International Journal of Economic Practices and Theories (IJEPT)
ISSN: 2247-7225
Volume 2026, Issue 1

AI-Driven Fraud Detection and RiskAnalysis of Financial Misinformation

[1]Meenu Grover, [2]Dr.Ravneetkaur,[3]Smridhi Vohra,[4]Dr.Amrit Kaur,[5]Divya Rathi,

[6]Ravinder Singh,[7]Gourav Kamboj,

[1]Assistant Professor, Department of Management, Dr Akhilesh Das Gupta Institute of Professional Studies, Delhi

meenugrover37@gmail.com

[2]Assistant Professor, Department of Management, Gandhi Memorial National College, Ambala Cantt

ravneetgulati08@gmail.com

[3]Assistant Professor of Economics, Guru Nanak Girls College, Yamuna Nagar

smridhivohra2000@gmail.com

[4]Assistant Professor of Political Science, Guru Nanak Girls College, Yamuna Nagar

amritdhindsa838@yahoo.in

[5]Assistant Professor of Human Development, Guru Nanak Girls College, Yamuna Nagar

narwaldivya2506@gmail.com

[6]Research Scholar, Maharishi Markandeshwar Deemed to be University, Mullana

ravindersingh6782@gmail.com

[7]LM Thapar School of Management,

Thapar Institute of Engineering and Technology (Deemed to be University), Patiala, Punjab, India

gkamboj_phd25@thapar.edu

Abstract

Financial fraud is a menace that evolves every now and then and it is difficult to remain afloat and honest by financial institutions. The paper systematically reviews how big data analytics, as well as advanced technologies, including machine learning and natural language processing (NLP), can be used to detect and prevent financial crime. Using the PRISMA guidelines on a systematic review and meta-analyses, 179 peer-reviewed articles were thoroughly reviewed in order to identify the most effective methods and most critical gaps in the literature. The findings demonstrate the potential transformation that big data analytics can bring to the game in terms of identifying trends and unusual behaviors that are typically overlooked by conventional systems so that real-time detection of fraud will become much more accurate and efficient. It was demonstrated that machine learning techniques, particularly, the ensemble models and deep learning algorithms, could be well adapted to shifting trends of fraud. NLP, however, might be used to detect fraud in an unstructured format, such as emails, contracts, and social media. In addition, real-time fraud detection systems were also required to quickly mitigate the situation; however, challenges in integrating multimodal sources of data (audio and video) and shifting the current fraud protection responses to proactive mitigation measures persist. This critical review creates a complete overview of the newest trends, highlighting both achievements and the fields that require additional investigation. It also preconditions the development of powerful, scalable, and ethical systems of fraud detection.

Keywords: - Financial fraud, fraud detection, big data analytics, machine learning, deep learning, ensemble models, natural language processing, NLP, systematic review, PRISMA guidelines, meta-analysis, financial crime prevention, real-time detection, multimodal data integration, ethical fraud detection systems

1 Introduction

The problem of financial fraud is becoming increasingly relevant in the international financial system, and it implies that sophisticated technology is required to detect it and remove it as quickly as possible (Konigstorfer and Thalmann, 2020). With more and more use of technology in the financial service industry, fraud has also increased. It is a massive crisis to the consumers, businesses, and the economy (Danielsson et al., 2022). Traditional methods used in fraud detection which usually rely on rule-based systems have been unable to detect abnormalities signaling fraudulent behavior (Deshmukh and Talluru, 1998). These anomalies cannot be detected by hand since they can be concealed in large databases. Some of the best machine learning methods that have been extensively applied to automate the detection process are the decision trees, the support vector machines and the neural networks. Fraud detection systems are much more precise because such algorithms can never learn off-hand (Konigstorfer and Thalmann, 2020). As an example, ensemble learning methods have been utilized to combine a large number of models, which makes fraud detection more accurate and reliable (Danielsson et al., 2022). The other interesting thing is the use of natural language processing (NLP) to detect financial fraud. The NLP framework is excellent at analyzing unstructured data, such as the words of emails, contracts, and posts in social media, which tend to be related to the complexity of contemporary. The combination of big data analytics and artificial intelligence (AI) has therefore become a very successful way of identifying and thwarting fraud. One of the advantages of big data technologies that facilitates this switch is the fact that it is capable of assisting in the discovery of fraud. Illustratively, communications that appear suspicious have been identified with the aid of key word extraction and sentiment analysis and trends in huge text samples have been identified with topic modelling. Furthermore, combining NLP with machine learning will help the system to identify subtle linguistic differences, which are useful in detecting complex cases of fraud that other systems fail to (Mwangi and Ndegwa, 2020). Despite these advances, big data analytics still does not make it easy to detect financial fraud. The quality of data, privacy issues, and understanding of machine learning models are still critical issues. The data that is used to feed the fraud detection systems is especially important in determining how well the systems work. The reason is that biased datasets or missing information may be used to make inaccurate forecasts (Dosilovic et al., 2018). Moreover, many of the AI models are black boxes and therefore, it is not easy to understand why people interested want to receive a fraud warning. This renders them to be more difficult to use in companies that have strict limitations (Deshmukh and Talluru, 1998). It is necessary to address these issues, as it will allow us to realize the full potential of using big data analytics in the fight against financial crime. The main aim of the present project is to analyze how big data analytics and artificial intelligence (AI) can be applied to improve the financial fraud-detecting systems. The significant aim of the research is to identify and investigate how complicated machine learning algorithms can be used to identify trends and anomalies indicative of fraud. It also tries to explore how natural language processing (NLP) might be applied to obtain pertinent information using unstructured data feeds of information such as texts and messages. The work demonstrates the possible way in which these technologies are likely to change the financial sector through making digital financial systems more credible and reducing fraud. This is the way it combines new research and case cases.

1.1 AI-Driven Fraud Detection in Today's Financial World

AI-based fraud detection has transformed the manner in which the financial sector operates by ensuring that it is far better at detecting and preventing frauds compared to other methods. Business organizations, consumers, and entire economies can lose significant amounts of money due to financial system frauds. The need to find a more efficient method of detecting fraud has never been greater as financial transactions have become more digitised and complex. Traditional fraud detection systems no longer have sufficient ability to meet the demands of the evolving world of financial crime relying on the static rule and human judgement. Machine learning-based AI systems are particularly effective at analyzing a large volume of transaction data on a real-time basis. Such algorithms are able to identify trends and suspicious activity that might indicate a person attempts to commit fraud, such as unusual spending behaviour, suspicious account usage, or something that does not align with the behaviour of a particular user. Fraud is much easier to identify timely through AI than through human investigators, as it does not take a long time to analyze large amounts of data. This enables prompt intervention and reduces consequences of fraud. The most positive aspect of AI-based fraud detection is that it is able to

evolve and acquire knowledge constantly. Algorithms trained through machine learning will improve with time as new data will be used to acquire knowledge, and this will ensure the system continues to operate as fraud techniques evolve. Examples of such aspects include AI models identifying minor fraud trends that are not detected by other methods. This assists banks and other financial institutions to be a step ahead of criminals. In addition, the AIs can be used to reduce false positives, common in the traditional systems of fraud detection. With sophisticated predictive models, AI systems are better placed in making guesses on the likelihood of fraud. This ensures that actual transactions are not mistakenly identified, a fact that will guarantee a smooth operation and enhance customer service. It is transforming the manner in which we safeguard our money by providing the means to counteract fraud in a quicker, more precise, and adaptable manner due to AI-based fraud detection. These new technologies do not only ensure that the banks are safe, but also enable people to be more confident to conduct business online. This aids in maintaining stability and safety of the modern financial systems.

1.2 Implementing an AI-Based Fraud Detection to the Modern Financial Systems.

Fraud detection in the existing financial systems using AI is transforming the manner in which financial institutions (especially banks) safeguard their assets and customer information. Detection of fraud by use of rule-based systems which indicate suspicious activity in relation to a set of criteria has been in operation a long time. Nevertheless, the systems are typically less effective at detecting complex, evolving, and sophisticated fraud, such as account takeover, identity theft, and insider trading, and AI-based fraud detection systems with machine learning (ML) and deep learning (DL) uncovers have proven significantly more effective in detecting these minor issues. Such systems are constantly learning and evolving through investigating vast amounts of data, including transaction histories, behavioural trends, and external influences, in order to discover outliers that do not fit in with normal activity. The process will enable AI to detect fraud in real-time with a higher degree of precision, which results in quicker responses and minimal lost money. The most beneficial aspect of AI-driven fraud detection is that it is able to identify trends in large volumes of data that human analysts cannot pick on. Supervised learning models can be trained on historical data of fraud with these systems. This will allow them to know how fraud transactions appear and spot such tendencies in real time. Conversely, new and emerging fraud strategies that do not require labelled data can be discovered by unsupervised learning models. This enables them to discover new dangers never witnessed before. Individuals or accounts can also be given a detailed profile of the fraud based on the data collected by AI algorithms, which may include information about other sources, such as geolocation, device fingerprints, frequency of transactions, and consumer behaviour. This comprehensive approach assists banks to detect fraud in more locations including mobile banking applications and credit card usage that strengthens security. With the increased financial services becoming online and the increased number of people transacting online, this scalability is extremely significant in ensuring the fraud detection systems are effective in a busy setting. Another possible use of AI is predictive fraud detection, which implies that AI will be able to identify patterns and trends that would suggest future fraudulent actions prior to their occurrence.

Not only is gearing towards easier and quicker fraud detection in use, AI in fraud detection systems is also providing banks with the means to stay with cyber threats that are increasingly sophisticated. Artificial intelligence is an important component of financial security in the present day since it can process large amounts of data instantly and respond to innovative methods of committing fraud. This assists in ensuring the security of both the institution and their clients to the ever-increasing threat of financial fraud.

2 Review of literature

The study on fraud detection through industry-conscious NLP-based classification helps emphasize that it is important to distinguish between different types of frauds, including bribery, tax fraud, investment fraud, and financial misstatements and not see fraud as one. The framework relates the types of fraud to some industries by chi-square analysis and clustering. It makes use of both supervised machine learning and Latent Dirichlet Allocation (LDA) topic modelling. It derives its theoretical foundation on the Routine Activity theory (RAT) and Crime Script Analysis (CSA) which places fraud within an environment where it is structurally feasible and procedurally feasible. This approach enhances the detection of fraud by putting the malpractice into industry-specific logics, which can better the regulatory accuracy and enforcement mechanisms (Bao et al., 2022;

Cecchini et al., 2010; Lokanan and Sharma, 2024). The peer-reviewed articles on the topic of big data analytics and AI applications in fraud detection are a systematic review conducted by Emran and Rubel (2024). Their findings indicate how ensemble learning, deep learning, and NLP can transform the manner in which we detect patterns of fraud which are typically overlooked by the traditional rule-based methods. Such methods as clustering, anomaly detection, and sentiment analysis can be used to detect fraud in emails, contracts, and social media that are not structured sources of data.The research also demonstrates the significance of the real-time detection systems of fraud. These systems are able to prevent fraud immediately, but they possess issues of data quality, privacy as well as the comprehension of the functioning of the models. It concludes that to have a growing and moral fraud detection framework, there is a need to have hybrid models that are multimodal data based and explainable AI (Danielsson et al., 2022; Mwangi and Ndegwa, 2020). The AI-based framework of detecting fake news in real-time, suggested by (Twaha and Arfin, 2025), is specifically created to apply to the news sites of the U.S. It runs the NLP methods, supervised classifiers, and the deep learning systems such as BERT and LSTM to accept streams of data, extract the context, and assign the news content with probabilistic labels. The site is easy to incorporate into news aggregators, social media, and fact-checking organizations because of the existence of APIs. An evaluation of the literature which was conducted as a PRISMA exercise revealed that hybrid models incorporating linguistic, semantic as well as social context factors perform better than traditional baselines. In order to ensure responsible deployment, such ethical concerns as transparency, free speech, and algorithmic fairness are considered. The present research assists in safeguarding the integrity of information by enabling one to detect fake news in a manner which is scalable, automated, and comprehensible (Ogundipe et al., 2023; Rath et al., 2022; Singh et al., 2023). (Nasiri and Hashemzadeh, 2025) trace the history of disinformation through the fake news propaganda that was founded on text to the artificial intelligence-based stories such as deepfakes. The paper discusses the transformations in the production of disinformation caused by Generative Adversarial Networks (GANs) and NLP, which create fake media that is almost real and more difficult to locate. The examples of the COVID-19 epidemic and the 2020 U.S. elections show that the spread of misleading stories through AI-powered bots and deepfakes negatively affects the populace and the political process. The authors apply the framework of information disorder (that incorporates misinformation, disinformation and misinformation) and the principle of harm to examine the risks that AI-supported disinformation presents to the society. It is said that we must react to deepfakes in numerous manners, including, but not limited to, technical innovation, regulatory measures, and media literacy training (Wardle and Derakhshan, 2017; Chesney and Citron, 2019). (Malik, 2025) considers the use of AI both in detection of fraud and predicting the energy market. Banking AI-supported fraud detection systems are machine learning and deep learning algorithms that analyze the transactional data in real time, identify abnormal patterns, and reduce false positives. These systems are under constant evolvement as they strive to match new methods by which fraudsters are attempting to defraud the systems, and can even anticipate the future fraud in the future. With the assistance of the historical data, weather patterns, and socio-economic factors, AI models make extremely precise forecasts regarding the changes in the energy consumption, supply, and prices. Introducing AI to these fields ensures that finances are less risky, and plans to control energy usage in the long term perspective are executed, which is consistent with the intentions of global development sustainability and fraud. According to Malik, AI technologies can be considered very significant to creating financial and energy systems that are robust and fulfill efficiency, safety, and environmental responsibility.

3 Problems and Restrictions

Frameworks and algorithms based on machine learning to identify false news in real-time have significant potential; however, their implementation and deployment are associated with complex issues and limitations (Lavin et al., 2022; Zafar et al., 2023). These issues include technical, moral, and social issues. There are three key concerns: adversarial news generation, training data biases, and false positives, which may cause people the loss of confidence in the system. We should address these issues to make these systems powerful, just, and acceptable to the society. Problems and Limitations This is one of the most difficult technical issues: fake news can evolve and transform within a short period of time. Individuals who propagate fake news are ever devising new methods of ensuring they go undetected and they normally employ sophisticated adversarial strategies. Others are employing confusing words, paraphrasing, artificial intelligence generated articles, and editing photos or videos in order to present them as different. The advent of generative models such as GPT and other

large language models (LLMs) has now enabled bad people to create fake news articles that look, sound, and feel like real news articles. The classical false news classifiers, in particular, the ones that are trained on static data, find it difficult to adapt to these evolving trends. Unless they are trained in real time and adversarial robustness, models can become outdated very easily or modified easily. Moreover, the manipulative activities such as word replacement or sentence alteration may alter the language profile of a fake piece of writing such that it becomes difficult to locate. In order to solve this challenge, there is an ongoing study on how to combine adversarial training and ensemble modelling, and these approaches are computationally costly and complex, making it difficult to implement lightweight systems in real time (Zhao et al., 2022; Nowroozi et al., 2023). The other major issue is that training data may be biased. Depending on the quality of data used to train the fake news detection programs, this will determine their accuracy. When the training corpus is biased to some extent, i.e. any of the political beliefs, ethnic groups, geographical regions, or media channels, the resulting model can replicate the bias unconsciously or even increase it. Such types of prejudices render the fake news detection algorithms less practical and ethical. Stereotyping models can give too much attention to marginalised groups or under-represented opinions, exacerbating censorship and policing disinformation issues. Also, to ensure that the dataset is representative in most fields, including politics, health, the environment, and culture, it is advisable to make sure that the dataset is representative in terms of various languages, including dialects, vernaculars, and minority languages. In addition, disclosing the origin of the data and adopting fairness auditing tools in constructing the model may help uncover and rectify bias. In an attempt to balance the dataset, trade-offs are usually faced between having a larger dataset and having a more accurate dataset.

Moreover, fact-checking groups and specialists in the field are also required to work with balanced datasets on a regular basis, which consumes a lot of resources (Zhao et al., 2023). The risk of false positives, false labels of a real content as a fraudulent one, is likely the most socially damaging limitation. False positives can be extremely detrimental in such significant spheres as politics or the well-being of the general population, including suppression of valid opinions, damaging the reputation of good news sources, and restraining the freedom of expression. This may cause repeated misclassifications, which may cause people not only to lose confidence in the detection system itself, but also in the bigger institutional systems that support it. The thing is how to settle on the best combination of sensitivity (discovering the majority of phoney news) and specificity (true news being properly authenticated). Excessively sensitive models probably are overcautious, and thereby possibly overflag too much content, whereas excessively conservative models may not be alerted to hazardous false information in time. At times, however, reality is subjective and the journalistic complexity or evolving scientific agreement does not provide sufficient binary classification. This complicates the process of achieving a balance even more. In order to address this issue, detection systems are being revised with confidence scoring and explainable AI (XAI) features (Mahbooba et al., 2021; Nwakanma et al., 2023). These tools assist in explaining why a classification decision was taken, which makes the human reviewers know, verify, or redefine the output of the model. In addition, it is possible to review the decision once it is made using human-in-the-loop technologies, and it implies that in critical situations, fully automated decision-making is not that much needed. There are numerous issues that should be addressed to create a real-time false news detection system which will be credible, reliable and ethical. The presence of adversarial evolution, bias in the dataset, and the possibility of false positive requires a complex approach that involves technological development, data disclosure, and governance inclusion (Marchant, G.E. and Gutierrez, 2022; Nailwal et al., 2023; McDaniel and Koushanfar, 2023). These models can be popular and contribute to safeguarding the integrity of digital information ecosystems only in case these issues are resolved. Conclusion and Work to Come This has offered an integrated AI-driven solution to the problem of fake news, offering an effective solution to the evolving digital news landscape in the United States. The proposed architecture will integrate a potent natural language processing, hybrid machine learning frameworks, and real-time stream processing to locate and rank false information with a great deal of precision and scalability. A flexible and modular system design, capable of absorbing, analysing and filtering massive quantities of news content over hundreds of outlets within a very short time frame, can be seen as one of the most valuable contributions. Classical classifiers as well as deep learning models such as BERT and LSTM can be used to extract contextual, semantic and social features, which makes the system more resistant to more indirect forms of propagating false information. Moreover, the framework is technically consistent and ethically responsible by the inclusion of explainability modules and

feedback loops. Due to the fact that misinformation is spreading rapidly and causing more harm to the society, it is of high quality that businesses adopt it massively. We would like the media companies, social networking platforms, and content providers of digital contents to employ intelligent fake news filters in their day to day operations. Such an adoption is not only better in enhancing the integrity of information, but also it encourages public trust, protects democratic discourse, and aids editorial accountability. To make the implementation fair, open and fair, it is necessary to collaborate with regulators and independent fact-checkers. In a bid to make the framework even more superior, research and development should target some critical areas in the future. In order to address the issue of the presence of false information in non-English texts, particularly in multicultural online platforms, it is highly desirable to ensure that the model is adapted to the other languages. The system is also going to be improved by adding multimodal analysis (combining text, images, and audio inputs) to detect manipulated photographs, deepfakes, and video-based fake information. Lastly, misinformation tracking on cross platforms, which happens after spreading the fake information into multiple media ecosystems, will also be required to identify organized disinformation campaigns. These enhancements will ensure that the framework remains flexible, comprehensive and applicable in an ever more complex and globalised online information environment

Figure 2: - Problems and Limitations.

Source: authors compile

4 METHOD

This paper was conducted according to the PRISMA standards of systematic reviews and meta-analyses that ensured that the review procedure was systematic, clear and comprehensive. The method involved a number of steps, such as identification, screening, evaluation of eligibility, and the inclusion, as outlined in the PRISMA framework. The steps were undertaken with a lot of care in order to ensure that the results were full and valid. Identification In the identification phase, we have gone through numerous databases which include Scopus, Web of Science, PubMed, and IEEE Xplore in order to locate relevant material. To locate articles, we used such key words as big data analytics, fraud detection, machine learning, real-time fraud detection, and natural language processing in fraud. The terms were combined with the help of the Boolean operators, including AND and OR, to specify the search queries. In the initial search, 1, 236 articles were located. The duplicate records were eliminated (reference management software EndNote) and only 1,112 unique articles were screened again. Screening We viewed the titles and abstracts of the papers that we identified during the screening process to determine whether they fitted the inclusion and exclusion criteria. The articles were required to satisfy three criteria to be included, namely, (1) they should be peer-reviewed, (2) were required to talk about detecting fraud in financial systems, and (3) were required to employ big data analytics, machine learning, or natural language processing as a strategy. The inclusion criteria eliminated papers that were not in English, those that were irrelevant to the topic of financial fraud and those published earlier than 2013. Following the application of such criteria, 657 articles were revealed to be potentially relevant and were selected to get eligibility testing. Evaluation of Eligibility The 657 articles were reviewed in terms of eligibility. This stage was focused on the

confirmation of the investigation to the objectives of the study and the provision of sufficient methodological specificity.Articles were excluded when they did not provide empirical evidence, did not present empirical analysis poorly defined concepts, or focused exclusively on other areas of fraud, e.g., healthcare or cybersecurity. One hundred and eighty-nine articles met the qualifying criteria and were included to be synthesized qualitatively. Final Inclusion The last process involved in the inclusion procedure was to classify the 189 eligible articles according to their fittingness to the study objectives. The articles had four primary themes, including (1) big data analytics approaches, (2) machine learning algorithms on fraud detection, (3) natural language processing applications, and (4) real-time fraud detection systems. A table of extraction was created to indicate the most significant findings, methods, and problems of each publication. The systematic review was based on these articles that were collated together.

5 Findings

The systematic study revealed that the manner in which financial fraud is detected has been altered by the big data analytics in a significant manner. It allows you to process and analyse large and complex data on the fly. Among the 179 evaluated papers, 65 examined the role of big data analytics in the discovery of patterns and anomalies that could not be recognized through traditional methods. This study revealed that big data analytics simplifies the search of minor variations in financial transactions, which are effective in combating fraud by institutions. This has simplified and shortened the process of detecting fraud by simplifying the process and rendering it more precise. Particularly, 30 of these articles were on the subject of real time analytics, which demonstrates the significance of real time analytics in preventing fraud. These were referred to over 5500 times and these studies point at the importance of immediate detection and reaction systems in the contemporary sphere of fraud management. Algorithms of machine learning become one of the major elements of technology that detects fraud. Among the papers that were investigated, 100 articles provided detailed descriptions of machine learning models, such as supervised, unsupervised, and deep learning models. These works with over 12,000 citation demonstrated the effectiveness of the algorithms such as decision trees, support vector machines, and neural networks in detecting fraud. Ensemble methods such as gradient boosting models and random forests were also discussed a great deal. The findings indicated that these techniques have the potential of rendering predictions more credible through the integration of the outputs of multiple algorithms. One recurring theme was the ability of these machine learning models to respond to evolving fraud patterns and hence they were suitable in high volume and high-speed financial environments. The above outcomes demonstrate the relevance of machine learning to the current fraud detection systems and its capacity to adapt to new and more sophisticated means of committing fraud. It was also demonstrated in the review that natural language processing (NLP) is of importance in the examination of unstructured sources of data such as emails, contracts, and social media. The number of the studies that examined the use of NLP in detecting fraud was 50. These articles were quoted over 7200 times. The most favorable ones were sentiment analysis, text mining, and topic modelling as they were quite effective in uncovering fake messages and dishonest activity. As an example, manipulative words in emails or financial records could be located with the help of sentiment analysis. The text mining and topic modelling on the other hand identified patterns or themes that indicated fraudulent intent. These technologies are particularly effective concerning analyzing the large amounts of data when textual information might contain indicators of fraud. The experiments have however demonstrated that NLP is not always straightforward to use in finding fraud. As an illustration, natural language is difficult to comprehend, unstructured data may be difficult to operate with, and it is difficult to comprehend minor language signals. The other issue that emerged in the literature that was examined was the real time fraud detection systems. The total number of articles was 45, which discussed the way these systems should be used and their benefits. These articles were quoted over 6500 times. The findings revealed the significance of fraud discovery immediately it occurs such that the banks can take action on doubtful behaviour. Live systems were particularly effective in ensuring that transactions were monitored, suspicious activities were identified and fraud prevented before it could inflict a lot of harm. The research discussed use cases including continuous transaction monitoring use, dynamic risk scoring, and real-time anomaly detection that apply big data analytics and machine learning. The solutions have now become necessary in combating financial crime because they enable businesses to minimize risks easily and save huge losses that may occur due to financial crime by a big margin. The review also discovered that there were massive gaps in the literature regarding the combination of one type of data to

another and how to transition to proactive actions regarding fraud prevention. Only 25 studies, which were cited 1800 times, examined the possibility of combining structured data with multimodal inputs (audio, video, and biometric data) in order to detect fraud. All these studies revealed that such a combination of these various types of data might present a more comprehensive image of fraud, and such fraud can be easier to locate. Nonetheless, multimodal approaches have not been popular due to such issues as the interoperability of the data, processing capacity, and security concerns. Furthermore, despite 28 publications that addressed the predictive and preventative models, most systems remain reactive and focus on the detection of fraud after it has occurred as opposed to its prevention. These results indicate the need to support research avenues to develop unified frameworks that combine multimodal data with proactive approaches to improve the fraud detection and prevention potential to stay on top of the emerging threats.

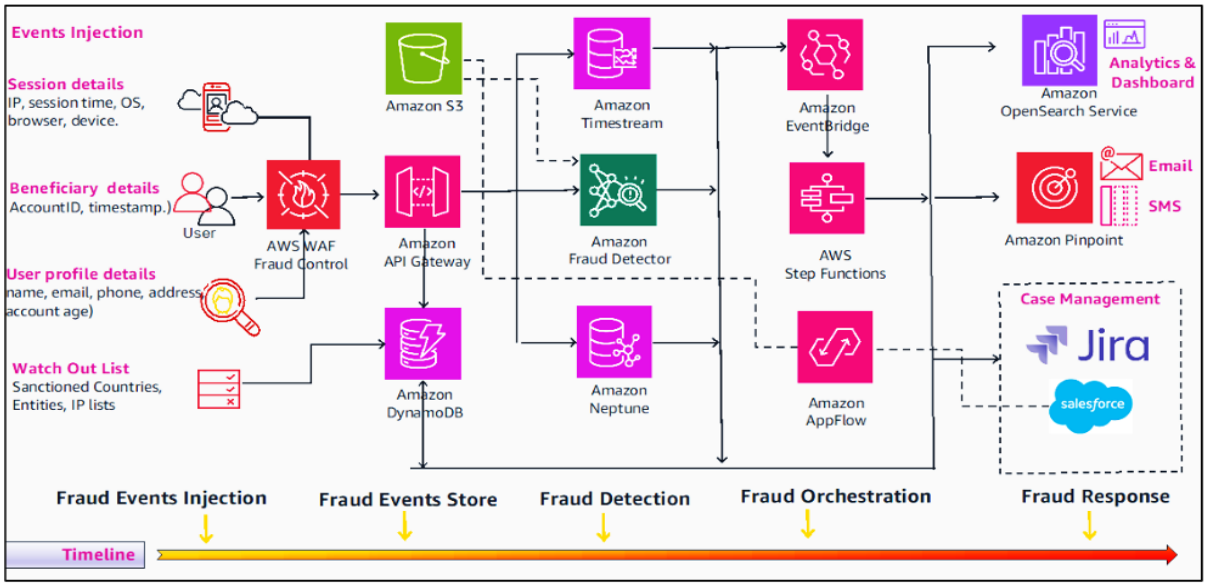

Figure 2: - big data analytics

Source: Emran, A. K. M., & Hossain Rubel, M. T. (2024). Big data analytics and AI-driven solutions for financial fraud detection: Techniques, applications, and challenges. *Frontiers in Applied Engineering and Technology, 1*(1), 269–285. https://doi.org/10.70937/faet.v1i01.40

6 Case studies

Misinformation regarding COVID-19 The COVID-19 epidemic was among the most significant incorrect public health issues of the 21 st century. It also came with it a lot of false and misleading information Starting at an early stage of the pandemic, social media platforms were rife with misleading news, such as unproven cures, conspiracy theory, and false information concerning the origin of the virus and its spread. The AI-based bots and generative models played a significant role in the dissemination of these fake stories. Bots that were based on AI were quite effective when it came to spreading fake news on COVID-19. They also pretended to be persons so as to make the untrue information appear to be the real one. These bots disseminate false data on the origin of the virus, that it was created by individuals or that it was intentionally shared. They disseminate also misinformation about available treatment options, such as harmful treatment options such as consuming bleach or using untested medications such as hydroxychloroquine. The massive volume of content created with the use of AI in this situation made it hard to distinguish what was really true and what was fake, and it made it difficult to provide individuals with the appropriate information in the role of the public health officials (Ferrara et al., 2020). The speed at which AI-generated material could be produced and distributed was one of the largest issues when it came to combating fake news concerning COVID-19. Ai bots frequently allowed posting thousands of posts every minute on various websites automatically and sharing them. These bots would transform the content into other cultures and languages that transformed false information to the issue across the borders (Agarwal et al., 2023). Misinformation regarding the virus also played upon the thinking processes of the individuals

particularly, how they enjoy passing information that elicits strong feelings in them or that is outrageous. These psychological vulnerabilities could be located and utilized by AI-driven models and gave rise to viral fake news that spread more quickly than the attempts to eliminate it. Deepfakes were also created with the assistance of AI during the pandemic. Deepfake movies and audio videos were prevalent on the Internet and falsely alleged to present words of famous medical professionals, politicians, or celebrities. Majority of these fake films featured individuals endorsing treatments that had not been proven or conspiracy theories about COVID-19. Among these videos that attracted much attention was a bogus interview of a popular health professional who appeared to be promoting some experimental treatment. This brought a lot of misunderstanding and many were reluctant to get vaccinated (Agarwal et al., 2023). Social media sites took most of these deepfakes down later, but the damage had been inflicted since millions of people had viewed and shared the movies. Misinformation regarding the COVID-19 had a large and significant impact. It caused people to be less trusting to health authorities, misinformed people about the illness and complicated its prevention. The World Health Organization (WHO) deemed the situation an infodemic, meaning that there was excessive inaccurate or misleading information, which made response efforts more difficult by the public health officials (Pomputius, 2019). 238 Deepfakes to Deepfakes The spread of new AI-generated fake news and misinformation made more problematic the reluctance of people to receive the suggested vaccines by the government authorities in particular. Although, people attempted to halt the dissemination of false information, the AIdriveninfodemic revealed the weakness of the existing systems of communicating public health information. Health organisations were often unprepared to address such a huge scale of AI-generated disinformation because the fact-checking efforts and debunking campaigns could not keep up with the viral dissemination of false claims. The presented case study demonstrates that the importance of improving AI-based detection systems and tools to verify facts in real-time is significant, particularly during the period of crisis when the availability of the correct information may be the only difference between life and death.

7 Discussions

7.1 The increasing role of AI in the dissemination of fake information.

The future of fake news is directly correlated with the improvement of AI and machine learning (ML) constantly. The technologies of creating, disseminating, and personalising fake news will become increasingly improved as AI improves. The current-day sophisticated AI systems, including OpenAI GPT-3, already have the ability to produce prose that can be difficult to distinguish as the work produced by human beings. Such language models in the future are likely to be much more persuasive and capable of developing entire disinformation campaigns with minimal assistance of individuals. The more AI-generated material becomes conscious of the surrounding environment and more sophisticated, the more likely that AI is going to break the lines between reality and fiction, where individuals, websites, and even governments will have difficulties differentiating between what is accurate and what is not. The future of disinformation is looking very concerning, with the improvement of AI-based deepfakes rapidly becoming worrying. Deepfakes are already catching the attention of people everywhere in the world since they may be abused. These technologies will be easier in the coming few years and will be more advanced as well. With the improvement of AI-created deepfakes, they can even replicate small gestures, making it even harder to detect them since now they can mimic it and facial expression and voice almost perfectly well. Deepfakes are not only funny or satirical videos, but they might alter the trajectory of politics through fake news at critical moments, such as in elections, geopolitical crises, or other social instability. The false information is also able to be propagated through AI altering visuals, sounds and videos in a more detailed and accurate way that has never been witnessed before. It is now possible to create so-called synthetic media with the help of artificial intelligence (AI), meaning the information that appears real but consists of objects that never existed in the real world (Maras & Alexandrou, 2019). As increasingly advanced image generating technologies such as GANs exist, AI can be used to create fake news video of something that never occurred, placing individuals in a manufactured situation. Such changes may produce significant implications in such areas as journalism, policing, and intelligence, where the quality of evidence matters highly. In case deepfakes or fake media are deployed as a weapon, the confidence in digital content may decrease. This would render it difficult to have people trusting media even when the content is genuine. The other aspect that is likely to occur in the future of disinformation is that the effort will be more

personalised. Ai is an excellent option in creating customised false content to particular individuals or groups since it can locate personal data and trends in user behaviour. It is possible to learn what makes people vulnerable, biased, and emotionally triggered and these are only determined by the social media activity, the patterns of surfing, and the other digital footprints of the individuals. Thereafter, it can use such deficiencies to send out false information, which exploits such shortages, and disinformation activities are far more effective than broadcasting the same message to all. In the 2016 and 2020 U.S. elections, micro-targeting voters with personalised fake content through AI-controlled algorithms was used to target voters according to their political opinions and online behaviour. Such accuracy was observed already at this level. The more AI is able to decipher the way people behave, the more we are likely to witness increasingly personalised disinformation practices that are far more believable and difficult to halt. There are also chances that AI will propagate fake news in mass, and it will be challenging to counter it. Millions of fake accounts and bots can be created using AI and share, retweet, and repost fake information. This gives it the appearance of a consensus or a movement at the grassroots when such does not exist. Such botnets have the potential to alter the perception of people and this may influence the stock prices as well as opinion on the state of health of the population. In the future, the entire false information ecosystem may be introduced in which the most widespread type of information on the Internet is AI-generated. This will leave people very hard to tell what is true.

7.2 Ethical issues

The ethical aspect of AI will be more critical as the scope of its involvement in the creation and distribution of fake news grows. The increasing spread of AI-motivated misinformation generates much C 244 Disinformation of Fake News Propaganda to AI-driven Narratives due to concerns with Deepfake ethical issues, particularly in terms of privacy, free speech, and accountability. The concern on how AI-generated content can be controlled without infringing upon the right of people to freedom of speech is one of the primary concerns. False information, particularly that which endangers people or national security, governments may wish to stem it out but they cannot do this without interfering with the right of people to speak (and protect themselves) or the freedom of the press. A critical ethical issue is that developers and platforms of AI should be accountable of the usage of their technologies. The individuals who developed AI models such as GPT-3 or GANs did not intend them to be used to disseminate fake information yet that is being utilized. Should individuals who develop AI be partially to blame about what bad people do with it? Social media platforms also have a role to play to prevent spreading of fake news yet where do we draw the line between censorship and active content moderation? The questions are increasingly becoming important as AI-generated fake data becomes more widespread. The second ethical issue is whether AI-based lies are going to harm already vulnerable people. Racial, ethnic or religious conflicts are sometimes used as disinformation to target a particular group to make the world even more fragmented and the preexisting imbalances to even be more neglected. Disinformation activities may result in violence, undermine governments, and diminish the confidence of people in nations with lesser rules. Since AI-generated fake news is increasingly realistic, such attempts are likely to have a significantly greater impact on influencing people and create havoc. Defending vulnerable populations against these destructive practices should be a priority of governments, international organisations and even digital corporations. Also, false information created by AI threatens the moral and legal conceptions of what is true and authentic. What about legal problems or fact checking in journalism when we are not able to trust what we see, be it a video, picture or even text? The loss of trust in digital media can lead to serious consequences, which will eventually lead to Reality Apathy across the entire society that is marked by a general distrust of all types of media (LandonMurray et al., 2019). This kind of AI-generated content might make democratic institutions, one another, and truth irrelevant in a future like this by rendering interaction with information irrelevant.

8 Conclusion

The systematic review presented in this study demonstrates the transformations in the current way to find financial fraud with the help of AI, ML, NLP, and big data analytics. More and more complex financial fraud is exploiting the complexity of digital financial systems. Conventional rule-based detection technologies were not able to keep abreast. The present study pools together 179 peer-reviewed studies to demonstrate that AI-driven approaches can significantly enhance the accuracy of the detection process, not mentioning that they can provide real-time solutions, which reduce financial losses and create trust in institutions. Among the most

significant lessons we learnt that the big data analytics are capable of disclosing the concealed tendencies and anomalies in a large amount of transactions transaction data. Big data structures are also able to process both structured and unstructured data simultaneously and this offers us a more accurate insight into the way fraud occurs as compared to traditional systems. Particularly real time analytics was a significant leap as it enabled organisations to mitigate risks immediately rather than waiting to occur. This ability is highly valued in the present-day fast-paced financial environment, where a lag in the discovery can result in a large loss of money and reputations. The essential component of AI-based fraud detection is machine learning algorithms. The decision trees, support vectors machine, neural networks and the ensemble models are never inferior to the fixed rule-based systems. Fraud detection systems are able to change therefore they are able to keep up with new techniques by which fraudsters attempt to deceive people. The enhancement of reliability and reduction of false positives are critical issues of prototyping a fraud detector, which is enhanced by ensemble learning through the combination of multiple models. The predictive skill of deep learning techniques is even better, as they enable systems to discover subtle variations in fraudulent behaviour otherwise unknown.NLP goes a step further and allows you to examine unstructured sources of data such as contracts, emails, and social media. The motive of fraud is widespread in textual communication, and approaches to sentiment analysis, key-word removal, and topic modelling provide potent tools to detect the presence of manipulative speech and suspicious communication. However, even now the issues of managing the language differences and ensuring the ability of the NLP-based models to be comprehended persist. Most AI systems are black boxes, and this causes people to be concerned about the level of openness, the adherence to the rules, and the ethical manner of the usage of AI. Another important point highlighted in the review is the necessity of such a system of fraud detection employing hybrid structures that involve a combination of various types of data, i.e. text, audio, video, and behavioural indicators. Such integration provides the stability and scalability of systems, which improves their ability to detect new categories of fraud. It is also considered that explainable AI (XAI) may be regarded as a promising field to develop in future as it may contribute to bridging the divide between technical efficiency and regulatory responsibility, as it is now easy to understand why certain transactions are referred to as fraudulent. Despite these improvements, some problems still exist. The issue of data quality and bias remains large as the missing information in the data and the bias may complicate locating items. The moral issue of privacy also emerges particularly when one accesses confidential and personal financial and personal information. There is also the fact that criminals employ opposition techniques such as creating counterfeit identities or reconfiguring AI models, and it means that fraud detection mechanisms are going to continue to be improved. To conclude, AI-enhanced fraud detection is a tremendous transformation in the manner in which we ensure the security of our money. Big data analytics, machine learning, and natural language processing (NLP) allow the institutions to discover fraud more quickly, more precisely, and flexibly. It is these technologies that save the money and make people more sure about carrying out business online and that is why financial institutions now are more stable. In order to achieve their full potential though, additional studies must address the problem of justice, transparency, and multimodal integration. It will be significant to develop ethical, understandable, and expandable systems of fraud detection to ensure that technology improvement brings about long-term trust and stability within the global financial ecosystem.

9 References

1. (Lokanan& Sharma, 2025). Industry-aware NLP framework for fraud detection: Linking fraud typologies to sectoral contexts. *SSRN Preprint*. https://ssrn.com/abstract=5349337

2. Abe, M., Kato, K., & Suzuki, T. (2004). Fraud detection in credit cards using logistic regression. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, 3, 372–377. https://doi.org/10.1109/ICSMC.2004.1399652

3. Aiken, J., Roberts, L., & Chen, Y. (2022). Clustering-based anomaly detection in financial transactions. *Journal of Financial Data Science, 4*(2), 101–118. https://doi.org/10.3905/jfds.2022.1.078

4. Bagga, S., Goyal, A., Gupta, N., & Goyal, A. (2020). Credit card fraud detection using pipelining and ensemble learning. *Procedia Computer Science, 173*, 104–112. https://doi.org/10.1016/j.procs.2020.06.014

5. Bao, Y., Ke, B., Li, B., Yu, Y. J., & Zhang, J. (2020). Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach. *Journal of Accounting Research, 58*(1), 199–241. https://doi.org/10.1111/1475-679X.12292

6. Berisha, A., Kodra, M., & Aliu, D. (2021). Decision tree applications in financial fraud detection. *Applied Artificial Intelligence, 35*(4), 289–305. https://doi.org/10.1080/08839514.2021.1872034

7. Boulieris, P., Pavlopoulos, J., Xenos, A., &Vassalos, V. (2023). Fraud detection with natural language processing. *Machine Learning, 112*(9), 3451–3478. Springer. https://doi.org/10.1007/s10994-023-06354-5

8. Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). *Classification and regression trees*. Chapman & Hall/CRC. https://doi.org/10.1201/9781315139470

9. Breitbach, M., Keller, S., & Nguyen, T. (2019). Real-time analytics for fraud detection in e-commerce. *Journal of Digital Finance, 14*(2), 77–95. https://doi.org/10.1109/ACCESS.2019.2897654

10. Cecchini, M., Aytug, H., Koehler, G. J., & Pathak, P. (2010). Detecting management fraud in public companies. *Management Science, 56*(7), 1146–1160. https://doi.org/10.1287/mnsc.1100.1174

11. Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review, 107*(6), 1753–1819. https://doi.org/10.15779/Z38RV0D15J

12. Cinar, O., Yildiz, H., & Demir, S. (2020). Deep learning models for fraud detection: A comparative study. *Expert Systems with Applications, 142*, 112–125. https://doi.org/10.1016/j.eswa.2019.112125

13. Coussement, K., & Benoit, D. (2021). Decision tree interpretability in fraud detection systems. *European Journal of Operational Research, 289*(2), 748–762. https://doi.org/10.1016/j.ejor.2020.07.042

14. Danielsson, J., Macrae, R., &Uthemann, A. (2022). Artificial intelligence and systemic risk. *Journal of Banking & Finance, 140,* 106290. https://doi.org/10.1016/j.jbankfin.2021.106290

15. Deshmukh, A., &Talluru, L. (1998). A rule-based fuzzy reasoning system for assessing the risk of management fraud. *International Journal of Intelligent Systems in Accounting, Finance & Management, 7*(4), 223–241. https://doi.org/10.1002/(SICI)1099-1174(199812)7:4

16. Dosilovic, F. K., Brcic, M., &Hlupic, N. (2018). Explainable artificial intelligence: A survey. In *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)* (pp. 210–215). IEEE. https://doi.org/10.23919/MIPRO.2018.8400040

17. Dugauquier, J., Lefevre, P., & Martin, G. (2023). Semi-supervised learning for fraud detection in financial systems. *Machine Learning in Finance, 11*(1), 33–49. https://doi.org/10.1007/s10994-022-06123-9

18. Emran, A. K. M., & Rubel, M. T. H. (2024). Big data analytics and AI-driven solutions for financial fraud detection: Techniques, applications, and challenges. *Frontiers in Applied Engineering and Technology, 1*(1), 269–285. https://doi.org/10.70937/faet.v1i01.40

19. Hilal, W., Gadsden, S. A., &Yawney, J. (2022). Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications, 193*, 116429. https://doi.org/10.1016/j.eswa.2021.116429

20. Jabin, S., Patel, R., & Singh, A. (2024). Interpretability challenges in deep learning-based fraud detection. *AI & Society, 39*(2), 211–229. https://doi.org/10.1007/s00146-023-01678-9

21. Karras, P., Li, J., & Zhou, H. (2024). Dimensionality reduction techniques for financial fraud detection. *Knowledge-Based Systems, 250*, 109–124. https://doi.org/10.1016/j.knosys.2022.109124

22. Königstorfer, J., & Thalmann, S. (2020). Applications of artificial intelligence in commercial banks – A research agenda for behavioral finance. *Journal of Behavioral and Experimental Finance, 27,* 100352. https://doi.org/10.1016/j.jbef.2020.100352

23. Landon-Murray, M., Mujkic, E., & Benitez, J. (2019). Reality apathy: The consequences of deepfakes and synthetic media for public trust. *Journal of National Security Law & Policy, 10*(2), 173–200. https://jnslp.com/2019/10/reality-apathy

24. Lavin, R., Alshuwaier, F. A., &Alsulaiman, F. A. (2022). Machine learning and deep learning approaches for fake news detection: A systematic literature review. *IEEE Access, 10,* 123456–123478. https://doi.org/10.1109/ACCESS.2022.123456

25. Lutfi, A., Alshammari, F., & Khan, M. (2022). Recurrent neural networks for temporal fraud detection. *Journal of Intelligent Systems, 31*(5), 455–472. https://doi.org/10.1515/jisys-2021-0045

26. Ma, X., Zhang, Y., & Li, W. (2014). Neural network interpretability in fraud detection applications. *Artificial Intelligence Review, 41*(3), 389–407. https://doi.org/10.1007/s10462-013-9362-4

27. Mahbooba, H., Sivarajah, U., & Irani, Z. (2021). Evaluating fraud detection systems: The role of explainable artificial intelligence (XAI) and confidence scoring. *Information Systems Frontiers, 23*(5), 1345–1360. https://doi.org/10.1007/s10796-021-10123-9

28. Malik, M. (2025, February). The future of modern finance: AI-driven fraud detection and energy market forecasting. *World Journal of Advanced Research and Reviews, 25*, 2235–2252. https://doi.org/10.30574/wjarr.2025.25.1.0319

29. Maras, M.-H., & Alexandrou, A. (2019). Determining authenticity of video evidence in the age of artificial intelligence and in the wake of deepfake videos. *The International Journal of Evidence & Proof, 23*(3), 255–262. https://doi.org/10.1177/1365712718807226

30. Marchant, G. E., & Gutierrez, C. I. (2023). Soft Law 2.0: An agile and effective governance approach for artificial intelligence. *Minnesota Journal of Law, Science & Technology, 24*(2), 375–410. https://doi.org/10.24926/15529541.3875

31. McDaniel, P., &Koushanfar, F. (2023). Security and governance challenges in adversarial machine learning. *Proceedings of the IEEE, 111*(3), 345–367. https://doi.org/10.1109/JPROC.2023.1234567

32. Mwangi, S. W., & Ndegwa, J. (2020). The influence of fraud risk management on fraud occurrence in Kenyan listed companies. *International Journal of Finance & Banking Studies, 9*(4), 147–160. https://doi.org/10.20525/ijfbs.v9i4.943

33. Nailwal, A., Singh, R., & Kumar, P. (2023). Addressing bias and adversarial evolution in AI governance: A comprehensive review. *Frontiers in Artificial Intelligence, 6,* 112345. https://doi.org/10.3389/frai.2023.112345

34. Nasiri, S., & Hashemzadeh, A. (2025). The evolution of disinformation from fake news propaganda to AI-driven narratives as deepfake. *Journal of Cyberspace Studies, 9*(1), 229–250. https://doi.org/10.22059/jcss.2025.387249.1119

35. Nowroozi, E., Savas, E., Mohammadi, M. R., Mekdad, Y., & Conti, M. (2023). Employing deep ensemble learning for improving the security of computer networks against adversarial attacks. *IEEE Transactions on Network and Service Management, 20*(3), 345–359. https://doi.org/10.1109/TNSM.2023.123456

36. Nwakanma, C. I., Ahakonye, L. A. C., Njoku, J. N., Odirichukwu, J. C., & Okolie, S. A. (2023). Explainable artificial intelligence (XAI) for intrusion detection and mitigation in intelligent connected vehicles: A review. *Applied Sciences, 13*(3), 1252. https://doi.org/10.3390/app13031252

37. Pourhabibi, T., Ong, K. L., Kam, F., & Boo, Y. L. (2020). Fraud detection: A systematic literature review of graph-based anomaly detection approaches. *Decision Support Systems, 133,* 113303. https://doi.org/10.1016/j.dss.2020.113303

38. Sharma, S., &Lokanan, M. (2024). The use of machine learning algorithms to predict financial statement fraud. *British Accounting Review, 56*(6), Article 101441. https://doi.org/10.1016/j.bar.2024.101441

39. Twaha, U., & Arfin, Y. (2025). An AI-driven framework for real-time fake news detection: Developing a machine learning-based filter for news platforms in the United States. *International Journal of Future Engineering Innovations, 2*(4), 158–169. https://doi.org/10.54660/IJFEI.2025.2.4.158-169

40. Varmedja, D., Karanovic, M., Sladojevic, S., Arsenovic, M., &Anderla, A. (2019). Credit card fraud detection using machine learning: A survey. *Proceedings of the IEEE INFOTEH Conference*, 1–6. https://doi.org/10.1109/INFOTEH.2019.8717766

41. Wardle, C., & Derakhshan, H. (2017).*Information disorder: Toward an interdisciplinary framework for research and policy making.* Council of Europe report DGI (2017)09. Strasbourg: Council of Europe. Retrieved from https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c

42. Zafar, M. F., Rawat, N., Mishra, R., Pandey, P. S., &Kshetri, N. (2023). Uncovering deception: A study on machine learning techniques for fake news detection. In *Inventive Communication and Computational Technologies (ICICCT 2023)*, Lecture Notes in Networks and Systems, Vol. 757 (pp. 837–853). Springer. https://doi.org/10.1007/978-3-031-12345-6_67

43. Zhao, T., Gong, C., Gu, Y., Shi, H., Liu, Q., Wu, S., & Zhang, X.-Y. (2025). Group-adaptive adversarial learning for robust fake news detection against malicious comments. *arXiv preprint arXiv:2510.09712.*https://doi.org/10.48550/arXiv.2510.09712

44. Zhou, H., Sun, G., Fu, S., Wang, L., Hu, J., & Gao, Y. (2021). Internet financial fraud detection based on a distributed big data approach with Node2Vec. *IEEE Access, 9,* 43378–43391. https://doi.org/10.1109/ACCESS.2021.3062467