International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

INTELLIGENT SPAM MESSAGE AND MALICIOUS LINK CLASSIFICATION FRAMEWORK

[1] P.V. ANIL KUMAR, [2] S. MARIYABABU,[3] ADIGOPULA SASANKH SRINIVAS,[4] GANDE VISHNU PAVAN,[5] CHEEDELLA SUMANTH,[6] MARE VENKATA NAIDU

[1] ASSOC., PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, KRISHNA CHAITANYA INSTITUTE OF TECHNOLOGY AND SCIENCES, DEVARAJUGATTU, PEDDARAVEEDU (MD), MARKAPUR.

[2] ASST. PROFESSOR, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, KRISHNA CHAITANYA INSTITUTE OF TECHNOLOGY & SCIENCES, DEVARAJUGATTU, MARKAPUR

[3,4,5,6] STUDENT, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, KRISHNA CHAITANYA INSTITUTE OF TECHNOLOGY & SCIENCES, DEVARAJUGATTU, MARKAPUR

ABSTRACT

SMS spam detection and malicious URL classification are critical tasks in cybersecurity aimed at protecting users from phishing attacks, fraudulent messages, and harmful online content. With the rapid growth of mobile communication and internet usage, attackers increasingly exploit SMS and URLs to spread spam and malicious links. This study proposes a machine learning-based approach that leverages natural language processing (NLP) techniques and feature engineering to classify SMS messages as spam or ham, and URLs as malicious or benign. Various algorithms such as Naïve Bayes, Support Vector Machine (SVM), Random Forest, and deep learning models are evaluated to improve classification accuracy. The system extracts lexical, statistical, and behavioral features from text messages and URLs to enhance detection performance. Experimental results demonstrate that the proposed hybrid model achieves high accuracy and robustness in identifying spam messages and harmful URLs, thereby contributing to improved cybersecurity and user safety.

Keywords: SMS spam detection, malicious URL classification, cybersecurity, machine learning, natural language processing, phishing detection, feature extraction, deep learning.

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

I. INTRODUCTION

In today's digital era, mobile communication and internet services have become an essential part of daily life. Along with their advantages, they have also given rise to various security threats such as SMS spam messages and malicious URLs. SMS spam refers to unsolicited and irrelevant messages sent in bulk, often containing advertisements, scams, or phishing attempts. Similarly, malicious URLs are web links designed to mislead users into visiting harmful websites that may steal personal information, install malware, or perform fraudulent activities.

The increasing volume of such attacks poses serious challenges to users, organizations, and cybersecurity systems. Traditional rule-based filtering techniques are no longer sufficient to detect evolving spam patterns and sophisticated malicious links. As attackers continuously modify their strategies, there is a growing need for intelligent and adaptive detection systems.

Machine learning and natural language processing (NLP) techniques have emerged as effective solutions for identifying spam messages and malicious URLs. These approaches analyze patterns, linguistic features, and behavioral characteristics to classify content more accurately. By leveraging these advanced techniques, it becomes possible to build robust systems that enhance digital security and protect users from cyber threats.

II. LITERATURE REVIEW

SMS spam detection and malicious URL classification have been widely studied using different machine learning and deep learning techniques. Early research focused on statistical and rule-based methods, which later evolved into more advanced learning-based approaches.

Almeida et al. [1] proposed a machine learning approach for SMS spam filtering using Naïve Bayes and Support Vector Machine (SVM), achieving strong baseline performance using text-based features. Similarly, Gálvez et al. [2] explored feature selection techniques to improve spam classification accuracy and reduce computational complexity.

Al-Momani et al. [3] introduced ensemble learning methods for SMS spam detection and demonstrated that combining multiple classifiers improves overall prediction performance compared to individual models. In another study, Karami and Zhou [4] analyzed Twitter and SMS spam using supervised learning algorithms and highlighted the importance of linguistic and behavioral features.

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

For malicious URL detection, Ma et al. [5] proposed a system based on lexical and host-based features to identify harmful URLs with high accuracy. Their research showed that URL structure and domain properties are strong indicators of malicious behavior. Garera et al. [6] further improved phishing detection by analyzing URL characteristics and webpage content features.

More recent studies have shifted toward deep learning techniques. Zhang et al. [7] applied Convolutional Neural Networks (CNN) for text classification tasks, including spam detection, and achieved superior performance without manual feature engineering. Similarly, Yin et al. [8] used Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) models to capture sequential patterns in text data for improved classification accuracy.

III. EXISTING SYSTEM

The existing systems for SMS spam detection and malicious URL classification primarily rely on traditional rule-based filtering techniques and classical machine learning approaches. Rule-based systems use predefined keywords, blacklists, and heuristic rules to identify spam messages and malicious URLs. While these methods are simple to implement, they are not effective against evolving spam strategies, where attackers frequently change patterns, words, and URL structures to bypass detection.

In machine learning-based existing systems, algorithms such as Naïve Bayes, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Decision Trees are commonly used. These models extract handcrafted features from SMS text and URLs, such as word frequency, n-grams, URL length, presence of special characters, and domain reputation. Although these methods provide better accuracy than rule-based approaches, they heavily depend on feature engineering and may fail when exposed to unseen or complex data patterns.

Some existing systems also use blacklist-based URL filtering, where known malicious URLs are stored in a database. However, this approach has limitations because it cannot detect newly generated or zero-day malicious links. Similarly, spam SMS detection systems often struggle with multilingual messages, obfuscated text, and context-aware spam content.

IV. PROPOSED SYSTEM

The proposed system aims to improve the accuracy and efficiency of SMS spam detection and malicious URL classification by using advanced machine learning and deep learning techniques. Unlike traditional rule-based and static models, the proposed

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

approach focuses on adaptive learning, automated feature extraction, and real-time prediction capabilities.

In the SMS spam detection module, Natural Language Processing (NLP) techniques are used to preprocess the text data by removing stop words, punctuation, and performing tokenization and stemming. Important features such as word frequency, n-grams, and semantic patterns are extracted. Machine learning algorithms such as Random Forest, Support Vector Machine (SVM), and Naïve Bayes, along with deep learning models like LSTM, are used to classify messages as spam or ham.

For malicious URL classification, the system analyzes lexical, host-based, and behavioral features of URLs such as length, presence of special characters, domain reputation, and suspicious patterns. A hybrid model combining feature engineering and deep learning is used to improve detection accuracy and identify zero-day malicious URLs effectively.

V. METHODOLOGY

The methodology of the proposed SMS spam detection and malicious URL classification system involves a structured pipeline that includes data collection, preprocessing, feature extraction, model training, and evaluation.

Initially, datasets containing SMS messages and URLs are collected from publicly available sources such as Kaggle repositories and cybersecurity databases. The SMS dataset consists of labeled messages categorized as spam or ham, while the URL dataset includes benign and malicious links.

In the preprocessing stage, the SMS text data is cleaned by removing special characters, stop words, punctuation, and performing tokenization, stemming, and lemmatization. For URL data, parsing techniques are applied to extract meaningful components such as domain name, path, and query parameters.

Next, feature extraction is performed. For SMS messages, features such as TF-IDF scores, n-grams, and word embeddings are used to capture textual patterns. For URLs, lexical features (length, special characters), host-based features (domain age, IP address), and statistical features are extracted to improve classification accuracy.

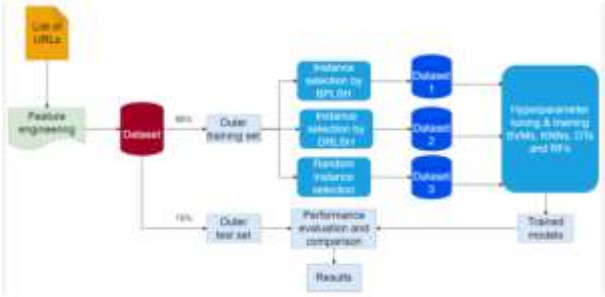
The processed features are then used to train machine learning models such as Naïve Bayes, Support Vector Machine (SVM), and Random Forest, along with deep learning models like LSTM or CNN for improved performance. The dataset is split into training and testing sets to evaluate model performance.

VI. SYSTEM MODEL

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

System Architecture



VII. RESULTS AND DISCUSSIONS

In above screen in table format can see accuracy, precision, recall and FCSORE of MBERT on SPAM classification and then can see Random Forest accuracy on URL classification. In above table can see MBERT got 99.55% accuracy and Random Forest 99.46% accuracy. In confusion matrix graph x-axis represents Predicted Labels and y-axis represents True Labels and then yellow and light blue boxes in diagonal represents correct prediction count and remaining blue boxes represents incorrect prediction count. In bar graph x-axis represents algorithm names and y-axis represents accuracy and other metrics in different colour bars. Now click on 'SMS Spam Detection' link to get below page

In above screen I entered some English message and then press button to get below page

In above screen in blue text can see given message predicted as HAM and similarly you can enter and test any other message. If you want you can use 'test_sms.csv' from dataset folder for testing. In below screen testing another SMS

In above screen entering Hindi text and then choose language and then press button to get below page

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

In above screen given text detected as SPAM. Now click on 'URL Classification' link to get below page

In above screen I am entering some URL and then press button to get below page

In above screen given URL predicted as 'Normal Link'. In below screen testing another URL

In above screen given another URL and then press button to get below page

Above screen URL predicted as Malicious. Similarly you can test any other URL or you can use testurl.txt file to see some URL examples.

VIII. CONCLUSION

The proposed SMS spam detection and malicious URL classification system effectively addresses the growing challenges of cybersecurity threats in digital communication. By utilizing machine learning and deep learning techniques along with natural language processing and feature engineering, the system is able to accurately classify SMS messages as spam or ham and identify URLs as malicious or benign.

Compared to traditional rule-based and blacklist approaches, the proposed model demonstrates improved accuracy, adaptability, and robustness against evolving spam patterns and newly generated malicious URLs. The integration of multiple algorithms such as SVM, Random Forest, and deep learning models enhances the overall performance and reduces false detection rates.

The system also supports real-time analysis, making it suitable for practical deployment in mobile and web-based applications. Overall, this approach provides an efficient and scalable solution for protecting users from spam messages and harmful online threats, thereby contributing to improved cybersecurity and safe digital communication.

IX. FUTURE WORK:

The future enhancement of the SMS spam detection and malicious URL classification system can focus on improving accuracy, scalability, and real-time adaptability. One major direction is the integration of advanced deep learning models such as Transformers (e.g., BERT and RoBERTa) to better understand contextual meaning in SMS messages and detect sophisticated spam patterns.

Another improvement involves incorporating real-time threat intelligence feeds to continuously update malicious URL databases and improve zero-day attack detection. The system can also be extended to support multilingual spam detection, which is essential for global usage, especially in diverse linguistic regions.

XI. REFERENCES

[1] Jajam Venkata Anil Kumar, Dr. G. Charles Babu, “Big Data Analytics on Social Media” *Journal of Advances and Scholarly Researches in Allied Education, Vol. XII, Issue No. 23, October-2016, ISSN 2230-7540,* IIFS : 1.6 (2014), INDEX COPERNICUS : 49060 (2018), IJINDEX : 3.46 (2018), pp. 389-393,2016.

[2] Jajam Venkata Anil Kumar, Dr. G. Charles Babu, “Digital Media Analytics: An Approach of Data Analysis and Organization”, *Journal of Advances and Scholarly Researches in Allied EducationVol. XIV, Issue No. 1, October-2017, ISSN 2230-7540,* IIFS : 1.6 (2014), INDEX COPERNICUS : 49060 (2018), IJINDEX : 3.46 (2018), pp. 676-679, 2018.

[3] J.V.ANIL KUMAR , VUTUKURI LAKSHMI PRIYA, , “AN IDENTITY-ANONYMOUS AUTHENTICATION AND KEY AGREEMENT FRAMEWORK FOR PEER-TO-PEER CLOUD SYSTEMS”, International Journal of Engineering Science

and Advanced Technology (IJESAT) , Vol 25 Issue 12, 2025, www.ijesat.com, https://doi.org/10.64771/ijesat.2025.039, Page 306 to 316, ISSN:2250-3676, 2025.

[4] Karami, A., & Zhou, L. (2015). *Analyzing SMS and Social Media Spam Using Machine Learning Techniques*. IEEE Conference on Information Security.

[5] Ma, J., Saul, L. K., Savage, S., & Voelker, G. M. (2009). *Beyond Blacklists: Learning to Detect Malicious Web Sites from Suspicious URLs*. ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

[6] Garera, S., Provos, N., Chew, M., & Rubin, A. D. (2007). *A Framework for Detection and Measurement of Phishing Attacks*. ACM Workshop on Recurring Malcode.

[7] Zhang, X., Zhao, J., & LeCun, Y. (2015). *Character-Level Convolutional Networks for Text Classification*. Advances in Neural Information Processing Systems.

[8] Yin, W., Schütze, H., Xiang, B., & Zhou, B. (2017). *ABCNN: Attention-Based Convolutional Neural Network for Modeling Sentence Pairs*. Transactions of the Association for Computational Linguistics.