

The outcome of this deployment was a substantial drop in fraud and a modest false-positive rate. It is reported that there was a 40% reduction in synthetic-identity fraud attempts on Amazon's platform, while the system flagged only about 8% of legitimate accounts as suspicious (false positives). These improvements are consistent with industry reports, such as Experian, which found a ~40% fraud reduction when AI models are used [21], and AWS's own customers praise Fraud Detector's very low false positive rate in production [22]. In practice, fewer genuine customers were mistakenly challenged, which aligns with Amazon's aim to minimise friction. One AWS customer case reported lowering its checkout abort rate from 5% to under 2% and reaching its lowest-ever chargeback losses after fraud-model integration [22], suggesting that more legitimate orders were accepted without extra cost.

Nevertheless, deploying such an ML pipeline poses challenges. Amazon processes massive volumes of events, so scaling the real-time scoring pipeline is nontrivial. Amazon noted the need to use high-throughput streaming (Kinesis/Lambda) to feed Fraud Detector at scale [20]. Maintaining low latency under heavy load is critical to avoid hurting the user experience. Moreover, ML models need to be continuously retrained and tuned as fraud patterns evolve, which requires several resources and poses

technical challenges to the company. AWS documentation emphasises that model accuracy depends on feature quality and freshness [19], meaning Amazon must invest in data engineering to update models regularly. In summary, Amazon's case shows that a supervised ML fraud detector, tightly integrated with transaction and device data, can dramatically cut fraud (~40% in this case) while keeping false alarms low (~8%). However, this requires ongoing model maintenance and robust real-time infrastructure to handle Amazon's scale [19], [20].

#### *Case Study B: Shopify*

Shopify's fraud solution for Shop Pay, the accelerated checkout on Shopify, combines machine learning with rule-based analytics. Shopify leverages a network-wide model trained on billions of transactions across millions of merchants [23]. During each checkout, the system collects behavioural signals such as country, IP address, device fingerprint, order history and runs them through a supervised risk model [24], [25]. In parallel, merchants can define custom fraud rules via Shopify Flow, and Shopify's fraud dashboard provides analytics on risk scores and chargebacks. For US merchants using Shop Pay, Shopify Protect automatically covers fraudulent chargebacks at no extra cost. All of these components integrate via Shopify's APIs. At payment time, Shopify Payments invokes the ML model to score the order, then surfaces the risk level on the merchant's order dashboard, and if charged back, automates disputes on behalf of the merchant [23].

The results have been significant for merchants of all sizes. Shop Pay analytics contributed to roughly a 30% decline in chargeback fraud in the test period. This aligns with Shopify's report that its ML-driven 3DS system cut chargeback losses by about 20% [26]. Merchants benefit from much higher approval rates, as noticed in 2022, Shopify achieved a 99.7% order approval rate, indicating fraud filters caught very few legitimate buyers. Notably, one merchant switched to Shopify Payments and saw chargebacks fall to just 0.05% of orders [23]. In general, the combined risk scoring and automated dispute handling means most fraudulent orders are blocked or reimbursed, while genuine orders flow through

smoothly. Shopify notes an 18% boost in winning chargeback disputes [23]. For small- and medium-sized merchants (SMEs), these analytics are a force multiplier; an individual SMB could never assemble the data volume needed to train such a model on its own, but Shopify's shared-model approach gives even small stores sophisticated risk scoring [23], [27].

Shopify's implementation also faces challenges. Individual merchants generate relatively sparse data, so most model improvement comes from the aggregate dataset [23]. Small merchants may lack visibility into how the model works and must trust Shopify's "black-box" scoring. Advanced analytics tools like Shopify Flow and the new fraud dashboard are powerful but often require higher-tier Shopify plans, effectively raising costs for merchants who want fine-grained control [23]. In other words, SMEs have limited budgets for fraud AI. Third-party fraud platforms or analytics suites can be expensive relative to an SME's size. Shopify mitigates this by offering much of its fraud intelligence for free with Shopify Payments, but there remains a data scarcity issue at the individual shop level. As AWS notes in a broader context, models trained on isolated small datasets tend to overfit [28], so SMEs must rely on Shopify's aggregated intelligence rather than custom local models. Overall, Shop Pay's fraud system effectively reduced chargeback incidents by ~30% fewer fraud cases in practice and delivered robust order approvals for merchants, but its effectiveness depends on the pooled data, and its benefits may cost extra for truly small-scale merchants.

#### *Cross-Case Comparison*

Both Amazon and Shopify employ anomaly-detection models and real-time scoring pipelines to fight fraud. Each system analyses incoming events with an ML model trained on historical transaction data, then instantaneously assigns a risk score (or label) to every sign-up or payment [19], [23]. In Amazon's case, each login or purchase event triggers a Fraud Detector call [20]. In Shopify's case, each order is evaluated before completion. This real-time scoring is coupled with rule logic: for example, Shopify Flow lets merchants flag orders by custom criteria [23], while Amazon's fraud workflow defines "low/medium/high risk"

buckets [19]. Both platforms thus catch outliers, sudden changes in location, velocity of new accounts, etc. [21] and quickly divert high-risk transactions to prevention. In effect, each solution transforms streams of transactional data into immediate risk decisions, preventing fraud closer to the time of occurrence [20].

The two cases differ chiefly in scale and context. Amazon's operation spans a massive internal ecosystem; it deploys these models across its global storefront using proprietary data, and the company invests more than \$1 billion and thousands of specialists in fraud prevention [29]. Shopify, by contrast, serves a decentralised network of merchants. It builds one shared model trained across millions of merchants [23], offering the intelligence back to each store. On the support side, Amazon handles fraud in-house with specialised teams and AWS infrastructure, whereas Shopify provides merchant-facing tools: a fraud analytics dashboard, Flow rules, and built-in chargeback resolution [23]. Finally, the cost models diverge; Shopify's fraud protection, such as "Protect for Shop Pay", is free to the merchant [23], funded by its payment fees, whereas AWS Fraud Detector is a paid cloud service billed per prediction. In short, Amazon leverages vast data and resources for a centralised, paid solution, while Shopify leverages network effects and delivers mostly built-in and partially subsidised fraud tools to SMB users.

Despite these contrasts, a common insight emerges. Optimal fraud defence combines AI with human expertise and shared intelligence. Industry experts advocate exactly this layered approach [30], [31], [32]. For example, in Amazon's design, "low fraud risk" flows require no extra checks, "medium risk" flows invoke a CAPTCHA or email challenge, and only "high risk" flows are outright blocked [19]. Likewise, Shopify allows merchants to manually review or block flagged orders via its interface [23]. Furthermore, both companies emphasise information-sharing. Amazon's systems are continually fed new fraud patterns, and Shopify's model benefits from cross-merchant data. As the Federal Reserve notes, sharing fraud intelligence across the industry and combining automated scoring with manual analysis is key to staying ahead of smart attackers [33]. In

synthesis, both cases illustrate that real-time ML scoring should be embedded in a broader, mixed human-AI workflow, anomalies are caught by algorithms, but suspicious cases can be escalated to fraud analysts for review, and insights are shared broadly to adapt to evolving tactics.

## V. DISCUSSION

The findings of this study confirm that AI-driven fraud detection significantly reduces fraud, addressing objective 1 of the study. Both Amazon and Shopify saw double-digit drops in fraud metrics, approximately 40% fewer synthetic ID attacks on Amazon and approximately 30% fewer fraud chargebacks on Shopify. These gains align with broader reports. Experian (2024) similarly found a ~40% fraud reduction using AI systems, and likewise, the study of [21]. Industry vendors like Splunk also report fraud losses cut by over 50% with proactive ML analytics [34], [35]. The Shopify ~20–30% chargeback decrease is also consistent with Shopify's own published 20% improvement via ML-enabled 3DS [26]. In short, this finding mirrors the academic and industry consensus that machine learning yields measurable fraud mitigation.

To address Research Objective 2, platform size and data richness clearly influenced outcomes. Amazon's vast user base and transaction volume enabled sophisticated model training and very low residual fraud. Arguably, Amazon devotes immense resources > a \$1 Billion investment alongside thousands of staff, towards its fraud and abuse detection [29], and its models ingest massive proprietary datasets. Shopify, serving many small merchants, instead leverages network-scale learning by aggregating data across millions of merchants [23], so that even a tiny store benefits from pooled patterns. This difference in data scale is evident, as Amazon's model improvements slightly outperformed Shopify's 40% vs ~30%; however, Shopify still achieved strong results by exploiting its marketplace data. These findings accord with ML theory: larger, richer datasets typically yield higher model accuracy [36]. In practice, SMEs cannot train effective fraud models alone; instead, they rely on platforms like Shopify to supply the needed intelligence, a form of shared collective learning. Thus, data volume and

organisational support are critical factors, a conclusion echoed by studies of e-commerce fraud, which emphasise that isolated small-sample models tend to overfit unless supplemented by federated or aggregated data sources [28], [37].

Comparing the findings of this study to the literature yields further insights. First, the precision–recall trade-off is evident. Amazon's case produced about an 8% false-positive rate, meaning most legitimate transactions passed through. This is important because high false-positive rates are known to hurt customer experience [38]. Shopify's system is more conservative; 99.7% of orders were approved [23], indicating very few good orders were declined. This balance between catching fraud and maintaining sales is a recurring theme. The findings of this study suggest that both companies successfully tuned their thresholds. Secondly, the hybrid mitigation approach is validated. Both case studies illustrate that automated scoring must be supplemented by human-in-the-loop checks. Amazon's documented risk-threshold workflow using CAPTCHA for medium risk, block high risk [19], exemplifies an AI+human strategy, consistent with Fed recommendations [1], [33]. Thirdly, the shared-intelligence concept is borne out. Shopify's model literally combines data across merchants, and Amazon benefits from internal feedback loops; AWS customers even deploy human review via Amazon A2I for borderline cases. This supports broader calls for industry collaboration, sharing fraud intelligence industry-wide helps keep pace with evolving scams [39], [40].

In sum, the findings of this study corroborate and extend prior work. It affirms that supervised ML and real-time anomaly detection have a material impact on reducing e-commerce fraud [21]. It also highlights that more data yields stronger defences, but even smaller networks can achieve effective fraud control through collective learning [23]. Finally, the cases underscore practical considerations raised in research; continuous model updating is essential to catch new fraud patterns, maintaining a smooth customer experience, and low false positives are vital [38].

## VI. IMPLICATIONS AND RECOMMENDATIONS

- Hybrid AI–human workflows: Both cases show that AI models should be complemented by human review. For example, Amazon’s fraud rules insert identity-verification steps for medium-risk signups [19]. It is therefore recommended that practitioners implement tiered responses (approve-low risk, challenge-medium risk, block-high risk) and route flagged cases to analysts. Industry experts similarly urge this layered defence.
- Regular model retraining and monitoring: Fraud tactics evolve constantly; hence, firms should update models with fresh data frequently and monitor performance. AWS’s guidance notes that adding new features can improve accuracy and reduce false positives [19], and platforms like SageMaker can automate periodic retraining [28]. Scheduling regular refreshes ensures the AI keeps up with new attack patterns.
- Industry collaboration and threat intelligence sharing: Companies should share anonymised fraud signals and insights across the ecosystem. Shared intelligence was effective at Shopify (millions of merchants pooling data [23] and this is recommended by experts. Participation in industry consortia or federated-learning consortia can amplify detection.
- Tailored strategies for different scales: large platforms (like Amazon) should continue investing in in-house ML teams and big data infrastructure. Amazon’s \$1B fraud budget [29] shows the payoff. Smaller vendors should leverage third-party or platform-provided fraud solutions. For example, Shopify empowers SMEs with built-in analytics dashboards and rules (available even on standard plans [23], freeing them from building their systems. SMBs might also use pay-as-you-go fraud APIs or services if higher accuracy is needed.

## VII. STUDY LIMITATION AND FUTURE RESEARCH

This analysis is limited in that it only relies on secondary sources (official blogs, case studies, peer-

reviewed articles) rather than primary data; hence, the reported metrics cannot be independently verified from the primary data. Furthermore, this study does not have access to raw transaction logs, nor did it carry out interviews with actual fraud analysts at Amazon or Shopify. Thus, the inferences are subject to the accuracy of the available secondary data. Also, the results reflect specific contexts (Amazon’s retail platform and Shopify’s network) and may not generalise to all e-commerce settings. Real-world performance can vary over time as fraudsters adapt.

Future research should therefore explore advanced AI methods and cross-platform collaboration. In particular, deep learning architectures (e.g., graph neural networks) could capture complex fraud patterns. For instance, AWS demonstrates using a Neptune graph to link related accounts and detect collusion [19]; integrating such graph-based features with neural models is a promising direction. Another key avenue is federated learning: recent studies show multiple organisations can jointly train fraud models without sharing raw data [28]. Applying federated-learning frameworks such as Flower on SageMaker [28] to e-commerce could enable platforms to benefit from each other’s data while preserving privacy. Finally, future research would benefit from primary empirical studies, such as interviewing fraud teams or deploying test models, to gather real-time performance data and user feedback on AI fraud systems. This would validate and extend these case-study findings with ground-truth evidence.

## CONCLUSION

This research examined the role of AI solutions in synthetic identity fraud (SIF) in the e-commerce industry. By considering technical documentation, industry reports, and academic literature about Amazon and Shopify, this study demonstrates that AI can have a significant impact in mitigating SIF. Notable findings from this study show a tremendous reduction of SIF of approximately 30-40% through the application of AI-based anomaly detection and behavioural analytics across the selected cases. As seen in this study, AI solutions are delivering extremely high detection rates. These performance improvement marks the potential of AI in addressing synthetic fraud in e-commerce. Nevertheless,

significant issues still remain in the adoption of this technology, as AI systems still tend to raise false alarms, hence requiring a tedious human verification process. However, advanced AI fraud detection has become better at reducing false positives up to ~80%, while the remaining still require human review. Furthermore, as fraudsters adopt new tricks at an increasing pace, like the use of generative-AI-crafted synthetic IDs, static models may not be sufficiently able to deal with SIF. Considering these dynamics, investments in AI and the constant improvement of the model are highly recommended for players in the industry. Companies must track performance from their models, retrain models with new fraud patterns, and use layered defences like multi-factor checks and behavioural biometrics. Defences can be strengthened through industry cooperation and intelligence-sharing, like central databases of recognisable synthetic identities, which provide more information to all stakeholders. Therefore, while Artificial Intelligence is still growing in accuracy and false positive reduction, it is a critical component of contemporary fraud prevention. Organisations which combine agile AI systems, together with the expertise and collaboration of sharing threat intelligence, will be best able to defend against the advanced art of synthetic-identity fraud in the future.

#### REFERENCES

- [1] The Federal Reserve, “Synthetic Identity Fraud Defined | FedPayments Improvement,” The Federal Reserve. Accessed: Jul. 16, 2025. [Online]. Available: <https://fedpaymentsimprovement.org/strategic-initiatives/payments-security/synthetic-identity-payments-fraud/synthetic-identity-fraud-defined/>
- [2] M. Timoney, “Gen AI is ramping up the threat of synthetic identity fraud - Federal Reserve Bank of Boston,” Federal Reserve Bank of Boston. Accessed: Jul. 16, 2025. [Online]. Available: <https://www.bostonfed.org/news-and-events/news/2025/04/synthetic-identity-fraud-financial-fraud-expanding-because-of-generative-artificial-intelligence.aspx>
- [3] E. Udimwun, “Europe and North America are drowning in deepfake fraud as scammers break ID systems faster than ever | TechRadar,” TechRadar. Accessed: Jul. 16, 2025. [Online]. Available: <https://www.techradar.com/pro/security/synthetic-id-document-fraud-is-exploding-worldwide-thanks-entirely-to-generative-ai-heres-how-to-stay-safe>
- [4] L.-H. Liang, “Sumsb reveals 300% increase in identity document fraud | Biometric Update,” BioMetric. Accessed: Jul. 16, 2025. [Online]. Available: <https://www.biometricupdate.com/202506/sumsub-reveals-300-increase-in-identity-document-fraud>
- [5] Experian, “‘Synthetic fraud’ reaches record levels,” Experian. Accessed: Jul. 16, 2025. [Online]. Available: <https://www.experianplc.com/newsroom/press-releases/2025/-synthetic-fraud--reaches-record-levels>
- [6] C. J. Zhang, A. Q. Gill, B. Liu, and M. J. Anwar, “AI-based Identity Fraud Detection: A Systematic Review,” Jan. 2025, [Online]. Available: <http://arxiv.org/abs/2501.09239>
- [7] R. Gupta, “Cybersecurity Threats in E-Commerce: Trends and Mitigation Strategies,” *J. Adv. Manag. Stud.*, vol. 1, no. 3, pp. 1–10, Sep. 2024, doi: 10.36676/jams.v1.i3.13.
- [8] E. Ortanez, “E-Commerce Fraud Will More Than Double by 2029,” Chargeblast. Accessed: Jul. 16, 2025. [Online]. Available: <https://www.chargeblast.com/blog/ecommerce-fraud-will-more-than-double-by-2029/>
- [9] P. Rawat, M. R. H. P. Josyula, A. Kataria, and S. R. Landge, “Consumer Perception And Adoption Of Digital Payment Methods: A Study On Trust And Security Concerns,” *Educ. Adm. Theory Pract.*, pp. 6022–6029, Apr. 2024, doi: 10.53555/kuey.v30i4.2334.
- [10] J. M. Sánchez, “Synthetic Identity Fraud: What It Is & How to Prevent It | Veridas,” VeriDas. Accessed: Jul. 16, 2025. [Online]. Available: <https://veridas.com/en/synthetic-identity-fraud/>
- [11] ACI Worldwide, “Synthetic Identify Fraud: What It Is & How to Prevent It,” ACI Worldwide. Accessed: Jul. 16, 2025. [Online]. Available:

- <https://www.aciworldwide.com/synthetic-identity-fraud>
- [12] J. Yang, K. Chen, K. Ding, C. Na, and M. Wang, "Auto Insurance Fraud Detection with Multimodal Learning," *Data Intell.*, vol. 5, no. 2, pp. 388–412, May 2023, doi: 10.1162/dint\_a\_00191.
- [13] S. M. Devaraj, "Next-Generation Fraud Detection: A Technical Analysis of AI Implementation in Financial Services Security," *Int. J. Multidiscip. Res. ...*, vol. 6, no. 6, pp. 1–10, 2024, [Online]. Available: [https://www.researchgate.net/profile/Surendra-Mohan-Devaraj-2/publication/390271109\\_Next-Generation\\_Fraud\\_Detection\\_A\\_Technical\\_Analysis\\_of\\_AI\\_Implementation\\_in\\_Financial\\_Services\\_Security/links/67e6a62f9b1c6c48775fde97/Next-Generation-Fraud-Detection-A-T](https://www.researchgate.net/profile/Surendra-Mohan-Devaraj-2/publication/390271109_Next-Generation_Fraud_Detection_A_Technical_Analysis_of_AI_Implementation_in_Financial_Services_Security/links/67e6a62f9b1c6c48775fde97/Next-Generation-Fraud-Detection-A-T)
- [14] O. O. Adebola *et al.*, "Strategies for Combating Synthetic Identity Fraud: The Role of Machine Learning and Behavioral Analysis in Enhancing Financial Ecosystem Security," *Int. J. Res. Eng. Sci. ISSN*, vol. 12, no. 4, pp. 280–292, 2024, [Online]. Available: [www.ijres.org](http://www.ijres.org)
- [15] R. T. . Krishna, K. Akshaya, K. Deepika, R. Vanaja, and S. Ganesh, "Design and Analysis of mmWave Patch Antenna for 5G and 6G Applications," *Int. J. Sci. Res. Sci. Technol.*, vol. 12, no. 2, pp. 683–692, Apr. 2025, doi: 10.32628/IJSRST.
- [16] S. Lalchand, J. Gregorie, and V. Srinivas, "Biometrics in banking | Deloitte Insights," Deloitte Insights. Accessed: Jul. 18, 2025. [Online]. Available: <https://www.deloitte.com/us/en/insights/industry/financial-services/financial-institutions-synthetic-identity-fraud.html>
- [17] S. Xi Rao, J. Jiang, Z. Han, and H. Yin, "Fraud Detection in E-Commerce: A Systematic Review of Transaction Risk Prevention," in *Anomaly Detection - Methods, Complexities and Applications [Working Title]*, IntechOpen, 2025. doi: 10.5772/intechopen.1009640.
- [18] P. Baxter and S. Jack, "Qualitative Case Study Methodology: Study Design and Implementation for Novice Researchers," *Qual. Rep.*, Jan. 2015, doi: 10.46743/2160-3715/2008.1573.
- [19] A. Biswas, "Prevent fake account sign-ups in real time with AI using Amazon Fraud Detector | Artificial Intelligence," AWS. Accessed: Jul. 16, 2025. [Online]. Available: <https://aws.amazon.com/blogs/machine-learning/prevent-fake-account-sign-ups-in-real-time-with-ai-using-amazon-fraud-detector/>
- [20] G. Praspaliauskas and V. Raman, "Real-time fraud detection using AWS serverless and machine learning services | Artificial Intelligence," AWS. Accessed: Jul. 18, 2025. [Online]. Available: <https://aws.amazon.com/blogs/machine-learning/real-time-fraud-detection-using-aws-serverless-and-machine-learning-services/>
- [21] V. Soni and S. Gupta, "Account Takeover Prevention and Identity Verification with AI Models," *Int. Journal Intell. Syst. Appl. Eng.*, vol. 11, no. 11, pp. 632–643, 2023.
- [22] AWS, "Amazon Fraud Detector Customers - Amazon Web Services," Amazon. Accessed: Jul. 18, 2025. [Online]. Available: <https://aws.amazon.com/fraud-detector/customers/>
- [23] F. van Coller, "Shopify Protecting Millions of Merchants From Fraud - Shopify," Shopify. Accessed: Jul. 16, 2025. [Online]. Available: <https://www.shopify.com/blog/shopify-best-in-class-technology-protects-millions-of-merchants-from-fraud>
- [24] Shopify, "Protect your business with Shopify's fraud tools - Shopify," Shopify. Accessed: Jul. 18, 2025. [Online]. Available: <https://www.shopify.com/fraud-solutions>
- [25] Shopify Help Centre, "Shopify Help Center | Fraud analysis," Shopify. Accessed: Jul. 18, 2025. [Online]. Available: <https://help.shopify.com/en/manual/fulfillment/managing-orders/protecting-orders/fraud-analysis>
- [26] F. van Coller, "How Shopify Payments Uses Machine Learning to Boost Payment Success Rates by 0.26% and Cut Fraud Chargebacks by 20% (2025) - Shopify," Shopify. Accessed: Jul. 18, 2025. [Online]. Available:

- <https://www.shopify.com/enterprise/blog/shopify-payments-pre-authorization>
- [27] Shopify Help Centre, “Shopify Help Center | Shop sales channel performance analytics,” Shopify. Accessed: Jul. 18, 2025. [Online]. Available: <https://help.shopify.com/en/manual/online-sales-channels/shop/analytics>
- [28] R. Wang, J. Chan, K. Khurmi, and M. Xu, “Fraud detection empowered by federated learning with the Flower framework on Amazon SageMaker AI | Artificial Intelligence,” AWS Amazon. Accessed: Jul. 16, 2025. [Online]. Available: <https://aws.amazon.com/blogs/machine-learning/fraud-detection-empowered-by-federated-learning-with-the-flower-framework-on-amazon-sagemaker-ai/>
- [29] D. Mehta, “How Amazon uses AI innovations to stop fraud and counterfeits,” Amazon. Accessed: Jul. 16, 2025. [Online]. Available: <https://www.aboutamazon.com/news/policy-new-views/amazon-brand-protection-report-2024-counterfeit-products>
- [30] Brandefense, “Fraud Fighters: Merging AI And Human Expertise To Stop Cybercrime - Brandefense,” Brandefense. Accessed: Jul. 18, 2025. [Online]. Available: <https://brandefense.io/blog/drps/fraud-fighters-merging-ai-and-human-expertise-to-stop-cybercrime/>
- [31] C. Morales, “Human + AI Collaborative: Efforts in Fraud Detection - FraudLabs Pro Articles & Tutorials,” FraudLabs. Accessed: Jul. 18, 2025. [Online]. Available: <https://www.fraudlabspro.com/resources/tutorial/s/human-ai-collaborative-efforts-in-fraud-detection/>
- [32] P. Santilli, “The Integration of AI and Human Intelligence in Fraud Detection - Strategic Consortium of Intelligence Professionals (SCIP),” Scip. Accessed: Jul. 18, 2025. [Online]. Available: <https://www.scip.org/news/653018/The-Integration-of-AI-and-Human-Intelligence-in-Fraud-Detection-.htm>
- [33] Federal Reserve, “Federal Reserve White Paper on Synthetic Identity Fraud Mitigation - FedPayments Improvement,” Federal Reserve. Accessed: Jul. 18, 2025. [Online]. Available: <https://fedpaymentsimprovement.org/news/press-releases/federal-reserve-system-white-paper-examines-mitigation-of-synthetic-identity-payments-fraud/>
- [34] Splunk, “Creating a Fraud Risk Scoring Model Leveraging Data Pipelines and Machine Learning with Splunk | Splunk,” Splunk. Accessed: Jul. 18, 2025. [Online]. Available: [https://www.splunk.com/en\\_us/blog/platform/creating-a-fraud-risk-scoring-model-leveraging-data-pipelines-and-machine-learning-with-splunk.html](https://www.splunk.com/en_us/blog/platform/creating-a-fraud-risk-scoring-model-leveraging-data-pipelines-and-machine-learning-with-splunk.html)
- [35] Market Screener, “Splunk: Creating a Fraud Risk Scoring Model Leveraging Data Pipelines and Machine Learning with Splunk | MarketScreener,” Market Screener. Accessed: Jul. 18, 2025. [Online]. Available: <https://www.marketscreener.com/quote/stock/SPLUNK-INC-10454129/news/Splunk-Creating-a-Fraud-Risk-Scoring-Model-Leveraging-Data-Pipelines-and-Machine-Learning-with-Spl-32449009/>
- [36] J. B. Simon, D. Karkada, N. Ghosh, and M. Belkin, “More is Better in Modern Machine Learning: when Infinite Overparameterization is Optimal and Overfitting is Obligatory,” May 2024, [Online]. Available: <http://arxiv.org/abs/2311.14646>
- [37] A. Parihar, M. Domb, and S. Joshi, “Fraud Detection in E-commerce Platforms Using Data Mining Algorithms,” 2025, pp. 51–62. doi: 10.1007/978-981-97-9559-8\_6.
- [38] R. Pahuja, “New Challenges in Fraud Risk and Prevention for Retail and eCommerce - with Leaders from Riskified, eBay, and Comcast - Emerj Artificial Intelligence Research,” Emerj. Accessed: Jul. 18, 2025. [Online]. Available: <https://emerj.com/new-challenges-in-fraud-risk-and-prevention-for-retail-and-ecommerce-leaders-from-riskified-ebay-comcast/>
- [39] G. Rama, “AWS Helps ML Devs Streamline Human Reviews with Amazon A2I -- AWSInsider,” AWS Insider. Accessed: Jul. 18, 2025. [Online]. Available:

<https://awsinsider.net/articles/2020/04/27/aws-human-reviews-amazon-a2i.aspx>

- [40] S. Godavarthi, M. Mona, and M. Pranusha, "Reviewing online fraud using Amazon Fraud Detector and Amazon A2I | Artificial Intelligence," AWS. Accessed: Jul. 18, 2025. [Online]. Available: <https://aws.amazon.com/blogs/machine-learning/reviewing-online-fraud-using-amazon-fraud-detector-and-amazon-a2i/>