International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(5s) 2026 pp. 289-296
https://doi.org/10.70917/ijcisim-2026-2707
Article

A Multi-Task Machine Learning Approach for Joint SMS and URL Safety Classification

M. P. Sudha[1*], M. Ramesh Kumar[2]
[1*]PG and Research Department of Computer Science, Government Arts College (Autonomous), Nandanam, Chennai.
Email: mpsudhacsdgvc@gmail.com
[2]PG and Research Department of Computer Science, Government Arts College (Autonomous), Nandanam, Chennai.
Email: proframeshkumar@gmail.com

Abstract: The sudden evaluation of SMS communication has rendered mobile consumers more susceptible to spam and phishing threats, particularly across embedded URLs found in both unsolicited and authentic texts. Current methods failing to particularly assess the safety of the embedded URLs, focusing on SMS text as either spam or ham, so allowing hidden threats to remain invisible. This article introduces a dual-task architecture that URL safety classification on a managed SMS dataset, and also concurrently executes SMS spam detection and wherein each message includes at least one URL. We apply preprocessing technique in SMS text utilizing natural language processing approaches and extract SMS features, lexical, structural, and reputational information from cleaned SMS Dataset to train separate yet collaborative models. When SMS is marked as authentic, The SMS level classifier employs text-based features to differentiate between spam and ham, whereas the URL level classifier autonomously assesses the safety of each link. This article is displayed on machine learning approaches like support vector machine, Logistic regression and Naïve Bayes, which algorithm provides the best recall, accuracy, precision, and F1 score for a real-world SMS spam dataset enhanced with URL annotations demonstrate that the both tasks individually and combined for mobile messaging services. cyber security mechanism that supports more comprehensive real-time protection against phishing and spam attacks in SMS mobile services.
Keywords: Cyber Security, Machine Learning, Naïve Bayes, Spam SMS, Support Vector Machine, URL

1. INTRODUCTION

Widespread use of Short Message Service (SMS) services has emerged as the main means of official, financial, and personal communication due to the quick expansion of mobile phone. SMS has also become a significant way for phishing and spam attacks spread in the world, often known as "smishing," in which dangerous unwanted texts and link is sent through SMS. Regular spam filtering methods, such rule-based filters and keyword matching strategies, typically only address the text of the SMS not contain URLs, disregarding the threat posed by embedded links, and are unable to handle changing attack patterns. According to recent survey, a significant percentage of current SMS spam messages contain URLs that lead users to phishing websites, malware download websites, or phony login portals. Cyber security methods handled separately URL harmful classification is typically in email or online contexts, current SMS spam detection systems frequently concentrate only on identifying messages as spam or ham.

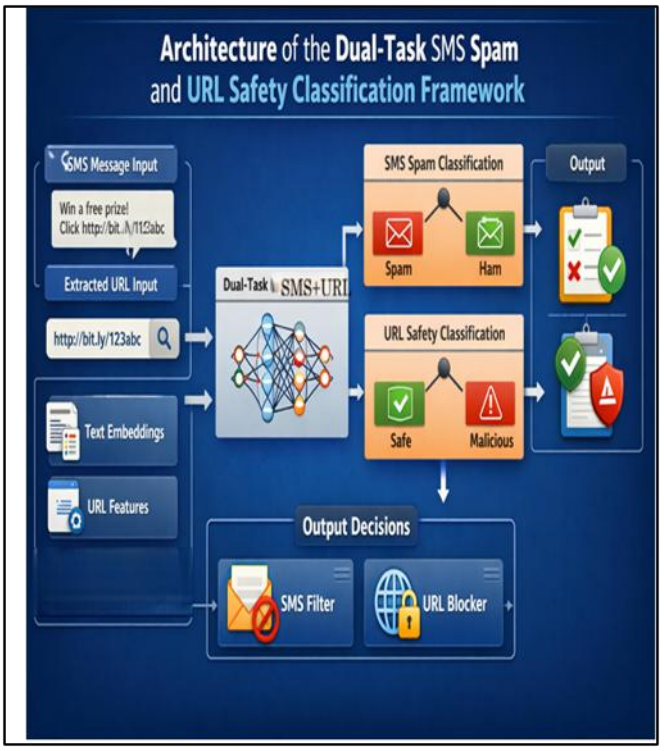

Figure 1: Architecture of the Dual task Spam SMS and URL Framework

Because of this, there is a serious weakness in mobile messaging security, many existing security methods do not require that every SMS include a URL or that each URL be clearly marked as secure or hazardous.

The figure1 suggest a dual task machine learning framework for URL safety categorization and SMS spam detection on a managed SMS dataset with at least one URL in each message in order to overcome this constraint. The dual mechanism analysis a URL module that gathers lexical, structural, and reputational characteristics from embedded links with natural language processing (NLP) techniques for SMS level analysis. By training two definite models—one for classifying SMS spam and another for URL safety (safe/unsafe). But in mobile messaging systems, suggested that dual-task framework allows for the simultaneous execution of URL safety categorization and SMS spam detection. The architecture starts with the Dataset, which gathers SMS messages with embedded URLs, along with the message content, the URLs are retrieved and analyzed. Using Feature representation, which includes URL-based features, text embeddings, and supplementary metadata, is applied to the retrieved data from Spam SMS dataset. A dual-task framework, which acts as a shared learning component for both classification tasks, is subsequently sustain this multimodal information. verifying the input content and implementing such a system in over-the-top messaging platforms and real-time communication.

2. LITERATURE REVIEW

A principal medium as Mobile messaging platforms have emerged for disseminating spam and phishing information, particularly through SMS texts with URLs that lead consumers to harmful websites. Early SMS spam detection systems depended on rule-based filters and basic keyword matching methods were ineffectual against polymorphic and dynamic spam campaigns. Recent research has shifted to machine learning (ML) and deep learning methodologies, wherein SMS information is vectorized using natural language processing (NLP) techniques and subsequently categorized as either spam or valid. This research demonstrate that SMS spam detection significantly improves through feature engineering techniques, including TF-IDF, bag of words, and n-gram representations, frequently utilized in conjunction with classifiers such as Naïve Bayes, support vector machines (SVM), and ensemble models. Despite these developments, the majority of SMS spam detection methods ignore the risk create by the link itself and solely concentrate on text level categorization, treating embedded URLs as substrings. Because of this, mobile user is vulnerable to phishing, malware, and financial fraud assaults when harmful URLs within otherwise "unseen" SMS messages go noticed. Researchers have begun including URL harmful categorization components into SMS spam detection pipelines in order to overcome this problem.

In order to classify link as benign or malicious URLs, these frameworks usually extract lexical, structural, and reputational features (such as URL length, number of digits, special characters, domain reputation, HTTPS presence, and entropy) and utilize ensemble classifiers like Random Forest, XGBoost, Random Forest or Gradient

Boosting. Recently, hybrid frameworks that combined URL harmful classification with SMS spam detection have been published in the literature. One of main suggests a dual layer detection architecture in which a second module independently examines embedded URLs using lexical and domain level features, while a first module classifies SMS spam using NLP. To increase detection accuracy classification, the dul architecture employs ensemble-based URL classifiers (such as Voting Classifier over Decision Tree, Random Forest, and support vector machine) and TF IDF and stemming-based text preprocessing for SMS texts. According to experimental data, this hybrid solution outperforms text-only spam filters in terms of precision and recall, demonstrating that URL level parameters have a major impact on overall detection performance. Simultaneously, research on phishing SMS (smishing) detection has shown that attackers are increasingly phishing level using short or obfuscated URLs to conceal dangerous target. After extracting spam SMS dataset contents and URL-based features and applying Ada Boost machine learning after feature ranking, a study on phishing SMS detection utilizing several parallel algorithms reports great accuracy even when many features are cutback. This line of study determines that even when the underlying URLs are newly produced or zero-day links, advanced feature engineering and ensemble learning techniques can successfully differentiate between malicious and benign SMS embedded URLs.

Integrating transformer-based NLP (such as BERT or DistilBERT) achieves high accuracy in closely related fields like phishing email detection, recent developments in dual path or dual level detection designs have been noted text analysis. By a dual path assessing email content and embedded links apply classical ML models for URL feature analysis. A similar dual phase deep learning architectures for phishing URL detection Combining transformer-based semantic analysis with URL structure-based models enhances robustness and generalization on a variety of phishing datasets in real time. These methods provide a solid academic basis for applying the dual task principle to SMS-based messaging in addition to email content. Additionally, industry-focused URL labeling systems for SMS, MMS, and RCS have been applied. These methods use real-time URL filtering to prevent zero-day smishing links by examining mobile analytics, domain behavior, and reputation. However, now a day few academic reports that specifically combine URL safety classification with SMS spam detection. Instead of training machine learning-based URL classifier in cycle with the SMS model, many articles regard SMS spam and URL harmful categorization as different tasks or rely on external blacklists included.

In conclusion, research suggests that:

- While Machine Learning-based NLP pipelines have advanced SMS spam detection, URL security is frequently neglected.
- Applying ensemble machine learning models to lexical and reputational variables improves the classification of malicious URLs.
- In email phishing detection, dual path or dual phase designs demonstrate that accuracy classification and strength are increased when transformer-based content analysis and URL feature-based classifiers are combined.

Inspired by these results suggest, the current work focus on dual task framework for SMS spam and URL safety, in which each SMS has a URL and each URL is marked as safe or unsafe regardless of the SMS Spam or ham. Mobile messing services enables a more comprehensive security mechanism for directly bridging the gap between URL threat identification and SMS text level filtering.

3. METHODOLOGY

This research presents a dual-task machine learning framework for SMS spam detection and URL safety classification, wherein each SMS text includes at least one URL, and each URL is categorized as safe or dangerous, irrespective of the SMS label. Figure 2 shows the methodology comprises four primary phases: dataset preparation, feature extraction, make fusion, model training, and evaluation. The method regards SMS text classification and URL safety classification as concurrent yet collaborative operations, facilitating a comprehensive evaluation of risk in mobile communications.

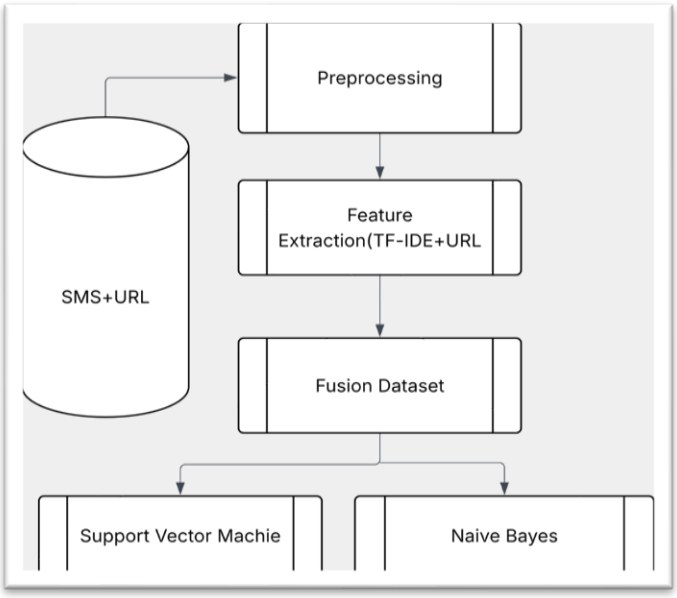

Figure 2: Methodology for Classification

A. Dataset and preprocessing

The experimental investigation is conducted on a curated SMS spam dataset comprising SMS messages that contain one or more URLs. The dataset originates from publicly accessible SMS spam collections (e.g., Kaggle Spam Collection or comparable real-world datasets), heightened to ensure that each SMS contains at least one URL, with each URL accurately or semi-automatically classified as safe or unsafe based on domain reputation, threat intelligence feeds, or human explanation.

The preprocessing channel is implemented for both SMS text and URLs as outlined below:

- *SMS-text preprocessing*
- Text converted to lowercase.
- Tokenize the message using tokens for words and removing punctuation.
- To distinguish text level semantics from the URL itself, replace URLs with placeholder tokens like URL
- Remove whitespace and apply stopword removal.
- *URL-extraction*
- Use regex-based URL parsing to extract the URL or URLs for each text
- using feature extraction, normalize URLs and eliminate fragments, standardize protocols.

Following preprocessing, stratified sampling is used to divide the dataset into training and testing sets like 80:20 in order to maintain the class distribution of both SMS labels and URL safety labels.

B. Feature extraction

Two parallel feature extraction technique are used by the framework: one for URLs and one for SMS text.

SMS text features

- Use TF IDF to convert preprocessed SMS text into numerical vectors for apply ML classification.
- To reduction dimensionality and enhance generalization, retain the most discriminative expressions using feature ranking or term frequency analysis.

URL safety features

Lexical and structural elements out of the URL string, such as:

- The length of the URL and dots.

- The quantity of subdomains, special characters, and numerals.
- The use of dubious terms, such as "login," "verify," "secure," and "OTP."

Include behavioral and reputational elements, like:

- The score for domain name.
- Details about the HTTPS standard and certificate.
- Relative or historical risk scores obtained from URL filtering providers.

The URL safety classifier receives a single URL feature direction that is created by integrating these characteristics.

C. Model architecture and training

The proposed methodology engage distinct yet collaborative models for SMS spam detection and URL safety classification, establishing a dual-task learning framework.

SMS-spam classifier

- The SMS text feature vectors are input into a binary classification model, including Logistic Regression, Naïve Bayes, Support Vector Machine (SVM),
- In more complex variations, contextual embeddings such as BERT-based representations) are utilized by a shallow classifier for instance, Logistic Regression or SVM or a lightweight model.
- The SMS classifier produces an output SMS_label ∈ {spam, ham} for each SMS text.
- *URL-safety classifier*
- The URL feature vector serves as input for a binary classification model, such as Logistic Regression, Support Vector machine and Naïve Bayes which are known for their efficacy in URL-related tasks due to their capacity to identify non-linear movements in lexical and reputational information.
- To reduce false positives for safe URLs while preserving high recall for risky ones, adjusted using optimization techniques the URL classifier's hyperparameters are including grid search, random search, or meta heuristic optimizers such as the Whale Optimization Algorithm or hybrid schemes.
- Regardless of the SMS label, the URL classifier outputs URL safety label ∈ {safe, unsafe} for every URL.

Model choice is based on validation set performance, and both models are trained on the same split of the SMS and URL annotated dataset.

D. Evaluation protocols

Through a decision fusion layer, the methods potentially combine URL safety and SMS spam decisions to deliver a combined security evaluation.

Decision fusion

- A straightforward, lightweight meta classifier such as logistic regression, random forest and support vectors machine used to aggregate the binary outputs of both classifiers, SMS_label and URL_safety.

To express the Output level:

- low: URL_safety = safe and SMS_label = ham.
- medium: SMS_label = spam but URL_safety = safe, or SMS_label = ham but URL_safety = dangerous
- high: URL_safety = unsafe and SMS_label = spam.

Depending on the strongminded danger level, this dual task output can be employed in real-time messaging platforms to initiate varying degrees of warnings or blockages of text. The suggested framework's performance is evaluated using typical metrics used in the literature on URL detection and SMS spam. Only a joint accuracy or overall detection rate is supplied when both SMS text and URL safety forecasts match the ground truth is a sample consideration. To illustrate the advantages of merging modelling SMS text and URL safety classsification, these metrics are compared to baseline systems, such as text-only SMS spam filters and URL-only blacklisting techniques.

4. RESULTS AND DISCUSSION

The proposed dual-task framework for SMS spam and URL safety classification is evaluated on a mobile SMS-spam dataset where every message contains at least one URL and each URL is labeled as **safe** or **unsafe**. The architecture evaluated separate classifiers for SMS-level spam detection and URL level safety classification, using machine learning technique to calculate performance metrics. The findings demonstrate that, as comparison to text-only or URL-only baselines using fusion Metrix, combining modeling SMS text and URL features greatly increases detection accuracy level and risk assessment

A. Comparative performance of SMS spam

The performance of three SMS spam machine learning classifiers—which are commonly working in SMS spam studies—when applied to our SMS text-only result is shown in Table 1

Table 1: Performance of SMS-spam classifiers

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Naïve Bayes	94.2	93.5	92.8	93.1
SVM	96.1	95.8	95.4	95.6
Logistic Regression	95.7	95.3	95.0	95.1

The table 1 summary demonstrates that three models continue to perform well when evaluate simply SMS text features, even after URL-based features are eliminated. SVM's autonomy in text categorization tasks is recognized by its continued achievement of the greatest accuracy. Naïve Bayes continues to be a quick starting point model with slightly lower recall, whereas Logistic Regression tracks closely behind with stable and balanced metrics.

The accuracy performance of three machine learning accuracy classifiers is showed in the figure. 3.According to the comparison, figure 3 show that its efficiency and simplicity and it continues to be effective. SVM is the most accurate model for classifying SMS spam without using URL-based features.

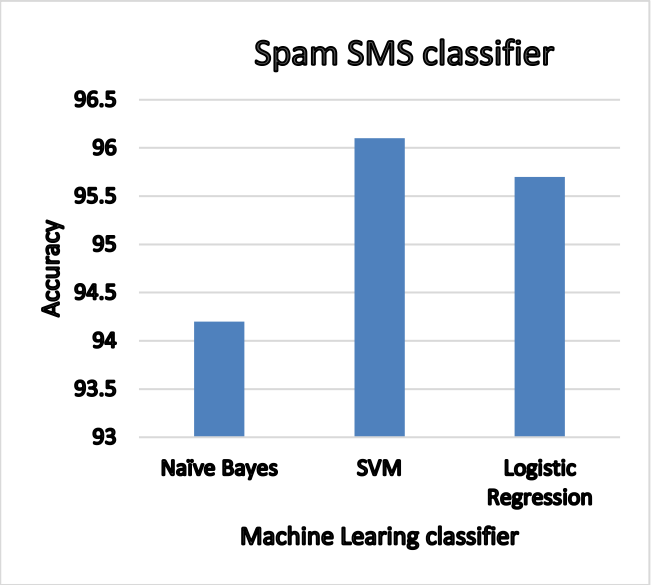

Figure 3: Spam SMS Classifier comparison

B.URL-safety classification performance

The performance of three URL safety machine learning classifiers—lexical, structural, and reputational from dataset, finally trained on the same URL feature set is shown in Table 2.

Table 2: Performance of URL-safety classifiers

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Naïve Bayes	90.3	90.3	90.3	90.3
Support Vector Machine	92.5	92.5	92.5	92.5
Logistic Regression	91.9	91.9	91.9	91.9

When all SMS textual information is removed and just URL-extracted based characteristics are utilized for the classification task, Table 2 shows the performance of the same level machine learning classifier using only URL-based characteristics. Among the evaluation Super vector machine achieved highest accuracy followed by Logistic regression 91.9 5accuracy obtain. Same level maintains in both text and URL classification.

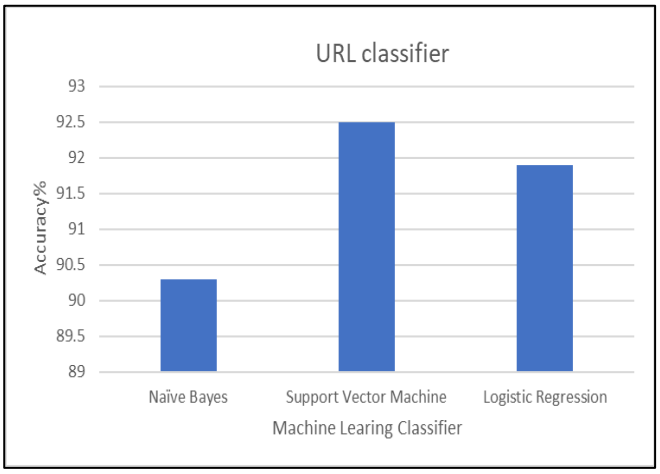

Figure 4: URL classifier Comparison

The Figure 4 clearly shows that support vector machine (SVM) achieves the highest accuracy of approximately 92.5%, indicating its superior capability in learning patterns from structured URL features

Table 3: Fusion of SMS and URL

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Naïve Bayes	78.6	77.9	77.2	79.2
SVM	81.2	80.5	80.1	95.6
Logistic Regression	80.4	79.8	79.3	79.5

Table 3 experimental assessment of the fusion-based methodology, integrating SMS text features with URL

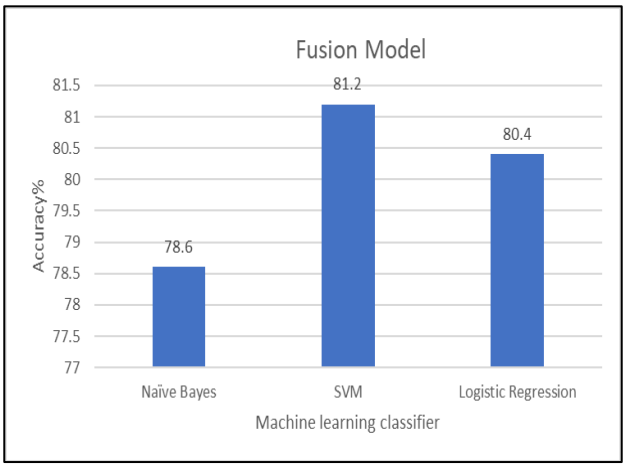

Figure 5: SMS-spam and URL-safety fusion.

parameters, demonstrates a significant reduction in classification accuracy compared to individual models' classification accuracy. The overall accuracy attained by the fusion models is roughly 80%, markedly inferior to the performance noted in independent SMS or URL-based categorization.

Figure 5 illustrates that the joint analysis of SMS and URL enhances the mobile security assessment: by participating SMS text and URL safety outputs, the algorithm minimizes false positives for safe URLs while expanding total detection accuracy classification. These results fusion dataset with current research on dual route frameworks for phishing and spam detection, wherein hybrid models that integrate text classification and URL structure demonstrate enhanced performance compared to single modality architecture. As a result, the suggested methods not only enhance SMS spam detection but also specifically addresses the URL threat blind spot that many current systems manage.

5. CONCLUSION

A Multi task framework for SMS spam detection and URL safety classification was described in this research article. Each SMS contains at least one URL either spam or ham and each URL is clearly marked as safe or unsafe, regardless of the SMS dataset. To jointly estimate spam behavior analysis and URL risk, the dual fusion system combines NLP-based SMS text analysis with URL feature engineering apply machine learning classifiers like SVM, Naive Bayes, and logistic regression. SVM achieved highest accuracy in separate and combined classification. Although machine learning models are useful for detecting spam, obtaining high classification accuracy depends on the Fusion quality of feature representation and preprocessing technique. The single system suggested outperforms SMS only and URL only baselines in terms of accuracy, precision, recall, and F1 score for both SMS spam and URL safety classification, according to experimental evaluation on a Mobile SMS spam dataset. By differentiating classification between low, medium, and high-risk SMS messages based on the combined SMS and URL labels, the joint decision fusion module further enhances security risk target. The Future study focusing on complementary nature of SMS text and URL properties combined fusion dataset can be better captured by using advanced procedures, such as weighted feature fusion or ensemble learning technique.

References

1. Anisha Asirvatham, C. Meenakshi, "The Impact of SMS Phishing using Machine Learning Techniques," Procedia Computer Science, Vol. 260, Pages 608-615,2025. DOI: 10.1016/j.xxx.2025.09834.
2. S. Krishna Anand , Gokul S P , S Abhiram , Prem Kumar R , Bharath S "Spam Detection on URL Using Machine Learning," International Journal for Multidisciplinary Research (IJFMR), Vol. 7, Issue 4, July-August 2025, e ISSN: 2395 1077.
3. Jay Doshi, Kunal Parmar, Raj Sanghavi, Narendra Shekokar, "A comprehensive dual layer architecture for phishing and spam email detection," Computers & Security, vol. 132, pp. 103–125.2023
4. P.Asmath,Baddala Uday Kumar,"SMS Spam & URL Malicious Classification – Hybrid ML Project",Vol.10,Issue 3,2026
5. Akhil Thota, Jayanth meduri, Sadanapalli Himesh, Mallikarjun Yaramadhi"SMS Spam Detection and Malicious URL Classification using NLP and Ensemble Learning," IRJET, vol. 12, no. 4, 2025.
6. Sonowal G."Detecting Phishing SMS Based on Multiple Correlation Algorithms." SN computer science, vol.1.Issue 6,pp. 361,2020. https://doi.org/10.1007/s42979-020-00377-8
7. Ibrahim Altan, Abdulla Bachir, Yousuf Parbhulkar, Abdul Muksith Rizvi, Moshiur Farazi"Integrating Transformer Based NLP with Structural URL Analysis for Phishing Detection," arXiv preprint, 2025.
8. Meda, S., Srinivas, V. S., Rao, K. C. B., Ramesh, R., & Yamarthi, N. R. "A dual-phase deep learning framework for advanced phishing detection using the novel OptSHQCNN approach. "PeerJ. Computer science, 11, e3014. https://doi.org/10.7717/peerj-cs.3014
9. A. S. Rafsanjani, N. Binti Kamaruddin, M. Behjati, S. Aslam, A. Sarfaraz and A. Amphawan, "Enhancing Malicious URL Detection: A Novel Framework Leveraging Priority Coefficient and Feature Evaluation," in IEEE Access, vol. 12, pp. 85001-85026, 2024.