Open Access Research Journal of
Engineering and Technology
Journals home page: https://oarjpublication/journals/oarjet/
ISSN: 2783-0128 (Online)

OARJ OPEN ACCESS RESEARCH JOURNALS

(REVIEW ARTICLE)

AI and machine learning for detecting social media-based fraud targeting small businesses

Blessing Austin-Gabriel [1, *], Adeoye Idowu Afolabi [2], Christian Chukwuemeka Ike [3] and Nurudeen Yemi Hussain [4]

[1] *Montclair State University, Montclair, New Jersey, USA.*
[2] *CISCO, Nigeria.*
[3] *Globacom Nigeria Limited.*
[4] *Department of Computer Science, Texas Southern University, Texas, USA.*

Open Access Research Journal of Engineering and Technology, 2024, 07(02), 142-152

Publication history: Received on 09 November 2024; revised on 22 December 2024; accepted on 24 December 2024

Article DOI: https://doi.org/10.53022/oarjet.2024.7.2.0067

Abstract

Social media has become an essential tool for small businesses in the digital age, offering unprecedented marketing and customer engagement opportunities. However, this widespread use also exposes these businesses to a growing threat of social media-based fraud. This review paper explores the role of Artificial Intelligence (AI) and machine learning in detecting and mitigating such fraud. It examines the common types of social media fraud targeting small businesses, the evolving tactics employed by fraudsters, and the challenges in detecting these fraudulent activities due to the dynamic nature of social media platforms. The paper delves into various AI and machine learning approaches for fraud detection, including the use of advanced algorithms, Natural Language Processing (NLP), and real-time anomaly detection. Furthermore, it discusses the integration of AI tools with social media platforms, highlighting the role of APIs, data privacy concerns, and the benefits of automation and continuous learning systems. Finally, the paper outlines future trends and provides recommendations for small businesses, emphasizing the importance of adopting AI-based solutions and the roles of policymakers and platform providers in supporting these technologies. By synthesizing current knowledge and offering actionable insights, this paper aims to enhance the understanding of AI-driven fraud detection in social media and provide guidance for small businesses seeking to safeguard their operations.

Keywords: Social Media Fraud; AI Fraud Detection; Machine Learning; Natural Language Processing (NLP); Real-Time Anomaly Detection; Data Privacy

1. Introduction

1.1. Overview of Social Media-Based Fraud Targeting Small Businesses

The rise of social media as a primary platform for business promotion and customer engagement has created new growth opportunities, especially for small businesses (Agnihotri, 2020). Through social media platforms like Facebook, Instagram, and X (formally Twitter), small businesses can reach large audiences without the need for costly marketing campaigns. However, this digital transformation also brings new challenges, one of the most pressing being social media-based fraud. Fraudsters exploit the anonymity, reach, and accessibility of these platforms to target small businesses, often with devastating consequences (Devereux, Grimmer, & Grimmer, 2020).

Social media-based fraud targeting small businesses takes many forms. Fraudsters might create fake accounts or pages impersonating legitimate businesses, tricking customers into making payments or sharing sensitive information. Other schemes include phishing attacks through direct messages or fraudulent promotions that mislead customers or steal their financial details. In many cases, small businesses may fall victim to such fraud by being impersonated or engaging

* Corresponding author: Blessing Austin-Gabriel.

with malicious accounts. These scams not only result in financial loss but also tarnish the reputation of the businesses, sometimes irreparably (Zamir, 2020).

Social media's dynamic and fast-paced nature makes it an ideal environment for fraud to proliferate. Fraudsters continually adapt their tactics, making it difficult for businesses to keep up with the evolving threats. Small businesses, often with limited resources and technical expertise, are particularly vulnerable to these risks, and traditional methods of fraud detection struggle to keep pace with the complexity and scale of these digital threats (Banerjee, 2024).

1.2. Importance of Fraud Detection in Safeguarding Small Businesses

The need for effective fraud detection systems is more critical than ever for small businesses navigating the social media landscape. Social media-based fraud can have far-reaching impacts beyond immediate financial loss. The reputational damage caused by being associated with fraud can undermine customer trust, which is especially valuable for small businesses that rely heavily on word-of-mouth marketing and positive customer interactions (Nicholls, Kuppa, & Le-Khac, 2021).

Fraud also disrupts business operations, diverting time and resources away from growth and innovation to damage control. For example, businesses may need to deal with customer complaints, correct misinformation, or regain control of hijacked social media accounts. In extreme cases, fraud could lead to legal liabilities if customer data is compromised. In light of these risks, the importance of implementing robust fraud detection systems cannot be overstated (Banerjee, 2024).

Effective fraud detection helps small businesses identify and mitigate threats before they can cause significant harm. It ensures business continuity, protects customer relationships, and upholds the integrity of the business's brand. However, small businesses often lack the resources to build complex fraud detection mechanisms on their own, which is why leveraging advanced technologies like Artificial Intelligence (AI) and machine learning has become an essential strategy (Taherdoost, 2021).

1.3. Role of AI and Machine Learning in Combating Fraud

AI and machine learning have emerged as powerful tools in combating social media-based fraud. Unlike traditional rule-based detection systems that rely on predefined patterns, AI and machine learning algorithms can learn from vast datasets and identify hidden patterns that may not be immediately apparent to human analysts. These technologies can analyze large volumes of social media activity in real-time, detecting anomalies that could indicate fraudulent behavior. For instance, machine learning models can be trained to recognize suspicious changes in account activity, such as a sudden increase in messages or posts, unusual login locations, or behavior inconsistent with typical user patterns. AI-powered tools can also detect fake accounts by analyzing account creation data, interaction patterns, and profile inconsistencies. Natural Language Processing (NLP), a subset of AI, can also analyze the language used in messages and posts to identify phishing attempts or other types of fraudulent communication (Smith, Haworth, & Žitnik, 2020).

One of the most significant advantages of AI and machine learning is their ability to adapt to new and evolving threats. As fraudsters change their tactics, these systems can continuously learn from new data and update their detection methods accordingly. This makes AI-driven fraud detection far more flexible and scalable compared to traditional methods, which may require frequent manual updates to stay effective. For small businesses, AI and machine learning offer a scalable, automated, and efficient solution to combat the growing threat of social media-based fraud (Bokolo & Liu, 2024).

1.4. Purpose and Scope of the Paper

This paper aims to explore the growing threat of social media-based fraud targeting small businesses and examine how AI and machine learning can be leveraged to combat these fraudulent activities. As small businesses increasingly depend on social media platforms for their operations, the risks of fraud become more significant. This paper aims to provide a comprehensive understanding of the various types of fraud that exist on social media, the challenges small businesses face in detecting them, and how AI-driven technologies can offer a solution.

The scope of the paper covers several key areas: an exploration of the different types of social media fraud targeting small businesses, a discussion of the evolving tactics used by fraudsters, and the role of AI and machine learning in detecting and mitigating these threats. The paper also delves into the integration of AI tools with social media platforms, examining how Application Programming Interfaces (APIs), real-time monitoring, and automation can enhance fraud

detection efforts. Finally, the paper provides insights into emerging trends and offers recommendations for small businesses, policymakers, and platform providers to support the adoption of AI-driven fraud detection systems.

By synthesizing current research and offering practical guidance, this paper seeks to contribute to the ongoing conversation on how small businesses can safeguard their operations in the face of increasingly sophisticated social media fraud. The ultimate objective is to provide small businesses with actionable insights and tools to protect themselves from fraud and to encourage the development of more robust and accessible AI-driven fraud detection solutions across social media platforms.

2. Challenges of Social Media-Based Fraud

2.1. Common Types of Social Media Fraud Targeting Small Businesses

Social media-based fraud can manifest in various forms, exploiting the ease of creating fake accounts, the viral nature of posts, and the open interaction between businesses and customers. Some of the most prevalent types of fraud affecting small businesses include phishing scams, fraudulent accounts, identity theft, fake reviews, and false promotions (Springfeldt).

Phishing scams on social media are particularly dangerous because they trick small business owners or their employees into providing sensitive information. Fraudsters typically send direct messages or create fake posts that appear to come from trusted sources, such as a business partner, customer, or even the platform itself. These messages often contain malicious links that prompt users to enter login credentials, payment information, or other private data (Alkhalil, Hewage, Nawaf, & Khan, 2021; Bodini, Rivolta, & Sassi, 2021).

Fraudulent accounts are another common tactic used to target small businesses. Fraudsters may create accounts that mimic a legitimate business by using similar logos, usernames, and even customer interactions. These fake accounts are often used to scam customers by promoting false sales or stealing personal information. Small businesses may also fall victim to fraudsters who impersonate their own customers, leading to fake orders or refund requests that drain resources (Saporta & Maraney, 2022).

Fake reviews are a growing problem, as fraudsters use bots or paid reviewers to leave negative or positive feedback to manipulate a business's reputation. Positive fake reviews may be used to deceive consumers into purchasing substandard products, while negative reviews can harm a business's credibility, driving customers away and damaging trust (Paul & Nikolaev, 2021). Finally, false promotions are frequently used to target both businesses and consumers on social media. Fraudsters create enticing offers—such as discounted products, free giveaways, or exclusive access—that lure users into providing sensitive information or making payments for products that do not exist. Small businesses may be unwittingly drawn into such scams, believing they are engaging in legitimate partnerships or promotional activities.

2.2. The Evolving Tactics Used by Fraudsters

Fraudsters are continuously adapting their methods to stay ahead of detection systems. As small businesses become more aware of certain scams, new tactics emerge that exploit the latest trends, technologies, and social media features. One major shift in fraud tactics is fraudsters' increased use of automation and artificial intelligence. Bots can now create thousands of fake accounts, post reviews, and even engage in conversations with potential victims, all while appearing human-like and convincing (Pope, 2023).

Deepfakes and AI-generated content have also become tools for fraudsters. Fraudsters can impersonate real people with high accuracy by using AI to create realistic images, videos, and even voice recordings. This has led to more sophisticated phishing attempts, where fraudsters may send fake video messages or live streams pretending to be a known customer or business partner (Jones, 2020).

Fraudsters are also taking advantage of new social media features like live streaming, stories, and ephemeral content, which are harder to monitor due to their temporary nature. Live streams can be used to promote fraudulent schemes in real time, while stories and disappearing posts provide limited opportunities for detection before they vanish (Helmus, 2022). Moreover, fraudsters have become adept at leveraging social engineering tactics to exploit the trust of small business owners and employees. By engaging in extended conversations, forming connections, and building trust over time, fraudsters can extract sensitive information gradually, making their actions harder to detect (Linnell).

2.3. Challenges in Detecting Fraud Due to the Dynamic Nature of Social Media Platforms

One of the primary challenges in detecting social media-based fraud is the sheer volume and speed at which content is generated. Social media platforms host billions of users, with millions of posts, messages, and interactions occurring every minute. For small businesses, monitoring every interaction for signs of fraud is overwhelming, especially without the resources for advanced cybersecurity measures (Rao, Verma, & Bhatia, 2021).

The constant evolution of platform features and user behaviors further complicates fraud detection on social media. New tools, such as chatbots, live shopping, and influencer marketing, open up fresh avenues for fraudsters to exploit. The rapid development of these features means that small businesses must constantly adapt their fraud detection strategies, often with limited technical expertise (Guo, Ding, Yao, Liang, & Yu, 2020).

The anonymity provided by social media platforms also makes fraud detection challenging. Fraudsters can create accounts with false identities, mask their locations using VPNs, and delete accounts as soon as fraudulent activities are completed. This makes it difficult for small businesses to trace or hold fraudulent actors accountable (Shu & Liu, 2022).

Another significant challenge is the global nature of social media. Fraudsters often operate from different countries, using language barriers and jurisdictional issues to their advantage. Cross-border fraud complicates the legal recourse available to small businesses, as pursuing international fraud cases can be costly and time-consuming. Finally, many social media platforms prioritize user engagement over security, leading to inadequate fraud detection systems. While major platforms like Facebook and Twitter have improved their security measures, smaller platforms or newer features may lack robust fraud prevention protocols, leaving small businesses vulnerable (Meel & Vishwakarma, 2020).

2.4. Impact of Social Media Fraud on Small Businesses

The impact of social media-based fraud on small businesses can be devastating, affecting them financially, reputationally, and operationally. Financially, fraud can lead to direct losses through stolen funds, fraudulent transactions, or the costs associated with responding to a cyberattack. Small businesses may also face additional expenses for legal action, cybersecurity tools, and insurance claims (Egejuru, 2023).

Reputation is another significant casualty of social media fraud. When a business falls victim to fraud, its customers may lose trust, especially if their data has been compromised. Negative reviews, fraudulent promotions, and phishing attempts can damage the brand's image, leading to customer churn and declining sales. In the highly competitive social media space, regaining lost trust can be a slow and difficult process for small businesses (Wolff, 2022).

Operationally, the time and effort required to investigate and mitigate fraud can take a toll on small businesses. Employees may need to divert attention from core business functions to handle fraud-related issues, such as resolving customer disputes, securing compromised accounts, and dealing with the fallout from a cyberattack. This can lead to decreased productivity and delayed business growth (Scott, 2020).

3. AI and Machine Learning Approaches for Fraud Detection

3.1. Overview of AI and Machine Learning Techniques Used in Fraud Detection

AI and machine learning are increasingly being employed to combat social media fraud due to their ability to process large volumes of data, identify subtle irregularities, and predict potential fraudulent activities. These technologies learn from historical data—such as previous fraud incidents, user behavior, and transaction patterns—and apply this knowledge to identify new threats (O. A. Bello, Folorunso, et al., 2023).

Traditional rule-based systems, which require predefined conditions to trigger fraud alerts, are no longer effective in combating sophisticated fraud schemes. AI and machine learning techniques are far more flexible, adapting to new types of fraud as they emerge. For example, fraudsters frequently change their tactics to avoid detection, making static rules ineffective. In contrast, AI-driven systems can continuously learn from new data, enabling them to detect evolving fraud patterns more effectively. Furthermore, AI-based models can handle unstructured data—such as text, images, and videos—commonly found on social media platforms, enhancing their ability to analyze a wider range of inputs (Bao, Hilary, & Ke, 2022).

Several AI and machine learning techniques are applied in fraud detection, including supervised, unsupervised, and deep learning. In supervised learning, algorithms are trained using labeled datasets containing known fraudulent and legitimate transactions. Based on this training, the model then classifies new activities as fraudulent or non-fraudulent.

In contrast, unsupervised learning does not rely on labeled data but instead identifies anomalies or outliers in user behavior, which may indicate fraud. Deep learning, a subset of machine learning, uses artificial neural networks to process complex datasets, such as images or videos, allowing businesses to detect fraud in more diverse formats (Bin Sulaiman, Schetinin, & Sant, 2022).

3.2. Machine Learning Algorithms for Detecting Fraudulent Patterns in Social Media Activities

Machine learning algorithms are crucial in identifying fraudulent patterns in social media activities. These algorithms can detect subtle deviations from normal behavior that may indicate fraud by analyzing vast amounts of historical data. One of the most commonly used algorithms in fraud detection is the decision tree, which builds a model of decisions based on input features and classifies new instances accordingly. For example, a decision tree might analyze a user's posting frequency, engagement metrics, and follower growth to flag suspicious behavior (Nwaimo, Adegbola, & Adegbola, 2024b; Nwaimo, Adegbola, Adegbola, & Adeusi, 2024; Okoli, Obi, Adewusi, & Abrahams, 2024).

Another popular algorithm is the random forest, an ensemble learning method combining multiple decision trees to improve accuracy and reduce false positives. Random forests are particularly effective in fraud detection because they can handle large datasets and reduce the likelihood of overfitting, making them ideal for analyzing complex social media data (Abbass, Ali, Ali, Akbar, & Saleem, 2020).

Support Vector Machines (SVM) are also widely used in fraud detection. SVMs classify data points by finding the hyperplane that best separates fraudulent activities from legitimate ones. This algorithm is effective when dealing with high-dimensional data, such as the multitude of variables present in social media interactions. For instance, an SVM might analyze a combination of user location, posting time, and content type to determine whether an account is engaging in fraudulent activities (Balaji, Annavarapu, & Bablani, 2021).

Clustering algorithms, such as k-means clustering, are useful in unsupervised fraud detection. These algorithms group similar data points together and flag those that do not fit into any cluster as potential fraud. For example, suppose most legitimate accounts follow a certain posting pattern and an outlier suddenly deviates significantly. In that case, it may be identified as suspicious (Huang, Zheng, Li, & Che, 2024). Finally, neural networks and deep learning models are increasingly being employed in fraud detection. These models can process large amounts of unstructured data, such as images, text, and videos, making them particularly useful for detecting fake profiles, doctored images, and suspicious comments on social media. Convolutional neural networks (CNNs), for instance, are effective in analyzing images for signs of manipulation, while recurrent neural networks (RNNs) can be used to analyze time-series data, such as the sequence of a user's posts or interactions (Shen et al., 2020).

3.3. Natural Language Processing (NLP) and Sentiment Analysis for Identifying Suspicious Behaviors

Natural Language Processing (NLP) is a branch of AI that deals with the interaction between computers and human language. In the context of fraud detection, NLP is used to analyze the content of social media posts, comments, and messages to identify potential fraud. NLP algorithms can detect phishing attempts, fake promotions, or messages containing language associated with scams. For instance, a phishing message on social media might contain certain keywords or patterns commonly associated with fraudulent behavior, such as offers that seem too good to be true or requests for personal information (Shu & Liu, 2022; Smith et al., 2020).

Sentiment analysis, a technique within NLP, is particularly valuable for identifying suspicious behaviors on social media. It involves analyzing the tone and sentiment of posts and comments to determine whether they exhibit unusual or negative patterns. For instance, a sudden surge of overly positive reviews for a business, coupled with generic language, may indicate the presence of fake reviews intended to manipulate the business's reputation. Conversely, a flood of negative comments, possibly generated by bots, could be part of a coordinated attack to harm the business's credibility (Blum, 2020).

NLP also helps in detecting identity fraud by analyzing the language patterns used by individuals. Fraudsters may attempt to impersonate legitimate customers or businesses, but their language usage, tone, and style may differ from the norm. NLP models can flag inconsistencies that suggest fraudulent activity by comparing new interactions to historical language patterns (Amare, 2023).

3.4. The Importance of Real-Time Data Processing and Anomaly Detection

In the fight against social media fraud, real-time data processing is crucial. The dynamic nature of social media means that fraudulent activities can occur and spread rapidly, making it essential for businesses to detect and respond to

threats as they happen. AI and machine learning systems equipped with real-time data processing capabilities can continuously monitor social media platforms, providing instant alerts when suspicious activities are detected. This allows businesses to take immediate action, such as blocking fraudulent accounts, removing malicious posts, or notifying affected users (Adekunle et al., 2024).

Anomaly detection is a core component of real-time fraud detection. Machine learning models are trained to recognize normal patterns of behavior for a given business or user and flag deviations as potential fraud. For example, suppose an account suddenly experiences a spike in follower growth or begins posting content that deviates from its usual style. In that case, anomaly detection algorithms can identify these patterns as suspicious and trigger an investigation. The ability to detect anomalies in real-time allows businesses to prevent fraud from escalating before it causes significant damage (Elmrabit, Zhou, Li, & Zhou, 2020).

Real-time fraud detection also enhances customer trust. Social media is a fast-paced environment where news of fraud can spread quickly, harming a business's reputation. By employing AI and machine learning to detect and mitigate fraud in real-time, businesses can maintain their credibility and reassure customers that their information is secure (Liu et al., 2021).

4. Integration of AI Tools with Social Media Platforms

4.1. How AI-Driven Tools Can Be Integrated into Social Media Platforms for Monitoring Fraudulent Activities

Integrating AI tools directly into social media platforms is vital in detecting and mitigating fraud. AI-driven tools are designed to monitor vast amounts of data, assess user behavior, and detect anomalies in real-time. These tools work by analyzing posts, comments, direct messages, and other interactions that occur on social media platforms. AI algorithms can then flag suspicious behaviors, such as phishing attempts, fake accounts, or coordinated fraudulent schemes, allowing businesses to take preventive action swiftly (Prabhu Kavin et al., 2022).

AI integration into social media platforms can occur through several mechanisms. For example, platforms can adopt pre-existing AI-powered fraud detection tools or develop their in-house solutions tailored to the specific threats they face. These tools can be embedded into the backend systems of the platforms, where they continuously analyze user activities and interactions. Moreover, social media platforms can employ machine learning models that evolve and improve over time, learning from new data inputs and adapting to the changing tactics used by fraudsters (H. O. Bello, Ige, & Ameyaw, 2024).

Another critical integration aspect is enabling small businesses to access these AI-driven tools easily. Businesses can monitor suspicious activities by creating user-friendly dashboards and interfaces without requiring deep technical knowledge. This accessibility empowers small businesses to remain vigilant against fraud and reduces their dependency on third-party services for fraud detection (Goyal et al., 2023).

4.2. The Role of APIs and Data Extraction in Social Media Monitoring

Application Programming Interfaces (APIs) are essential in integrating AI tools for fraud detection on social media platforms. APIs enable the extraction and analysis of data seamlessly, allowing AI systems to access the vast volumes of information generated by social media users. APIs can pull data related to user interactions, content, account activities, and more, providing the necessary inputs for AI systems to function effectively (Nwobodo, Nwaimo, & Adegbola, 2024).

Through APIs, social media platforms can grant businesses and AI developers access to specific datasets that can be used for fraud detection. These datasets might include user metadata, such as location and login times, or content metadata, like the frequency and type of posts. AI systems can analyze this information to detect unusual patterns that may signify fraudulent activities. For example, suppose an account is created and immediately begins sending direct messages with phishing links to hundreds of other users. In that case, the API can flag this unusual behavior for further investigation (Agrawal, 2022).

Data extraction is also a key aspect of AI integration for fraud detection. With the help of APIs, AI tools can extract and analyze structured and unstructured data, including text, images, and videos, to identify fraud-related patterns. This allows small businesses to detect a wide range of fraudulent activities, from fake product promotions to manipulated images or fraudulent profiles (Dhieb, Ghazzai, Besbes, & Massoud, 2020).

However, while APIs facilitate the seamless integration of AI systems, they also present challenges. Not all social media platforms offer the same level of access to data through their APIs, with some platforms placing restrictions on data extraction due to privacy concerns. This limitation can hinder the effectiveness of AI-driven fraud detection systems. Nevertheless, API development is expected to evolve in line with the increasing demand for sophisticated fraud detection, providing more robust and secure data access in the future (H. O. Bello, Idemudia, & Iyelolu, 2024).

4.3. Addressing Data Privacy and Ethical Considerations

While the integration of AI tools into social media platforms for fraud detection offers significant benefits, it also raises important data privacy and ethical considerations. Social media platforms collect vast amounts of user data, which can be highly sensitive. Businesses must be cautious when using AI tools to monitor this data, ensuring they comply with privacy laws and ethical standards (Abdul-Azeez, Ihechere, & Idemudia, 2024).

One of the primary concerns is user consent. Social media platforms must ensure that users are fully informed about how their data is being monitored and used, especially when it comes to AI-driven systems. This can be achieved through clear privacy policies and opt-in mechanisms, which allow users to give explicit consent for their data to be analyzed. Additionally, businesses should avoid collecting unnecessary data, instead focusing on the information required to detect fraud effectively (Adewusi et al., 2024; Nwaimo, Adegbola, & Adegbola, 2024a).

Data security is another critical ethical concern. AI tools rely on large datasets to function, making them potential targets for hackers seeking to exploit the collected information. To mitigate these risks, businesses and social media platforms must implement strong security protocols, such as encryption, access controls, and regular audits. By safeguarding user data, businesses can maintain trust while protecting themselves from legal liabilities associated with data breaches (Pombal et al., 2022). Furthermore, AI systems must be designed to avoid bias. Fraud detection algorithms should be fair and impartial, ensuring that they do not unfairly target specific groups of users based on factors such as ethnicity, gender, or geographic location. This is particularly important in social media, where diverse user populations are involved. To address this, businesses should regularly test their AI systems for potential biases and take steps to rectify any unfair practices (Orwat, 2020).

4.4. Enhancing Detection through Automation and Continuous Learning Systems

One of the key advantages of AI-driven tools is their ability to automate fraud detection processes, reducing the need for human intervention. Automation allows AI systems to scan social media platforms around the clock, identifying potential fraud with speed and accuracy. For small businesses, this means they no longer have to rely on manual monitoring, which is time-consuming and prone to errors. AI tools can sift through millions of social media interactions in real-time, flagging suspicious activities that would be impossible for humans to catch at the same scale (Hassan, Aziz, & Andriansyah, 2023).

Automation also enables AI tools to respond to fraudulent activities quickly. For instance, if an AI system detects an account engaging in a phishing scheme, it can immediately freeze or flag the account for further investigation, minimizing the damage caused to the business. This level of responsiveness is crucial in the fast-paced social media environment, where fraud can escalate rapidly (Odeyemi, Mhlongo, Nwankwo, & Soyombo, 2024).

Continuous learning systems further enhance the capabilities of AI-driven fraud detection. Unlike static rule-based systems, AI tools with machine learning capabilities can adapt to new types of fraud by learning from historical data. As fraudsters change their tactics, AI systems continuously update their algorithms to recognize emerging threats. For example, suppose a new type of phishing message gains popularity. In that case, a continuous learning system will analyze similar messages and refine its detection criteria accordingly. This adaptability makes AI tools highly effective at combating ever-evolving social media fraud. Moreover, continuous learning systems reduce the number of false positives—instances where legitimate activities are mistakenly flagged as fraudulent. AI tools can fine-tune their models to improve accuracy by learning from past detection errors. This is particularly important for small businesses, as excessive false positives can disrupt their operations and may erode trust with their customers (O. A. Bello, Ogundipe, Mohammed, Adebola, & Alonge, 2023).

5. Future Trends and Recommendations

5.1. Emerging Trends in AI and Machine Learning for Combating Fraud in Social Media

One of the key emerging trends in AI and machine learning for fraud detection is the development of more sophisticated algorithms that can detect subtle patterns in social media interactions. Deep learning, a subset of machine learning, is

becoming instrumental in detecting complex fraudulent behaviors that may evade traditional detection systems. For example, convolutional neural networks (CNNs) are being increasingly used for image analysis, allowing AI tools to spot fake product promotions or counterfeit brands on social media platforms. Similarly, generative adversarial networks (GANs) are being utilized to improve fraud detection by simulating real-world scenarios and identifying potential fraud before it occurs.

Additionally, the integration of AI with blockchain technology is gaining traction. Blockchain offers a secure and transparent method of tracking transactions and user activities on social media, making it harder for fraudsters to manipulate data. AI systems can leverage blockchain's immutable records to cross-verify the legitimacy of transactions and accounts, thus improving fraud detection accuracy.

While current AI-based fraud detection systems are effective, there are areas for improvement. One promising avenue is the enhancement of real-time detection capabilities. AI systems that can process data instantly and identify fraudulent activities as they happen will be vital in reducing the impact of social media fraud. Furthermore, hybrid models that combine machine learning with human oversight could lead to more accurate and context-aware detection. These systems would allow AI to handle large-scale data analysis while human experts provide nuanced judgment for complex or ambiguous cases.

Another area for improvement lies in reducing false positives. AI systems sometimes flag legitimate user activities as fraudulent, leading to disruptions for small businesses. Fine-tuning machine learning models to better differentiate between genuine and fraudulent behaviors will make fraud detection more reliable and less invasive.

5.2. Recommendations for Small Businesses to Adopt AI-Based Solutions

For small businesses, adopting AI-based fraud detection systems can be a game-changer in protecting their operations from social media scams. Small businesses must explore affordable AI tools that offer automated social media accounts and transaction monitoring. Many third-party platforms offer fraud detection services tailored to small business needs, reducing the cost and complexity of integrating such systems.

Small businesses should also consider using AI-powered sentiment analysis tools and monitoring customer interaction. These tools can help identify potential threats by flagging unusual customer reviews or comment patterns. Additionally, collaborating with cybersecurity experts to regularly audit AI systems ensures businesses stay updated on the latest fraud detection techniques and respond effectively to emerging threats.

6. Conclusion

The findings of this study highlight the growing risks posed by social media-based fraud to small businesses and the transformative potential of artificial intelligence (AI) and machine learning (ML) in combating these challenges. By exploring fraud tactics such as phishing, fake accounts, and false promotions, this research underscores the complexity and evolving nature of digital threats. AI and ML emerge as vital tools for detecting and mitigating fraud through techniques like real-time anomaly detection, natural language processing, and automated systems. The integration of AI-driven tools with social media platforms offers scalable solutions, enabling small businesses to monitor and address fraudulent activities efficiently, even with limited resources. Furthermore, this paper identifies significant areas of improvement, such as enhancing the accessibility of AI technologies and addressing ethical considerations, including data privacy and algorithmic biases. The study also emphasizes the importance of collaborative efforts among small businesses, policymakers, and social media platform providers to develop robust fraud detection systems. Emerging trends, such as the combination of AI with blockchain technology and advancements in real-time fraud detection, present exciting opportunities to strengthen the digital defenses of small businesses. By offering actionable recommendations, this research equips small businesses with insights to protect their operations while encouraging innovation in fraud prevention strategies. In conclusion, the adoption of AI-based fraud detection tools is crucial for safeguarding small businesses from the increasing sophistication of social media fraud. This study not only provides a comprehensive analysis of current challenges but also paves the way for the development of secure and resilient systems that can adapt to future threats. By fostering a secure digital ecosystem, the insights presented here contribute to a broader societal benefit, ensuring trust and sustainability in the ever-evolving digital landscape.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Abbass, Z., Ali, Z., Ali, M., Akbar, B., & Saleem, A. (2020). A framework to predict social crime through twitter tweets by using machine learning. Paper presented at the 2020 IEEE 14th International Conference on Semantic Computing (ICSC).

[2] Abdul-Azeez, O., Ihechere, A. O., & Idemudia, C. (2024). Digital access and inclusion for SMEs in the financial services industry through Cybersecurity GRC: A pathway to safer digital ecosystems. Finance & Accounting Research Journal, 6(7), 1134-1156.

[3] Adekunle, T. S., Alabi, O. O., Lawrence, M. O., Ebong, G. N., Ajiboye, G. O., & Bamisaye, T. A. (2024). The use of ai to analyze social media attacks for predictive analytics. Journal of Computing Theories and Applications, 1(4), 386-395.

[4] Adewusi, A. O., Okoli, U. I., Olorunsogo, T., Adaga, E., Daraojimba, D. O., & Obi, O. C. (2024). Artificial intelligence in cybersecurity: Protecting national infrastructure: A USA. World Journal of Advanced Research and Reviews, 21(1), 2263-2275.

[5] Agnihotri, R. (2020). Social media, customer engagement, and sales organizations: A research agenda. Industrial marketing management, 90, 291-299.

[6] Agrawal, S. (2022). Enhancing payment security through AI-Driven anomaly detection and predictive analytics. International Journal of Sustainable Infrastructure for Cities and Societies, 7(2), 1-14.

[7] Alkhalil, Z., Hewage, C., Nawaf, L., & Khan, I. (2021). Phishing attacks: A recent comprehensive study and a new anatomy. Frontiers in Computer Science, 3, 563060.

[8] Amare, L. (2023). Scamming Detection from social media using Deep Learning Approach.

[9] Balaji, T., Annavarapu, C. S. R., & Bablani, A. (2021). Machine learning algorithms for social media analysis: A survey. Computer Science Review, 40, 100395.

[10] Banerjee, R. (2024). Corporate Frauds: Now Bigger, Broader and Bolder: Penguin Random House India Private Limited.

[11] Bao, Y., Hilary, G., & Ke, B. (2022). Artificial intelligence and fraud detection. Innovative Technology at the Interface of Finance and Operations: Volume I, 223-247.

[12] Bello, H. O., Idemudia, C., & Iyelolu, T. V. (2024). Integrating machine learning and blockchain: Conceptual frameworks for real-time fraud detection and prevention. World Journal of Advanced Research and Reviews, 23(1), 056-068.

[13] Bello, H. O., Ige, A. B., & Ameyaw, M. N. (2024). Adaptive machine learning models: concepts for real-time financial fraud prevention in dynamic environments. World Journal of Advanced Engineering Technology and Sciences, 12(2), 021-034.

[14] Bello, O. A., Folorunso, A., Onwuchekwa, J., Ejiofor, O. E., Budale, F. Z., & Egwuonwu, M. N. (2023). Analysing the impact of advanced analytics on fraud detection: a machine learning perspective. European Journal of Computer Science and Information Technology, 11(6), 103-126.

[15] Bello, O. A., Ogundipe, A., Mohammed, D., Adebola, F., & Alonge, O. A. (2023). AI-Driven Approaches for real-time fraud detection in US financial transactions: challenges and opportunities. European Journal of Computer Science and Information Technology, 11(6), 84-102.

[16] Bin Sulaiman, R., Schetinin, V., & Sant, P. (2022). Review of machine learning approach on credit card fraud detection. Human-Centric Intelligent Systems, 2(1), 55-68.

[17] Blum, D. (2020). Rational cybersecurity for business: the security leaders' guide to business alignment: Springer Nature.

[18] Bodini, M., Rivolta, M. W., & Sassi, R. (2021). Opening the black box: interpretability of machine learning algorithms in electrocardiography. Philosophical Transactions of the Royal Society A, 379(2212), 20200253.

[19] Bokolo, B. G., & Liu, Q. (2024). Artificial Intelligence in Social Media Forensics: A Comprehensive Survey and Analysis. Electronics, 13(9), 1671.

[20] Devereux, E., Grimmer, L., & Grimmer, M. (2020). Consumer engagement on social media: Evidence from small retailers. Journal of Consumer Behaviour, 19(2), 151-159.

[21] Dhieb, N., Ghazzai, H., Besbes, H., & Massoud, Y. (2020). A secure ai-driven architecture for automated insurance systems: Fraud detection and risk measurement. Ieee Access, 8, 58546-58558.

[22] Egejuru, K. C. (2023). The Role of a Corporate Governance Mechanism in Mitigating Fraud: A Case Study of The Nigerian Banking Industry. University of Salford,

[23] Elmrabit, N., Zhou, F., Li, F., & Zhou, H. (2020). Evaluation of machine learning algorithms for anomaly detection. Paper presented at the 2020 international conference on cyber security and protection of digital services (cyber security).

[24] Goyal, B., Gill, N. S., Gulia, P., Prakash, O., Priyadarshini, I., Sharma, R., . . . Yadav, K. (2023). Detection of fake accounts on social media using multimodal data with deep learning. IEEE Transactions on Computational Social Systems.

[25] Guo, B., Ding, Y., Yao, L., Liang, Y., & Yu, Z. (2020). The future of false information detection on social media: New perspectives and trends. ACM Computing Surveys (CSUR), 53(4), 1-36.

[26] Hassan, M., Aziz, L. A.-R., & Andriansyah, Y. (2023). The role artificial intelligence in modern banking: an exploration of AI-driven approaches for enhanced fraud prevention, risk management, and regulatory compliance. Reviews of Contemporary Business Analytics, 6(1), 110-132.

[27] Helmus, T. C. (2022). Artificial intelligence, deepfakes, and disinformation. RAND Corporation, 1-24.

[28] Huang, Z., Zheng, H., Li, C., & Che, C. (2024). Application of machine learning-based k-means clustering for financial fraud detection. Academic Journal of Science and Technology, 10(1), 33-39.

[29] Jones, V. A. (2020). Artificial intelligence enabled deepfake technology: The emergence of a new threat. Utica College,

[30] Linnell, K. A case study of how SMEs protect themselves against fraud. University of Portsmouth,

[31] Liu, M., Zhang, Y., Liu, B., Li, Z., Duan, H., & Sun, D. (2021). Detecting and characterizing SMS spearphishing attacks. Paper presented at the Proceedings of the 37th Annual Computer Security Applications Conference.

[32] Meel, P., & Vishwakarma, D. K. (2020). Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. Expert Systems with Applications, 153, 112986.

[33] Nicholls, J., Kuppa, A., & Le-Khac, N.-A. (2021). Financial cybercrime: A comprehensive survey of deep learning approaches to tackle the evolving financial crime landscape. Ieee Access, 9, 163965-163986.

[34] Nwaimo, C. S., Adegbola, A. E., & Adegbola, M. D. (2024a). Data-driven strategies for enhancing user engagement in digital platforms. International Journal of Management & Entrepreneurship Research, 6(6), 1854-1868.

[35] Nwaimo, C. S., Adegbola, A. E., & Adegbola, M. D. (2024b). Predictive analytics for financial inclusion: Using machine learning to improve credit access for under banked populations. Computer Science & IT Research Journal, 5(6), 1358-1373.

[36] Nwaimo, C. S., Adegbola, A. E., Adegbola, M. D., & Adeusi, K. B. (2024). Evaluating the role of big data analytics in enhancing accuracy and efficiency in accounting: A critical review. Finance & Accounting Research Journal, 6(6), 877-892.

[37] Nwobodo, L. K., Nwaimo, C. S., & Adegbola, A. E. (2024). Enhancing cybersecurity protocols in the era of big data and advanced analytics. GSC Advanced Research and Reviews, 19(3), 203-214.

[38] Odeyemi, O., Mhlongo, N. Z., Nwankwo, E. E., & Soyombo, O. T. (2024). Reviewing the role of AI in fraud detection and prevention in financial services. International Journal of Science and Research Archive, 11(1), 2101-2110.

[39] Okoli, U. I., Obi, O. C., Adewusi, A. O., & Abrahams, T. O. (2024). Machine learning in cybersecurity: A review of threat detection and defense mechanisms. World Journal of Advanced Research and Reviews, 21(1), 2286-2295.

[40] Orwat, C. (2020). Risks of Discrimination through the Use of Algorithms. Berlín: Federal Anti-Discrimination Agency.

[41] Paul, H., & Nikolaev, A. (2021). Fake review detection on online E-commerce platforms: a systematic literature review. Data Mining and Knowledge Discovery, 35(5), 1830-1881.

[42] Pombal, J., Cruz, A. F., Bravo, J., Saleiro, P., Figueiredo, M. A., & Bizarro, P. (2022). Understanding Unfairness in Fraud Detection through Model and Data Bias Interactions. arXiv preprint arXiv:2207.06273.

[43] Pope, K. R. (2023). Fool Me Once: Scams, Stories, and Secrets from the Trillion-dollar Fraud Industry: Harvard Business Press.

[44] Prabhu Kavin, B., Karki, S., Hemalatha, S., Singh, D., Vijayalakshmi, R., Thangamani, M., . . . Kshirsagar, P. R. (2022). Machine Learning-Based Secure Data Acquisition for Fake Accounts Detection in Future Mobile Communication Networks. Wireless Communications and Mobile Computing, 2022(1), 6356152.

[45] Rao, S., Verma, A. K., & Bhatia, T. (2021). A review on social spam detection: Challenges, open issues, and future directions. Expert Systems with Applications, 186, 115742.

[46] Saporta, G., & Maraney, S. (2022). Practical Fraud Prevention: " O'Reilly Media, Inc.".

[47] Scott, G. (2020). Go Phish: Circuit Split in Policy Interpretation for Social Engineering Fraud Losses Creates Ambiguity for Insurers and Insureds. Vill. L. Rev. Online, 65, 1.

[48] Shen, Q., Wu, Y., Jiang, Y., Zeng, W., Alexis, K., Vianova, A., & Qu, H. (2020). Visual interpretation of recurrent neural network on multi-dimensional time-series forecast. Paper presented at the 2020 IEEE Pacific visualization symposium (PacificVis).

[49] Shu, K., & Liu, H. (2022). Detecting fake news on social media: Springer Nature.

[50] Smith, G. G., Haworth, R., & Žitnik, S. (2020). Computer science meets education: Natural language processing for automatic grading of open-ended questions in ebooks. Journal of educational computing research, 58(7), 1227-1255.

[51] Springfeldt, S. Social Media Manipulation in India–.

[52] Taherdoost, H. (2021). A review on risk management in information systems: Risk policy, control and fraud detection. Electronics, 10(24), 3065.

[53] Wolff, J. (2022). Cyberinsurance policy: Rethinking risk in an Age of ransomware, computer fraud, data breaches, and cyberattacks: MIT Press.

[54] Zamir, H. (2020). Cybersecurity and social media. In Cybersecurity for Information Professionals (pp. 153-171): Auerbach Publications.