Systematic Literature Review on CNN-LSTM Model for Smishing Detection in Hausa and English Messages

[1]Innocent Omale Ocheme
[1]Information Systems and Technology Department,
Faculty of Computing,
National Open University of Nigeria, Jabi, Abuja. Nigeria.
Email: rossinno [AT] yahoo.com

[2]Olawale Surajudeen Adebayo
[2]Cybersecurity Department, Faculty of Computing,
National Open University of Nigeria, Jabi, Abuja. Nigeria.
Email: waleadebayo [AT] noun.edu.ng

[3]Adenrele Afolorunso
Computer Science Department, Faculty of Computing,
National Open University of Nigeria, Jabi, Abuja. Nigeria.
Email: aafolorunso [AT] noun.edu.ng

Abstract— **Despite advancements in machine learning and cybersecurity, traditional rule-based and machine learning (ML) techniques struggle to keep pace with the continuously evolving tactics of cybercriminals. Deep Learning (DL) models such as hybrid Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) models, have demonstrated improved performance in smishing detection in a mixed mobile environment supporting Hausa and English messages. This work provides a systematic literature review (SLR) following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. Synthesizing the 41 selected papers using the SLR approach by analyzing the DL - based smishing detection methods, identified previous research efforts, datasets type for models training, their effectiveness, limitations and operational challenges such as computational and algorithms complexities. This review provides a clearer understanding of smishing attacks, refinement of detection algorithms and discusses research gaps and future directions to address current challenges and improvement of smishing detection systems.**

Keywords - Cyberattacks, Deep Learning, PRISMA, Smishing, Spam, Vulnerability

I. INTRODUCTION

The advancement of broadband internet technology and rapid digitalization has unprecedentedly promoted mobile communication such as increased usage of smart phones. Smartphones are becoming increasingly luxurious, and users are connecting to the internet worldwide (Muhammad et al, 2024). In mobile communication, Short Message Service (SMS) has gained popularity due to its ubiquitous communication channel. Short message service which can also be called text messaging, has revolutionized mobile communication by enabling rapid and convenient information exchange (Turban et al, 2010). SMS is the most widely used feature on mobile phones, more than 5 billion people send and receive text messages globally, with around 40 billion messages sent every day around the world (Dobson et al, 2024). The number of SMS users is projected to reach 5.9 billion by 2025 (Al-Kabbi et al, 2024). Ogunsanwo et al. (2025) observed that despite its widespread use, it comes with some flaws that have made it a target for spammers to send SMS spam. SMS spam, which is also known as mobile spam or text spam refers to as unsolicited and unwanted messages sent to mobile devices. Smishing, a form of phishing that uses text messages, leverages social engineering to deceive users into clicking malicious links or divulging sensitive personal identifiable information (SPII). Smishing or SMS phishing poses serious cybersecurity and usability challenges to mobile communication system ranging from privacy violation, financial losses and phishing attacks. For instances, in 2023 SMS spam accounted for 45% of global mobile message traffic, with smishing attacks surging by 62% compared to previous year which resulted to around $10 billion in global financial loses (GSMA, 2023; Symantec, 2024). According to Abdallah et al. (2020) A survey exposes that 68% of mobile phone users are affected by SMS Spam. The world sent 8.3 trillion SMS messages in the year 2017, the number of SMS messages sent monthly is 690 billion, so SMS is important for business communications.

Current smishing detection methods have limitations, with high false-positive rates (FPR) and difficulty adapting to new attack strategies. This highlights the need for new and more accurate techniques capable of distinguishing smishing messages from legitimate ones. Simon (2025) proposed a novel hybrid DL models framework for the detection and blocking of hate speech on Facebook English-Hausa code-mixed language. The model combines the strengths of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to form a Convolutional – Long Short-Term Memory (C-LSTM) with glove word embeddings to efficiently identify and block hate speech content. The datasets used in this work were collected from Facebook and were

labeled manually as hate speech or non-hate speech. Their paper also analyses the model's architecture, data preprocessing techniques, and the experimental results show that the hybrid C-LSTM model generated 0.90 accuracy which outperformed the CNN, LSTM, and the baseline-models. The findings suggest that hate speech in Facebook English-Hausa code-mixed language can be detected and blocked instantly before posting. Ogunsanwo et al. (2025) developed a Hybridized Deep learning Techniques for Enhanced SMS Spam Detection system. They developed Term Frequency-Inverse Document Frequency (TF-IDF) and Bidirectional Encoder Representations from Transformers (BERT), Long Short Time Memory (LSTM), Convolutional Neural Network (CNN), CNN-LSTM and Linear Regression (LR). The CNN-LSTM model achieves the highest accuracy of 99%, followed by LSTM 98% and Logistic Regression 94%. CNN-LSTM also recorded superior performance in precision 98%, Sen 91%, and F1-score 94%. However, the study was based on limited dataset for model training and poses real-time detection complexity. formatter will need to create these components, incorporating the applicable criteria that follow. Abdallah et al. (2020) developed a hybrid CNN-LSTM Model for SMS Spam Detection in Arabic and English Messages. DL CNN and LSTM intended to deal with mixed text messages that are written in Arabic or English. For the comparative evaluation, other well-known machine learning algorithms were tested. The experimental results show that their CNN-LSTM model outperforms the other algorithms. It achieved a very good accuracy of 98.37%. However, it is observed that there is the gap of imbalanced/noisy datasets and real-time application challenges.

Despite advancements in machine learning and cybersecurity, conventional smishing detection systems struggle to keep pace with the continuously evolving tactics of cybercriminals. This challenge is particularly significant when identifying Smishing on mobile devices, which often have limited computational resources. Many existing models focus solely on binary classification, neglecting the need to detect and classify the diverse variants of phishing within multiclass datasets (Miriam et al, 2025). Muhammad et al. (2024) states that the insufficiency of traditional phishing detection methods such as user education and rule-based methods against sophisticated phishing attack techniques has led researchers to exploration of possible artificial intelligence - AI-based solutions. While several ML-based models have been proposed, the fact that attackers use advanced innovative methods that are continuously changing to carry out phishing attacks renders previously proposed machine learning models ineffective against sophisticated attacks (Tosin et al, 2024).

The purpose of this review is to provide a clearer understanding of Smishing attacks detection, including the techniques used in the field of phishing detection and mitigation. Additionally, this review examines previous research efforts to facilitate the improvement of existing Smishing detection methods, assess the empirical evidence regarding the efficacy of DL algorithms in smishing detection, summarize the advantages and drawbacks of adapted datasets and current DL algorithms in detecting smishing attacks, and to identify the emerging challenges and potential improvements in current smishing detection research. Hence, a systematic literature review needed to understand how deep learning algorithms were developed, what kinds of datasets were used to train the model, and what the gaps are, and what the future direction is. The remainder of this paper is structured as follows: Section II reviews analysis of related works, Section III details the methodology, Section IV discusses results and findings. Section V provides the conclusion with future research directions for advancement of smishing detection.

A. Research Questions Concerns

1. What are the deep learning methods being applied recently to detect smishing attacks and their efficiency?
2. How does the different types of adapted datasets impact the performance of smishing detection models?
3. What are the emerging challenges and proffered future solutions in enhancing smishing detection systems?

II. RELATED WORKS

Ghourabi et al. (2020) proposed a hybrid CNN+LSTM architecture and evaluated on multilingual SMS corpora (Arabic/English mixed) to capture local and sequential patterns. The dataset sizes for non-English languages were small, indicating that results may not generalize across dialects. Roy et al. (2020) compared CNN and LSTM models on SMS spam datasets and showed deep models can outperform classical baselines in-domain. Evaluations were mostly on standard public corpora, cross-dataset generalization and adversarial tests were limited.

Ghourabi et al. (2021) proposed a hybrid security model called “SM-Detector” aiming to detect smishing messages in mobile environments. To increase the efficiency of “SM-Detector,” combined three different detection methods: (i) identification of malicious (Uniform Resource Locators) URLs, (ii) identification of suspected words, phone numbers and emails with regular expression analysis, and (iii) classification of messages using (Bidirectional Encoder Representations from Transformers) BERT-based algorithms to distinguish spam messages. “SM-Detector” also includes a mobile application allowing the user to check their SMS and report smishing messages. Its strength is that it can deal with mixed text messages written in Arabic or English. The experimental evaluation conducted on English and Arabic datasets showed a remarkable accuracy of 99.63%. Gadde et al. (2021) in SMS Spam Detection using ML and DL Techniques applied various machine learning and deep learning techniques for SMS spam detection. They used a dataset from UCI and build a spam detection model and also used python for all implementations. Their experimental results show that LSTM model

outperforms previous models in spam detection with an accuracy of 98.5%.

Stephen et al. (2022) analyses how combining various deep learning algorithms, such as CNN and LSTM, leads to improved performance in detecting malicious activity compared to single-algorithm methods. Their work examined the characteristics that make these hybrid models effective, including their ability to approximate complex functions, minimize empirical risk, and manage model complexity. Aliza et al. (2022) A Comparative Analysis of SMS Spam Detection; Conducted a systematic comparison of ML models on public SMS corpora and produced strong baseline results. The result reinforces cross-dataset performance drops, emphasizes need for standardized datasets and protocols. Emad et al. (2022) proposed two primary deep learning models, one combining CNN with LSTM layers, and another utilizing a custom (Artificial Neural Network) ANN with multiple layers. Through rigorous testing on the (University of New South Wales, for network intrusion and detection systems) UNSW-NB15 dataset, both models demonstrate superior performance when compared to traditional Logistic Regression (LR), with the ANN model achieving slightly higher accuracy in detecting anomalies. The author's research brings to notice the critical need for robust security in expanding Internet of Things (IoT) infrastructure and highlights the efficiency of deep learning in order to develop adaptable and better intrusion detection systems (IDS) solutions.

Kar & Debbarma (2023) introduced an effective feature extraction and hybrid methodology for deriving diverse features from code-mixed texts. A quantum search optimization approach was subsequently utilized to enhance the recovered attributes which decreases data complexity during the detecting phase. They utilized a hybrid diagonal-gated recurrent neural network (RNN) to identify hate speech and assess the sentiment conveyed in the phrase. The authors evaluated the feasibility of the (Federated Graph Recurrent Neural Networks) FEDGRNN approach utilizing the Hate Speech and Offensive Content (HASOC) 2019 dataset. The dataset facilitates the multilingual assessment of the approach's efficacy, as it comprises postings in Hindi, English, and German. The simulation results indicated that the accuracy of FEDGRNN for Task-1, Task-2, and Task-3 on the multilingual code-mixed texts dataset was 87.74%, 88.98%, and 84.74%, respectively. The FEDGRNN method significantly enhances accuracy, precision, recall, and F-measure relative to existing classifiers, suggesting its potential as a reliable and efficient instrument for sentiment analysis and hate speech detection in multilingual and code-mixed texts. Sayan et al. (2023) research paper focuses on the detection of multilingual spam SMS using the Naive Bayes classifier. The proposed approach leverages the Naive Bayes algorithm's simplicity and efficiency to classify SMS messages as spam or non-spam, while specifically addressing the multilingual aspect. A comprehensive dataset of labeled SMS messages was collected and preprocessed, including text normalization, and tokenization, and language translation. Feature extraction techniques such as word frequency, keyword presence, and message length are employed to represent the SMS messages in a numerical format suitable for the Naive Bayes classifier. The conditional probabilities of features occurring in spam and non-spam messages are calculated using maximum likelihood estimation. The system's performance was conducted, including accuracy assessment, precision computation, recall, and F1 score calculation. The results demonstrate the efficacy of the novel approach in detecting multilingual spam SMS. The Naive Bayes classifier exhibits high accuracy and efficiency in classifying spam messages across various languages, offering users protection from unwanted and potentially harmful content. Furthermore, to facilitate easy utilization and accessibility, a user interface was developed to interact with the spam detection system. The user interface allows users to input messages and the language of the message and receive real-time feedback on the likelihood of a message being spam. The research findings highlight the significance of considering language diversity in spam detection systems and provide insights into the difficulties and opportunities associated with multilingual spam SMS detection. The over fitting and imbalanced dataset are noted challenges. Robert et al. (2023) applied various machine-learning models to develop an optimized model that effectively, reliably, and precisely identifies and filter out spam or junk message from a genuine SMS text. The dataset used is a combination of self-acquired data and internet collected dataset with 60–40 ham to spam partitions. With regards to the accuracy of the model, the Bernoulli Naive Bayes achieved the highest performance with 96.63% accuracy upon optimization.

In Edward et al. (2023), they attempted to avoid spam by making an SMS filtering machine using various algorithms. This study, aiming to compare the accuracy percentage of SMS spam detection using Naive Bayes, Support Vector Machine (SVM), LSTM and CNN. These algorithms have been implemented and tested over a dataset consisting of 5.574 records. This testing gave specific results of mean, which are 96.95% by using the Naive Bayes Algorithm, followed by 97.93% by using SVM, 98.57 by CNN, and 99.1% by LSTM leads in this testing. Dare et al. (2023) in their study introduced a novel approach utilizing Natural Language Processing (NLP) and machine learning models, particularly BERT, for SMS spam detection and classification. Data preprocessing techniques, such as stop word removal and tokenization, are applied, along with feature extraction using BERT. Machine learning models, including SVM, Logistic Regression, Naive Bayes, Gradient Boosting, and Random Forest, are integrated with BERT for differentiating spam from ham messages. Evaluation results revealed that the Naïve Bayes classifier + BERT model achieves the highest accuracy at 97.31% with the fastest execution time of 0.3 seconds on the test dataset. This approach demonstrates a notable enhancement in spam detection efficiency and a low false-positive rate. The developed model presents a valuable

solution to combat SMS spam, ensuring faster and more accurate detection. This model not only safeguards users' privacy but also assists network providers in effectively identifying and blocking SMS spam messages.

Uddin et al. (2024) in Explainable Detector: Transformer-based SMS spam detection with explainability analysis fine-tuned BERT/RoBERTa (Robustly Optimised BERT Pretraining Approach) variants on standard SMS corpora and applied XAI (feature attributions) to explain model decisions. The results show high in-domain accuracy reported, but explanations for very short texts remain noisy and need human-factor validation. Muhammad et al. (2024) developed an Enhancing Smishing Detection using a DL approach for Improved Accuracy and Reduced False Positives. A multi-layered DL architecture - CNN-LSTM, which leverages the benefits of both CNN and LSTM structures. The CNN-LSTM architecture shows robust performance by achieving a 0.9974 accuracy score and a high precision score, indicating a low number of false positives in detecting smishing attacks. There is still lagging of datasets insufficiency and enhanced methods to significantly reduce high false-positives and difficulty adapting to new attack strategies. Ibrahim et al. (2024) proposed a technique that detects Arabic phishing messages using natural language processing and a random forest classifier. The performance of the random forest classifier is compared with other machine learning algorithms, namely, K-Nearest Neighbors (KNN), AdaBoost, and Logistic Regression. According to all evaluation matrices, the random forest classifier has outperformed other classifiers. The model was trained with 638 phishing messages and 4844 legitimate ones. The experimental outcomes indicate that the proposed approach has obtained an accuracy of 98.66%, 99.10% precision, 98.23% recall, and 98.67% F1 score. Shdefat et al. (2024) in ML-Based Solution for SMS Spam Detection Problem, used deep learning models and advanced image processing techniques, they demonstrate high accuracy and efficiency in detecting face modifications and malicious content within digital images. The authors illustrate the proposed technique by evaluating its performance on a unique dataset of unaltered and modified images. The method allows for improving detection accuracy by 93.99% and demonstrates robustness in real-time applications. The new method's effectiveness is confirmed by calculating precision and recall metrics. New research results develop advanced security solutions that can be used to enhance digital forensics, content authentication, and cybersecurity applications. Kyaw et al. (2024) conducted a systematic literature review on Natural Language Processing (NLP) and ML-based phishing email detection, following PRISMA guidelines.

Ahmed et al. (2025) applied the PRISMA guidelines to systematic literature review on Distributed Denial of Service (DDOS) Attacks, by presenting a detailed assessment of the approaches and methodologies taken throughout the nine years, emphasizing ML and DL techniques. However, these works did not specifically focus on the use of DL techniques for smishing detection. To the best of our knowledge, there are a few articles that have systematically reviewed DL-based smishing detection using the PRISMA guidelines. The improved hybridized CNN-LSTM model architecture fills the research gaps by effectively analyzing the intricate structure and meaning of messages, leading to a more robust and accurate detection of smishing threats in English and Hausa diverse linguistic in mobile environments. Miriam et al. (2025) in Deep Learning Approaches for Multi-Class Classification of Phishing Text Messages, utilized chain transformer model, (Generative Pre-Trained Transformer) GPT-2 for synthetic data generation and BERT for embedding to detect smishing within a multiclass dataset, including minority smishing variants. The precision rate exceeding 97% validation accuracy, demonstrating strong performance in detecting minority phishing types. The study concentrated mainly on data generation and classification. Akshata et al. (2025) proposed a structured framework involving pre-processing, feature extraction, and model training to enhance detection accuracy. They evaluated algorithms such as decision trees, ensemble models, and neural networks, examining their strengths and limitations in classifying spam. Their findings show that advanced ML and DL techniques out perform traditional methods, providing more efficient and accurate spam detection. Johari et al. (2025) in their key insights into recommended SMS spam detection, evaluated performance of models across ten SMS datasets and analyzed dataset heterogeneity and its impact on detectors. The study demonstrates benchmarking/data heterogeneity problem; calls for community standardization.

The Figure 1 below, shows the authors and years of publication of the related works.

Figure 1: A Visual Summary Chart of the Related Works

III. METHODOLOGY.

This systematic review provides an unbiased assessment of a research topic with reliable, thorough, and transparent methodologies. The study reviewed existing literature on DL-based smishing detection techniques by adopting PRISMA guidelines. Figure 2 illustrates the adopted PRISMA flow diagram for selecting relevant studies. The pre-existing studies were chosen from the vast Goggle scholar repository.

A. Overview of CNN-LSTM Model Framework

The reviewed hybridized DL architecture-based model for detecting SMS spam in a mixed mobile environment that

supports Hausa and English messages. Evaluation and comparison of the reviewed model with other ML algorithms.

B. Protocol and Phases of the Study

The PrBeferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), (Moher et al., 2009) and the established guidelines in the work of (Kitchenham et al., 2009) were adopted in this review.

C. Keywords for Search

In this study, the literature search methodology of Kitchenham et al. (2009) was applied. The core search terms were carefully selected to identify the most relevant search phrases.

D. Source of Data

Reputable academic research databases covering engineering and computational disciplines with a good publication reputation are necessary to obtain data from trustworthy sources. As indicated in Table 1, these databases included, but were not limited to, Science Direct, Elsevier, Nature, Springer, (Institute of Electrical and Electronics Engineers) IEEE, Future Internet, Academia and Research-gate. These databases were chosen because they are widely used in technology, engineering and computer science. The study's main source of data came from conference proceedings and journals published by these organizations. Papers from unverified and unreliable sources, such as Wikipedia etc, were not considered.

Table 1. Database Sources

S/N	Database Sources	No. of Articles extracted
1.	Elsivier	**2**
2.	Science Direct	**4**
3.	IEEE	**7**
4.	Nature	**5**
5.	Future Internet	**3**
6.	Academia (University based)	**15**
7.	Systems	**5**

E. Inclusion and Exclusion Criteria for Paper Selection

A set of criteria was used to choose the papers for this investigation. Only papers that fit and fulfilled these requirements were chosen.

E.1 Inclusion Criteria

i. Published using the English language
ii. Published within 5years. However, some literature which influenced the development of DL for smishing detection were included in the study.
iii. The relevance to the topic or answer to the research questions.

E.2 Exclusion Criteria

i. The studies which are not related to the research questions such as general cyberattacks.
ii. Traditional machine learning techniques or signature-based solutions for phishing detection.
iii. Non research article, books, chapters, editorials, summaries of workshops, duplicated publication on the same study. For duplicated publications on the same topic, the latest but known publication will be selected. Table 2 displays the criteria together with the corresponding rationale.

Table 2: Inclusion/Exclusion Criteria of research publications

S/N	Criteria	Justification
1.	The original research publication was not a survey or review paper.	The research papers are focused on the Smishing detection techniques.
2.	The proposed solutions must be on the methodologies/techniques/processes of Smishing attack detection.	This research aims to aid newer and expert researchers in the development of better techniques and approaches.
3.	The publication must be a full-length paper.	Short papers are insufficient in providing relevant information on the proposed solution.
4.	The language chosen for writing the research paper must be written in the English language.	The publication must be written in English language.
5.	The paper must be published between 2020-2025	The coverage of the Systematic Literature Review is 5 years, from 2020-2025.

(Ahmed et al. (2025), P57)

F. Study Selection Process

A large set of papers were identified by the initial keyword searches i.e. CNN – LSTM Smishing Detection on the chosen database platforms. After eliminating the duplicate research, the research papers that match the inclusion/exclusion criteria were thoroughly examined, about 60 publications were left for reading. Only papers written in English and published between 2020-2025 were chosen. 274 articles were initially identified from the five databases, after sorting and duplicates

elimination, 79 were eventually selected. The selected articles were categorized as journal and conference proceedings. The distribution of journal and conference proceedings, and years of selected papers are illustrated in Table 3. After reading through the papers, using the inclusion/exclusion criteria again, 41 papers were left for the systematic literature review. A total of forty-one (41) articles were assessed and listed as follows. There were six (6) reviews in 2020, four (4) reviews in 2021, five (5) reviews in 2022, seven (7) reviews in 2023, nine (9) reviews in 2024 and ten (10) reviews in 2025. The identification, screening, eligibility and included phases are presented below in Figure 2.

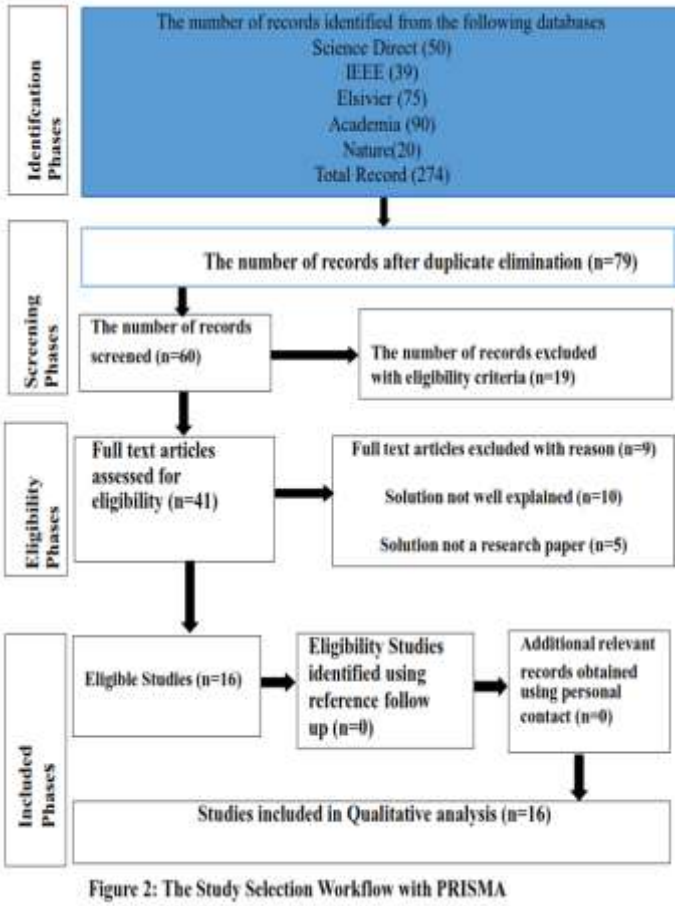

Figure 2: The Study Selection Workflow with PRISMA

IV. RESULT AND DISCUSSION.

Mostly peer-reviewed works published between 2020 and 2025 were the subject of this study. This was due to the fact that the objective was to determine which models and frameworks were utilized in the past years for detection of Smishing attacks. Table 3 summarizes the selected reviewed articles, including the authors, the year and title of the publications, the problems addressed, the methodology or framework they used to address those problems, the datasets used and the research gaps of the study. It was challenging to describe the methods employed in some of the research publications, and several of the researchers failed to explicitly explain the limitations of their study. According to the study survey, more works are coming up due to the novelty of smishing attack and the introduction of AI's DL approaches in its detection and mitigation.

RQ1. ML and DL architectures such as BERT, CNN, LSTM and hybridized CNN-LSTM are being deployed for efficient smishing detection. Results of high accuracy, ranging from 98% to 99.70% etc, indicate that they are highly efficient in detection of smishing attacks. Therefore, there are minimal false negatives and false positives rates.

RQ2. The fact that algorithms learn from data and the quality of the dataset used in training an ML/DL model directly impacts the performance of that model, making AI to be data-centric. The more relevant dataset scopus used in the model training, the better the efficiency, divergence and self-learning of the trained model.

RQ3. Emerging challenges are very dynamic due to research and advancement in technology. Some of the challenges such as insufficient, unbalanced/noisy datasets, AI weaponized smishing attacks etc, are minimized by deployment of modern datasets preprocessing techniques such as Synthetic Minority Oversampling Technique (SMOTE) function of Python Scikit Learn library, rigorous testing and retraining of models with updated and larger datasets scopus for model self-learning and autonomy in smishing detection. More solutions are recommended in Future Work.

The contributions of this study are,

1. A systematic literature review using the PRISMA approach with transparency and no bias.
2. Conduct a qualitative analysis of 41 selected papers to categorize and present various DL approaches for smishing detection.
3. Discover emerging challenges and proffer future solutions in enhancing smishing detection.

Table 3: Number of Smishing attack detection reviewed published articles per year

S/N	Number of Published Papers	Publication Year
1.	6	2020
2.	4	2021
3.	5	2022
4.	7	2023
5.	9	2024
6.	10	2025

With respect to the evaluation methods, different researchers, use different techniques to assess their models' effectiveness i.e. accuracy, precision, recall, F1 score, false positive rate (FPR), receiver operating characteristic (ROC) curve, and area under the curve (AUC) etc. Therefore, due to the deployment of different datasets and different evaluation methods, it is not possible to compare the effectiveness of each research work. However, this study will present their accuracy with the F1 score etc, to assess the balanced performance of the model. Table 4 presents the results, main contributions and limitations of smishing detection systems in the selected articles.

Table 4: Summarized the above selected reviewed article

Authors	**Problem Addressed**	**Methodology**	**Datasets Used**	**Performance Metrics**	**Research Gaps**
Ogunsanwo et al. (2025)	SMS Spam Detection system	Hybridized DL BERT, CNN, LSTM, CNN-LSTM Models	Kaggle datasets repository	Accuracy Recall Precision F-1 Score	Limited datasets, Computational complexity
Muhammad et al. (2024)	Datasets insufficiency and to significantly reduce high false-positives	A multi-layered DL architecture - CNN-LSTM	On-line datasets repositories	Accuracy Precision Recall F-1Score	Lagging datasets insufficiency, larger and more diverse datasets needed.
Miriam et al. (2025)	Detect smishing within a multiclass dataset	GPT-2 for synthetic data generation and BERT for embeddings	Kaggle datasets repository	Precision	The study concentrated mainly on data generation and classification.
Abdallah et al. (2020)	SMS Spam Detection in Arabic and English Messages.	DL CNN and LSTM Model	Arabic datasets from Telcos and English from Kaggle.com	Accuracy	Noisy datasets and real-time application challenges.
Stephen et al. (2022)	Examined the characteristics that make these hybrid models effective.	DL CNN-LSTM Model	English Datasets from Kaggle.com	Accuracy	Concentrated only on improving model performance in detecting malicious activity compared to single-algorithm methods.
Emad et al. (2022)	DL efficiency in developing adaptable IDS solutions for IoT infrastructures.	Multiple layers of CNN with LSTM layers, and Artificial Neural Networ k (ANN)	UNSW-NB15 dataset	Accuracy	Limited study scope to only DL IDS efficiency.
Akshata et al. (2025)	Examining ML & DL strengths and limitations in classifying spam	Decision trees, Ensemble models, and Neural Networks		Accuracy Precision	High false positive rate. Computational complexity.
Uddin et al. (2024)	SMS spam detection with explainability analysis	Explainable Detector BERT/ RoBERTa	Variants on standard SMS corpora and applied XAI (feature attribution)	Accuracy Precision	Need explanations for very short texts remain noisy and need human-factor validation.
Aliza et al. (2022)	Comparative Analysis of SMS Spam Detection	Systematic comparison of ML models	Public SMS corpora	Accuracy	The result reinforces cross-dataset performance drops, emphasizes need for standardized datasets and protocols.
Ghourabi et al. (2020)	To capture local and sequential patterns	Hybrid CNN+ LSTM architecture and evaluated on multilingual SMS corpora (Arabic/English mixed)	SMS corpora (Arabic/ English mixed)	Accuracy Precision Recall F-1Score	The dataset sizes for non-English languages were small, indicating that results may not generalize across dialects.
Roy et al. (2020)	Evaluating Deep models & standard public corpora.	CNN-LSTM Model	SMS spam datasets	Accuracy	Cross-dataset generalization and adversarial tests were limited.
Ibrahim et al. (2024)	Evaluation of random forest classifier with other ML algorithms, namely, KNN, AdaBoost and LR.	Random Forest Classifier and ML KNN, LR and AdaBoost	638 Arabic phishing messages and 4844 legitimate ones.	Accuracy Precision, Recall and F-1 score	Explore the possibility of implementing the classifier in real time and developing a

					smartphone application that lets users quickly detect fake information.
Dare et al. (2023)	SMS spam detection and classification	NLP & BERT (Bidirectional Encoder Representations from Transformers)	Test datasets	Accuracy Precision, Recall and F-1 score	Developed model ensures faster and more accurate detection, but need balanced datasets.
Edward et al.(2023)	Compare the accuracy percentage of SMS spam detection	Naive Bayes, Support Vector Machine (SVM), LSTM and CNN	Test dataset consisting of 5.574 records	Accuracy Precision,	Non-generalised detection on inadequate datasets. High FPR, TNR issues.
Gadde et al. (2021)	SMS spam detection	Various ML & DL Techniques	Dataset from UCI	Accuracy Precision, Recall and F-1 score	Computational complexity problems
Eppakayala et al. (2025)	Developed a real-time email automated spam filtering system	5 ML classifiers: LR, Decision tree, KNN, Gaussian naive Bayes and AdaBoost.	Two different spam email datasets	Precision, Recall, and F1-score.	Imbalanced datasets and computational dexterity.

V. CONCLUSION

This study presents literature evaluation on Smishing attack detection. Articles from 2020 to 2025 were the only ones considered. Researchers' various approaches, methods, and strategies were examined and documented. The study ended with a summary of all the findings from the previous review, which would serve as a guide to researchers interested in smishing attacks detection and mitigation. The study is also expected to advance research and development of comprehensive solutions such as improvement in low-resource languages datasets and model's real-time incorporation in telecommunications systems and much more.

A. Future Work

This study summarized some of the future work from the papers reviewed for future researchers in Smishing attacks detection. These are to:

1. extend the research results to other types of Phishing attack with adapted different multilingual smishing messages that could improve the detection process.
2. recommend more improved datasets quality in low-resource Hausa language to enhance model robustness.
3. investigate self-generated (synthetic) datasets for industrial & regulatory standards compliance.
4. improve AI - DL techniques for advance models' algorithms through rigorous testing and retraining to guarantee higher robustness, effective self-learning and efficacy.
5. Smishing attacks leverage on IP phones as veritable tools. Incorporate future advanced technology studies to evaluate IP source addresses, acknowledgement, reset, transmission control protocol/internet protocol (TCP/IP), internet control message protocol (ICMP) segments and ports more effectively to ameliorate vulnerabilities and curb smishing attacks.

B. Funding

There is no funding for this study. I am a self-sponsored doctoral student seeking to build advance cybersecurity capacity.

C. Conflict of Interests

There are no conflicts of interests in this research work and its publication.

VI. ACKNOWLEDGMENT

With profound gratitude to the Lord, God Almighty for the gift of life and good health, the Dean, Faculty of Computing, Prof. Olayemi Mikail Olaniyi, Head of Information Systems and Technology Department, Dr. Vivian O. Nwaocha, my supervisor, Dr. Olawale Surajudeen Adebayo, for his mentorship, fellow students for their knowledge sharing and cooperation.

My sincere appreciation to my family for their support in prayers, understanding and cooperation. Also worthy of mention are the management of Nigeria Atomic Energy Commission (NAEC), my subordinates and colleagues for their support. God bless you richly, Amen.

REFERENCES

[1] Aaron, Z., & Katongo, O. P., (2025). An Enhanced Machine Learning with NLP Modelling Technique for Smishing Attacks Detection in Low-Resourced Languages. http://dx.doi.org/10.2139/ssrn.5195337.

[2] Abdallah, G., Mahmood, A. M., & Qusay, M.A., (2020). A Hybrid CNN-LSTM Model for SMS Spam Detection in Arabic and English Messages. *Future Internet 2020, 12, 156;* https://doi:10.3390/fi12090156.

[3] Abiramasundari, S., & Ramaswamy, V., (2025). Cacography Based Ransomware Email Phishing Attack Prevention using Language Pack Tuned Transformer Language Model. *Scientific Reports (2025) 15(1).* https://doi.org/10.1038/s41598-025-06530-8.

[4] Adebayo, O. S., Udo-Nya, E., (2021). Comparative Analysis of Machine Learning Algorithms for the Detection of Android Malware. *International Journal of Innovative Research in Advanced Engineering (IJIRAE),* Issue 09, Volume 8. doi: https://doi.org/10.26562/ijirae.2021.v0809.004.

[5] Adebayo, O. S., Aziz, N. A., (2019). Improved Malware Detection Model with Apriori Association Rule and Particle Swarm Optimization. *Hindawi Security and Communication Networks* Volume 2019, Article ID 2850932. https://doi.org/10.1155/2019/2850932.

[6] Adebayo, O. S., Mabayoje, M. A., Mishra, A., & Oluwafemi, O., (2012). Malware Detection, Supportive Software Agents and Its Classification Schemes. *International Journal of Network Security & Its Applications (IJNSA),* Vol.4, No.6, November 2012. DOI: 10.5121/ijnsa.2012.460333

[7] Ahmed, R. B., Adebayo, O. S., Ahmad, S., Anyaora, P. C., Atiku, M., Adeleke, N. D., & Baba, M., (2025). Systematic Literature Review on Distributed Denial of Service Attack. *UNIABUJA Journal of Engineering and Technology*, Volume 2, Issue 1, 2025; 51-68. https://ujet.uniabuja.edu.ng.

[8] Akshatha, P. S., Sonia, D'Souza., (2025). Advanced Spam Detection Techniques in Social Media: A Comprehensive Analysis of Machine Learning and Deep Learning Approaches. *IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS).*

[9] Aliza, H. Y., (2022). A Comparative Analysis of SMS Spam Detection (academic comparative study / TUNI thesis).Tampere University Research Portal.

[10] Al-Kabbi, H, A., Derakhshi, M, R, F., Pashazadeh, S., (2024). Hierarchical Two-Level Feature Fusion Approach for SMS Spam Filtering. *Intelligent automation and soft computing (IASC).* DOI: 10.32604/iasc.2024.050452.

[11] Alzahrani, A., (2024). Explainable AI-based Framework for Efficient Detection of Spam from Text using an Enhanced Ensemble Technique. *Engineering, Technology & Applied Science Research* Vol. 14, No. 4, 2024, 15596-15601 15596 DOI: https://doi.org/10.48084/etasr.7901.

[12] Aqsa, Shaikh., Mariya, Shaikh., Srivaramangai, Ramanujam., (2025).Smishing Detection: Combating SMS Phishing Attacks by Utilizing Machine-Learning Algorithms. *International Journal of Innovative Technology and Exploring Engineering 2025.*

[13] Dare, O. A., Ojo, A., (2023). SMS Spam Detection and Classification to Combat Abuse in Telephone Networks Using Natural Language Processing. *Journal of Advances in Mathematics and Computer Science.*

[14] Dobson, R., Whittaker, R., Abroms, L. C; Bramley, D; Caroline, F., McRobbie, H; Melanie Stowell, M., & Rodgers, A., (2024). Don't Forget the Humble Text Message: 25 Years of Text Messaging in Health. *Journal of Medical Internet Research.* https://www.jmir.org/2024/1/e59888.

[15] Edward, W., Ghinaa, Z. N., (2023). Spam Detection in Short Message Service (SMS) Using Naïve Bayes, SVM, LSTM, and CNN. *International Conference on Information Technology, Computer, and Electrical Engineering.*

[16] Eppakayala, S., Kumar, S., & Ankula, G., (2025). Intelligent Spam Filtering and Analysis Using Web Automation, Report Lab, and Ensemble Machine Learning Techniques. DOI:10.21203/rs.3.rs-5941111/v1.

[17] Ersin, E. E., Ondokuz, M. D., & Ozkan, S., & Erdal, K., (2020). Filtering Turkish Spam using LSTM from Deep Learning Techniques. https://doi:10.1109/ISDFS49300.2020.9116440

[18] Gadde, S., Lakshmanarao, A., Satyanarayana, S., (2021). SMS Spam Detection using Machine Learning and Deep Learning Techniques. *7th International Conference on Advanced Computing and Communication Systems (ICACCS).*

[19] Ghourabi, A., Mahmood, M., Alzubi, Q., (2020). A hybrid CNN-LSTM model for SMS spam detection in Arabic and English Messages. *Future Internet (2020) 12(9).*

[20] Gopalsamy, M., (2024). Identification and Classification of Phishing Emails Based on Machine Learning Techniques To Improvise Cyber Security. *IJSART - Volume 10 Issue 10.*

[21] GSMA. 2023. Global mobile messaging trends: SMS spam and security risks [Annual Report]. GSM Association. Retrieved February 1, 2025, from [redacted link].

[22] Ibrahim, A., Alyousef, S., Alajmi, H., Aldossari, R., & Masmoudi, F., (2024). Phishing Detection in Arabic SMS Messages using Natural Language Processing. *2024 Seventh International Women in Data Science Conference at Prince Sultan University (WiDS PSU).* DOI:10.1109/wids-psu61003.2024.00040.

[23] Ige, T., Kiekintveld, C., Piplai, A., (2024). Deep Learning-Based Speech and Vision Synthesis to Improve Phishing Attack Detection through a Multi-layer Adaptive Framework.

[24] Johari, M. F., (2025). Key insights into recommended SMS spam detection. *Nature Scientific Reports / comparative dataset study.*

[25] Kar, P., Debbarma, S., (2023). Multilingual Hate Speech Detection Sentimental Analysis on Social Media Platforms using Optimal Feature Extraction and Hybrid Diagonal Gated Recurrent Neural Network. *Springer* Journal of Supercomputing, 2023.

[26] Kyaw, P. H., Gutierrez, J., and Ghobakhlou, A., (2024). A Systematic Review of Deep Learning Techniques for Phishing Email Detection. *Electronics 2024, 13, 3823.* https://doi.org/ 10.3390/electronics13193823.

[27] Mahmud, T., Prince, M.A.H., Ali, M.H., Hossain, M.S., & Andersson, K., (2024). Enhancing Cybersecurity: Hybrid Deep Learning Approaches to Smishing Attack Detection. *Systems (2024) 12(11).* https://doi.org/ 10.3390/systems12110490.

[28] Mambina, I. S., Ndibwile, J., & Andkisangiri, F. M., (2022). Classifying Swahili Smishing Attacks for Mobile Money Users: A Machine-Learning Approach. *IEEE Access (2022) 10 83061-83074.*

[29] Miriam, L. M., & Muhammad, F. I., (2025). Deep Learning Approaches for Multi-Class Classification of Phishing Text Messages. https://doi:10.20944/preprints202508.1703.v1.

[30] Mohammad, R.F.D., Elnaz, Z.M., Hussein, A. A., & Ahmed, H. J. A., (2024). PCLF: Parallel CNN-LSTM Fusion Model for SMS Spam Filtering. *BIO Web of Conferences (2024) 97.* https://doi.org/10.1051/bioconf/20249700136.

[31] Muhammad, R. R., Fadhli, M., Muin, Y., (2025). Comparison of Statistical and Linguistic Feature in K-Nearest Neighbors (KNN) & Neural Network Algorithms for SMS Spam Classification. *The 8th International Joint Conference on Science and Technology.* http://dx.doi.org/10.11594/nstp.2025.4812.

[32] Muhammad, K. M., Humaira, A., Moatsum, A., & Abid, M., (2024). Enhancing Smishing Detection: A Deep Learning Approach for Improved Accuracy and Reduced False Positives. *IEEE Access (2024) 12 137176-137193.*

[33] Ogunsanwo, G. O., Wycliff, O. J., Owoade A. A., Alaba, O. B., & Odulaja, G. O., (2025). Hybridized Deep learning Techniques for Enhanced SMS Spam Detection system. *UNIZIK Journal of Engineering and Applied Sciences (2025) 5(2) 2536-2551.*

[34] Robert, G., Luna, D., Redondo, J., (2023). A Machine Learning Approach for Efficient Spam Detection in Short Messaging System (SMS). *IEEE Region 10 Conference.*

[35] Roy, P. K., (2020). Deep learning to filter SMS Spam. *Journal/Elsevier.* Science Direct.

[36] Sayan, A. K. H., (2023). Detection of Multilingual Spam SMS Using Naïve Bayes Classifier. *IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA).*

[37] Simon, H., (2025). Identification and Blocking of Hate Speech in English-Hausa Code-mixed Language on Facebook: A Hybrid Deep Learning Framework. *Journal of Networking and Communication Systems*, Vol.8 No.1 Jan 2025.

[38] Symantec. 2024. Internet security threat report: Mobile threats and smishing trends. Symantec Retrieved from https://www.symantec.com/reports.

[39] Tian, Y., Dai, X., Li, Z., Guo, H., Mao, X., (2025). Improving the accuracy of cybersecurity spam email detection using ensemble techniques: A stacking approach Machine learning for spam email

detection. PLoS One 20(9): e0331574. https://doi.org/10.1371/journal.pone.0331574.

[40] Tosin, I., Kiekintveld, C., & Piplai, A., (2024). Deep Learning-Based Speech and Vision Synthesis to Improve Phishing Attack Detection through a Multi-layer Adaptive Framework.

[41] Uddin, M. A., (2025). Transformer / BERT fine-tuning for SMS Transformer-based language modeling for SMS spam (fine-tuning BERT/variants). arXiv+1.

[42] Wael, H. G., (2020). The Impact of Deep Learning Techniques on SMS Spam Filtering. *International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 11, No. 1, 2020.*

[43] Yuyun, Y. L., Nuali, A. A., & Mahule, D. G. M., (2025). BERT Sentiment Analysis for Detecting Fraudulent Messages. https://doi.org/1069916/jkbti.v4i2.24.