Deep Science Publishing
https://doi.org/10.70593/978-93-49910-16-4_8

DeepScience
Open Access Books

Chapter 8: Artificial intelligence-driven fraud detection systems and the future of security in financial networks

8.1. Introduction

In the digitized era of rapid technological advancements and very high-speed interactions across digital platforms, financial fraud schemes have also evolved from traditional applications to online webs. Financial fraud poses a severe risk to investors and financial institutions and is among the most common fraud types. Fraud detection can be regarded as a classification task that requires samples of fraud and non-fraud as training data. In practice, however, fraud data is sparse and hard to identify, as fraud activities are increasingly hidden. Hence, non-fraud samples may contain unrecognized fraud samples, leading to natural errors among training samples. The time-sensitive nature of fraud detection lashed action post-boom of fraud activities, making the addition of the sampled transaction history of merchants rapidly posted on the network more critical than ever (Ajayi et al., 2024; Ismaeil, 2024; Islam & Rahman, 2025).

As the framework of fraud activities is getting more complex, fraud detection benefits from a broader investigation perspective beyond transactions is profoundly examined. Financial activities are concerned with a wider range of business, leading to massive but heterogeneous information, and are accompanied by lower-value density. In addition, fraud typically occurs in a sophisticated new mode consisting of various financial activities, making it more challenging and complicated to detect. Meanwhile, the involved information is scattered across different institutions or operators and can be difficult to integrate if not well managed. Therefore, the usage of knowledge graphs as a suitable and effective data structure is constructed to store and analyze massive data. Due to the few-shot nature of fraud activities, it highlights the construction of an involved knowledge graph containing essential business deputies to address knowledge missing for the minority or new type of fraud activities.

In the era of digital banking, ensuring the security and integrity of financial activities is paramount for banking institutions. Financial frauds, specifically frauds committed via online banking and credit card transactions, pose serious threats to an institution's customer base and reputation, along with their associated financial losses. Financial institutions incur losses into billions annually due to fraudulent activities conducted over the internet, posing the need for robust fraud detection mechanisms. As incidences of financial fraud continue to grow, leading financial institutions continue to undertake a wide berth of research and experimentation to combat financial fraud and identify their occurrence. Bank account fraud differs from other types of financial deception in its methods, impacts, and perceptual patterns (Varga, 2024; Yuhertiana & Amin, 2024).

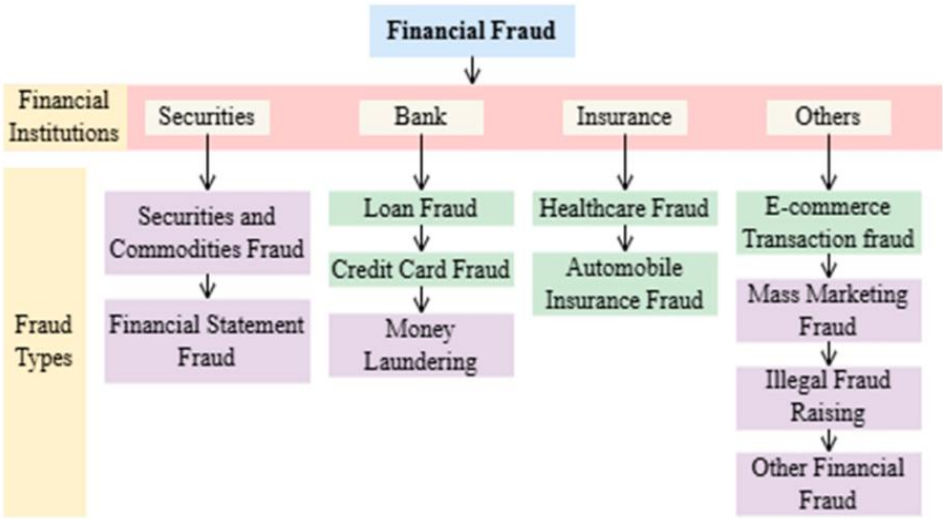

Fig 8.1: AI-Driven Financial Fraud Detection

8.1.1. Research design

This research aims to explore federated learning with explainable AI methods for fraud detection in Real-Time Online Banking networks. The motivation is to enhance current transaction fraud detection mechanisms for web-banking with technologies that protect the privacy of customer data while improving the results over existing technologies. In academic literature, federated learning and explainable AI for transaction FD have separately been described and analyzed, but here the aim is to advance work that combines them. Thus, while theory on each area is analyzed, the focus is on developing an overarching architecture and algorithm where customer data never leaves the banks, results are communicated out to the banks to ensure privacy and security for customers, and explainable AI is provided to bank transactions.

The proposed architecture is composed of three subsystems. The Transaction Fraud Detection System (TFDS) is located at the federated server and combines the models built by the banks to result in a federated model that can analyze the banks' transactions. For this purpose, a novel framework is proposed that uses random select batch training based on old transaction data to ensure circumvention of the bank's private customer transaction data, as well as the models would otherwise not capture diverse fraud activities. The second subsystem is the Federated Learning System (FLS), the main aim of which is to protect and secure the privacy and integrity of customer data. The FLS is described in detail regarding ecosystems such as the communication model, secure aggregation mechanism, encryption protocols, and differential privacy technologies. The explainable AI methods employed in the TFDS are the third subsystem. They provide the banks an understanding of the features used by the AI to support its results. Several XAI methods are described and analyzed to select the most effective techniques for this architecture.

Following the architecture's analysis, the methodology employed in the architecture design and the system development process is described. The architecture and the methods used in the architecture are inspected and the implementation process is documented. The conclusion evaluates the work's contribution to the field and outlines avenues for future research. Banks provide their customers an array of services that are continuously being digitized to improve customer convenience and access to a bank's services. However, Internet banking has an inherent risk of fraud with new transactional methods injecting new ways for criminals to deceive banks. The increasing speeds of online funds transfers and purchases crack down on banks' capabilities to analyze transactions and detect suspicious patterns. Current methods used to detect fraudulent transfers can give false positives for unknown standards and need to be retrained every time a new strategy is implemented.

8.2. Understanding Fraud in Financial Networks

In finance, fraud refers to the acts conducted by fraudsters in order to deceive others for purposes of unauthorized benefit. Generally, fraud can be classified into three categories: financial fraud, non-financial fraud, and technical fraud. Financial fraud refers to any false financial actions or transactions and/or any unfair activities regarding coins or tokens in the crypto ecosystem. As a rapid-growing financial approach to manage assets and raise money, decentralized finance (DeFi) aims to remove intermediaries in traditional finance protocols and provides users with various financial services leveraging blockchain technology. However, without a central authority, the DeFi space is at the whim of malicious actors. Though there have been attempts to employ machine learning and data mining approaches to tackle DeFi fraud, there is still a lack of

comprehensive overview that describes the fraud landscape and important attributes specific to the DeFi ecosystem.

The emergence and rapid development of the blockchain network and cryptocurrency technology has spawned a new financial landscape under the name of smart contracts and decentralized finance (DeFi). DeFi preserves traditional financial services like stocks trading, token exchange and benchmarks, which have evolved into various forms, descriptions and implementations, e.g., DEX, lending platforms, synthetic derivatives and more. By removing intermediaries, DeFi allows a more open and inclusive secondary market for retail traders. However, for the same reason of lacking central authorities and regulators, DeFi is vulnerable to malicious actors. Greedy actors exploit contracts' vulnerabilities in the early days of the ecosystem to gain a large supply of coins/tokens, or to create fake non-existent projects to deceive investors' assets into the smart contracts where they are irrevocably minted. With its fast-growing adoption, there have been devastating attacks on different protocols causing the loss of multi-billion dollars. Fraud, which is defined as false actions conducted by deceivers for gaining unauthorized benefits, can be classified into three categories: financial fraud, non-financial fraud and technical fraud.

8.2.1. Types of Financial Fraud

In recent years, concern about financial crime has increased dramatically, prompted on the one hand by the unprecedented levels of financial complexity and reconfigurations of financial networks and on the other by the amplified scrutiny of financial crime in the wake of the horrific consequences of financial crises and scandals. Financial fraud is defined as the intentional act of deception involving the misrepresentation of material facts or the concealment of the truth in order to gain unfair financial or personal advantage or to induce another party to act to its disadvantage or risk. The multitude of fraud types and their varied ways of detection will be presented using as an example the fifteen most pertinent fraud types identified from a recent literature review. For each fraud type, five properties related to i) how detection is performed, ii) its data requirements, iii) what features are usually exploited in the process, iv) what data formats are feasible, and v) the algorithms used, are specified. Recommendations for the selection, implementation and tuning of fraud detection systems are made, facilitating the selection of the most appropriate existing and future methods for fraud detection applications.

As networks of financial transactions are reconfigured online, more financial crime occurs in the domain of data and systems. Therefore, in order to combat and prosecute financial criminals, it is crucial to be able to comprehend as well as detect their mode of operations and the data input to their modus operandi. When this is done, it is easier to

develop systems that can automatically avert, halt, flag for forensic investigation or prosecute financial crimes and criminals. Since costs and risks increase with a crime's impact, it is deemed of relevance to study fraud types considered most pertinent by experts. To this end, insights into expert opinion will be derived and the fraud types are specified in detail, along with their processing requirements. Additionally, a heuristic evaluation of performance is provided. To facilitate the open and constructive integration of scientific and public efforts to combat financial crime, metadata about at-present-available and future fraud detection systems (and their implementation and tuning) is collected. By presenting the fraud types, this knowledge repository is deemed beneficial for their detection from an applied perspective.

8.2.2. Impact of Fraud on Financial Institutions

Financial fraud has become a common phenomenon at domestic and international levels, harming not only the public but also financial institutions. For banks, financial fraud can lead to huge loss of direct cash flow, as well as loss of customers, reputations, brand names and share prices. Meanwhile, huge financial losses may lead to tighter regulations imposed by authority and even loss of licenses. Thus, how to accurately detect financial fraud is critical for financial institutions and is of significant importance to academic and practical fields. Compared with other non-financial fraud domains, the characteristics of financial frauds differ. A multi-level detection paradigm is necessary for comprehensively detecting financial fraud based on heterogeneous and massive data. Capital market, financial report and transaction frauds are investigated to propose holistic solutions for comprehensive detection frameworks. Artificial intelligence (AI) refers to machines behaving intelligently, typically as: behaving like humans; behaving rationally; taking insightful actions. While "intelligent" computer programs were developed, the "intelligence" remained limited compared to humans. While AI is still in its infancy stage, algorithms, data and computing power continue to rapidly advance AI's capabilities. AI, particularly the data-driven approaches, has achieved excellent performance in the financial fraud detection domain. However, key issues remain unsolved as financial fraud schemes are rapidly evolving. One of the severe difficulties for financial fraud detection is that the fraud is hidden in complex financial activities. Financial fraud is harder to identify due to its increasing secretiveness and complexity. The increased motives and the accelerated digital transformation caused by the pandemic lead to more intelligent fraud schemes, making fraud more difficult to identify. Machine learning based detection models are usually data-driven, assuming that frauds occur as in the past. However, fraudsters have access to more computing power and employ more intelligent algorithms to better identify and exploit loopholes in current detection models, significantly harming banks while hard to be detected.

8.3. The Role of AI in Fraud Detection

As regulatory watchers discuss the future of regulations in finance and governments announce the roll-out of digital currencies, the past is a guide to what tomorrow may hold and what sorts of economic issues may arise. Buckle up for a deep dive into a world of finance where users can create assets subject to few rules, identities remain anonymous, and transactions are validated without the involvement of brokers or governments. Welcome to the "Wild West of finance": decentralized finance (DeFi) where open-source smart contracts on public blockchains handle around 200 billion US dollars of assets with transaction volumes of around 100 billion US dollars a month, and with more than 10 000 yield pools in dozens of different protocols. Bitcoin and blockchains were invented to operate and transact currency without a centralized party or a government. These are open, permissionless, publicly viewable databases, a ledger, accessible from anywhere not controlled by a single party. But security is reliant on the impossibility of tampering with a distributed ledger when its one copy is held by thousands of computers. A successful hack affects the entire system, and identification of the hacker is a public, offline process, leaving blame on either the code's programmers or the software's users or botched convergence. While the absence of intermediaries and control reduces costs and intermediaries, it increases the costs of digging deeper.

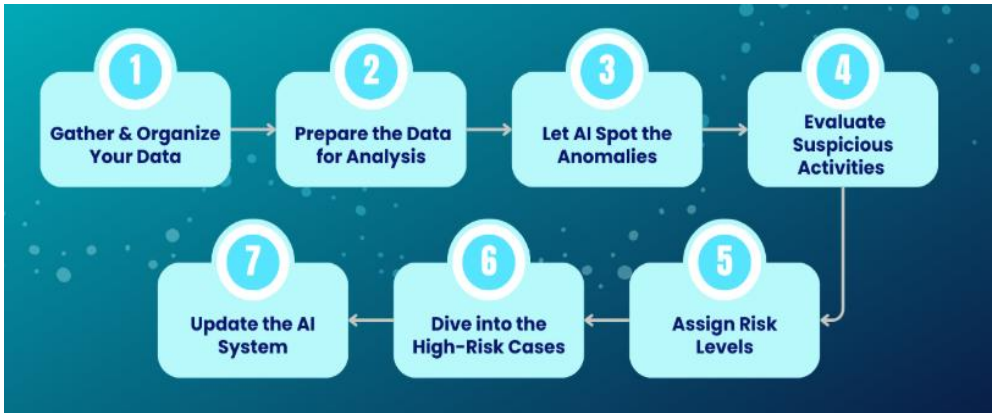

Fig 8.2: AI-Driven Fraud Detection Reshaping Financial Security

8.3.1. Machine Learning Algorithms

Automated payment transaction processing is one of the functions carried out by the financial industry. It consists of buying tickets from the web, paying field bills for

customers, etc. Credit, debit and other types of cards facilitate it. If these cards are lost/stolen, fraudulent access is obtained. The current traditional systems use around 400 different methods to validate a transaction. In view of a hierarchy & excellent will be burdened & biased. The algorithms, which are used to verify transactions known as rule based, require adding more scenarios physically & can barely detect uncorrelated relationships.

Machine learning can create algorithms, which can process big data sets with different variables to predict/identify the correlations between user behavior and fraudulent actions. The objects of the financial institutions (banks) are to protect their financial security and prevent the loss to banks and the customers. Fraud detection is a part of financial security. Major financial institutions are already using machine learning technology to tackle fraud. For example, MasterCard is using it to protect itself and also its customers. It has combined AI and machine learning to track variables like time, transaction size, location, and purchase browse data. Algorithms are trained on pushing vast amounts of variables. The industry's objective is to reduce the number of incorrect declines at merchant payments on behalf of the financial institution. Clients and banks both suffer as a result. False declines made the loss of around $118 billion per year to the merchants, and the client’s loss is around $9 billion per year.

8.3.2. Natural Language Processing in Fraud Detection

Leveraging Financial Textual Data for Fraud Detection In a world of large financial networks where abundant textual data is available, consequently identifying fraudulent events falls as a great challenge for financial monitoring systems. Financial journals contain continuous reporting on financial frauds with diverse natures, including fraud, embezzlement, accounting discrepancy, market manipulation, insider trading, and Ponzi scheme to name a few. Providing defensive measures at financial institutions has thus become a task for domain experts, including compliance officers, legal and forensic auditors, educated reliance on either the review supports of voice-assisted systems or the pure-text outputs from AI text reactors. The information contained in the textual data in newspapers and journals can further better facilitate the understanding of related fraudulent incidents either by professional analysts or general readers. As a consequence of the unstructured nature, however, effectively extracting key features from journalism texts for analysis has become an open challenge [6]. The National Institute of Standards and Technology (NIST) formed a consortium in 2017 to improve the fraud detection, reporting, and investigation ecosystem. That initiative led to an annual public challenge to academia focusing on a series of tasks including adversarial voice investigation, structured data fraudulent behavior detection, and textual data fraud prevention. This presentation thus focuses on the task of detecting fraudulent articles published in a

financial journal. Specifically, textual features are extracted from the inquiry-response pairs of dense articles. A low-complexity spatio-temporal deep learning model called Semantics-averaged Dilated Convolution Layer (SDCL) is proposed to perform multi-class fraud detection. The collaborative discounting learning process is used to capture the inherent spatial-temporal correlations with low complexity [7]. In extensive experiments, the proposed SDCL model is evaluated and compared with several state-of-the-art methods using a newly constructed real-world dataset of financial journals, demonstrating its potential as an effective and insightful tool for detecting fraudulent articles.

8.4. Current AI-Driven Fraud Detection Systems

Fraud detection is an essential task in decentralized finance (DeFi) for protecting users from malicious activities and promoting the sustainable development of the market. Recent studies have examined the application of AI approaches to DeFi fraud detection. However, there are still immense opportunities for future research. The fraud detection in DeFi is a highly explored but under-studied topic. The diversity and expansion of DeFi protocols result in a high variety of fraud types, which requires more comprehensive and systematic exploration of the DeFi fraud landscape. More efforts are also needed to enhance AI techniques in detecting DeFi fraud. Despite AI growing as a new data-driven approach for DeFi fraud detection, existing studies mostly fell short in comprehensively adopting advanced AI techniques. One possible future direction is to leverage the use of pre-trained models for fraud detection in DeFi, as pre-trained models have been applied in a wide range of NLP tasks, especially for enhancing generalizability with less domain-specific training data. Transfer learning is also a potential research direction. Scholars can apply transfer learning in DeFi to borrow knowledge from other related fields. In DeFi fraud detection, the researcher community could consider transferring models developed in traditional finance fraud detection, as there is a lack of value transfer mechanisms in the off-chain world. Apart from the potential of transformer-based pre-trained models, large language models can also help advance fraud detection in DeFi. To date, large language models are showing great potential in various human tasks. More importantly, recent studies have developed finance large language models that can be applied in DeFi research too. Generative agents are another interesting possibility. Using generative agents in collaboration with AI tools can explore the project parameters in an agent mode. In each step, the generative agents can call various tools for evaluating whether the project is potentially suspicious, which can bring new insights to project-based fraud detection in DeFi.

8.4.1. Overview of Leading Solutions

This section introduces AI-based systems for fraud detection and prevention in financial services. The adoption of digital banking has brought a wide spectrum of benefits. However, in the view of digitization, ensuring security and integrity has become paramount. Financial frauds are perceived as a great threat in the form of online banking and credit card frauds. 1.3 billion of losses had occurred in the year 2021 alone due to fraudulent activities of different types including online banking and credit card transactions from which account takeovers were the costliest with 515 million. Banking frauds are also considered criminal offenses due to loss of personal data and violation of privacy. Financial institutions have been undertaking an extensive range of rigorous research and analysis in combating and identifying fraud, with continuous improvements to knowledge discovery in databases (KDDs) and automated machine processing in knowledge engineering. A prevalent and an extremely ripe domain of research considered in this paper is analysis of bank-related fraud. Bank account fraud differs significantly from other deceptions in finance with respect to the method of deception, the impacts of perpetrating or becoming a victim and the detection and prevention. A bank account fraud occurs when a cheating person employed a stolen identity and statement to obtain a bank account successfully without an intention of recovering the login credential for withdrawal of money. This method of deception is deemed successful only if the genuine customer loses their account and transaction actions are reproducible by the fraudster. This action results in multiple claims from the victims to the bank or financial institution from where the fraud account had been created using their information.

Detecting and mitigating this form of fraud is severely difficult and requires a thorough research effort starting from abundant and assorted datasets of account actions to the modeling of smart analytical and preventive tools. With the rise of various operations in finance, machine Learning (ML) techniques have been extensively adopted by leading banks, financial institutions and applied scientists/investors for atypicality detection and processing of banking fraud. Data sets of bank account transactions comprise a string of time-stamped behavioral moments which are in most cases employed to devise a robust transaction rejection system. In these approaches, rich data and sufficient representative feature vectors need to train a ML model to achieve the utmost performance possible.

8.4.2. Case Studies of Successful Implementations

Numerous organizations have successfully adopted AI-driven fraud detection systems and demonstrated effectiveness in fraud detection. They built a predictive model that could alert organizations in seconds before cases escalated. They used hidden Markov models to identify loyalty program account takeover and statistical analysis to identify originators' suspicious behavior while reducing false positives. As a result, they had a

35% success rate in detecting fraud cases when a customer complaint was raised compared to before implementation. Furthermore, this approach made the service relatively easy for operators and low-cost in manpower and training as operations would demand just configuring the threshold values.

A public housing online auction system that built and managed a question-answer webpage was hacked and used by scammers. The company employed machine learning technologies on the data structure of message flows while constructing a robust and reliable anti-fraud system. Their system embedded large-scale parallel processing engines, which included statistical modeling to extract suspicious users, hidden Markov models to detect collusion warm-ups and bundle links, and clustering models to identify colluding groups through message flows. Feedback from operations personnel and quality assurance was obtained later to improve the results. Their fraud detection system was able to leverage spinning independent servers to satisfy performance requests as they had long-term privacy-preserving requirements of mining. Their model was evaluated through random cases from operations personnel, showing that the model successfully found several substantial underground chain links and key scammers who collected millions of Yuan.

8.5. Challenges in AI-Driven Fraud Detection

Detecting fraudulent transactions in banks has been an almighty challenge, owing to the dynamic and sophisticated nature of fraudsters. Machine learning (ML) is capable of screening multiple risk variables in a large volume of data and is able to identify complex

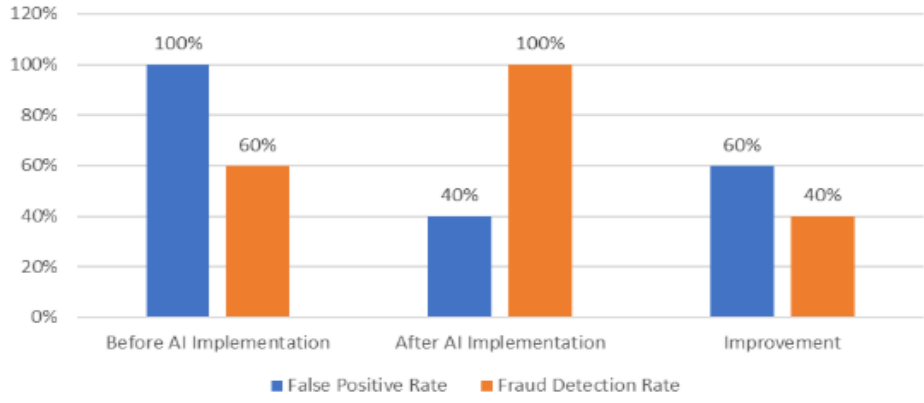

Fig : AI-Driven Fraud Detection System

non-linear relationships. However, the data sets of bank account transactions are often faced with a privacy breach due to the confidential data they hold. The mismatch between the quantity of fraudulent transactions and the legitimate ones leads to an

imbalance, which is another critical challenge in devising a robust fraud detection system. The banks, losing customers or revenues owing to frauds, are always looking for better methods of detection. In the financial domain, different banks often have different fraudulent patterns. Therefore, the current robust methods of fraud detection are unable to spot the new types of fraud taking place in a bank that is uneven to the other banks, and it can further result in severe financial chaos in the network. A single central bank can detect and spot large-scale distributed frauds, but collusion among banks by sharing transaction logs can leak the sensitive data of the customers, whose transactions are enclosed in the logs.

8.5.1. Data Privacy Concerns

It is widely recognized that privacy is a right. In line with this imperative, and enabled by the rapid development of privacy-preserving technologies, federated learning (FL) systems have begun to emerge in the financial domain and other sectors [1]. Spread across an array of centers of investment, financial firms are naturally inclined to turn towards collaborative machine learning systems. However, firms are wary of putting sensitive transaction-level data in the hands of off-the-shelf technology partners. Even in instances where on-premise installation is a possibility, concerns about unintentional information leaks underlying the models or the infrastructure loom large. This makes identification of innovative business objectives, which can be provisioned with promising outcomes with local computation only, paramount. Therefore, a new kind of embedding model architecture utilizing federated representations has been developed in their financial fraud prevention ecosystem.

A collaborative learning framework for fraud prevention systems has been proposed to maintain the privacy of sensitive financial data. To share insightful transaction history information privately, as a first step, this framework has proposed embedding the transaction descriptions generating representations which are furnished with local differential privacy. Conceptually, a range of security privacy ends may well be able to be satisfied leveraging differential privacy frameworks. There has been a significant focus by the computer science community on federated solutions to machine learning, to continue providing international technology and algorithm co-development in this area, hybrid systems leveraging federated compute nodes or other federated principles will have to be considered. This can be contrasted against other co-developed platforms based on open-ledger or blockchain systems that, in contrast, can be much harder to deploy or communicate with governments or regulatory bodies.

With the proposed federated data publication mechanism, representations of transaction histories may be used to train shared fraud prevention models while ensuring that more detailed transaction descriptions cannot be recovered from shared representations.

Because the scheme conveys dummy values across a larger space than the number of actual transactions and with the addition of randomness, only probabilistic inferences can be drawn from any representations shared across other firms.

8.5.2. Bias in AI Algorithms

With the increasing prominence of machine learning in high-stakes decision-making processes, its potential to exacerbate existing social inequities has been a reason of growing concern. Financial services have been no exception, with multiple works in the field warning against potential discrimination. By leveraging complex information from data to make decisions, these models can also learn biases that are encoded within. Using biased patterns to learn to make predictions without accounting for possible underlying prejudices can lead to decisions that disproportionately harm certain social groups. The goal of building systems that incorporate these concerns has given rise to the field of Fair ML, which has grown rapidly in recent years. Fair ML research has focused primarily on devising ways to measure unfairness and to mitigate it in algorithmic prediction tasks. Mitigation is broadly divided in three approaches: pre-processing, in-processing, and post-processing, which map respectively to interventions on the training data, on the model optimization, and on the model output. Pre-processing assumes that the cause is bias in the data, while in- and post-processing shift the onus to modeling choices and criteria. Research seems to be divided along the same lines in what concerns uncovering the source of bias in the ML pipeline. There is work defending that bias in the data is at the root of downstream unfairness in predictions. Some researchers have advertised the crucial role that model choices have in algorithmic unfairness. However, the consequences of different sources of bias on unfairness produced by ML algorithms remains unclear. They maintain that the two views are complementary, not mutually exclusive. In fact, the landscape of algorithmic bias and fairness does change dramatically with the specific bias patterns present in a dataset. Under the same data bias conditions, different models incur in distinct fairness-accuracy trade-offs. This work has two overarching goals. First, to provide empirical evidence that predictive unfairness stems from the relationship between data bias and model choices, rather than from isolated contributions of either of them. Second, to steer the discussion towards relating algorithmic unfairness to concrete patterns in the data, allowing for more informed, data-driven choices of models and unfairness mitigation methods.

8.6. Future Trends in Fraud Detection Technology

In January 2023, the Federal Bureau of Investigation (FBI), the United States Secret Service, and the Cybersecurity and Infrastructure Security Agency (CISA) issued a joint

cybersecurity advisory, highlighting potential attacks related to the exploitation of vulnerabilities in the Zimbra Collaboration Suite (ZCS) and HCL Digital Solutions to gain access to email accounts. Reportedly, Threat Actors (TAs) could exploit vulnerabilities to deploy ransomware and steal sensitive emails and files. Generally, war on fraud and financial cybercrimes is difficult to win by traditional technologies and mechanisms. While more data-driven intelligent techniques have been explored or applied in practice, they also encounter great challenges as fraud schemes are evolving with inappropriate behavior . Resistance against financial fraud is hard to keep pace with the intelligence of fraud schemes because of great exploitation in social engineering techniques. Data-driven AI fraud detection systems with ethical considerations have been introduced beyond the prevention of fraud occurrence, which require potential intrusion of such systems by TAs. Privacy issues of MI and non-retrofitting of explainable methods are the most challenging tasks for current architectures. Adversarial learning could effectively evade fraud detection while retraining incremental learning models to handle real-time financial networks. A competitive edge against advanced fraud schemes also lays in adopting AI technology in the provision of financial services because adopting similar mechanisms could decrease the time-to-value. Assessments of AI-driven applications on fraud detection systems do not satisfy the systems' market deployment, explainability, and robustness.

8.6.1. Advancements in AI and Machine Learning

Fraud detection is an essential task in decentralized finance (DeFi) applications that has gained increasing attention in recent years. Fraud detection approaches in DeFi can be summarized from the project life cycle. The development phase refers to the design and deployment of smart contracts. Newly deployed smart contracts may contain vulnerabilities, which can be exploited by attackers through logic flaws or protocol violations. Security audits are thus important to detect vulnerabilities before exploitation, and several specialized fraud detection tools have been proposed. The illicit financial flows phase refers to the behaviour involving the movement of funds after fraud. DeFi-based frauds usually start with money laundering, which conceals the proceeds of fraud through illicit financial flows. Address clustering or off-chain behaviour detection is often used for laundering detection. Some researchers propose using proactive measures to prevent all behaviors relating to address clustering or even malicious actions pulling liquidity and price oracle manipulation or influencing the market by flash loans . Additionally, using advanced AI techniques to推动 DeFi fraud detection is worthy of exploration. One possible direction is to leverage the use of pre-trained models. In the area of transaction modelling, scholars can apply more pre-trained techniques to the field of fraud detection in DeFi for improvement of detection performance to address data

sparsity. Transfer learning is another direction that focuses on transferring knowledge across domains. In the area of fraud detection, researchers could consider transferring learning from TradFi fraud detection or adapting methods from one fraud type to another. Another powerful tool is large language models (LLMs). LLMs could provide immense potential for the advance of fraud detection with on-chain transactions being represented more comprehensively or the off-chain financial market social media data being utilized in fraud detection systems. One possible exploration is to leverage an LLM specialized in DeFi fraud detection or other search engines. Extending LLMs to generative agents is another direction that could benefit fraud detection, which could simulate multiple market conditions in a DeFi-based game environment or forecast potential fraud in emerging DeFi markets. In fraud detection across the DeFi project life cycle, tree-based, graph-related, and deep learning models have been extensively explored and proven effective. Explainable AI has also gained traction as a promising research topic in machine learning in recent years. It can benefit fraud detection significantly in terms of regulatory compliance and customer trust, and many techniques have been developed to tackle this challenge. With the rapid development of the DeFi ecosystem, continuously exploring the DeFi fraud landscape is critical to maintaining the safety of the DeFi system.

8.6.2. The Role of Blockchain Technology

Blockchain technology itself is an electronic digital ledger for cryptocurrency transactions. The ledger creates copies of data on multiple nodes even in untrustworthy environments. Each updated record is formed into a block. Adding a new block to the chain requires verification by trustworthy validators, preventing data alteration. Any alteration disrupts remaining blocks in the chain and requires agreement to restore. Blockchain has multiple applications in various fields, providing users with fast yet secure transactions. Blockchain can also be used for secure data storage, track digital product ownership, and provide proof against credential forgery.

Blockchain forensics, a sub-field of blockchain technology, can help banks inspect cryptocurrency transactions sent to wallets that emanate from suspicious activities. Blockchain can also provide data verification for financial transactions. Blockchain technology is increasingly being utilized to establish trust between digital financial institutions and has prompted research on approaches to enhance the security of blockchain networks.

The central aim of most financial institutions is to securely remit and receive money through digital platforms. In recent times, multiple online financial platforms have appeared and accumulated great masses of data regarding sale purchases, transfer of wealth, and social usage records on various networks. These platforms have paved the

way for several illegal and sinful activities. Financial fraud has increased tremendously over the years, both in person and digitally. To avoid being trapped and robbed by these financial criminals, users have started using Digital Financial Networks more frequently. Bitcoin was created in 2009 as the first decentralized cryptocurrency.

8.7. Conclusion

To sum up, fraud detection systems in financial networks are of critical significance given the rampant expansion of fraud activities over the past decades and the vulnerability of the financial network to an abundance of various fraud schemes. In particular, unsupervised AI-driven detection systems are becoming a rigorous threat to the security of financial networks. The exploitation of these systems presents an extensible strategy to maximize the profit of fraud activities at the expense of the loss and costs of financial organizations. Despite existing studies on analyzing misleading intelligent agents, several major issues remain open, such as the realtime model update under an extremely dynamic environment and the battle between the conventional classifiers and AI-driven fraud detection systems. Facing the security threat, a defensive model based on a transformer+GNN ensemble is established to effectively curb the capabilities of unsupervised AI-driven fraud detection systems, reflecting a robust perturbation pattern. Beyond these studies, potential future works are suggested, such as the investigation of novel defensive models against pre-trained models, the effectiveness evaluation of perturbation patterns against both GNNs and DNNs, and the consideration of incomplete records and unlabeled edges in real-world applications. Overall, this work delivers a new lens on the detection of fraud activities, which is vital for the financial organizations to protect themselves against this potential security threat by exploiting unsupervised methods [10]. In the future, geometric deep learning is emerging as a compelling paradigm for processing geometric data on nonlinear domains. The need for geometry-induced multi-disciplinary approaches has continued to grow exponentially in recent years and it is expected to play a key role in manipulating enormously complex data residing in complicated high-dimensional spaces. Moreover, the utilization of 3D geometric information in the analysis of social problems is receiving an increasing focus and in-depth study within the communities of computer vision, social science, and machine learning. Now, a complete and up-to-date overview of geometric deep learning will address its theoretical background and solutions in the areas of graph learning, mesh processing, and geometric computer vision. A vision of future perspectives will be presented from the ends of theory, models, applications, and techniques.

8.7.1. Future Trends

Although data-driven artificial intelligent techniques have achieved excellent performance in the financial fraud detection domain, there are still key issues remaining unsolved, as financial fraud schemes are rapidly evolving to adapt to this new digital environment. Financial fraud is harder to identify due to its increasing secretiveness and complexity. One of the severe difficulties for financial fraud detection is that the fraud is hidden in complex financial activities. The increased motives and the accelerated digital transformation caused by the pandemic lead to more intelligent fraud schemes. The secretiveness of financial fraud leads to the natural error in samples. Fraud detection can be regarded as a classification task, which requires fraud samples and non-fraud samples as training data. However, as the fraud activities are increasingly hidden, fraud usually cannot be fully identified. Consequently, the non-fraud samples may contain some unrecognized fraud samples. The complexity of financial activities leads to massive information involved. The financial activities are related to a wider range of business. The involved information is massive but heterogeneous, accompanied by lower-value density. The multi-source information will be difficult to play with if it is not well integrated. Financial data for fraud detection is massive but scattered.

Scholars can apply more pre-trained techniques for fraud detection in DeFi for improvement of detection performance. Another related direction is transfer learning, which focuses on transferring knowledge across domains. In DeFi fraud detection, researchers could consider transferring models from TradFi fraud detection or adapting methods from one fraud type to another. Large language models (LLMs) can help advance fraud detection with comprehensive on-chain transactions and off-chain social media data. Researchers can leverage an LLM specialized in DeFi fraud detection. Additionally, extending LLMs to generative agents could simulate market conditions and forecast potential fraud in emerging DeFi markets. In fraud detection across the DeFi project life cycle, tree-based, graph-related, and deep learning models have proven effective. Explainable AI is another promising direction that has gained prominence, contributing to regulatory compliance and customer trust. Scholars can use techniques like LIME for DeFi fraud detection. It is important to continuously explore the DeFi fraud landscape to maintain the safety of the DeFi system. The rapidly evolving DeFi landscape is giving rise to novel fraud types.

References

Ismaeil, M. K. A. (2024). Harnessing AI for Next-Generation Financial Fraud Detection: A DataDriven Revolution. Journal of Ecohumanism, 3(7), 811-821.

Yuhertiana, I., & Amin, A. H. (2024). Artificial Intelligence Driven Approaches for Financial Fraud Detection: A Systematic Literature Review. KnE Social Sciences, 448-468.

Islam, M. S., & Rahman, N. (2025). AI-Driven Fraud Detections in Financial Institutions: A Comprehensive Study. Journal of Computer Science and Technology Studies, 7(1), 100-112.

Ajayi, A. M., Omokanye, A. O., Olowu, O., Adeleye, A. O., Omole, O. M., & Wada, I. U. (2024). Detecting insider threats in banking using AI-driven anomaly detection with a data science approach to cybersecurity.

Varga, G. (2024). Data-Driven Methods for Machine Learning-Based Fraud Detection and Cyber Risk Mitigation in National Banking Infrastructure. Nuvern Machine Learning Reviews, 1(1), 33-40.