Journal of Cyber-Physical Security and Robotics
Volume 01, Issue 2, 23-31, 2025

Research Article

Data-Driven Artificial intelligence (AI)-based System for Identifying Fraudulent Telephony Activities

Jenitha Pilli[1], Prathik Kumar Jannu[2], Javed Ali Mohammad[3], Sri Harsha Panchali[4], Usha Mohani Kavirayani[5], Krishna Bhardwaj Mylavarapu[6]

[1]MS in Computer Science, University of Louisiana at Lafayette
[2]Computer Science Engineering, JNTU Hyderabad
[3]Masters in Data Science, New England College
[4]Information Systems Engineer, CrowdStrike Inc
[5]MS in Computer Science, Kent State University
[6]MS in Computer Science, University of Illinois Springfield
[*]Corresponding author email: jenitha.pilli@gmail.com

Received: 03rd November, 2025 Accepted: 14th December, 2025

Abstract

The telecommunications industry is an essential part of the modern world of communication tools as it links billions of users across the globe. Nevertheless, telephony fraud has become a widespread issue, and it includes an assortment of fraud types, including subscription fraud, SIM cloning, toll fraud, and voice phishing (vishing), which cost people significant amounts of money and compromised the credibility of the service. The traditional number-based fraud detection techniques feel restricted by the dynamic strategies used by the fraudsters such as spoofing of caller IDs and spoofing of numbers, thereby causing high rate of false positive and late response. This paper suggests a machine-driven Artificial Intelligence (AI) model to detect suspicious telephony transactions with a big amount of textual information. The Baidu search engine was used to gather textual information about actual telecom fraud cases, which was preprocessed by removing noise, segmenting, and tagging part-of-speech and converted into TF-IDF feature vectors. This data was trained in an Artificial Neural Network (ANN) to learn complicated patterns of fraud. Measures of evaluation such accuracy, precision, recall, and F1-score indicate a strong performance of the system with an accuracy of 98.53. The findings indicate that AI-based models have the potential to identify the constantly changing telephony fraud, which is scalable and can deliver timely information to defend users and service providers. It is a new and innovative adaptive scheme of telecom fraud detection, as opposed to the conventional schemes.

Keywords: Telecommunication fraud, Artificial Intelligence (AI), Machine Learning, Fraud Detection, Textual Data, Real-Time Detection, Telephony activity.

Introduction

Telecommunications industry is a fundamental part of the information and communication technology sector that covers all telecommunication corporations and internet service providers and is one of the key elements in the development of mobile communications and the information society as a whole. The industry has remained at the focal point of growth, innovation, and disruption in almost all sectors offering services that directly or indirectly affect billions of individuals across the world [1]. Telephony services are a part and parcel of everyday life, and they have made it possible to effectively use services like voice calls, SMS, mobile banking, and VoIP services to not only communicate personally, but also to carry out business transactions, financial services, and other valuable services [2]. The accessibility, scalability, and reliability of telephony networks bring them to be among the pillars of the contemporary communication infrastructure [3].

The telecommunication sector is under a sustainable and dynamic threat of fraud despite its significance. Telecommunication fraud is a major issue in the world,

which has cost countries a lot of finance and compromised the reputation and efficiency of the service providers [4]. Telephony is a broad field of illegal activity, the fraud of which includes subscriptions and SIM-card cloning, toll fraud, and voice phishing (vishing). These activities usually entail theft of services or also intentional abuse of voice and data networks in a bid to avoid or minimize charges of legitimate use [5]. The implications of fraud go beyond the financial losses, and include marketing policies, relationships with customers, and feelings of the shareholders. Telecommunication fraud is a rather hard to fight as it is extremely dynamic. Fraudsters are constantly modifying their techniques to circumvent the security systems and are using techniques like caller ID spoofing and number-changing software. The development of Internet telephony has opened up additional opportunities to engage in fraudulent activities, allowing cybercriminals to make mass automated calls, interfere with services, and target people as a source of monetization plans [6]. This has contributed to the increases in unsolicited telemarketing as well as interactive voice Response based vishing campaigns and thus effective detection mechanisms are more vital than ever before. Traditional methods of detecting a fraud in the telephony involve the labelling of the caller numbers which a user has labelled as fraudulent. Although such techniques are somehow effective, they can be constrained by the dynamism of fraud [7][8]. Since fraudsters keep on changing or masking their telephone numbers, conventional methods of detecting frauds based on numbers may not work thus false positives are high and the detection is delayed. Telecommunication data is even more difficult to manage as it is large and complex in nature and requires more advanced, flexible, and more automated methods.

Machine learning (ML) and artificial intelligence (AI) have become strong instruments in fraud detection within telecommunication networks. Analyzing historical call data and extracting the key features, AI-based systems can find highly-complex patterns and predict the possible fraudulent behavior with a high level of accuracy [9][10]. The use of AI in banking, e-commerce, and cybersecurity has proven the possibility of machine learning models that are able to enhance the detection of fraud and minimize false alarms. Regarding the telephony sector, AI-based data techniques may be used to fit the changing fraud trends, surmount the constraints of traditional fraud detection tools, and offer scalable and real-time solutions to secure service providers and consumers.

Motivation and Contribution of the Study

This study is motivated by the rising rate of telecommunication fraud, such as vishing, subscription fraud, and SIM cloning, which severely affect service providers and users with both financial and operational consequences. The use of conventional methods of fraud detection, which are highly reliant on caller number conditions, cannot be used to combat the ever-changing and dynamic nature of fraudsters as they tend to delay the detection time and have a high number of false positives. The goal of this research is to create a smart, information-sensitive, and flexible system of the real-time fraudulent telephony detection. Key contributions include:

- Collected real world text data on telephony fraud in the Baidu search engine to assure of relevant and diverse contributions.
- Conducting the thorough data preparation, such as noise reduction, elimination of duplicates, relevance filtering by humans, stop-word removal, keyword optimization, and text segmentation, to improve the quality and semantic transparency of the data.
- To replace textual data with meaningful numerical data, a Natural Language Processing (NLP) pipeline that consists of part-of-speech tagging and TF-IDF representation should be implemented to extract features.
- Adopt Artificial Neural Network (ANN) model that would capture evolving and nonlinear patterns of fraud.
- Separating the data where training and testing sets are created to assess the model objectively.
- Evaluating the system using accuracy, precision, recall, and F1-score to validate its effectiveness in detecting telephony fraud.

Novelty of the Paper

This study is novel because of the combination of text-data analysis and neural network-based pattern detector so that real-time detection of complex and dynamic patterns of telephony fraud can be enabled. This approach learns semantic patterns of frauds reports as opposed to traditional approaches based. The noise elimination, the keyword refinement and the extraction of the TF-IDF features used as representations to the input data improves the quality of the data representation and the ANN is effective in representing the nonlinear relationships which guarantees a high level of detection. The framework provides a generalizable and scalable solution to telecommunication providers, which is superior to the conventional machine learning frameworks to identify dynamic fraud patterns.

Structure of the Paper

The paper will be organized in the following way: Section II will be the review of related work. Section III outlines the methodology, comprising of data collection, pre-processing and model design. Section IV presents results and comparative analysis of the experiment. Section V explains the end of the study with the conclusion as the findings are summarized, limitations are discussed, and the future research directions are presented.

Literature Review

The literature shows varied machine learning, statistical and signal-processing strategies of telecom fraud detection, with a focus on accuracy, real-time operation, ensemble models and behavior based analysis.

Arafat, Qusef and Sammour (2019) propose an intelligent automated system for detecting telecom fraud. Telecom fraud, known in Japan as "wangiri" or "one ring and cut fraud," is based on this one-ring technique to generate money quickly. Unknown callers' attempts to trick subscribers into calling premium numbers, where they are tricked into staying on the line for an excessive amount of time, ultimately leading to an inflated bill. In order to efficiently and effectively create more accurate classifications, this research suggests using a variety of ensemble classifiers to overcome the dataset's significant bias. The most effective and accurate method was determined to be the Extreme Gradient Boosting algorithm [11].

Zhong et al. (2019) present a Broad Learning System-based technique for identifying fraudulent phone calls. assessed the text data of fraudulent phone calls using the initial fifteen seconds of call content identification monitoring, built the TF-IDF model, and subsequently used it to train a neural network based on the BLS to detect fraudulent phone calls. Also, comparable incremental learning algorithms allow for rapid model updates without retraining using BLS, making them ideal for fraud detection systems that have limited data characteristics but need accurate predictions in real time. A thorough analysis and experimentation with the aforementioned procedure is conducted. The results demonstrate that this strategy outperforms the others in terms of training speed and accuracy when applied to fraud data [12].

Li et al. (2018) present a successful and widely applicable fraud user detection approach based on the user's Call Detail Record (CDR). Machine learning and template detection are the two main components of the suggested approach. The machine learning module employs a supervised learning-based Support Vector Machine (SVM) method to categorize users based on summary attributes. The template detection module employs a Finite State Machine (FSM) trained on fraudulent user behavior to screen out potentially malicious individuals. In the third and final module, fraudulent users will be identified. apply the method and test it on a dataset that actually exists in the real world [13].

Wang and Sang (2018) present a fingerprint data hiding speaker identification tracing approach to combat telecom fraud. When making a phone conversation, the first step is to encrypt the speaker's fingerprint data as identification information into the recorded voice signal. Then, the OFDM technology used in 4G wireless communication is employed to modulate the speech signal that contains the speaker's identity information. Once telephonic fraud has taken place, the perpetrator can be identified by deriving fingerprint information from the audio stream. The results of the experiments demonstrate that the suggested method can successfully identify the source of remote speech communications and mitigate telecommunication fraud to a certain point [14].

Hussain, Du and Ren (2017) use the massive amounts of data generated by the 4G LTE-A core network, called call detail records (CDRs), to identify abnormalities in the network and thereby solve the problems outlined before. showcase a statistically-based anomaly detection method that is semi-supervised and can detect two types of anomalies in near-real-time: first, a region with abnormally low user activity, which represents a sleeping cell (a type of cell outage) and is hard to spot because it doesn't raise an alarm and stays hidden until subscribers start complaining; and second, an area with abnormally high user traffic, which corresponds to a situation that may require special action (e.g., resource allocation, fault avoidance solutions, etc [15].

Yulianto et al. (2017) offer a technique that integrates the benefits of hybrid NBTree with Kullback Leibler divergence (KL-divergence or KLD) to identify potential fraudsters on IDD call services. When compared to Decision Tree and Naive Bayesian, NBTree performs better in terms of accuracy and tree size, and it can handle data sets with a lot of dimensions. Also, KL-divergence has been around for a while and has been used for a variety of proven and practical purposes, including feature selection, similarity measurement, and fraud detection. When compared to the prior methods—Naive Bayesian Classifier, hybrid Naive Bayesian—KL-divergence, and Support Vector Machine (SVM)—the experimental findings reveal that the combination of the two produces greater accuracy and F1-measure [16].

The Table 1 comparing telecom fraud detection literature by year, data source, techniques, and performance, summarizes the methodological variety, strengths, and results of supervised, semi-supervised, and hybrid methods

Methodology

In this research, a data-based artificial intelligence (AI)-based system is demonstrated to detect fraud telephony actions through the use of large-scale textual content as demonstrated in Figure 1. The methodology starts by collecting textual data on telecom related issues that were found on the Baidu search engine that includes actual descriptions of fraud, reports and scam stories. It is then processed to guarantee the quality and relevance of data through refining the work, eliminating duplications, filtering through irrelevant data, eliminating commonly utilized words that are not informational and organizing the information through segmentation and linguistics

Table 1: Comparative Analysis of Ai-Driven Telephony Fraud Detection

Reference	*Focus Area*	*Key Findings*	*Challenges*	*Key Contribution*	*Limitations*
Arafat, Qusef & Sammour (2019)	Wangiri (one-ring-and-cut) telecom fraud	Ensemble learning improved detection on highly imbalanced datasets; XGBoost outperformed other classifiers	Highly biased datasets; rapidly changing fraud behavior	Proposed intelligent automated ensemble-based ML solution for Wangiri fraud detection	Scalability and real-world deployment considerations not discussed
Zhong et al. (2019)	Fraud phone call identification using call content	Broad Learning System (BLS) with TF-IDF features achieved high accuracy and fast training; supported incremental learning without retraining	Limited textual features in short calls; early-stage call content dependency	Introduced real-time, incrementally updatable fraud detection model suitable for fast prediction environments	Performance depends on availability and quality of call content data
Li et al. (2018)	Fraud user detection using Call Detail Records (CDR)	Hybrid SVM + FSM framework effectively detected fraud users on real-world datasets	Dependence on handcrafted summary features; behavior rules may become outdated	Combined machine learning classification with behavior-based template filtering	Adaptability to evolving fraud patterns not fully evaluated
Wang & Sang (2018)	Speaker identity tracing for telecom fraud prevention	Fingerprint data hidden in speech enabled accurate speaker tracing in fraud scenarios	Added communication and processing overhead; assumes access to speech signals	Introduced biometric-based tracing to support fraud investigation and prevention	Does not directly detect fraud; limited scalability for large telecom systems
Hussain, Du & Ren (2017)	Anomaly detection in telecom networks using CDR big data	Semi-supervised statistical method detected abnormal low and high traffic regions in near real-time	Does not explicitly label fraud types; anomaly ≠ fraud	Demonstrated feasibility of near real-time anomaly detection using core network data	Indirect fraud detection; lacks supervised ML classification
Yulianto et al. (2017)	IDD call fraud detection	Hybrid NBTree + KL-divergence achieved higher accuracy and F1-score than Naïve Bayes and SVM	Feature selection sensitivity; dependency on historical fraud behavior	Showed effectiveness of hybrid probabilistic and divergence-based ML methods	Evaluation limited to specific IDD datasets; generalization not validated

annotation. Key terms have also been narrowed down to improve semantic clarity. The processed text undergoes meaningful features to represent patterns that are related to fraud. The dataset that is prepared is split into 30% testing and 70% training parts to facilitate objective learning and validation. A model of Artificial Neural Network (ANN) is used to acquire knowledge of intricate fraud indicators based on the training sample. Accuracy, precision, recall and F1-score are used to measure model performance, and reliably identify fraudulent telephony activities.

Data Description

The dataset utilized in the present study is the textual data that is associated with telecommunication fraud and is presented by the search engine Baidu, which provides elaborate and excellently organized information of frauds that occur in reality. The data gathered comprise news articles, case descriptions and recorded scam threads relating to fraudulent telephone calls. In the first place, 3,234 samples of texts were collected The data set represents a variety of fraud forms such as impersonation, financial extortion, and misleading information requests.

Data Preparation

Data preparation involved cleaning up telephony records to remove inconsistencies and irrelevant records and thus enhance the quality of the data. The ready dataset was subsequently organized in such a manner that it facilitated successful learning and analysis of the proposed model. The steps are given below:

- **Text Formatting:** The raw text data was noisy, in the form of emojis, special characters, URLs, and non-regular characters that are not at all relevant to the semantics. These factors were eliminated systematically in order to make the text standardized so that all the samples are consistent.
- **Duplicate Removal:** The text similarity measurement was used to remove any duplicate and near-duplicate text instances. In cases where the resemblance between two texts was above 80, a single case was assumed to be redundant and was eliminated so as to minimize the aspect of bias and redundancy.
- **Manual Relevance Filtering:** A manual sifting was done to only keep the texts which specifically talked about the telecommunication fraud calls. Relevant discussions, vague sections, and disconnected

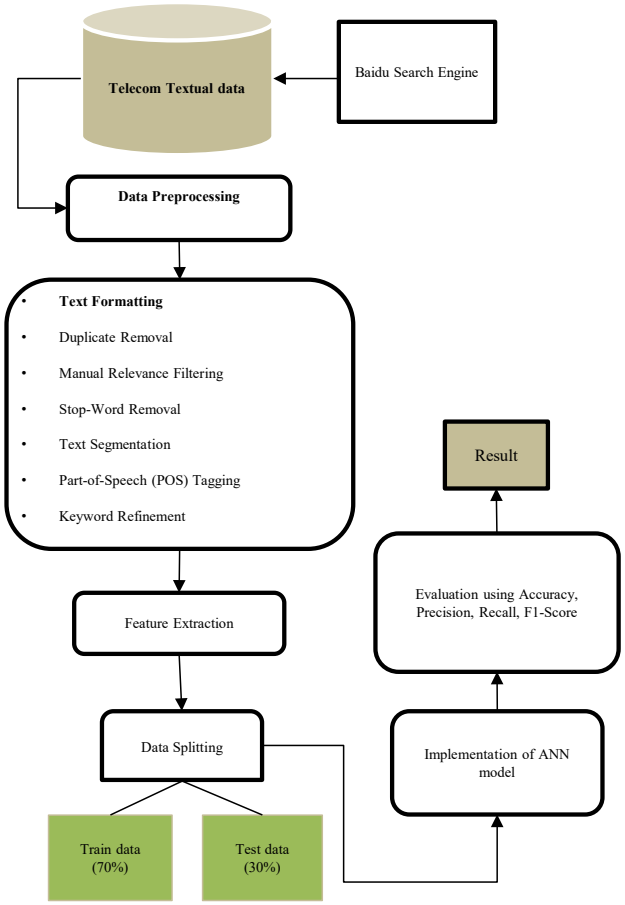

Figure 1: Workflow of the AI-Based System for Detecting Fraudulent Telephony Activities Using Textual Data

complaints were avoided to ensure that the data-sets were relevant.

- **Stop-Word Removal:** A personal stop-word list, which included 1,601 words, was used to eliminate semantically irrelevant words, such as common conversational phone expressions, like "hello" and "please hold and so on).
- **Text Segmentation:** The Chinese word segmentation was used to segment continuous text into meaningful lexical units, which allowed appropriate feature extraction and semantic interpretation.
- **Part-of-Speech (POS) Tagging:** Tagging was done on each segmented word with its grammatical category. Nouns, verbs and action oriented words were kept and pronouns, prepositions, and fillers were deleted.
- **Keyword Refinement:** A last refining step in the manual was conducted to drop non-semantic terms like personal names and geographical locations so that the extracted key words centered on the language patterns about fraud.

Feature Extraction

Following preprocessing, Natural Language Processing techniques were applied to extract discriminative textual features to elicit the semantic patterns of telecommunication fraud. Extracting keywords was done using frequency analysis of words and then by weighting words using Term Frequency-Inverse Document Frequency (TF-IDF) in order to be informative and relevant to the domain and also to minimize the effects of the frequently used words. Consequently, every text example was converted to a high-dimensional numerical feature vector which is an effective representation of the inherent linguistic properties of fraud related content and can be used as a input to machine learning classification models.

Data Splitting

The pre-processed data was split into 70:30 for training and testing set, so that the model would acquire enough data to learn, and also to have a reasonable amount of data to evaluate the performance of the model.

Implement the ANN Model

The artificial neural network (ANN) was trained with a back-propagation learning technique. To reduce error, the synaptic weights were adjusted using gradient descent in the transformation function. The training process started by assigning a value between 0 and 1 to each variable [17][18]. Effectively, this model's hyperbolic tangent function modification yielded values ranging from -1 to +1. Hyperbolic tangent and soft max functions are shown by Equations (1&2):

Hyperbolic Tangent Function Output

$$O_j = \tanh\left(H_j\right) = \frac{e^{Hj} - e^{-Hj}}{e^{Hj} + e^{-Hj}} \quad (1)$$

Soft max Function Output

$$O_j = \acute{o}\left(H_j\right) = \frac{e^{Hj}}{\sum_{k=1}^{m} e^{Hk}} \quad (2)$$

A probability distribution function, which may be defined as the network's pseudo-probability or estimated probability of input classification function, produces the output Oj [19]. The more common of the two optimization modes offered by the gradient descent technique, batch mode updates all synaptic weights of all records in the training data set.

Evaluation Metrics

This section identifies the primary assessment metrics used to measure model performance. In machine learning, confusion matrices are a typical way to evaluate algorithm performance. Table II shows the sum of all values predicted by the machine learning algorithms, including both positive and negative predictions.

The key terms of the confusion matrix:

- **Positive (P):** is actual positive [20].

Table 2: Confusion Matrix

	Predicted	*Predicted*
Actual	TP	FN
Actual	FP	TN

- **Negative (N):** is not shows the positive that's why it is negative.
- **True Positive (TP):** is both positive and expected as positive.
- **A false negative (FN):** occurs when a positive result is anticipated but is instead a negative one.
- **A True Negative (TN):** is one that is both projected to be negative and actually negative.
- **A false positive (FP):** occurs when a result that is actually negative is projected to be positive.

Accuracy measures how well a prediction matches the actual outcome and is calculated using the formula shown in Equation (3).

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{3}$$

The Precision measure is calculated by dividing the total number of detections classified as true by the number of frauds that were successfully identified [21]. It measures the system's accuracy in correctly identifying frauds, differentiating them from other frauds or normal flows. The formula for precision is defined in Equation (4).

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

The number of actual frauds divided by the number of predicted frauds is the Recall [22]. Consequently, it reveals the correct number of frauds detected, which is called the "true positive rate." There is a formula for recall in Equation (5).

$$Recall\left(Rc\right) = \frac{TP}{TP + FN} \tag{5}$$

The F1-Score, described in Equation (6), is a measure that finds a happy medium between recall and precision; it can take on values between zero and one.

$$F1score\left(F1\right) = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{6}$$

These performance measurements are used to test the effectiveness of the model based on its results on the test data.

Result and Discussion

This paper examines the performance of a data-driven artificial intelligence (AI)-based solution in detecting fraudulent telephony practices through advanced machine learning methods. Specifically, the ANN model was selected because of its high outlier ability to learn complex and nonlinear trends that can be found in the data of telephony usage. The model was developed and assessed using standard machine learning libraries in a Python environment, thus making sure that pre-processing, training, and validation of data were efficient. Table III also indicates that the ANN model had an accuracy of 98.53%, a precision of 97.97%, and a recall of 98.25%. The F1-score of 98.11% also demonstrates the strength and accuracy of the suggested system of fraud detection.

Table 3: Performance Evaluation of Ann-Based Telephony Fraud Detection System

Metrics	*ANN*
Accuracy	98.53
Precision	97.97
Recall	98.25
F1-Score	98.11

The receiver operating characteristic (ROC) curve, a popular method for assessing the accuracy of a proposed artificial neural network (ANN)-based telephone fraud detection system. Figure 2 displays the receiver operating characteristic (ROC) curve, which illustrates the compromise between specificity (the rate of false positives) and sensitivity (the rate of genuine positives) for various decision threshold values. The blue curve's sharp ascent towards the top left corner of the plot demonstrates the ANN model's capability to reliably differentiate between valid and fraudulent actions. This action portrays a strong detectability ability and low false alarm rate, and therefore, it is evident that the suggested system is strong and reliable in detecting telephony fraud. The confusion matrix that measures an ANN-based telephony fraud detection system presented in the Figure 3. This model checks its forecast against the real world 1,000 times. There is high performance on the main diagonal; 490 True Negatives and 491 True Positives. The non-diagonal entries show the smallest errors: 10 False Positives, i.e. legitimate actions were identified as fraud, and 9 False Negatives, i.e. the fraud eluded the DSS.

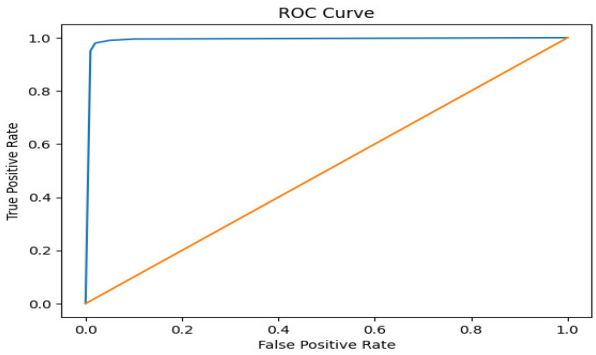

Figure 2: ROC Curve of ANN Model for Detect Telephony Fraud

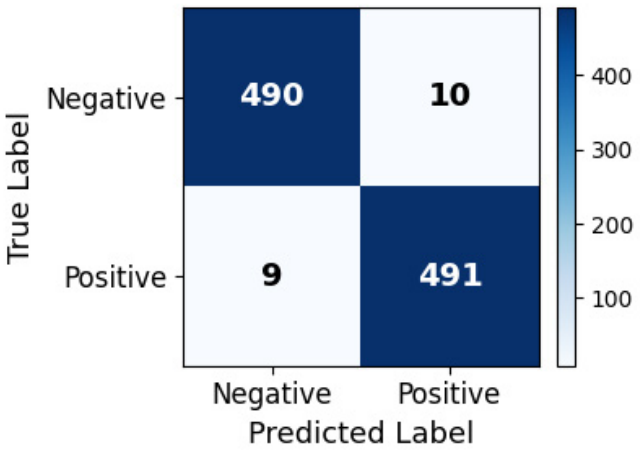

Figure 3: Confusion Matrix of ANN Model for Detect Telephony Fraud

The visual confirmation of the high level of accuracy and reliability is provided by the deep blue color of the diagonal cells that proves that the neural network can successfully identify standard telephony traffic and fraudulent one.

Comparative Analysis

In this section, a comparative analysis of various machine learning models will be provided in terms of accuracy in classifying fraudulent telephony activity. The Artificial Neural Network (ANN) model had the best accuracy rate of 98.53% as indicated in Table IV and it proved to be more effective in terms of picking up intricate patterns of fraud in telephony data. Comparatively, the LIBSVM model achieved an accuracy of 94.26% by which ANN is much stronger but lower, respectively. Comparison of Decision Tree (DT) and the K-Nearest Neighbor (KNN) models showed accuracy of 78.11% and 75.40% respectively, which is not very impressive in the management of complex and nonlinear fraud patterns. In general, the findings demonstrate ANN as the most valid model of the tested ones.

The study shows that AI-based solutions can successfully detect fraudulent telephony activities using textual descriptions of real-life events. The ANN model was highly accurate, precise, recalled, and F1-score, although it surpassed traditional machine learning algorithms. This result highlights the power of the approach of combining feature extraction via NLP with neural networks in order to identify complex, nonlinear fraud patterns. The ROC curve and the confusion matrix indicate that the false positives and false negatives are very low, and it is emphasized that detection is very strong. Comprehensively, the results suggest that textual data, when well processed and modeled, is a viable and dynamic method of detecting real-time frauds in telecommunication systems.

Table 4: Comparative Accuracy Performance of Models for Telephony Fraud Detection

Models	*Accuracy*
ANN	98.53
LIBSVM[23]	94.26
DT[24]	78.11
KNN[25]	75.40

Conclusion and Future Work

An AI-driven data-driven framework has been created to detect fraudulent telephony activities through textual data of real-world sources. The system makes use of the natural language processing methods to preprocess and derive meaningful features, which are in turn processed by an Artificial Neural Network to detect complex patterns of fraud. ANN model had proper performance with accuracy of 98.53, precision of 97.97, recall of 98.25, and F1-score of 98.11. Relative analysis with other machine learning models like Decision Tree, KNN and SVM showed that the ANN had a better chance to detect nonlinear and changing fraud patterns. The outcomes suggest that AI-powered textual analysis is an effective scalable and dependable solution to detect telecommunication fraud, as it overcomes the drawbacks of the traditional number-based methods. Future Research will include the inclusion of multilingual and multi-channel fraud reports to enhance the generalization of the models to different telephony systems. The inclusion of adaptive mechanisms of learning and real-time monitoring will enable ongoing identification of new patterns of fraud. Also, textual analysis can be combined with call metadata, voice features, and user behavioral analytics, which can further increase the accuracy of detection. The solutions may be stronger and more efficient as they explore hybrid AI models and incremental learning strategies that would guarantee the telecommunication networks to be more secure and reliable over time.

References

A. Gupta, K. Raghav, and P. Dhakad, "The Effect on the Telecom Industry and Consumers after the Introduction of Reliance Jio," *Int. J. Eng. Manag. Res.*, vol. 9, no. 3, pp. 118–137, Jul. 2019, doi: 10.31033/ijemr.9.3.16.

S. Dixit, S. Gupta, and C. V Ravishankar, "Lohit: An online detection \& control system for cellular sms spam," *IASTED Commun. Network, Inf. Secur.*, 2005.

I. A. Ibrahim, I. Mohammed, and B. Saidu, "Fraud management system in detecting fraud in cellular telephone networks," *Int. J. Innov. Res. Comput. Sci*, vol. 3, no. 3, pp. 92–99, 2015.

R. C. R. Karne, A. K. R. Pasham, and G. Pratibha, "Classification of Intrusion Detection System and its

Methodologies," in *International Conference on Research Challenges in Engineering and Technolog*, IEEE, 2016.

S. Achouche, U. B. Yalamanchi, and N. Raveendran, "Method, apparatus, and computer-readable medium for performing a data exchange on a data exchange framework," 2019

H. E. Bordjiba, E. B. Karbab, and M. Debbabi, "Data-driven approach for automatic telephony threat analysis and campaign detection," *Digit. Investig.*, vol. 24, pp. S131–S141, Mar. 2018, doi: 10.1016/j.diin.2018.01.016.

F. O. Aranuwa, "Hybridized intelligent data analysis model for fraud detection in mobile communication networks," 2013.

M. Alazab, S. Venkatraman, P. Watters, and M. Alazab, "Zero-day malware detection based on supervised learning algorithms of API call signatures," in *9th Australasian Data Mining Conference, AusDM-2011*, 2010, pp. 171–182.

Q. Zhao, K. Chen, T. Li, Y. Yang, and X. F. Wang, "Detecting telecommunication fraud by understanding the contents of a call," *Cybersecurity*, 2018, doi: 10.1186/s42400-018-0008-5.

S. Garg, "Predictive Analytics and Auto Remediation using Artificial Inteligence and Machine learning in Cloud Computing Operations," *Int. J. Innov. Res. Eng. Multidiscip. Phys. Sci.*, vol. 7, no. 2, 2019, doi: 10.5281/zenodo.15362327.

M. Arafat, A. Qusef, and G. Sammour, "Detection of Wangiri Telecommunication Fraud Using Ensemble Learning," in *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology, JEEIT 2019 - Proceedings*, 2019. doi: 10.1109/JEEIT.2019.8717528.

R. Zhong, X. Dong, R. Lin, and H. Zou, "An Incremental Identification Method for Fraud Phone Calls Based on Broad Learning System," in *2019 IEEE 19th International Conference on Communication Technology (ICCT)*, IEEE, Oct. 2019, pp. 1306–1310. doi: 10.1109/ICCT46805.2019.8947271.

R. Li, Y. Zhang, Y. Tuo, and P. Chang, "A Novel Method for Detecting Telecom Fraud User," in *2018 3rd International Conference on Information Systems Engineering (ICISE)*, IEEE, May 2018, pp. 46–50. doi: 10.1109/ICISE.2018.00016.

H. Wang and J. Sang, "Speaker Identity Tracing Using Fingerprint Data Hiding against Telecommunications Fraud," in *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, IEEE, Nov. 2018, pp. 554–559. doi: 10.23919/APSIPA.2018.8659749.

B. Hussain, Q. Du, and P. Ren, "Big data-driven anomaly detection in cellular networks," in *2017 IEEE/CIC International Conference on Communications in China (ICCC)*, IEEE, Oct. 2017, pp. 1–6. doi: 10.1109/ICCChina.2017.8330468.

A. Yulianto, Adiwijaya, M. A. Bijaksana, and K. M. Lhaksmana, "Fraud detection on international direct dial call using hybrid NBTree algorithm and Kullback Leibler divergence," in *2017 5th International Conference on Information and Communication Technology (ICoIC7)*, IEEE, May 2017, pp. 1–6. doi: 10.1109/ICoICT.2017.8074676.

Y. Khan, S. Shafiq, A. Naeem, S. Ahmed, N. Safwan, and S. Hussain, "Customers Churn Prediction using Artificial Neural Networks (ANN) in Telecom Industry," *Int. J. Adv. Comput. Sci. Appl.*, vol. 10, no. 9, 2019, doi: 10.14569/IJACSA.2019.0100918.

R. Sallehuddin, S. Ibrahim, A. Mohd Zain, and A. Hussein Elmi, "Detecting SIM Box Fraud by Using Support Vector Machine and Artificial Neural Network," *J. Teknol.*, vol. 74, no. 1, Apr. 2015, doi: 10.11113/jt.v74.2649.

A. U. S. Khan, N. Akhtar, and M. N. Qureshi, "Real-time credit-card fraud detection using artificial neural network tuned by simulated annealing algorithm," in *Proceedings of international conference on recent trends in information, telecommunication and computing, ITC*, 2014, pp. 113–121. doi: 02.ITC.2014.5.65.

F. Ogwueleka, "Fraud Detection In Mobile Communications Networks Using User Profiling And Classification Techniques," *J. Sci. Technol.*, vol. 29, no. 3, Jan. 2010, doi: 10.4314/just.v29i3.50052.

H. Farvaresh and M. M. Sepehri, "A data mining framework for detecting subscription fraud in telecommunication," *Eng. Appl. Artif. Intell.*, vol. 24, no. 1, pp. 182–194, Feb. 2011, doi: 10.1016/j.engappai.2010.05.009.

A. Mohamed, A. F. M. Bandi, A. R. Tamrin, M. D. Jaafar, S. Hasan, and F. Jusof, "Telecommunication Fraud Prediction Using Backpropagation Neural Network," in *2009 International Conference of Soft Computing and Pattern Recognition*, IEEE, 2009, pp. 259–265. doi: 10.1109/SoCPaR.2009.60.

S. Subudhi and S. Panigrahi, "Use of fuzzy clustering and support vector machine for detecting fraud in mobile telecommunication networks," *Int. J. Secur. Networks*, vol. 11, no. 1/2, p. 3, 2016, doi: 10.1504/IJSN.2016.075069.

W. Moudani and F. Chakik, "Fraud Detection in Mobile Telecommunication," vol. 1, no. 1, 2013, doi: 10.7763/LNSE.2013.V1.17.

N. Bajaj *et al.*, "Fraud detection in telephone conversations for financial services using linguistic features," *arXiv Prepr. arXiv1912.04748*, 2019.

Gupta, A. K., Polu, A. R., Narra, B., Buddula, D. V. K. R., Patchipulusu, H. H. S., & Vattikonda, N. (2024). Leveraging deep learning models for intrusion detection systems for secure networks. *Journal of Computer Science and Technology Studies*, *6*(2), 199-208.

Narra, B., Buddula, D. V. K. R., Patchipulusu, H., Vattikonda, N., Gupta, A., & Polu, A. R. (2024). The integration of artificial intelligence in software

development: Trends, tools, and future prospects. *Available at SSRN 5596472*.

Achuthananda, R. P., Bhumeka, N., Dheeraj Varun Kumar, R. B., Hari Hara, S. P., & Navya, V. (2024). Evaluating machine learning approaches for personalized movie recommendations: A comprehensive analysis. *J Contemp Edu Theo Artific Intel: JCETAI-115*.

Waditwar, P. (2024) The Intersection of Strategic Sourcing and Artificial Intelligence: A Paradigm Shift for Modern Organizations. Open Journal of Business and Management, 12, 4073-4085. doi: 10.4236/ojbm.2024.126204.

Bitkuri, V., Kendyala, R., Kurma, J., Mamidala, J. V., Attipalli, A., & Enokkaren, S. J. (2024). A Survey on Blockchain-Enabled ERP Systems for Secure Supply Chain Processes and Cloud Integration. *International Journal of Technology, Management and Humanities*, *10*(04), 126-135.

Mamidala, J. V., Bitkuri, V., Attipalli, A., Kendyala, R., Kurma, J., & Enokkaren, S. J. (2024). Machine Learning Approaches to Salary Prediction in Human Resource Payroll Systems. *Journal of Computer Science and Technology Studies*, *6*(5), 341-349.

Waditwar, P. (2024) AI for Bathsheba Syndrome: Ethical Implications and Preventative Strategies. Open Journal of Leadership, 13, 321-341. doi: 10.4236/ojl.2024.133020

Attipalli, A., Kendyala, R., Kurma, J., Mamidala, J. V., Bitkuri, V., & Enokkaren, S. J. (2024). Privacy Preservation in the Cloud: A Comprehensive Review of Encryption and Anonymization Methods. *International Journal of Multidisciplinary on Science and Management IJMSM*, *1*(1).

Tamilmani, V., Maniar, V., Singh, A. A., Kothamaram, R. R., Rajendran, D., & Namburi, V. D. (2024). A Review of Cyber Threat Detection in Software-Defined and Virtualized Networking Infrastructures. *International Journal of Technology, Management and Humanities*, *10*(04), 136-146.

Singh, A. A. S., Kothamaram, R. R., Rajendran, D., Deepak, V., Namburi, V. T., & Maniar, V. (2024). A Review on Model-Driven Development with a Focus on Microsoft PowerApps. *International Journal of Humanities, Science Innovations and Management Studies*, *1*(1), 43-56.

Padur, S. K. R. (2024). AI-augmented platform engineering: Redefining developer experience through autonomous, self-optimizing enterprise systems. *International Journal of Science, Engineering and Technology*.

Gangineni, V. N., Tyagadurgam, M. S. V., Pabbineedi, S., Penmetsa, M., Bhumireddy, J. R., & Chalasani, R. (2024). AI-Powered Cybersecurity Risk Scoring for Financial Institutions Using Machine Learning Techniques (Approved by ICITET 2024). *Journal of Artificial Intelligence & Cloud Computing*.

S. R. Sagili, C. Goswami, V. C. Bharathi, S. Ananthi, K. Rani and R. Sathya, "Identification of Diabetic Retinopathy by Transfer Learning Based Retinal Images," 2024 9th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2024, pp. 1149-1154, doi: 10.1109/ICCES63552.2024.10859381.

S. R. Sagili and T. B. Kinsman, "Drive Dash: Vehicle Crash Insights Reporting System," 2024 International Conference on Intelligent Systems and Advanced Applications (ICISAA), Pune, India, 2024, pp. 1-6, doi: 10.1109/ICISAA62385.2024.10828724.

Padur, S. K. R. (2024). Securing Oracle Integration Cloud ERP ecosystems, zero trust architecture, data governance, and compliance automation. *International Journal of Science, Engineering and Technology*, *12*(4), 10-5281.

S. R. Sagili, S. Chidambaranathan, N. Nallametti, H. M. Bodele, L. Raja and P. G. Gayathri, "NeuroPCA: Enhancing Alzheimer's disorder Disease Detection through Optimized Feature Reduction and Machine Learning," 2024 Third International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT), Trichirappalli, India, 2024, pp. 1-9, doi: 10.1109/ICEEICT61591.2024.10718628.

S. R. Sagili, V. K, B. Puli, P. Sundaramoorthy, M. R and K. N V, "Advancing Cervical Cancer Identification using Generative-based Adversarial Networks: An Integrative Learning Methodology," 2025 6th International Conference for Emerging Technology (INCET), BELGAUM, India, 2025, pp. 1-5, doi: 10.1109/INCET64471.2025.11140170.

Routhu, K. K. (2024). Beyond Automation: AI-Powered Employee Engagement Journeys in Oracle HCM Cloud. *KOS Journal of AIML, Data Science, and Robotics*, *1*(1), 1-6.

Routhu, K. K. (2024). The future of HCM: Evaluating Oracle's and SAP's AI-powered solutions for workforce strategy. *Journal of Artificial Intelligence, Machine Learning & Data Science*, *2*(2), 2942-2947.

Sannapureddy, R., Nadella, V. M., & Nelavelli, S. (2024). Edge–Cloud Continuums for Latency-Sensitive Tasks. *International Journal of AI, BigData, Computational and Management Studies*, *5*(4), 189-201.

Arigela, A. K., Brahmareddy, A., Sreenivas, T. S., Selvan, M. P., Venu, N., & Lal, D. K. (2024, December). Optimizing Energy Efficiency and Latency in IoT Devices Through AI-Based Adaptive Protocols in Fog-Edge Computing Environments. In *Congress on Smart Computing Technologies* **(pp. 595-607).** Singapore: Springer Nature Singapore.

Nadella, V. M. (2024). Al-Native 6G Network Management. *American International Journal of Computer Science and Technology*, *6*(1), 23-37.