A Review on Voice-based Scam Detection Using Machine Learning and Speech Analytics

Mr. Ram Pratap Singh
Department of Computer Science and Engineering,
Lakshmi Narain College of Technology,
Bhopal
rampratap@lnct.ac.in

To Cite this Article
Mr. Ram Pratap Singh, "A Review on Voice-based Scam Detection Using Machine Learning and Speech Analytics", Journal of Science Engineering Technology and Management Science, Vol. 03, Issue 05, May 2026, pp: 271-279, DOI: http://doi.org/10.64771/jsetms.2026.v03.i05.pp271-279
Submitted: 08-04-2026 Accepted: 12-05-2026 Published: 19-05-2026

***Abstract*—Voice-based scams have become a serious cybersecurity threat because people increasingly adopt artificial intelligence and communication technologies. Fraudsters use vishing, and robocall and deepfake voice attacks to trick victims into disclosing their personal information and authorizing fake transactions. The review presents an in-depth examination of the evolving operating environment of voice scams and the development of detection and prevention technologies. The study examines various types of voice-based fraud, with a focus on AI-generated voices and social engineering techniques. The research paper investigates methods for enhancing speech signals by removing noise and processing audio to improve sound quality and intelligibility for precise analysis. The review examines the various analytical features that detect scams through acoustic, linguistic and semantic patterns, as well as speaker verification and emotion analysis capabilities. The paper investigates how ML techniques, such as supervised learning, unsupervised learning, DL and hybrid ensemble models, detect fraudulent voice interactions. Modern voice scam detection systems use speech processing, NLP and cutting-edge machine learning techniques to detect suspicious calls in real phone conversations. The review presents information about existing obstacles, recent progress and upcoming research paths for voice-based scam detection systems.**

***Keywords*—Voice Spoofing Detection, Telecom Fraud, Speech Analytics, Natural Language Processing (NLP), Machine Learning.**

I. INTRODUCTION

Victims of voice phishing, a type of social engineering crime, are lured in by phone calls [1] and voice messages with the promise of financial and personal information by playing on their vulnerabilities and the confidence they have in others [2]. The second aspect of cybercrime, which includes online scams and fraud behaviour [3] represents the final stage of criminal activity [4]which begins when offenders establish connections with victims through social media platforms. Scammers and fraudsters, in particular, are able to access cardholder cash if they have successfully stolen bank card data through hacking, phishing, or skimming [5]. Cybercriminals use malicious software to steal identity information from their victims by operating from remote locations through hidden links[6][7]. The research demonstrates that criminals use multiple methods [8] of manipulation, which include caller ID spoofing and authority figures [9] impersonation and time pressure tactics, with personalized conversation scripts.

The ever-increasing volume of instances and the corresponding quantities of money involved in online fraud [10] offer a significant danger to the safety of individuals, their property, and the stability of society and the economy [11][12]. In 2021, there were about USD 39.89 million in losses to global telecoms income due to fraud, according to the Communications Fraud Control Association [13] [14]. The cybersecurity chain is being weakened by the rapid increase of social engineering assaults in modern networks [15][16]. For the benefit of cybercriminals, they seek to deceive individuals and businesses into disclosing sensitive and valuable information [17][18]. All networks are vulnerable to social engineering [19][20], and anti-virus software systems [21], regardless of the effectiveness of their firewalls, cryptographic approaches, and intrusion detection systems. People have a higher tendency to trust one another than they do computers or other forms of technology. Figure 1 shows the social engineering attack life cycle.

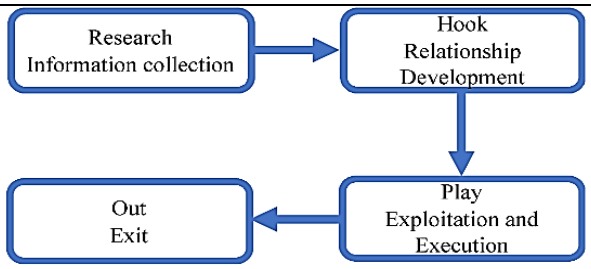

Fig. 1. Social engineering attack stages [22]

The rise of mobile devices as main tools for communication established a new way for people and companies to connect with each other. The digital transformation[23][24][25] which brought new technological advancements to the world has created fresh opportunities for fraudsters who target unprotected groups. Mobile communication fraud exists in various forms which include vishing (voice phishing), smishing (SMS phishing), caller ID spoofing, impersonation scams and advanced AI-generated voice scams [26][27]. The threat level of this danger becomes clear through the fact that mobile devices now account for 75% of all digital payment [28] fraud cases which includes 45% of global fraud incidents originating from Asia-Pacific regions.

The detection of fraud requires multiple methods for its detection. ML technologies[29] have become an effective weapon in the fight against telecom fraud[30] because they now operate in the present day [31][32]. The detection of deepfake voices requires advanced machine learning (ML) [33] and transfer learning, because synthetic audio forgeries have become more complex to detect[34]. The development of deepfake technology makes it more difficult to use traditional methods for distinguishing between authentic and fraudulent vocalizations[35]. Advanced ML models enable researchers to detect hidden audio patterns that typical people and basic algorithms cannot understand[36]. The organization can maintain its defense against fast-evolving deepfake technology through ongoing model development and updates, enabling it to protect against impersonation, fraud and misinformation attacks[37].

A. Structure of the paper

The paper's structure is as follows: Section II introduces environment of voice-based scams and deepfake detectives. Section III discusses speech signal enhancement techniques. Section IV presents ML methods of scam detection. Section V presents a literature review that includes a comparison of previous research. Section VI wraps up the work and provides a roadmap for future studies.

II. Voice-Based Scam Landscape and Deepfake Voice Detection

The system uses AI and ML together with voice biometrics to identify fraudulent activity and suspicious calling patterns through its voice-based scam detection system[38][39]. The system detects voice characteristics, which include pitch, rhythm and stress patterns, to distinguish between real human voices and deepfake artificial voices [40]. The advanced systems use natural language processing to analyse spoken content, which enables them to issue instant alerts and perform automatic responses during suspicious phone calls. The technology enhances cybersecurity [41][42] and in particular, in areas such as banking and call centers, by providing an adaptive layer of real-time protection against voice scams that change.

A. Types of Voice-based Scams

The term "telephone scams" has been used to describe one type of fraud that discussed in this section. Scammers' next-gen techniques for telephone scams that use AI were therefore investigated [43]. Various scams are discussed below:

1) Vishing (Voice Phishing)

Vishing is a type of social engineering attack that involves scammers that are performed over the phone and consequently are used to acquire personal or financial data by impersonating an entity of trust. They apply psychological tricks such as urgency and authority to coerce victims in live calls. Vishing may be either human or automatic, and AI [44] voice cloning is becoming more popular in producing convincing fake voices, attacking banks, agencies, and support services[45].

2) Robocalls and Automated Scams

Robocalls are automated systems that are used to send pre-taped messages to groups of victims. Such frauds usually express non-real threats, emergency demands, or fictitious prize wins. Although less interactive than AI-driven vishing, robocalls still deceive millions annually [46]. There is a high rate of people scamming using VoIP technology to generate multiple phone numbers and getting away with it.

3) Deepfake Voice Attacks

Deepfake voice attacks use AI to generate or imitate natural voices convincingly. Social engineering, financial fraud, extortion, and misinformation are promoted using these attacks. It is high-profile impersonation where criminals impersonate executives or other figures of prominence to approve fraudulent transactions or to otherwise manipulate their victims [47]. Deepfake audio disregards the normal verification through trust, like in a scam in the UK energy sector, and a false Elon Musk cryptocurrency video. Figure 2 shows the overview of the Deepfake Technology.

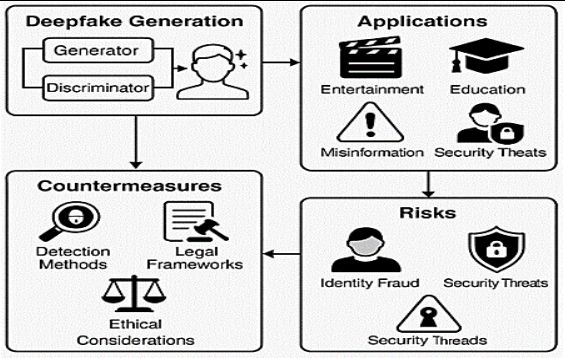

Fig. 2. Graphical Overview of Deepfake Technology

B. Key Characteristics of Voice-Based Frauds

The fast evolution of technology has led to an increase in the sophistication and complexity of telephone scams. The major characteristics of these types of fraud are thoroughly described below:

- **Urgency and Threats:** Scammers often use legal action or the suspension of accounts to make victims feel like they need to pay up quickly [48], to which the victim responds with immediate action without checking the authenticity of the email or the site carefully.
- **Spoofed Email Addresses and Websites:** Phishing email messages and spoofed websites are designed in a way that they look real. Fraudsters may also send emails or domains that seem nearly the same as the genuine organizations only with minor changes that one can easily afford to ignore.
- **Malware Deployment:** In other instances, when one clicks a link in the mail or opens an attachment, the malware installs itself and this results in the personal information or financial information being stolen without the knowledge of the victim.

The best way to fight this type of fraud is to ensure that people and organizations implement high levels of security, such as training to detect phishing attacks [49], ensuring emails are authentic before clicking links or entering information, and implementing MFA to protect accounts. Moreover, it is very important to maintain software and security systems to avoid installation of malware that may be acquired as a consequence of such fraudulent activities.

C. Applications and Implications of Deepfake

Deepfake speech recognition The concept of deepfake speech recognition describes the capability of AI [50] programs to recognize and evaluate fake or modified audio that sounds like a human voice [51]. Due to the growing development of deepfake voice technology, multiple applications and misuses of it are possible in the future, and detection of deepfakes is essential in various industries. Table I outlines certain applications of Deepfake.

TABLE I. PRACTICAL APPLICATIONS OF DEEPFAKE VOICE DETECTION SYSTEMS

Category	Application	Role of Deepfake Voice Detection
Security and Fraud Prevention	Voice Authentication	Detects fraudulent attempts in systems that use voice biometrics (e.g., banking services and smart devices) by identifying synthesized or manipulated voices.
	Phone Scam Detection	Identifies impersonation in social engineering attacks during phone calls by analyzing abnormal voice characteristics.
	Identity Theft Prevention	Helps prevent unauthorized impersonation of individuals by detecting deepfake or cloned voices.
Organizational Risk Management	Internal Communication Security	Verifies speaker identity in sensitive organizational communications such as fund transfers, access authorization, and confidential discussions.
	Executive Impersonation Detection	Prevents attacks targeting executives by detecting suspicious voice patterns generated through voice cloning.
	Risk Governance and Policy Enforcement	Supports security policies by enabling verification mechanisms such as multi-factor authentication for high-risk voice instructions.
Law Enforcement and Digital Forensics	Audio Evidence Authentication	Ensures the authenticity of voice recordings used in legal investigations and prevents the misuse of manipulated audio evidence.
	Criminal Investigation Support	Assists investigators in distinguishing between genuine and manipulated conversations in intercepted communications.
Media and Information Verification	Misinformation Detection	Identifies manipulated audio in news and social media to prevent the spread of fake information.
	Protection of Public Figures	Detects fake audio impersonating journalists or public figures, helping maintain credibility and public trust.
Voice-Activated Systems	Smart Assistants Security	Ensures requests come from authorized users, improving the security of voice-activated devices (like Alexa, Siri, and Google Assistant).
	Speech Recognition Reliability	Improves the reliability of speech recognition systems in sectors such as healthcare by verifying authentic voice inputs.

III. SPEECH SIGNAL ENHANCEMENT TECHNIQUES

Digital speech [52] signal processing includes the step of speech enhancement, which uses an algorithm or filter to make the speech signal better in terms of clarity, intelligibility, understandability, and comprehensibility. Background noise, such as babble or reverberation, can weaken the speech signal as it is being recorded. Certain speech-enabled applications, such those for mobile devices, hearing aids, and VoIP, utilize clear and uninterrupted voice signals. There are several strategies for improving speech. The methods for improving speech vary depending on the type of deterioration and the noise present in the obtained speech signal. The voice enhancement system's basic phases are depicted in Figure 3.

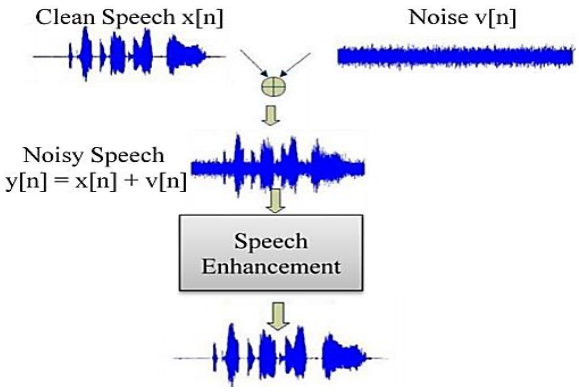

Fig. 3. Basic Steps of a Speech Enhancement System

A. Types of Noise and Its Removal Methods

This section reviews and describes several noise reduction strategies[53]. There are a variety of noises that can degrade the voice signal, such as periodic noise, wideband noise, and speech interference.

1) Periodic Noise & Its Removal Techniques

Use of transform-domain, adaptive, or stationary filters can eradicate periodic noise. One approach uses a comb filter made of a bank of notch filters, like twin T-filters, to remove periodic noise; this method is stationary. The second is adaptive filters, where periodic noise may be eliminated by using an inverse filter that is a forward prediction error filter. The third is a changing domain where the periodic noise spectrum is observable and modifiable. By examining the spectrum, the periodic components may be determined.

2) Wide Band Noise & Its Removal Techniques

The spectra subtraction method (SS) and adaptive cancellation can get rid of wideband noise. In order to remove the predicted noise spectrum from the noisy speech spectrum, spectral subtraction is employed. Adaptive cancellation can filter out background noise in a signal and can estimate the channel using the corresponding signal even when there isn't a signal. In order to remove noise, an adaptive filter can be fine-tuned by choosing an impulse response that makes the filtered channel noise equal to the signal noise. Changes are made to the coefficients until the output hits a minimum.

3) Interfering Speech & Its Removal Techniques

Speech enhancement methods are useless when two speech signals are interfering. It is possible to separate the voices of various speakers if one can distinguish between distinct pitches. Pitch separation requires the monitoring of vocal portions. A comb filter can be employed to recover the target speaker's harmonics if pitch values are known. Isolating several speakers' voices is another possible usage of a transform domain approach.

B. Speech Enhancement Methods

The following are only a few of the numerous strategies used to improve speech. They may be separated into two fundamental groups as follows: Techniques for Enhancing a Single Channel and a Multi-Channel.

1) Single Channel Enhancement Method

This technology is commonly used for real-time applications such as mobile communication and hearing aids because there is usually no second channel present [54]. While this method improves the signal's overall quality, it does so at the cost of some intelligibility and, thus, performance. Additionally, this technology is less complicated and more affordable than a multichannel system. This system typically employs various speech and unwanted noise statistics.

2) Multichannel Enhancement Method

These systems are more intricate than single-channel systems. The adaptive noise cancellation device serves as a noise reference for this system, which uses several signal inputs. Compared to single-channel systems, these systems perform better on non-stationary noises because they consider the spatial features of the signal and the noise source, in addition to the limits of the former.

C. Analytical Features Used in Voice-Based Scam Detection

Advanced AI speech creation systems are beyond the capabilities of present voice authentication and anti-spoofing defenses, which rely on acoustic structural or policy-based methods. In addition, most existing detection methods aren't practical since they don't have access to high-quality synthetic audio data.

1) Acoustic Feature Extraction (Pitch, Tone, Prosody, Spectral Features)

It implements a simulation of the simplest acoustics such as F0/pitch contours, energy, jitter/shimmer, MFCCs, spectral flux/centroid, and prosodic rhythm, and modern fraud screening can already be employed before words are recognized. According to studies, voicemail and call audio include enough discriminative signal to distinguish between spam and robocalls and human or non-spam calls [55].

2) Linguistic and Semantic Features (NLP for Scam Content)

This type of NLP pipeline scans language-based indicators of social-engineering after the speech is transcribed, including deception-style rhetoric, payment signs, entity co-occurrence patterns (brands gift cards), imperative/urgency words (as in, verify now or account locked), and imperative/urgency phrases [56]. To detect phishing and other message-based threats, reviews and comparative analyses show the effectiveness of token/character n-grams, POS-patterns, and contextual embeddings (BERT family); these can be transferred to call transcripts, chat, and chat makes them even more accurate with the assistance of keyword rules and topic models.

3) Speaker Identification and Verification

Voice biometrics enable a very strong defense against account takeover: watch-list detection can recognize a known fraudster over a conversation and speaker verification can immediately compare a voiceprint with a registered voiceprint [57]. Also, it is demonstrated in the literature that the score-level fusion and anti-spoofing front-ends are used to secure ASV against spoofing (replay, TTS/voice cloning, unfriendly examples).

4) Emotion and Sentiment Analysis in Voice Calls

Affective signals may be detected by both text (sentiment/emotion categorization) and acoustics (prosody, voice quality). Two-stream architectures (acoustic + lexical) of older note are precursors of the present-day multimodal architectures currently applied to live coaching and risk warnings when callers are agitated or manipulative[58]. Real call data has also been shown to be recognized in order to aid agent assistance, de-escalation and service-quality prediction in contact-center contexts.

IV. Machine Learning Approaches for Scam Detection

The advent of ML [59] as an efficient method of analysis has greatly contributed to the identification of voice-based scams. Through the examination of communication patterns and vast volumes of call audio data, ML methods [60] allow systems to determine suspicious behaviour and peculiarities that can signal abuse of power [61]. Such models have an opportunity to analyze acoustic characteristics, language patterns, and the conduct of callers in order to draw the line between valid and fraudulent appeals [62]. A proper implementation of the proper ML [63] algorithms is thus necessary in order to identify the growing voice scam patterns and reduce the number of false alarms [64], which may disrupt the normal organizational operation of communication or customer service [65].

Fraud detection methodologies based on ML may be categorized systematically under three major methodologies: supervised, unsupervised and hybrid methods. All methods have their own benefits and drawbacks depending on the particular situation of fraud detection and the nature of the data [66].

1) Supervised Learning

The majority of fraud detection research uses supervised learning. Models are trained using datasets that are carefully classified, with voice recordings or call exchanges being labeled as either authentic or fake. This allows the algorithm to master the patterns of distinction in the acoustic features, speech attributes, and conversation types, which it can successfully detect and identify suspicious voice calls and detect possible scamming needs.

2) Unsupervised Learning

Unsupervised learning techniques are solely employed in detecting voice scams. The methods can also detect patterns and anomalies without labeled training data, which is especially useful for detecting new fraud patterns not already observed in historical labeled data. These methods include Isolation Forest, Autoencoders, K-Means Clustering, Hidden Markov Models etc.

3) Deep Learning Approaches

DL approaches [67] belong to an area that is rapidly expanding in data scam detection studies [68]. These superior methods can acquire complicated patterns using the voice signals and conversation data and can correctly identify fraudulent calls. DL models [69] are frequently used separately or in combination with classical ML approaches to enhance the process of tracking scam-related voice patterns [70], such as artificial speech and social engineering techniques. These include LSTM [71], CNN, RNN, GAN etc.

4) Hybrid Ensemble Methods

There are numerous works that use hybrid methods that strategically use various ML methods to enhance overall detection and robustness [72]. Likely, Supervised-Unsupervised Methods, Ensemble Methods, DL Hybrids [73] etc.

V. LITERATURE OF REVIEW

This section presents earlier studies on voice spoofing, voice phishing, and telecom fraud detection. The systematic comparison of previous studies with an emphasis on the models of spoofing-detection, phishing defense, and system architecture, as well as their main advantages, is presented in Table II.

G. Lin *et al.* (2026) proposed a dual-level classification strategy: the frame-level classifier captures encoding discrepancies within individual frames, while the utterance-level classifier aggregates these frame-level features to learn global encoding patterns through global covariance pooling. The experimental results obtained from testing VoIP Phone Call Identification Database VPCID show that the proposed method achieves better results than all current methods because it maintains higher accuracy and better performance across various difficult testing situations. The ablation studies demonstrate that the proposed model architecture successfully achieves its intended goals and operates according to its design principles[74].

S. Chang *et al.* (2025) The study demonstrated that voice assistants that developers are creating for wearable devices without providing full human-computer interaction capabilities become vulnerable to voice spoofing attacks. Such attacks exploit pre-recorded or synthesized voice commands to trick assistants into executing unauthorized actions by legitimate users. In this work, they propose GyroTalk, a novel approach that extracts individual, reliable features from users' speech-movement sequences using built-in gyroscopes in wearables to differentiate between legitimate users and malicious attackers. GyroTalk is inspired by two critical insights. These collective muscle movements propagate throughout the body, providing unique movement signatures [75].

R. Sonwane *et al.* (2024). *The* project makes use of state-of-the-art technologies for voice feature extraction, processing, and pattern matching. By using React Native only for the Android application, "TrustCaller" plays a key role in enhancing the security of phone communications when combined with ML and Speech Synthesis. This program is a vital step in protecting people from possible dangers and promoting safer phone communication. Moreover, it provides scalable security settings to guarantee users' privacy and control over communication [76].

B. B. Gupta, A. Gaurav, and K. T. Chui (2024) proposed that smartphones have personal and private information about the user; hence, attackers target smartphones to access personal and confidential information. In this context, this paper proposed a GoogLeNet-based mobile phishing attack detection model. In their proposed model, whenever a user visits a webpage, its screenshot is analyzed by the GoogLeNet model, and if the website is malicious, the model alerts the user. They employed GoogLeNet as it is effective at identifying multiclass photos and has been trained on vast volumes of data[77].

F.-Y. Liang *et al.* (2023) introduced a new approach to telecom fraud detection using autoencoder and feature binning (TFD-FA). Customers are categorized into different telecom scenarios according to their unique qualities in TFD-FA's feature-binning system. In addition, data on nearby areas is assembled using an autoencoder component. An imbalance classifier component is also used to deal with the fact that there are more fraudsters than legitimate users. Comprehensive tests on an actual dataset show that TFD-FA is more successful than the baseline models that were examined[78].

R. H. J. and Mohana (2022) The project's overarching goal is to research different methods for preventing and detecting fraud in the communications sector. Various types of telecom fraud are outlined in this paper, along with obstacles to detection and possible solutions. The performance of the present techniques is stated at, and then suggestions and suggestions for selecting the best fit performance indicators are provided[79].

TABLE II. COMPARATIVE ANALYSIS OF MACHINE LEARNING LITERATURE ON VOICE SPOOFING AND TELECOM FRAUD DETECTION

Author (Year)	Voice Spoofing / Fraud Scenario	Detection Technique	System / Architecture	Outcome / Advantage
G. Lin et al. (2026)	Voice spoofing and fraudulent VoIP phone calls	Dual-level classification using frame-level and utterance-level analysis	Frame-level classifier + utterance-level classifier with global covariance pooling	Achieved higher accuracy and robustness in detecting spoofed voice calls on the VPCID dataset.
S. Chang et al. (2025)	Voice spoofing attacks on voice assistants in wearable devices	Motion-based voice authentication using gyroscope signals	GyroTalk system using wearable device gyroscopes to capture speech movement patterns	Effectively differentiates legitimate users from attackers by analyzing body movement signatures during speech.
R. Sonwane et al. (2024)	Fraudulent phone calls and unsafe phone communications	Voice feature extraction and machine learning–based pattern matching	TrustCaller mobile application built using React Native, integrated with ML and speech synthesis	Enhances phone call security and provides scalable privacy settings for safer communication.
B. B. Gupta et al. (2024)	Mobile phishing attacks targeting smartphone users	Deep learning–based webpage screenshot analysis	GoogLeNet-based phishing detection model	Efficiently identifies malicious websites and alerts users, improving mobile security.
F.-Y. Liang et al. (2023)	Identification of telecom fraud in extensive communication networks	Feature binning with autoencoder-based anomaly detection	TFD-FA (Telecom Fraud Detection with Feature Binning and Autoencoder) framework	Handles data imbalance and improves fraud detection accuracy on real-world telecom datasets.
R. H. J. & Mohana (2022)	Telecom fraud detection and prevention in communication networks	Survey-based analysis of fraud detection techniques	Conceptual framework reviewing telecom fraud classifications and detection methods	Provides insights into fraud types, detection challenges, and recommendations for selecting appropriate performance metrics.

VI. Conclusion And Future Work

Frauds that can be carried out over voice, such as spoofing, vishing, and telecom fraud, have become more advanced due to the development of AI-generated speech and deepfakes technology. In conclusion, voice-based fraud like vishing, robocalls, and deepfake voice assaults have become significantly more advanced and pervasive due to the rapid growth of communication and artificial intelligence. This paper has analyzed the changing nature of voice fraud, and the necessity of combining advanced technologies in voice fraud detection (speech signal processing, natural language processing, voice biometrics, and ML) and prevention (voice fraud). Different analysis characteristics such as acoustic patterns, linguistic hints, speaker verification, and emotion detection are significant to detect suspicious calling activities. The combination of supervised learning, unsupervised learning, DL, and hybrid systems in ML methods demonstrates their ability to improve the accuracy and reliability of voice scam detection systems. In spite of this progress, issues like paucity of datasets, the development of new attack methods, as well as the growth of deepfake technologies, still remain a serious challenge. Future work, therefore, ought to be done to ensure that more scalable, real-time, and adaptive detection structures are developed to promote security in telecommunications, banking and other voice-based service systems.

The next voice scam detection research must be carried out to ensure the creation of powerful real-time systems that can detect new threats like deepfake voice attacks. Detective accuracy can be enhanced by incorporating DL and multimodal analysis and larger sets of labeled information. Also, the joint structures with the telecom companies, banks, and cyber-deterrence measures will improve prevention strategies.

References

[1] H. Cyril, "AI-Driven Anomaly Detection, Outage Prediction, and Self-Healing in Telecom Provisioning Systems," *Int. J. Appl. Math.*, vol. 38, no. 12s, August, pp. 2817–2832, 2025, doi: 10.12732/ijam.v38i12s.1589.

[2] N. Prajapati, "Federated learning form privacy-presreving cybersecurity: A review on secure threat detection," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 5, no. 4, pp. 522–528, 2025.

[3] V. Pal and S. K. Chintagunta, "Transformer-Based Graph Neural Networks for Real-Time Fraud Detection in Blockchain Networks," *Int. J. Adv. Res. Sci. Commun. Technol.*, pp. 1401–1411, Jul. 2023, doi: 10.48175/IJARSCT-11978Y.

[4] S. Kumara, "Zero Trust Identity Fabric for Multi-Layer Telecom Networks: Implications for Secure and Scalable Digital Infrastructure," *Int. J. Curr. Eng. Technol.*, vol. 15, no. 6, 2025, doi: 10.14741/ijcet/v.15.6.7.

[5] H. T. Luong and H. M. Ngo, "Understanding the Nature of the Transnational Scam-Related Fraud: Challenges and Solutions from Vietnam's Perspective," *Laws*, vol. 13, no. 6, p. 70, Nov. 2024, doi: 10.3390/laws13060070.

[6] H. Cyril and S. Kumara, "Cybersecurity Architecture For Autonomous Telecommunication Networks," *Int. J. Adv. SIGNAL IMAGE Sci.*, vol. 12, no. 1s, pp. 618–639, Jan. 2026, doi: 10.29284/9admy374.

[7] R. Patel, "Security Challenges In Industrial Communication Networks: A Survey On Ethernet/Ip, Controlnet, And Devicenet," *Int. J. Recent Technol. Sci. Manag.*, vol. 7, no. 8, 2022, doi: 10.10206/IJRTSM.2025171772.

[8] K. M. R. Seetharaman and P. Yadav, "A Machine Learning Framework for Detecting and Mitigation of Cyber Threats in IoT Environments," in *2025 3rd International Conference on Inventive Computing and Informatics (ICICI)*, IEEE, Jun. 2025, pp. 1112–1119. doi: 10.1109/ICICI65870.2025.11069697.

[9] J. Kim, S. Gu, Y. Kim, S. Lee, and C. Kang, "A Multimodal Voice Phishing Detection System Integrating Text and Audio Analysis," *Appl. Sci.*, vol. 15, no. 20, p. 11170, Oct. 2025, doi: 10.3390/app152011170.

[10] V. Verma, "Deep Learning-Based Fraud Detection in Financial Transactions : A Case Study Using Real-Time Data Streams," vol. 3, no. 4, pp. 149–157, 2023, doi: 10.56472/25832646/JETA-V3I8P117.

[11] G. Sarraf, "Behavioral Analytics for Continuous Insider Threat Detection in Zero-Trust Architectures," *Int. J. Res. Anal. Rev.*, vol. 8, no. 4, pp. 596–602, 2021.

[12] S. Kumara, "AI-Driven Threat Identification and Response: Implications for Secure and Scalable Telecom Infrastructure," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 5, no. 4, p. 559, Dec. 2025, doi: 10.48175/IJARSCT-30567.

[13] G. Song, J. Liang, L. Wu, L. Liu, and C. Zhang, "The Impact of Residents' Daily Internet Activities on the Spatial Distribution of Online Fraud: An Analysis Based on Mobile Phone Application Usage," *ISPRS Int. J. Geo-Information*, vol. 14, no. 4, p. 151, Mar. 2025, doi: 10.3390/ijgi14040151.

[14] H. P. Cyril, "Event-Driven Provisioning Architectures For Modern Telecom Networks: Overcoming Legacy Limitations And Enabling Autonomous 6g Operations," *Int. J. Adv. Res. Comput. Sci.*, vol. 16, no. 6, pp. 75–82, Dec. 2025, doi: 10.26483/ijarcs.v16i6.7389.

[15] S. Kumara, "Identity-Driven IoT Security in Telecom Ecosystems: Implications for Scalable and Trustworthy Digital Infrastructure," *Int. J. Appl. Math.*, vol. 38, no. 12s, pp. 2797–2816, Dec. 2025, doi: 10.12732/ijam.v38i12s.1588.

[16] D. Patel, "Leveraging Blockchain and AI Framework for Enhancing Intrusion Prevention and Detection in Cybersecurity," *Tech. Int. J. Eng. Res.*, vol. 10, no. 6, pp. 853–858, 2023, doi: 10.56975/tijer.v10i6.158517.

[17] S. Singh, "Advancing Network Security in 5G: Leveraging the 5G-NIDD Dataset for Intrusion Detection and Mitigation," in *2025 IEEE 12th International Conference on Cyber Security and Cloud Computing (CSCloud)*, IEEE, Nov. 2025, pp. 1–6. doi: 10.1109/CSCloud66326.2025.00055.

[18] A. R. Bilipelli, "Forecasting the Evolution of Cyber Attacks in FinTech Using Transformer-Based Time Series Models," *Int. J. Res. Anal. Rev.*, vol. 10, no. 3, pp. 383–389, 2023.

[19] S. Singamsetty, "CyNet: Amalgam Deep Learning Model for Multi-Vector Cyber Intrusion Detection System (IDS)," in *2025 3rd International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*, IEEE, Sep. 2025, pp. 914–918. doi: 10.1109/ICoICI65217.2025.11253990.

[20] P. Nutalapati, J. R. Vummadi, S. Dodda, and N. Kamuni, "Advancing Network Intrusion Detection: A Comparative Study of Clustering and Classification on NSL-KDD Data," in *2025 International Conference on Data Science and Its Applications (ICoDSA)*, IEEE, Jul. 2025, pp. 880–885. doi: 10.1109/ICoDSA67155.2025.11157595.

[21] F. Salahdine and N. Kaabouch, "Social Engineering Attacks: A Survey," *Futur. Internet*, vol. 11, no. 4, p. 89, Apr. 2019, doi: 10.3390/fi11040089.

[22] P. Gallegos, J. F. Bravo-Torres, and P. E. V. Tapia, "Social engineering as an attack vector for ransomware," in *2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)*, IEEE, Oct. 2017, pp. 1–6. doi: 10.1109/CHILECON.2017.8229528.

[23] S. K. Davuluri, V. Challagulla, V. Mudapaka, and U. Konka, "Telcoformrix: An AI-Augmented Framework for Declarative and Scalable Provisioning of Real-Time Communication Infrastructure (Work in Progress)," in *2025 IEEE International Conference and Expo on Real Time Communications at IIT (RTC)*, 2025, pp. 1–4. doi: 10.1109/RTC66985.2025.11211725.

[24] S. P. Kalava, "Digital Transformation in the COVID Era: Evolution, Impact, and Future Pathways," *Int. J. Sci. Res.*, vol. 12, no. 4, pp. 1–7, 2023.

[25] J. W. Sajja, G. B. Komarina, and N. K. R. Choppa, "Enterprise Data Transformation in the Era of S/4HANA: Real-World Cloud Migration Architecture, Governance Strategies, and Lessons from the Field," *World J. Adv. Res. Rev.*, vol. 26, no. 2, pp. 3596–3619, May 2025, doi: 10.30574/wjarr. 2025.26.2.2038.

[26] V. K. Sharma, "Enabling Mission-Critical Communication via VoLTE for Public Safety Networks," *J. Adv. Dev. Res.*, vol. 10, no. 1, pp. 1–10, June, 2019.

[27] S. Garg, "Next-Gen Smart City Operations with AIOps & IoT : A Comprehensive Look at Optimizing Urban Infrastructure," *J. Adv. Dev. Res.*, vol. 12, no. 1, 2021, doi: 10.5281/zenodo.15364012.

[28] S. P. Kalava, "Revolutionizing Customer Experience: How CRM Digital Transformation Shapes Business," *Eur. J. Adv. Eng. Technol.*, p. 4, 2024.

[29] S. Singh, "Performance Evaluation of Machine Learning Regression Models for 5G Network Resource Allocation Optimization," in *2025 International Conference on Multimedia Computing, Networking and Applications (MCNA)*, 2025, pp. 47–52. doi: 10.1109/MCNA65829.2025.11124374.

[30] H. P. Cyril, "Serialization of Telecom Provisioning Transactions in Distributed Systems," *Int. J. Eng. Adv. Technol. Stud.*, vol. 15, no. 6, pp. 526–533, 2025, doi: 10.14741/ijcet/v.15.6.6.

[31] S. Singh, S. A. Pahune, P. Chatterjee, and R. Sura, "Advanced Machine Learning Methods for Churn Prediction and Classification in Telecom Sector," in *2025 IEEE 6th India Council International Subsections Conference (INDISCON)*, 2025, pp. 1–7. doi: 10.1109/INDISCON66021.2025.11252233.

[32] B. Yehya and N. Salhab, "Telecommunications Fraud Machine Learning-based Detection," in *2023 4th International Conference on Data Analytics for Business and Industry (ICDABI)*, Bahrain: IEEE, 2023, pp. 656–661, October. doi: 10.1109/ICDABI60145.2023.10629612.

[33] R. P. Mahajan, "Improvised Diabetic Retinopathy Detection Accuracy in Retinal Images Using Machine Learning Algorithms," *Int. Res. J.*, vol. 12, no. 3, pp. 155–161, 2025.

[34] R. Dattangire, R. Vaidya, D. Biradar, and A. Joon, "Exploring the Tangible Impact of Artificial Intelligence and Machine Learning: Bridging the Gap between Hype and Reality," in *2024 1st International Conference on Advanced Computing and Emerging Technologies (ACET)*, IEEE, 2024, pp. 1–6. doi: 10.1109/ACET61898.2024.10730334.

[35] R. Patel, "Optimizing Oil and Gas Pipeline Monitoring with AI Machine Learning Based Predictive Framework," in *2025 International Conference on Smart & Sustainable Technology (INCSST)*, IEEE, Jul. 2025, pp. 1–6. doi: 10.1109/INCSST64791.2025.11210432.

[36] M. Usama Tanveer Gujjar, K. Munir, M. Amjad, A. U. Rehman, and A. Bermak, "Unmasking the Fake: Machine Learning Approach for Deepfake Voice Detection," *IEEE Access*, vol. 12, pp. 197442–197453, 2024, doi: 10.1109/ACCESS.2024.3521026.

[37] V. K. Bollu, "Threat Landscape in Artificial Intelligence Systems: Taxonomy, Attack Vectors and Security Implications," *World J. Adv. Res. Rev.*, vol. 29, no. 1, pp. 285–294, 2026, doi: 10.30574/wjarr. 2026.29.1.0007.

[38] A. Katangoori, "The Role of Big Data in Advancing Artificial Intelligence: Methods and Case Studies," *Int. J. Artif. Intell. Mach. Learn.*, vol. 6, no. 1, pp. 37–54, Jan. 2026, doi: 10.51483/IJAIML.6.1.2026.37-54.

[39] D. Patel, "Explainable Risk Decision Systems Using Artificial Intelligence Models for Payment Fraud Identification with Mitigation," in *2026 14th International Symposium on Digital Forensics and Security (ISDFS)*, IEEE, Mar. 2026, pp. 01–06. doi: 10.1109/ISDFS69419.2026.11459006.

[40] C. Lee, B. Kim, and H. Kim, "The silence of the phishers: Early-stage voice phishing detection with runtime permission requests," *Comput. Secur.*, vol. 152, p. 104364, May 2025, doi: 10.1016/j.cose.2025.104364.

[41] A. R. Bilipelli, "AI-Driven Intrusion Detection Systems for Large- Scale Cybersecurity Networks Data Analysis : A Comparative Study," *TIJER – Int. Res. J.*, vol. 11, no. 12, pp. 922–928, 2024.

[42] S. Thangavel, "AI Enhanced Image Processing System For Cyber Security Threat Analysis," 2024.

[43] R. Karanjkar and S. Phalke, "AI Investment and the Pendulum Effect: The Crisis of Software Quality Infrastructure," *Int. J. Res. Appl. Innov.*, vol. 8, no. 6, pp. 12974–12977, 2025, doi: 10.15662/IJRAI.2025.0806018.

[44] S. A. Pushkala, "Generative AI in battling Fraud," in *2024 IEEE 4th International Conference on ICT in Business Industry & Government (ICTBIG)*, 2024, pp. 1–5. doi: 10.1109/ICTBIG64922.2024.10911802.

[45] P. Parida and N. Senguttuvan, "Responsible Utilization of Cloud in Retail Banking Ecosystem," *Int. J. Comput. Appl.*, vol. 187, no. 49, pp. 34–39, oct. 2025, doi: 10.5120/ijca2025925835.

[46] N. Yalçın[a] and B. Lale, "Types of cyber-attacks using voice," 2025. doi: 10.59313/jsr-a.1600934.

[47] L. Yasur, G. Frankovits, F. M. Grabovski, and Y. Mirsky, "Deepfake CAPTCHA: A Method for Preventing Fake Calls," Jan. 2023. doi: https://doi.org/10.48550/arXiv.2301.03064.

[48] S. Narang and A. Gogineni, "Zero-Trust Security in Intrusion Detection Networks: An AI-Powered Threat Detection in Cloud Environment," *Int. J. Sci. Res. Mod. Technol.*, vol. 4, no. 5, pp. 60–70, jun. 2025, doi: 10.38124/ijsrmt.v4i5.542.

[49] H. Yıldırım, Y. Bütüner, and C. Ünal, "Use of Artificial Intelligence Tools for Telephone Scams and Countermeasures," *J. Public Econ. Public Financ. Manag.*, vol. 5, no. 2, pp. 0–2, 2025.

[50] S. Amrale, "A Novel Generative AI-Based Approach for Robust Anomaly Identification in High-Dimensional Datasets," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 4, no. 2, pp. 709–721, 2024, doi: 10.48175/IJARSCT-19900D.

[51] M. T. T. Bajwa, F. Tehreem, Z. Farid, H. M. F. Tahir, and ..., "Deepfake Voice Recognition: Techniques, Organizational Risks and Ethical Implications," *Spectr. ...*, vol. 3138, pp. 106–121, 2025.

[52] P. R. Marapatla, "Intelligent APIs: AI-Powered Ecosystem for Nonprofit Digital Transformation," *J. Inf. Syst. Eng. Manag.*, vol. 10, no. 60s, pp. 605–618, Sep. 2025, doi: 10.52783/jisem.v10i60s.13174.

[53] P. P. S. Kulkarni, Devyani S., Ratnadeep R. Deshmukh, "A Review of Speech Signal Enhancement Techniques," *Int. J. Comput. Appl.*, vol. 139, no. 16, pp. 23–26, 2016.

[54] V. K. Sharma, "Federated Learning in Mobile and Edge Environments for Telecom Use Cases," *Int. J. Innov. Res. Creat. Technol.*, vol. 10, no. 1, pp. 1–10, 2024, doi: 10.5281/zenodo.17062956.

[55] M. K. Singh, S. Kumar, and D. Nandan, "Faulty voice diagnosis of automotive gearbox based on acoustic feature extraction and classification technique," *J. Eng. Res.*, vol. 11, no. 2, p. 100051, 2023, doi: https://doi.org/10.1016/j.jer.2023.100051.

[56] R. Lima, B. Espinasse, and F. Freitas, "The impact of semantic linguistic features in relation extraction: A logical relational learning approach," *Int.*

Conf. Recent Adv. Nat. Lang. Process. RANLP, vol. 2019-Septe, pp. 648–654, 2019, doi: 10.26615/978-954-452-056-4_076.

[57] S. Nakagawa, L. Wang, and S. Ohtsuka, "Speaker Identification and Verification by Combining MFCC and Phase Information," *IEEE Trans. Audio. Speech. Lang. Processing*, vol. 20, no. 4, pp. 1085–1095, May 2012, doi: 10.1109/TASL.2011.2172422.

[58] P. Badri, A. Nerella, R. Murugesan, and K. Sundravadivelu, "Deep Learning–Based Multivariate Models for Bankruptcy and Litigation Risk Prediction," *Adv. Consum. Res.*, vol. 2, no. 4, pp. 4442–4450, 2025.

[59] N. Prajapati, "The Role of Machine Learning in Big Data Analytics: Tools, Techniques, and Applications," *ESP J. Eng. Technol. Adv.*, vol. 5, no. 2, 2025, doi: 10.56472/25832646/JETA-V5I2P103.

[60] D. Patel, "Improving Software Performance Through Early Bug Detection Using Large-Scale Machine Learning Models," in *2025 3rd World Conference on Communication & Computing (WCONF)*, IEEE, Jul. 2025, pp. 1–6. doi: 10.1109/WCONF64849.2025.11233621.

[61] M. R. R. Deva and N. Jain, "Utilizing Azure Automated Machine Learning and XGBoost for Predicting Cloud Resource Utilization in Enterprise Environments," in *2025 International Conference on Networks and Cryptology (NETCRYPT)*, IEEE, May 2025, pp. 535–540. doi: 10.1109/NETCRYPT65877.2025.11102235.

[62] L. Ogbidi and B. Oteh, "Advances in Hybrid Machine Learning and Physics-Based Models for Enhanced Reservoir Simulation," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 10, no. 6, pp. 2533–2543, Dec. 2024, doi: 10.32628/IJSRCSEIT.

[63] P. B. Patel, "Predictive Maintenance in HVAC Systems Using Machine Learning Algorithms: A Comparative Study," *Int. J. Eng. Sci. Math.*, vol. 13, no. 12, pp. 118–125, 2019.

[64] K. C. C. Hemish Prakashchandra Kapadia, "Machine Learning Based Data Processing Equipment for Security in Cloud," 6449113, 2025

[65] S. B. Karri, S. Gawali, S. Rayankula, and P. Vankadara, "AI Chatbots in Banking: Transforming Customer Service and Operational Efficiency," in *Advancements in Smart Innovations, Intelligent Systems, and Technologies*, 2025, pp. 61–81. doi: 10.3233/FAIA251498.

[66] A. A. Compagnino *et al.*, "An Introduction to Machine Learning Methods for Fraud Detection," no. Ml, pp. 1–32, 2025.

[67] R. Patel, "Automated Threat Detection and Risk Mitigation for ICS (Industrial Control Systems) Employing Deep Learning in Cybersecurity Defense," *Int. J. Curr. Eng. Technol.*, vol. 13, no. 06, pp. 584–591, 2023, doi: 10.14741/ijcet/v.13.6.11.

[68] S. Kumara, "A Lightweight Deep Learning-Based Classification Model for Non-Human Identity Threat Detection," in *2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC)*, IEEE, Feb. 2026, pp. 1–6. doi: 10.1109/ICAIC67076.2026.11395886.

[69] R. P. Mahajan and N. Jain, "Optimizing CT Image Quality through AI-based Reconstruction and Deep Learning Models for Enhanced Diagnostic Accuracy," in *2025 4th International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, 2025, pp. 1–7. doi: 10.1109/ICDCECE65353.2025.11035138.

[70] R. V. S. S. B. R, Y. M. M. John, M. B. B. G, B. Karim, and G. Saritha, "Multi-Domain Cyber Threat Classification Using Enhanced Genetic Algorithm and Deep Neural Networks," in *2025 Third International Conference on Networks, Multimedia and Information Technology (NMITCON)*, IEEE, Aug. 2025, pp. 1–6. doi: 10.1109/NMITCON65824.2025.11187556.

[71] V. Verma, "Improving Product Recommendations in Retail with Hybrid Collaborative Filtering and LSTM Varun Verma Independent Researcher," *Int. J. Eng. Sci. Math.*, vol. 10, no. 8, pp. 113–128, 2021.

[72] T. A. Khan *et al.*, "Multi-Source Cyber Intrusion Detection Using Ensemble Machine Learning," *J. Comput. Sci.*, vol. 21, no. 1, pp. 111–123, Dec. 2024, doi: 10.3844/jcssp. 2025.111.123.

[73] K. B. Thakkar and H. P. Kapadia, "The Roadmap to Digital Transformation in Banking: Advancing Credit Card Fraud Detection with Hybrid Deep Learning Model," in *2025 2nd International Conference on Trends in Engineering Systems and Technologies (ICTEST)*, 2025, pp. 1–6. doi: 10.1109/ICTEST64710.2025.11042822.

[74] G. Lin, W. Luo, P. Zheng, and J. Huang, "VoIP Call Identification via a Dual-Level 1D-CNN With Frame and Utterance Features," in *IEEE Transactions on Information Forensics and Security*, IEE, 2026, pp. 2389–2402, February. doi: 10.1109/TIFS.2026.3667459.

[75] S. Chang, L. Zhou, W. Liu, H. Zhu, X. Hu, and L. Yang, "Combating Voice Spoofing Attacks on Wearables via Speech Movement Sequences," *IEEE Trans. Dependable Secur. Comput.*, vol. 22, no. 1, pp. 819–832, Jan. 2025, doi: 10.1109/TDSC.2024.3418908.

[76] R. Sonwane *et al.*, "TrustCaller- Voice-based Fraud Prevention System," in *2023 4th International Conference on Intelligent Technologies (CONIT)*, IEEE, Jun. 2024, pp. 1–6. doi: 10.1109/CONIT61985.2024.10626085.

[77] B. B. Gupta, A. Gaurav, and K. T. Chui, "Securing Smartphone from Mobile Phishing Attacks Using GoogLeNet Model," in *2024 IEEE International Symposium on Consumer Technology (ISCT)*, IEEE, Aug. 2024, pp. 522–527. doi: 10.1109/ISCT62336.2024.10791190.

[78] F.-Y. Liang *et al.*, "Telecom Fraud Detection Based on Feature Binning and Autoencoder," in *2023 IEEE International Conference on Data Mining (ICDM)*, IEEE, Dec. 2023, pp. 368–377. doi: 10.1109/ICDM58522.2023.00046.

[79] R. H. J. and Mohana, "Fraud Detection and Management for Telecommunication Systems using Artificial Intelligence (AI)," in *2022 3rd International Conference on Smart Electronics and Communication (ICOSEC)*, IEEE, Oct. 2022, pp. 1016–1022. doi: 10.1109/ICOSEC54921.2022.9951889.