AI-Driven Fraud Detection in UK Digital Payment Systems: Challenges and Solutions

ADEPEJU DEBORAH BELLO [1], OLUWASEYI BABATUNDE OGUNTOLA [2], AYODEJI TEMITOPE AJIBADE[3], AKINDAYO AKINDOLANI[4], OLUWADAMILOLA AYOOLA[5], AJIFOLAWE MUBARAQ BELLO[6]

[1] CyberFraud R&D Analyst, Barclays UK,
[2] Senior Auditor, Cybersecurity & Technology Infrastructure, Citi Bank, US,
[3] Fraud Analyst, Barclays UK,
[4] Director, McAnderson Institute of Technology,
[5] Doctoral Candidate, University of Fairfax,
[6] Controls Oversight Professional, Barclays, UK

Abstract- This study presents a systematic literature review on the use of Artificial Intelligence (AI) in fraud detection within the digital payment systems of the United Kingdom. As more people embrace the use of contactless cards, mobile wallets and peer-to-peer payment methods, the UK is on the edge of more sophisticated fraud. This study therefore consolidates peer-reviewed articles and authoritative grey literature on machine learning, deep learning and a hybrid-based model performance in fraud detection. It also investigates operational challenges, such as data asymmetry, model transparency, and regulations like the UK GDPR and FCA sandbox scheme. The findings from this study indicate the necessity of explainable AI, more diverse public datasets, and collaborative regulatory frameworks. This study recommends strategic testing of AI for fraud detection in controlled settings and implementation of a human-in-the-loop system. This paper enriches the discussion of fintech in the UK, explaining the existing possibilities, limitations, and future directions toward having a trustworthy and scalable AI fraud detection system.

Indexed Terms- Artificial Intelligence, Digital Payment Systems, Fraud Detection, Machine Learning, United Kingdom.

I. INTRODUCTION

Globally there has been a rapid shift towards cashless and online payment transactions in everyday business [1]. The UK consumers and businesses have made more than 48 billion payments in 2023, which is 5 per cent higher than the previous one, and contactless tap payments were nearly 38 per cent of all transactions [2]. More importantly, the usage of mobile contactless, e.g., Apple Pay, Google Pay, has seen high adoption, with approximately 1 in 3 UK adults making at least a monthly contactless payment using their mobile device [2]. In the same way, digital wallets, have gone mainstream. An industry report shows that 40 per cent of UK online purchases in 2024 were through digital wallets and is set to increase to 68 per cent by 2030) [3]. Faster payment rails and banking applications have also expanded peer-to-peer (P2P) payments, with the consumer desire to improve peer transfers. Overall, the payment behaviour in the UK is becoming progressively digitalised, with contactless cards and mobile apps, online checkouts and peer-to-peer payments gradually replacing cash-based transactions [2],[3].

The increased usage of digital payments, unfortunately, has been paralleled by an increased wave of fraud [4]. According to financial crime reports, digital payment fraud has become one of the most serious threats to the UK financial system [5]. In 2024, payment fraud alone resulted in fraud accounting to an estimated amount of £1–1.2 billion and a total of over 3.3 million crimes. Particularly, remote purchase fraud, in which stolen card information is used for online shopping, increased by 22% in 2024 with 2.6 million cases [6]. There has also been an increase in Authorised Push Payment

(APP) fraud, where victims are deceived and defrauded into sending money, with 222,000 instances recorded and losses of more than £340m in the UK in 2023 [7]. This type of fraud exploits online platforms like social media, email, and phishing websites etc. and has significant financial and psychological implications on consumers [6], [7]. Thus, as there is a growth in the use of digital payments in the UK, so also is the increase in fraud related to digital payment systems.

Conventional fraud detection techniques have not been able to keep up with this evolving environment. Historically, the ruleset definitions and manual review processes have been utilised by the banks and merchants as a way to identify fraud, e.g., writing a rule to flag transactions larger than a set amount, or querying a transaction manually [8]. Nevertheless, these systems are limited due to their inflexibility [8]. The main drawback of these traditional systems, as Odufisan et al. (2025) observe, is the level of flexibility; the fixed rules used are not responsive enough to adapt to the ever-changing tactics of fraudsters [9]. These techniques are also poor at capturing complex, multivariate patterns, as fraud patterns can be sophisticated which can be missed by rule engines [10]. Furthermore, humans and legacy systems cannot keep up with transaction volumes massively increasing as the massive data quantity and real-time streaming data make manual or low-logic-based checks impractical [11]. These limitations, in practice, imply that traditional fraud defences may result in a large number of false positives, marking legitimate users and failing to capture advanced attacks [9].

To address these challenges, researchers and industry experts have begun to advocate and adopt AI-driven solutions. Machine learning and artificial intelligence methods can quickly process large and intricate transaction data and detect changes in fraud patterns. Empirical evidence has shown that fraud models using AI may have exceedingly accurate results, typically ranging from 90-95%, compared to conventional methods, hence, lowering undetected fraud and unnecessary alerting significantly [12]. Although several reviews in this field have discussed the possibility of AI in the detection of fraud, they do not provide a holistic and comprehensive consideration for the UK digital-payment environment and typologies of fraud [13], [14]. This indicates a gap, that there is a need to have a comprehensive understanding of how AI techniques have been, and can be, applied specifically to UK digital payment fraud. By undertaking a systematic literature review focused on the UK digital payment system, this study aims to fill that gap, synthesising what is known about AI-driven fraud detection in the country's contactless, mobile, and peer-to-peer payment systems.

II. RESEARCH AIM AND OBJECTIVES

The aim of this study is to provide a systematic literature review (SLR) to examine how AI-driven methods are being applied to detect fraud in digital payment systems in the United Kingdom. This research seeks to assess not only the types and performance of AI models deployed but also the implementation challenges and emerging context-specific solutions relevant to UK stakeholders.

To achieve this aim, the following research objectives have been formulated:

1. Identify and categorise AI-based fraud detection techniques.
2. Evaluate existing solutions or frameworks proposed in the literature that address the identified challenges.
3. Analyse the operational, technical, and regulatory challenges that influence the adoption and deployment of these techniques in the UK.

III. SIGNIFICANCE OF THE STUDY

This research contributes to both theory and practice within the UK's growing digital financial ecosystem. With contactless cards, mobile wallets, and P2P platforms becoming ubiquitous, digital payment fraud losses in the UK reached £1.17 billion in 2024, and therefore is deemed a national security threat [15]. The findings from this study provide a better perspective to address lapses in the fraud detection system in the UK digital payment ecosystem. Theoretically, this systematic literature review fills a research gap by focusing on AI, especially machine learning (ML) and deep learning (DL), in the UK context. Therefore, this study provides conceptual clarity on UK-specific AI methods and challenges.

Practically, these findings provide a guide for UK fintechs, regulators, and cybersecurity professionals in selecting AI tools that are compliant, ethical, explainable, and aligned with evolving regulations, such as the FCA's mandatory APP fraud reimbursement from 7th October, 2024 [16]. Therefore, this paper supports the development of resilient, AI-driven fraud detection systems that uphold digital trust and user experience.

IV. METHODOLOGY

This study utilises a Mixed-Source Systematic Literature Review (MLR-SLR) methodology, combining both peer-reviewed academic research and grey literature (e.g., regulatory reports, white papers, vendor publications) to holistically examine AI-driven fraud detection in UK digital payment systems. While traditional SLRs rely primarily on peer-reviewed sources, applied fields such as fintech and cybersecurity often require grey literature to capture industry innovations and regulatory context [17]. Regulatory authorities like Cochrane acknowledge that legal frameworks in the case of UK regulations and technical deployments may not yet be published in academic outlets, making grey literature essential [18]. To ensure rigour, this research follows the PRISMA 2020 guidance by explicitly defining inclusion criteria and assessing all sources for credibility and relevance [19].

V. LITERATURE SEARCH STRATEGY

This study searches several academic databases, IEEE Xplore, Scopus, ScienceDirect, SpringerLink, and Google Scholar for studies published between 2015 and 2025. Key search strings included combinations such as "AI fraud detection UK", "machine learning digital payment fraud", and "UK Open Banking fraud AI". Grey literature sources were also identified via targeted searches on UK Finance, FCA, CDEI, Google search engine and company blogs (e.g., Visa, Mastercard, Reolut). Only documents with clear authorship, publication date, and traceable provenance were included, ensuring transparency.

A. Inclusion and Exclusion Criteria

To maintain the relevance and quality of the literature review, specific inclusion and exclusion criteria were established. Sources were included if; they focused on machine learning, deep learning, or hybrid AI methods; they pertained to fraud in UK digital payments (including card, mobile, and P2P systems); they were peer-reviewed or credible grey literature published by reputable entities; and were written in English and published between 2015-2025.

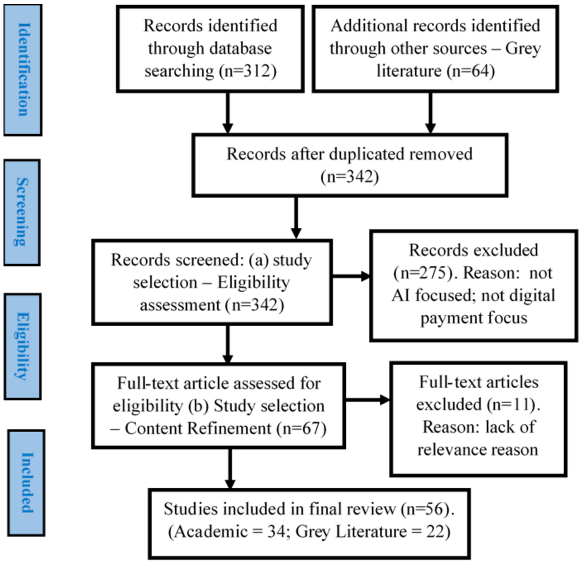
On the other hand, sources were excluded if they were purely rule-based systems, lacked AI application, were duplicates, or held unknown credibility.

B. Selection Process

Following PRISMA principles, this study executed a two-stage screening process. The first stage, screened the Title and Abstract to identify initial relevance. The second stage involved a full-text review to assess the depth of AI engagement and UK focus. Furthermore, backwards and forwards citation tracking (snowballing) was applied to capture relevant works beyond initial searches, as recommended by SLR best practices [20]. Both academic and grey sources were systematically logged using bibliographic tools, with quality criteria applied consistently across source types.

Following the PRISMA framework process of systematic literature review, 312 articles were identified through academic databases and 64 from grey literature. After deduplication, 342 records were screened. Of these, 67 full-text articles were assessed for eligibility, and 56 studies were included in the final review, comprising 34 peer-reviewed academic sources and 22 high-quality grey literature documents relevant to AI-driven fraud detection in the UK. The PRISMA flow diagram is shown below in Figure 1, illustrating the identification, screening, eligibility process and inclusion criteria applied during the literature selection phase.

Figure 1: The PRISMA flow Diagram


Identification
Screening
Eligibility
Included
Records identified through database searching (n=312)
Additional records identified through other sources – Grey literature (n=64)
Records after duplicated removed (n=342)
Records screened: (a) study selection – Eligibility assessment (n=342)
Records excluded (n=275). Reason: not AI focused; not digital payment focus
Full-text article assessed for eligibility (b) Study selection – Content Refinement (n=67)
Full-text articles excluded (n=11). Reason: lack of relevance reason
Studies included in final review (n=56). (Academic = 34; Grey Literature = 22)

VI. APPROACH TO ANALYSIS

A thematic coding approach was used to categorise findings into AI techniques, application domains, regulatory challenges, and governance issues. Both quantitative metrics (accuracy, precision) and qualitative indicators (regulatory adoption, data governance) were synthesised. Comparison across source types highlighted convergence and divergence between academic research and real-world UK implementation. The inclusion of grey literature enhances relevance without compromising methodological integrity, supporting both theoretical and practical contributions.

VII. THEORETICAL BACKGROUND

A. AI Techniques in Financial Fraud Detection

Machine Learning Methods

The conventional machine-learning (ML) algorithms are still identified as the workhorse of fraud detection, as they offer an explainable risk score (fraud score), and perform dependably on tabular transaction information. A probability or risk score is usually assigned to every transaction or account using classical models such as logistic regression and decision trees [21]. Support vector machines (SVM) and more complex classifiers are then been used to deal with non-linear patterns of fraud [22]. Researchers have observed that these ML algorithms play a significant role in the regulatory control and fraud detection in the UK financial industry [23]. These algorithms are in practice trained on historical fraud data to differentiate between fraud and legitimate transactions. To illustrate, A study documents an almost 99% credit-card fraud detection level with the help of a logistic regression, random forest, and linear discriminant analysis [24]. Nonetheless, scholars also point out shared problems such as a class imbalance and data privacy, which are inherent in the datasets of fraud [24].

Ensemble tree methods often further improve performance. In particular, Random Forests and boosting (e.g. XGBoost) are widely deployed [25]. An empirical study found that a random forest achieved about 96.8% accuracy on a credit-card fraud dataset, outstripping logistic regression (95.2%) and a single decision tree (91.1%) [26]. This aligns with industry experiences: for instance, Lloyds Bank reported that Random Forest models gave “very high discrimination” between fraudulent and normal cases, making them hard for fraudsters to game. Fraud-scoring models built on ensembles can therefore detect subtle patterns in transaction features. At the same time, experts caution that complex trees must be carefully managed, Langron (2015) notes ongoing challenges of model validation and overfitting in real deployments [27]. In summary, UK studies and practice indicate that a spectrum of supervised ML methods from logistic regression to ensemble trees, are central to fraud scoring and detection [23], [26], with ensemble methods typically yielding the highest accuracy.

Deep Learning Methods

Deep neural networks have emerged to capture more complex fraud patterns, especially in large or sequential transaction datasets [28]. These models automatically learn hierarchical features and can model temporal behaviours in account activity. For example, Convolutional Neural Networks (CNNs) have been applied to fraud data by treating transaction histories as structured inputs. Recent UK research (Edge Hill University) found that a CNN achieved a remarkable ~99% detection accuracy on credit-card data, outperforming Long Short-Term Memory (LSTM) networks, recurrent networks, multi-layer perceptrons (MLPs) and deep belief nets [24]. This proves that CNNs excel at picking out subtle patterns in high-dimensional financial data. Recurrent neural networks (RNNs) and LSTMs are another class tailored to sequential modelling: by remembering prior activity, they can detect anomalies in spending behaviour over time. An Edinburgh study reported that a deep RNN detected 35% of fraudulent debit-card transactions (vs. only 10% for a legacy rule-based system), demonstrating a clear gain from temporal modelling [29].

In principle, deep architectures are “adept at identifying fraud patterns and anomalies" by analysing large datasets [24]. Autoencoder networks, which learn to compress normal transactions and thus highlight outliers, are also listed among effective techniques [24]. In practice, researchers often pre-process transaction streams into time-series or sequence formats suitable for these networks. For example, one approach uses multi-day transaction sequences as input to LSTM or CNN encoders to flag abnormal spending bursts [30]. A UK survey highlights that both supervised (LSTM, SVM) and unsupervised (autoencoder) deep models have demonstrated strong fraud detection accuracy [24]. Overall, deep learning methods can significantly enhance detection rates by modelling the temporal and complex feature interactions in transaction data. The trade-off is the need for larger datasets and computational power, as well as careful handling of class imbalance and interpretability concerns [24].

Hybrid/Ensemble Models

Given the complementary strengths of different techniques, hybrid and ensemble approaches are increasingly advocated to maximise detection accuracy and robustness. Ensemble learning, combining multiple models, has strong theoretical support in fraud detection [26]. Talukder et al. (2024) noted that it reduces overfitting, enhances model generalisation, and improves overall performance by leveraging the strengths of different models [31]. In practice, this means using bagging such as Random Forests, or boosting such as XGBoost, or stacking frameworks that merge predictions from diverse classifiers. Recent research exemplifies this trend. Btoush et al. (2025) propose a stacked hybrid that integrates multiple ML algorithms (decision trees, SVM, CatBoost, etc.) with deep networks (CNN, Bidirectional LSTM). By consolidating all base-model predictions through a stacking ensemble, the hybrid system significantly outperformed any single model [32]. These experiments on imbalanced card transaction data show that the ML+DL ensemble achieved very high F1 scores (~0.95) while better handling class imbalance [32]. These studies underscore that carefully designed hybrid models, especially those that combine tree-based and neural components, can capture diverse fraud signals and boost predictive power.

Real-world applications echo these findings. Studies show that many UK banks layer rule-based filters for known fraud patterns with AI-based scoring to balance accuracy and transparency. In a recent UK case study, a bank deployed a new ML-driven fraud layer using LightGBM and XGBoost on top of its credit-application system. This hybrid solution led to an estimated 35% reduction in fraud incidents within one year [33]. The added ML layer flagged complex suspicious applications that static rules had missed, while rules still filtered out obvious invalid cases. Such hybrid systems tend to reduce false positives and adapt to emerging fraud tactics. Collectively, both academic and industry experience indicate that combined rule/AI or ML+DL ensembles deliver greater accuracy and resilience than single-method approaches [31], [32].

B. AI Applications in UK Digital Payment Systems

Contactless Card and Chip Fraud

UK banks and card schemes have long used layered analytics to secure contactless and chip-based payments [34]. The major card networks themselves embed advanced machine learning across the transaction flow. Visa reports that it invests heavily in fraud technology helping drive its global card-fraud rate to historic lows, and has now extended its AI tools to account-to-account (A2A) transfers. In a UK pilot with Pay.UK, Visa’s real-time AI identified 54% of fraudulent payments that traditional bank filters missed [35]. As Visa notes, its network successfully identified 54% of the fraudulent transactions that had previously been made, and is now offering Visa Protect for A2A Payments to catch scams before money leaves the account [35]. Similarly, Mastercard leverages AI at scale. Its new generative-AI system scans billions of card transactions to predict compromised account details. Mastercard reports that this technique doubles the detection rate of compromised cards while cutting false positives by up to 200% and flagging at-risk merchants 300% faster than before [36]. In practice, this means that the networks can rapidly block stolen cards or link partial card numbers from malware leaks. Collectively, contactless card payments in the UK are protected both by device-level checks (EMV chip rules, PIN on threshold) and by network AI that scores every transaction for unusual patterns [35], [36]. This two-tier approach, on-device authorisation plus back-end AI scoring, helps keep card fraud on Visa's network among the lowest of all payment forms [35].

Mobile Wallet and App-Based Payments

Digital banks and mobile wallets layer further intelligence on top of traditional card security. UK banks and payment apps use machine learning to flag suspicious activity within their platforms. For example, Monzo, a UK app-based bank reports it has dramatically reduced fraud on mobile top-ups by deploying a TensorFlow-based fraud model. Monzo's team built a classifier analysing many features, links between users and behavioural patterns, to predict risk [37]. The result was a collapse in fraud losses: monthly fraud went from ~0.84% of top-ups to under 0.01% after ML improvements [37]. Equally important, Monzo closely tuned its model to reduce false alarms. Initially, the system flagged six genuine users for every three fraudsters, but by refining the risk rules it eventually reached about one false positive per three actual fraud cases [37]. These use cases illustrate the technical trade-off inherent in app-based fraud detection, and also the power of data-driven monitoring.

Global wallets and payment processors also employ rich data models. PayPal, as used in the UK, explicitly combines myriad signals in its fraud engine. According to PayPal's documentation, its machine-learning systems scrutinise device attributes cookies, fingerprinting, account data email, login patterns, IP address/phone, and transaction history. By analysing device, email, IP, phone, and behavioural user data, PayPal's algorithms rapidly score each login or payment for risk [38].

Beyond transaction fields, many UK providers use device fingerprinting and behavioural biometrics. Device fingerprinting passively profiles the customer's hardware/software, browser type, OS version, graphics card features, installed fonts, etc [39]. These create a unique fingerprint for the smartphone or laptop that is difficult for a fraudster to mimic [40]. Once a user's device is profiled, the system can recognise it on return visits. According to this study, banks compare a device's stored fingerprint to a database of known-good or known-bad devices [40]. If a payment request comes from a device previously used by fraudsters, the bank can flag the transaction or require extra checks. Some vendors even bind devices cryptographically to a user's profile so that only that device can approve transactions, meeting PSD2 strong-authentication rules [40].

Behavioural biometrics add another layer, systems monitor how a person interacts with the app. It checks if the typing rhythm or swipe pattern is consistent with past usage. Although precise UK cases are confidential, industry publications note that firms are increasingly ingesting such data. Notably, Mastercard's recent UK product with partner Feedzai explicitly integrates device intelligence, network data, and behavioural biometrics into its fraud scoring [41]. Feedzai's platform analyses not only what is being done, the payment details but also the timing, location, keystroke rhythm, etc. to unmask imposters. In real deployments, UK banks using these advanced models have reported early success, One coalition-backed report found that introducing device+behavior signals in payment monitoring corresponded with a 12% drop in APP push-payment scam losses in 2023 [41]. In sum, mobile wallets and apps in the UK rely on ML models built from rich behavioural and device data, enabling banks and fintechs to profile each login and payment in real time and halt unusual transactions with minimal customer friction [37], [41].

Peer-to-Peer and Open Banking

The rise of open banking and instant push-payments has made peer-to-peer monitoring a priority. Under PSD2, banks expose APIs and share payment data, which can aid AI-driven surveillance across institutions. In the UK, networks now co-operate on real-time fraud scoring for app-based transfers [42]. Mastercard's Consumer Fraud Risk solution built with Feedzai gives both the sending and receiving banks a real-time risk score on every faster-payment transfer. Since its launch in early 2023, regulators report that the UK's APP scam losses have fallen, in part due to this cross-bank intelligence-sharing, by over 12% [41]. This system ingests device IDs and behavioural cues to spot high-risk payments instantly [41], effectively flagging a potential money mule account or abrupt payment outflow.

UK fintechs are significantly innovating in the AI space [43]. Revolut headquartered in the UK/EU now embeds ML into its payment workflows to intercept scams [44], [45]. In 2024, Revolut launched an AI-scam feature, where every outbound card payment or transfer is scored by a model trained on historical scam patterns. If the AI flags a likely scam, for instance, a sudden large payment to a newly-added recipient. Revolut automatically declines the transaction and routes the customer into an anti-scam flow [45]. In the in-app flow, users answer a few questions to test if they are being manipulated and are shown educational prompts. Revolut reports indicated that in the UK this intervention has already cut actual scam losses by about 30% in initial trials [45]. Crucially, they note, this lets genuine payments including legitimate investments proceed while stopping those made under the fraudster's spell.

The international payment providers exemplify similar AI use at scale. Wise, which was formerly known as TransferWise, a UK-founded cross-border payments firm, employs hundreds of ML models to screen every transaction. Wise's infrastructure runs 150 machine learning algorithms on seven million transactions a day, performing 80 checks per second [46]. These algorithms compare each payment to the user's typical patterns, e.g., amount, currency route, destination, and flag outliers. For example, an unusually large transfer to an unfamiliar country will be marked suspicious, especially if the phone's GPS shows the user is nowhere near home. By fusing this anomaly detection with regulatory watchlists sanctions, and KYC checks, Wise interrupts high-risk transfers before settlement. In practice, these automated models have dramatically improved their efficiency in catching fraud and compliance issues [46]. Many UK banks partner with such providers or build equivalent AI pipelines, so that open-banking and P2P payments are continuously monitored by adaptive, data-driven systems [47].

C. Challenges and Risks in AI Adoption

Technical Challenges

Implementing effective fraud AI in finance presents significant technical hurdles in the UK [13]. One core problem identified in the context of the UK is imbalanced data, where genuine transactions vastly outnumber fraud cases. As the FCA in the UK notes, fraudulent payments often comprise less than 0.2% of a bank's transactions [48]. Training a machine-learning model on such skewed data tends to bias it toward the majority class (legitimate). Financial, firms therefore resort to data augmentation, e.g. synthetically oversampling past frauds, to help the model learn rare patterns [48]. Even so, performance can be tricky. A hypersensitive model catches more fraud but at the cost of many false positives, hence flagging legit activity. For example, early in Monzo's fraud project, the system was flagging roughly six genuine users for every three fraud attempts; extensive feature engineering later improved this to about one false alert per three true frauds [37]. In practice, balancing recall i.e., catching true frauds, against precision, i.e., limiting false alarms is an iterative art. High false-positive rates frustrate customers, while false negatives let fraud slip through [49].

Another challenge noted in the literature is model complexity. A study shows that many modern fraud detectors use ensemble or deep-learning models that are essentially black boxes. As UK regulators caution, the more complex the AI, the harder it is to interpret its decisions and detect when it fails [50]. A neural net might catch subtle transaction patterns, but it offers little insight into why it flagged a transfer as suspicious. This opacity makes it difficult for risk teams to tune the model or to provide any human explanation to customers. It also complicates validation, as banks must constantly monitor models to avoid drift, e.g., new fraud tactics and ensure accuracy [50]. Moreover, inherent bias in training data can skew results, If the historical fraud data over-represents certain customer profiles or regions, the model may unfairly target those groups [21]. Hence, achieving low error rates with machine learning requires continuous retraining, advanced engineering, to handle imbalance, and extensive feature design, which is far from trivial.

Legal and Regulatory Concerns

Beyond engineering, legal and compliance issues pose critical barriers to the implementation of AI technologies in digital payment fraud detection in the UK. In the UK, data-protection laws restrict automated decision-making [51]. Under the UK GDPR, mirroring EU rules, Article 22 prohibits sole reliance on automated systems for decisions with legal or similarly significant effects on individuals [52]. In practice, this means a bank cannot reject a credit or insurance application purely by a black-box AI score without any human review. Similarly, customers have the right to information about automated profiling and a route to human appeal or correction [52]. Although much fraud detection is more advisory than binding, firms still need to ensure their processes are fair and transparent. Many institutions therefore incorporate human oversight analyst review of high-risk flags or clear customer disclosures to comply with these rules [53].

Regulatory sandboxes and pilot programs have helped financial firms test AI fraud tools, but they have limits. The FCA have a new AI Supercharged Sandbox, created in partnership with NVIDIA, aimed at giving firms technical support and secure environments for model testing [54]. However, the FCA emphasises that this support does not exempt firms from existing laws: participants must still follow all regulations, no special legal waivers [54]. In other words, even in a live-test setting, firms must maintain compliance with, PSD2 Strong Customer Authentication rules, anti-money-laundering (AML) standards, and consumer protection laws. PSD2 itself imposes multi-factor authentication and transaction limits on open-banking APIs, which can sometimes conflict with seamless AI operations e.g., a one-time password interrupting a fraud-screening flow [54]. Finally, cross-border data-sharing restrictions (both privacy and banking secrecy rules also pose significant challenges, which can limit pooling intelligence. In summary, while UK regulators encourage innovation and even offer guided sandboxing with the full rigour of GDPR, FCA/PRA rules and SCA requirements still apply to any deployed AI system. Firms must bake in compliance from the start or risk penalties which pose implementation challenges for financial institutions and industry experts.

Ethical and Social Implications

AI in the payments system also raises ethical and trust issues. A foremost concern noted in the literature is algorithmic bias. If the training data reflects historical prejudices or uneven demographics, the model can perpetuate unfair outcomes [55]. In the UK, regulators have warned that biased ML fraud detectors could disproportionately mark customers from certain socioeconomic or ethnic groups as suspicious, simply because those groups were over-represented in past fraud patterns [50]. Such bias may go unnoticed inside a complex model and only come to light after many false blocks. It is especially perilous if vulnerable populations, the elderly, language minorities, etc. are affected, for instance, if their atypical spending habits are mistaken for fraud. UK authorities explicitly highlight fairness and non-discrimination as core AI principles [50], meaning firms must vigilantly audit their models for disparate impacts.

Trust and accountability have also been highlighted in literature crucial aspect of ethical consideration in the implementation of AI. Customers expect financial decisions to be explainable and contestable. As the FCA's recent AI review noted, "consumers should not be left in the loop with an opaque algorithm, they should receive clear explanations and the ability to challenge outcomes" [56]. In practice, this might require showing customers why a login was blocked e.g., unusual device, or providing an easy appeals process. This demand for transparency can clash with proprietary AI models, creating tension between intellectual property and consumer rights. Moreover, unresolved questions of legal accountability arise when, say, a jointly developed AI system misbehaves. It poses the question of whether the bank, the vendor, or the data provider is responsible. UK regulatory guidance on AI underscores accountability and governance, firms hence must take ownership of their AI systems, even if parts are outsourced [50].

Finally, public trust can lag behind new tech. High-profile cases of AI bias or error can undermine confidence in digital payments. The industry knows this, as UK finance bodies and the FCA have stressed that strong safeguards and human oversight are needed to maintain trust [50], [56]. If users fear being wrongly frozen out of their accounts by a black-box system, they may resist the technology. Thus, beyond pure detection performance, payment firms must invest in governance, ethics training, and customer communication. Demonstrating explainability, fairness, and accountability is critical to ensuring that AI fraud solutions bolster rather than erode trust in the financial system [50], [56].

VIII. DISCUSSION

Synthesis of AI Techniques and Effectiveness

Findings from examined research, documentation, and industry development show a variety of AI models achieving very high fraud-detection accuracy. Deep learning approaches, particularly Convolutional Neural Networks (CNNs), often outperform other architectures by capturing complex transaction patterns. For example, one UK study found a CNN achieved ~99% accuracy, substantially beating alternative deep models (LSTM, RNN, MLP, DBN) on a large transaction dataset [24]. Other work highlights that gradient-boosted trees like XGBoost, CatBoost and ensemble methods are also extremely effective. In one benchmark, a two-stage CatBoost/XGBoost model reached ~99.96% accuracy [24]. Traditional algorithms such as Random Forests, SVMs and autoencoders remain competitive as well [24], especially when tuned with oversampling and ensemble techniques. In practice, the choice of model depends on context. CNNs excel when rich, multi-dimensional features are available; RNN/LSTM networks suit sequential transaction data by leveraging time dependencies; and tree-based or linear models offer faster training or easier explainability. Importantly, combining models often boosts performance, for instance, data resampling plus stacking or ensembling has been shown to improve fraud prediction [24].

UK-Specific Operational Barriers

In the UK, strong privacy and regulatory regimes pose significant barriers to AI deployment in fraud detection. Surveys of UK financial firms find that data protection (GDPR-related) and privacy rules are the top regulatory concerns [57]. The FCA's new Consumer Duty and other conduct rules further complicate data use and model deployment [57]. These constraints often limit access to the large, labelled datasets needed to train complex models. Also, the FCA has reported that data availability and quality are key challenges for financial AI, leading to initiatives on synthetic data generation to overcome privacy issues [48]. Beyond regulation, cultural factors also slow adoption. UK firms tend to be risk-averse and demand transparency, so black-box AI models are approached cautiously. Industry survey shows 40% of UK risk managers cited model explainability and governance as the biggest AI barrier [58]. This aligns with the BoE/FCA findings that 81% of UK firms using AI already implement explainability methods, such as feature importance or SHAP to satisfy oversight [57]. Talent and resource limitations further constrain progress, as roughly a quarter of UK financial respondents report insufficient skills or staff as major hurdles to AI adoption [57], [58]. Taken together, UK institutions need to balance innovation against a heavyweight compliance culture.

Global vs UK Comparison

Looking internationally, different useful lessons can also be learnt. In Europe (EU), there are strong data-sharing frameworks and regulations that drive fraud detection. For example, PSD2's Strong Customer Authentication has mandated richer transaction data now ~100+ security features per payment [59], forcing EU issuers to employ ML for real-time analysis. The UK adopted PSD2 pre-Brexit, so UK banks similarly benefit from these extra data signals, though they must apply them under UK regulatory oversight. Meanwhile, in the US, financial firms often leverage massive customer data and invest heavily in proprietary AI systems [60]. Large payment networks like Visa exemplify this global leadership, In a UK pilot, Visa's AI tool analysed billions of transactions and caught 54% of fraud cases missed by banks' existing systems [35]. This shows the value of high-volume data and state-of-the-art ML, which UK institutions can emulate through partnerships e.g., Visa's collaboration with Pay.UK.

In Asia-Pacific, adoption has lagged for institutional reasons. A 2024 study reported that only 15% of APAC firms actively use AI for AML/fraud, hindered by legacy systems and data issues [61]. These firms cite the same obstacles seen in the UK integration with old IT (58.6%), data quality gaps (58.6%), explainability (46.6%) and privacy (43.1%) were top challenges [61]. Notably, the report emphasises open collaboration between FIs, technology providers, and regulators as crucial to building trust in AI [61]. This mirrors UK initiatives, the BoE/FCA regularly engage industry via AI forums and expect firms to appoint accountable AI leads, 84% report having an AI responsible officer [57].

Gaps in Existing Practice

Despite promising models, important gaps remain. A chief concern is explainability. Many high-accuracy AI methods are opaque. As existing studies note, the opacity of AI models, suggests that practitioners have been slow to adapt, creating a need for XAI techniques tailored to fraud detection [59]. While tools like SHAP have been applied in some studies [24], overall there is a paucity of interpretable AI solutions in operational settings. Addressing this gap is critical given regulatory emphasis on transparency as an existing study identifies explainability as a top industry pain point [58].

Another limitation is data scarcity. Publicly available fraud datasets are few, and most are non-UK benchmarks [62]. Experts often have to rely on proprietary or synthetic data, which may not reflect local payment patterns. The FCA has acknowledged data availability as a blocker, hence its work on synthetic data to mitigate privacy risks while enabling model training [48]. In the meantime, ML models may overfit to the limited data they see. Future work should thus focus on creating or sharing large, realistic UK-specific transaction datasets possibly via privacy-enhancing techniques and validating models across diverse data.

Other research gaps include fairness and domain adaptation. Studies rarely address bias in fraud models or how well a model trained on, say, debit-card data generalises to open-banking payments. There is also little exploration of emerging architectures e.g., graph neural networks, federated learning in the UK context. Overall, literature reviews note that many ML/DL approaches still face class imbalance, scalability and overfitting issues [62]. Continued innovation is needed to keep models alert to the latest fraudulent activities [62], such as adapting to new fraud tactics or incorporating streaming analytics.

Strategic Implications

Strategically, firms should not rely on a single model or data source. A layered detection approach, combining rule-based systems, supervised classifiers, and unsupervised anomaly detectors, is widely advisable. Ensemble learning and multi-stage pipelines have been shown to significantly enhance model performance in fraud detection. In practice, this means blending existing expert rules with ML outputs, using ensemble scores or meta-learning to catch diverse fraud patterns. Such hybrid systems can leverage the high accuracy of AI while retaining human-understandable checks, hence, bridging the explainability gap.

RECOMMENDATIONS

Considering the noticeable gap in existing practice. It is strongly recommended that regulators and industry collaborate closely. The UK's Joint Fraud Taskforce exemplifies how public-private partnerships can tackle fraud holistically. Similar to Asia's call for open collaboration [61], UK banks, fintechs and the FCA must share anonymised fraud indicators and model insights (e.g., through federated learning consortia) to improve collective defence. Regulatory bodies have signalled a pro-innovation stance, the Bank and FCA actively support AI adoption under sound governance for instance via pilot sandboxes and AI principles [57]. Firms should therefore leverage these channels, by participating in FCA/BoE surveys and AI forums, to shape sensible oversight. Ultimately, combining complementary detection techniques and a cooperative regulatory regime can maximise fraud prevention. Furthermore, it will also be essential to keep refining models in line with research, pursuing XAI methods, expanding datasets, including synthetic data per FCA guidance [48], and integrating future technologies to stay ahead of increasingly sophisticated fraud.

To strengthen AI-driven fraud detection, UK fintechs need to adopt explainable AI (XAI) to meet regulatory transparency needs and reduce false positives. Collaborations between academia and industry should support shared, anonymised datasets tailored to UK payment contexts. Fintechs are encouraged to pilot AI models in FCA regulatory sandboxes, refining compliance in live environments. Finally, deploying hybrid detection systems that combine human oversight with machine learning can balance interpretability, adaptability, and operational efficiency in fraud prevention.

CONCLUSION

This paper provides a UK-centred systematic review on the way in which artificial intelligence is reshaping the detection of fraud in the digital payment environment, through a synthesis of evidence both in the academic and grey literature, demonstrating the deployment of machine learning, deep learning and hybrid-based models in the digital payment system. The prospects and restrictions of AI implementation rely on a series of unique UK challenges, including data protection laws, explainability requirements, and infrastructure. Although the existing models yield high accuracy rates, they have not been adopted extensively because they have a black-box structure and are based on limited datasets. Future research may focus on security and privacy in digital payment using federated learning, adversarial resilient algorithms, and graph-based models. In the future, a long-term and guided interaction among regulators, fintech developers, and academic researchers will be indispensable in the co-design of AI mechanisms that are safe, transparent, and adaptable to the ever-changing nature of the UK digital payment system.

REFERENCES

[1] A. Kirobo, J. Lissah, and M. M. Govella, “Adoption of Cashless Economy in the World: A Review,” *IOSR J. Econ. Financ.*, vol. 13, no. 2, pp. 37–48, 2022, doi: 10.9790/5933-1302083748.

[2] UK Finance, “One third of UK adults now use mobile contactless payments | Insights | UK Finance,” UK Finance. Accessed: Jul. 10, 2025. [Online]. Available: https://www.ukfinance.org.uk/news-and-insight/press-release/one-third-uk-adults-now-use-mobile-contactless-payments

[3] S. Barber and M. Boyle, “Digital wallet statistics: Usage and market size,” Finder. Accessed: Jul. 10, 2025. [Online]. Available: https://www.finder.com/uk/banking/digital-wallet-statistics

[4] A. K. Singh and K. K. Agarwal, “Overview of Digital Payment Frauds: Causes, Consequences, and Countermeasures,” *J. Informatics Educ. Res.*, vol. 5, no. 1, pp. 2297–2311, 2025, doi: 10.52783/jier.v5i1.2230.

[5] B. Cooper, “UK Finance Annual Fraud Report 2025 - TLT LLP,” TLT. Accessed: Jul. 10, 2025. [Online]. Available: https://www.tlt.com/insights-and-events/insight/uk-finance-annual-fraud-report-2025/

[6] R. Jones, “‘Remote purchase’ fraud in UK surges as customers tricked into disclosing passcodes | Scams | The Guardian,” The Guardian. Accessed: Jul. 10, 2025. [Online]. Available: https://www.theguardian.com/money/2025/may/28/remote-purchase-fraud-uk-surges-customers-tricked-passcodes

[7] Payments Intelligence, “Unveiling digital fraud: Insights into scam trends and prevention in the UK payment sector | The Payments Association,” Payments Intelligence. Accessed: Jul. 10, 2025. [Online]. Available: https://thepaymentsassociation.org/article/unveiling-digital-fraud-insights-into-scam-trends-and-prevention-in-the-uk-payment-sector/

[8] C. Bagwe, “Fraud Detection in Financial Institutions: AI VS. Traditional Methods,” *Int.*

J. Sci. Res. Eng. Trends, vol. 10, no. 6, pp. 3274–3279, 2024, doi: 10.61137/ijsret.vol.10.issue6.654.

[9] O. I. Odufisan, O. V. Abhulimen, and E. O. Ogunti, "Harnessing artificial intelligence and machine learning for fraud detection and prevention in Nigeria," *J. Econ. Criminol.*, vol. 7, p. 100127, Mar. 2025, doi: 10.1016/j.jeconc.2025.100127.

[10] M. N. M. Sunny, K. M. S. Hossain, M. M. Amin, S. N. Sadmani, and M. A.-A. Siddique, "Numerical analysis of multivariate data for fraud detection," *Nanotechnol. Perceptions*, vol. 15, no. 2024, pp. 325–335, 2024, doi: 10.62441/nano-ntp.vi.3486.

[11] S. Anchoori, "Optimizing Real-time Data Pipelines for Financial Fraud Detection: A Systematic Analysis of Performance, Scalability, and Cost Efficiency in Banking Systems," *Int. J. Comput. Eng. Technol.*, vol. 15, no. 6, pp. 878–894, 2024, [Online]. Available: https://doi.org/10.5281/zenodo.14264978

[12] O. Olowu *et al.*, "AI-driven fraud detection in banking: A systematic review of data science approaches to enhancing cybersecurity," *GSC Adv. Res. Rev.*, vol. 21, no. 2, pp. 227–237, Nov. 2024, doi: 10.30574/gscarr.2024.21.2.0418.

[13] PWC, "Impact of Artificial Intelligence on Fraud and Scams: Research in collaboration with Stop Scams UK," *PWC*, no. December, 2023, [Online]. Available: https://www.pwc.co.uk/forensic-services/assets/impact-of-ai-on-fraud-and-scams.pdf

[14] D. Mcloughlin, "Tackling fraud in 2025: Fighting AI fraud with new AI models | The Payments Association," Payment Association. Accessed: Jul. 10, 2025. [Online]. Available: https://thepaymentsassociation.org/article/tackling-fraud-in-2025-fighting-ai-fraud-with-new-ai-models/

[15] UK Finance, "Fraud continues to pose a major threat with over £1 billion stolen in 2024 | Insights | UK Finance." Accessed: Jul. 10, 2025. [Online]. Available: https://www.ukfinance.org.uk/news-and-insight/press-release/fraud-report-2025-press-release?

[16] Financial Conduct Authority, "ACTION REQUIRED: FCA EXPECTATIONS ON AUTHORISED PUSH PAYMENTS (APP) FRAUD REIMBURSEMENT," Financial Conduct Authority. Accessed: Jul. 07, 2025. [Online]. Available: https://www.fca.org.uk/publication/correspondence/dear-ceo-letter-expectations-app-fraud-reimbursement-banks-building-societies.pdf?utm_source=chatgpt.com

[17] V. Garousi, M. Felderer, and M. V. Mäntylä, "Guidelines for including grey literature and conducting multivocal literature reviews in software engineering," Sep. 2018, [Online]. Available: http://arxiv.org/abs/1707.02553

[18] R. J. Adams, P. Smart, and A. S. Huff, "Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies," *Int. J. Manag. Rev.*, vol. 19, no. 4, pp. 432–454, Oct. 2017, doi: 10.1111/ijmr.12102.

[19] M. J. Page *et al.*, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, p. n71, Mar. 2021, doi: 10.1136/bmj.n71.

[20] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," in *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, New York, NY, USA: ACM, May 2014, pp. 1–10. doi: 10.1145/2601248.2601268.

[21] V. Chang, L. M. T. Doan, A. Di Stefano, Z. Sun, and G. Fortino, "Digital payment fraud detection methods in digital ages and Industry 4.0," *Comput. Electr. Eng.*, vol. 100, pp. 1–31, 2022, doi: 10.1016/j.compeleceng.2022.107734.

[22] A. Ali *et al.*, "Financial Fraud Detection Based on Machine Learning: A Systematic Literature Review," *Appl. Sci.*, vol. 12, no. 19, p. 9637, Sep. 2022, doi: 10.3390/app12199637.

[23] B. G. Buchanan and D. Wright, "The impact of machine learning on UK financial services," *Oxford Rev. Econ. Policy*, vol. 37, no. 3, pp.

537–563, Sep. 2021, doi: 10.1093/oxrep/grab016.

[24] C. F. Onyeoma, H. Rafiq, D. Jeremiah, V. T. Ta, and M. Usman, "2024 International Conference on Frontiers of Information Technology, FIT 2024," *2024 Int. Conf. Front. Inf. Technol. FIT 2024*, 2024.

[25] M. Imani, A. Beikmohammadi, and H. R. Arabnia, "Comprehensive Analysis of Random Forest and XGBoost Performance with SMOTE, ADASYN, and GNUS Under Varying Imbalance Levels," *Technologies*, vol. 13, no. 3, p. 88, Feb. 2025, doi: 10.3390/technologies13030088.

[26] A. R. Khalid, N. Owoh, O. Uthmani, M. Ashawa, J. Osamor, and J. Adejoh, "Enhancing Credit Card Fraud Detection: An Ensemble Machine Learning Approach," *Big Data Cogn. Comput.*, vol. 8, no. 1, p. 6, Jan. 2024, doi: 10.3390/bdcc8010006.

[27] A. Langron, "A survey of Random Forest Usage for Fraud Detection at LLoyds Banking Group," LLOYDS BANK. Accessed: Jul. 06, 2025. [Online]. Available: https://cer.business-school.ed.ac.uk/wp-content/uploads/sites/55/2017/02/A-Survey-of-Random-Forest-Usage-for-Fraud-Detection-at-Lloyds-Banking-Group-Adam-Langron.pdf

[28] G. Zioviris, K. Kolomvatsos, and G. Stamoulis, "An intelligent sequential fraud detection model based on deep learning," *J. Supercomput.*, vol. 80, no. 10, pp. 14824–14847, Jul. 2024, doi: 10.1007/s11227-024-06030-y.

[29] A. Martini, "Deep Recurrent Neural Networks for Fraud Detection on Debit Card Transactions," Barclays. Accessed: Jul. 05, 2025. [Online]. Available: https://www.crc.business-school.ed.ac.uk/sites/crc/files/2020-10/E29-Deep-Recurrent-Neural-Networks-Martini.pdf

[30] I. Benchaji, S. Douzi, B. El Ouahidi, and J. Jaafari, "Enhanced credit card fraud detection based on attention mechanism and LSTM deep model," *J. Big Data*, vol. 8, no. 1, p. 151, Dec. 2021, doi: 10.1186/s40537-021-00541-8.

[31] M. A. Talukder, R. Hossen, M. A. Uddin, M. N. Uddin, and U. K. Acharjee, "Securing Transactions: A Hybrid Dependable Ensemble Machine Learning Model using IHT-LR and Grid Search," Feb. 2024, [Online]. Available: http://arxiv.org/abs/2402.14389

[32] E. Btoush, X. Zhou, R. Gururajan, K. C. Chan, and O. Alsodi, "Achieving Excellence in Cyber Fraud Detection: A Hybrid ML+DL Ensemble Approach for Credit Cards," *Appl. Sci.*, vol. 15, no. 3, p. 1081, Jan. 2025, doi: 10.3390/app15031081.

[33] Accedia, "Case Study: How a Leading UK Bank Leveraged AI Agents for Fraud Detection? - Accedia," Accedia. Accessed: Jul. 10, 2025. [Online]. Available: https://accedia.com/insights/case-study/how-a-leading-uk-bank-leveraged-ai-agents-for-fraud-detection

[34] O. Al-Maliki, "Analysing and Improving the Security of Contactless Payment Cards," University of Buckingham, 2020.

[35] Visa Navigate, "Visa's new AI tool could save the UK over £330m a year on fraud and APP scams | Visa Navigate," VISA NAVIGATE . Accessed: Jul. 11, 2025. [Online]. Available: https://navigate.visa.com/europe/security/visas-new-ai-tool/

[36] Master Card, "Mastercard accelerates card fraud detection with generative AI technology," Master Card . Accessed: Jul. 11, 2025. [Online]. Available: https://www.mastercard.com/us/en/news-and-trends/press/2024/may/mastercard-accelerates-card-fraud-detection-with-generative-ai-technology.html

[37] Monzo, "Fighting Fraud with Machine Learning," Monzo . Accessed: Jul. 11, 2025. [Online]. Available: https://monzo.com/blog/2017/02/03/fighting-fraud-with-machine-learning

[38] Paypal, "Machine Learning Fraud Detection Technologies | PayPal US," PayPal . Accessed: Jul. 11, 2025. [Online]. Available: https://www.paypal.com/us/brc/article/payment-fraud-detection-machine-learning

[39] T. Eglitis, R. Guest, and F. Deravi, "Data behind mobile behavioural biometrics – a survey," *IET Biometrics*, vol. 9, no. 6, pp. 224–

237, Nov. 2020, doi: 10.1049/iet-bmt.2018.5174.

[40] Callsign, "Detecting fraud with device fingerprinting - Callsign," Callsign. Accessed: Jul. 11, 2025. [Online]. Available: https://www.callsign.com/knowledge-insights/preventing-fraud-with-device-fingerprinting

[41] L.-H. Liang, "Mastercard and Feedzai partner in fraud prevention | Biometric Update," Biometric Update. Accessed: Jul. 11, 2025. [Online]. Available: https://www.biometricupdate.com/202502/mastercard-and-feedzai-partner-in-fraud-prevention

[42] M. Gounari, G. Stergiopoulos, K. Pipyros, and D. Gritzalis, "Harmonizing open banking in the European Union: an analysis of PSD2 compliance and interrelation with cybersecurity frameworks and standards," *Int. Cybersecurity Law Rev.*, vol. 5, no. 1, pp. 79–120, Mar. 2024, doi: 10.1365/s43439-023-00108-8.

[43] R. Jarvis and H. Han, "FinTech Innovation : Review and Future Research Direction," *Int. J. Banking, Financ. Insur. Technol.*, vol. 1, no. 1, pp. 79–102, 2021.

[44] AIX, "Case Study: Revolut's AI Revolution - AIX | AI Expert Network," AI Expert Network. Accessed: Jul. 15, 2025. [Online]. Available: https://aiexpert.network/case-study-revoluts-ai-revolution/

[45] Revolut, "Revolut launches AI feature to protect customers from card scams and break the scammers 'spell' | Revolut Netherlands," Revolut. Accessed: Jul. 11, 2025. [Online]. Available: https://www.revolut.com/en-NL/news/revolut_launches_ai_feature_to_protect_customers_from_card_scams_and_break_the_scammers_spell/

[46] Wise, "Delivering a secure, frictionless global payments experience with AI - Wise," Wise. Accessed: Jul. 11, 2025. [Online]. Available: https://wise.com/gb/blog/ai-for-global-payments

[47] A. Wolska, "Bridging the Gap: The Impact of Open Banking on Traditional Banking and FinTech Collaboration," *FinTech AI Financ.*, 2025, doi: 10.7190/fintaf.v2i1.414.

[48] Financial Conduct Authority, "Report on using synthetic data in financial services," *Financ. Conduct Auth.*, no. March, pp. 1–35, 2024, [Online]. Available: https://www.fca.org.uk/publication/corporate/report-using-synthetic-data-in-financial-services.pdf

[49] F. Tolulope, "Real-time fraud detection with reinforcement learning: An adaptive approach," *Int. J. Sci. Res. Arch.*, vol. 6, no. 2, pp. 126–136, Aug. 2022, doi: 10.30574/ijsra.2022.6.2.0068.

[50] N. Kerr-Shaw and W. Adams, "UK Regulators Publish Approaches to AI Regulation in Financial Services | Insights | Skadden, Arps, Slate, Meagher & Flom LLP," Skadden. Accessed: Jul. 13, 2025. [Online]. Available: https://www.skadden.com/insights/publications/2024/05/uk-regulators-publish-approaches-to-ai

[51] J. Krook, P. Winter, J. Downer, and J. Blockx, "A systematic literature review of artificial intelligence (AI) transparency laws in the European Union (EU) and United Kingdom (UK): a socio-legal approach to AI transparency governance," *AI Ethics*, vol. 5, no. 4, pp. 4069–4090, Aug. 2025, doi: 10.1007/s43681-025-00674-z.

[52] ICO, "Rights related to automated decision making including profiling | ICO," Information Commisioner Office. Accessed: Jul. 13, 2025. [Online]. Available: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/individual-rights/individual-rights/rights-related-to-automated-decision-making-including-profiling/

[53] J. Laux, "Institutionalised distrust and human oversight of artificial intelligence: towards a democratic design of AI governance under the European Union AI Act," *AI Soc.*, vol. 39, no. 6, pp. 2853–2866, Dec. 2024, doi: 10.1007/s00146-023-01777-z.

[54] L. Scanlon, "FCA AI 'sandboxing' strategy a positive step for financial services," Pinsent Masons. Accessed: Jul. 13, 2025. [Online]. Available: https://www.pinsentmasons.com/out-law/news/fca-ai-sandboxing-strategy-financial-services

[55] G. Iddenden, "Algorithmic gatekeepers: The hidden bias in AI payments | The Payments

Association," The Payment Association . Accessed: Jul. 15, 2025. [Online]. Available: https://thepaymentsassociation.org/article/algorithmic-gatekeepers-the-hidden-bias-in-ai-payments/

[56] Financial Conduct Authority, "AI Sprint summary | FCA," Financial Conduct Authority. Accessed: Jul. 13, 2025. [Online]. Available: https://www.fca.org.uk/publications/corporate-documents/ai-sprint-summary

[57] Bank of England, "Artificial intelligence in UK financial services - 2024 | Bank of England," Bank of England. Accessed: Jul. 13, 2025. [Online]. Available: https://www.bankofengland.co.uk/report/2024/artificial-intelligence-in-uk-financial-services-2024

[58] UK Finance, "New study: the state of AI risk adoption | Insights | UK Finance," UK Finance. Accessed: Jul. 13, 2025. [Online]. Available: https://www.ukfinance.org.uk/news-and-insight/blog/new-study-state-ai-risk-adoption

[59] E. Mill, W. Garn, N. Ryman-Tubb, and C. Turner, "Opportunities in Real Time Fraud Detection: An Explainable Artificial Intelligence (XAI) Research Agenda," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 5, pp. 1172–1186, 2023, doi: 10.14569/IJACSA.2023.01405121.

[60] Saritha Rai, "Fidelity's Vast Trove of Data Coveted by Tech Firms in Age of AI," Wealth Management . Accessed: Jul. 15, 2025. [Online]. Available: https://www.wealthmanagement.com/financial-technology/fidelity-s-vast-trove-of-data-coveted-by-tech-firms-in-age-of-ai

[61] SymphonyAI, "AI Adoption Lag Leaves Asian Financial Institutions Vulnerable Amid Rising Financial Crime - SymphonyAI," SymphonyAI . Accessed: Jul. 13, 2025. [Online]. Available: https://www.symphonyai.com/news/financial-services/ai-adoption-asian-financial-crime/

[62] I. Y. Hafez, A. Y. Hafez, A. Saleh, A. A. Abd El-Mageed, and A. A. Abohany, "A systematic review of AI-enhanced techniques in credit card fraud detection," *J. Big Data*, vol. 12, no. 1, p. 6, Jan. 2025, doi: 10.1186/s40537-024-01048-8.