IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Multi-Model Cyber Phishing Detection Using AI And Deep Learning

[1]Thumpiri Sai Prasuna, [2] Girish B G,

[1]Post Graduation Student Master of Technology, Department of Computer Science and Engineering SJCIT, Chickballapur , India.

[2]Assistant Professor, Department of Computer Science and Engineering, SJCIT, Chickballapur , India.

Abstract: The increasing use of digital communication has led to a rise in sophisticated phishing attacks targeting users through URLs, SMS messages, and manipulated images. Traditional detection methods often struggle to identify these evolving threats. This paper proposes a Multi-Model Cyber Phishing Detection framework using deep learning to analyze multiple data modalities. Dedicated preprocessing and feature extraction pipelines process lexical, semantic, and visual information independently before adaptive feature fusion. The integrated framework improves detection reliability and reduces false positives compared to single-modality approaches. Its modular architecture supports scalability, efficient feature learning, and real-time deployment. Experimental evaluation demonstrates consistent performance across different phishing scenarios, making the proposed system suitable for modern cybersecurity applications.

Index Terms - Cyber Phishing Detection, Deep Learning, Multi-Model Learning, URL Classification, SMS Spam Detection, Image-Based Phishing Analysis, Artificial Intelligence, Cybersecurity.

I. INTRODUCTION

The proliferation of digital services has changed how people and businesses communicate, perform financial activities, and navigate the Internet. While the digitization brings enhanced convenience and usability, it has made people more vulnerable to cyber threats. For instance, phishing continues to be one of the most prevalent and consequently destructive types of cyber crime. Attackers keep coming up with novel schemes aimed at the impersonation of trusted organizations or individuals to convince people into divulging their private data.

Traditional phishing attacks primarily relied on fraudulent websites and deceptive email messages. However, recent attack campaigns have become significantly more sophisticated, targeting users through multiple communication channels including malicious URLs, fraudulent SMS messages, QR codes, manipulated images, and AI-generated content. The availability of generative artificial intelligence has further accelerated the creation of highly convincing phishing material, making it increasingly difficult for users and conventional security systems to distinguish legitimate content from malicious attempts.

The majority of phishing detection systems today center on one major data source: methods based on URL either focus on lexical and domain information while text-based systems rely on a linguistic analyses of messages. Whereas image-based systems evaluate visual elements that may identify them as being modified or computer-generated. While all of the above techniques work fine within their operation domain, they tend to face difficulties when attackers use several types of deception in one campaign. In

view of this, the systems relying on a single type of input may experience increased false-positive rate or fail in recognizing more complex attacks.

Deep learning breakthroughs are proving to have huge applicability in cybersecurity because machine learning could comprehend complicated features directly from the raw data. Examples of this technology include neural networks such as convolutional networks, recurrent networks, transformers, and mechanisms of attention that achieved a lot of success in several classification tasks in the area of cybersecurity. These algorithms can help identify nonlinear dependencies and other context-related data, which allows them to detect new strategies of phishing without using only pre developed rules.

For this study secondary data has been collected. From the website of KSE the monthly stock prices for the sample firms are obtained from Jan 2010 to Dec 2014. And from the website of SBP the data for the macroeconomic variables are collected for the period of five years. The time series monthly data is collected on stock prices for sample firmsand relative macroeconomic variables for the period of 5 years. The data collection period is ranging from January 2010 to Dec 2014. Monthly prices of KSE -100 Index is taken from yahoo finance The integrated nature of the framework allows the system to analyze structural, semantic and visual features of the input simultaneously in generating a final security evaluation of the input.

The presented architecture is constructed in a manner that renders its deployment practical. Because of its modular structure, it is easy to make independent updates of each detection model as new phishing attacks become available while maintaining a unified prediction process through the fusion mechanism. This type of flexibility paves the way for future implementations in enterprise security systems, browser add-ons, mobile applications, and other types of cloud-based solutions in cybersecurity.

This study has made the following key contributions:

1. A unifying multi-model structure that is capable of detecting phishing across links, SMS messages, and images.
2. A modular deep-learning architecture that achieves separation of preprocessing, feature extraction, and decision-making stages.
3. An adaptive feature fusion technique that enables integration of different representations obtained from models to increase trustworthiness of detection.
4. Complete implementation that makes it possible to utilize it in real-time and with minimum inference times.

The proposed framework demonstrates how combining multiple complementary sources of information can strengthen phishing detection while providing a scalable foundation for next-generation intelligent cybersecurity systems.

II. RELATED WORK

Detecting phishing attacks is an interesting topic for researchers in view of the growing complexity of cybercrimes and the evolution of online media. Over the years, detection solutions have moved from the simpler, manually created rule-based systems to state-of-art machine learning and deep learning methods that can help in discovering complex attack schemes. However, detecting phishing relations remains a challenge.

In the early stages of detecting phishing attacks, techniques employed were mostly based on manual feature extraction from the URLs. Some of these included URL features such as domain length, lexical patterns, characters distribution, presence of special characters, or the hyperlink structure. Basic machine learning algorithms such as Decision Trees, Support Vector Machines, Naïve Bayes, Random Forests, and Logistic Regression were used because they are efficient to use and easy to implement. However, the effectiveness of these strategies in working with static datasets has its limits, which are related to their reliance on features that are manually extracted.

Deep learning has revolutionized feature learning through the use of unprocessed input data. RNNs (Recurrent Neural Networks), LSTMs (Long Short-Term Memory networks), and GRUs (Gated Recurrent

Units) have been able to efficiently make use of the sequential nature of character and token data, thereby improving phishing URL classification. In addition to the above, Bidirectional networks have been further used to improve interpretation of these data types because of their backward and forward processing capabilities, allowing for the reduction of reliance on lexical guidelines.

The study of phishing messages detection has developed from classical text classification techniques to the use of methods, such as Transformer-based models. Older methods were based on TF-IDF representation plus traditional classifiers to categorize messages as either authentic or fraudulent. Currently, using contextual embeddings created with different Transformer models facilitates the identification of common features inherent in messages from phishing campaigns.

The rise of image-based phishing detection has occurred in parallel with the increased employment of modified screenshots, imitation login pages, fake ads, and visuals produced by artificial intelligence. For the task of recognizing complex visual features, Convolutional Neural Networks are the most commonly used tool nowadays. Moreover, the technique of transfer learning contributes to the training of the models even in cases of limited availability of cybersecurity images. The cutting-edge backbone networks contribute to the high quality of using features in addition to the high computational efficiency of implementation.

While there has been considerable advancement in each individual detection domain, various existing solutions still work independently within the confines of a single modality. Applications like URL analysis, mail classification, and image validation are typically treated as independent pipelines, with minimal interaction between their learnt representations. Such isolated processing can lead to missing out on corroborative information that can be obtained by combining different types of data input.

The latest research is now exploring multimodal learning techniques that have applications in cybersecurity. Different methods like feature-level fusion, attention mechanisms, and ensemble learning have shown great promise due to their ability to combine information from various text, structural, and visual sources. By applying such methods, models can take advantage of complementary aspects and reduce uncertainty in terms of predictions. Nevertheless, despite the potential, a great number of methods used have proved to be computationally expensive, difficult to extend, and, in some cases, dependent on a certain dataset.

The system designed in this research paper solves these challenges by proposing a modular multi-model architecture that analyzes the URLs, SMS messages and digital pictures using specialized deep learning processes before integrating the output data through a flexible feature combining technique. In contrast to conventional ensemble methods that rely on final outputs, in our case, we have done the integration at the level of the features that make it possible for structural, semantic and visual information to participate in the process of classification together. Due to the modular nature of the design, it is possible to make independent upgrades of each detection system without changing the entire architecture enabling better scalability and maintanability as well as long-term adaptability to new phishing methods.

III. DESIGN AND METHODOLOGY

The multi-model cyber phishing detection using ai and deep learning system has been developed for the purpose of phishing attacks detection through reviewing different formats of digital data, for instance, urls, sms messages, images, etc. in contrast with the old method of detection that uses one technique, the new system employs specific deep learning models for the detection of phishing attacks through identifying its structural, semantic, and visual properties. the system is module-based, which enables the increase in the accuracy of the detection process, the possibility to scale the system up, and the implementation of the upgrades.

The overall architecture of the phishing detection system is divided into five principal phases: data acquisition, data preprocessing, feature extraction using deep learning methods, adaptive feature fusion, and finally, classification and decision making. all phases have a distinct role to play, along with smooth interaction with other.

3.1 SYSTEM DESIGN

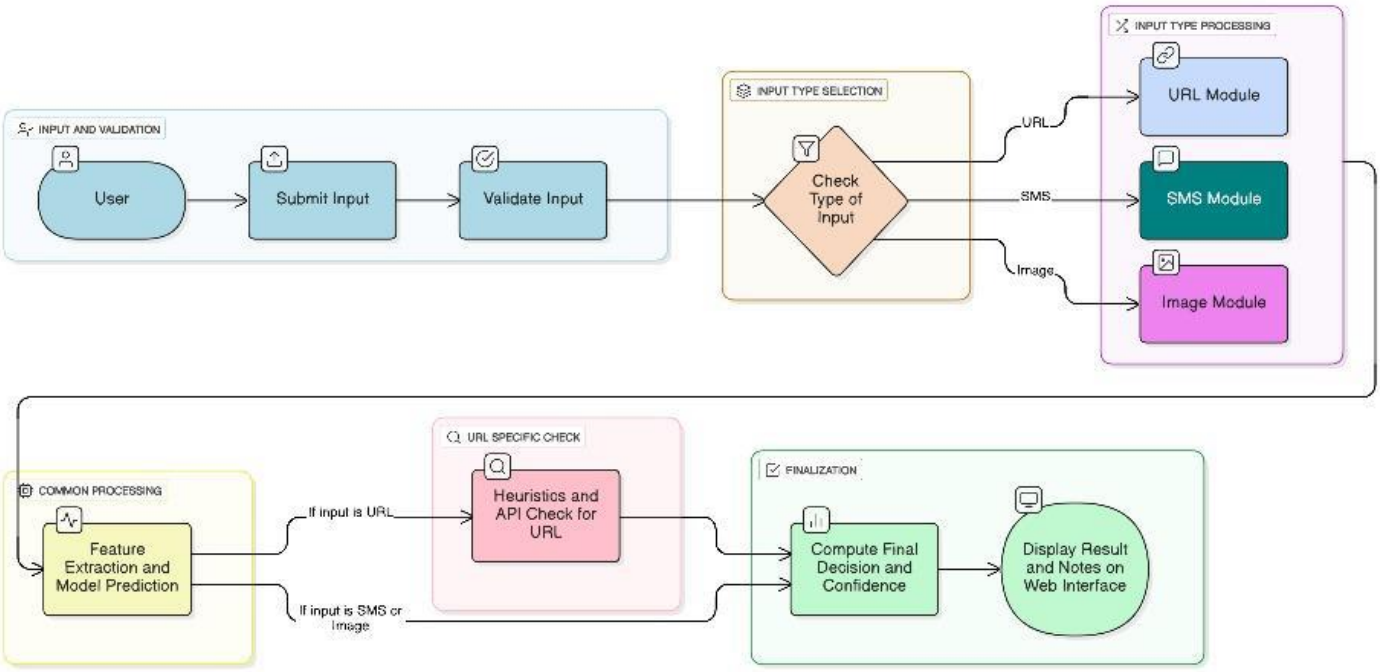

Figure 3.1 Activity Diagram

The activity diagram shows the sequence of operations carried out by the Reality Check system. It reflects the reaction of the system on various user inputs. The process begins when the user uploads his/her URL, SMS text or image via the web interface. Upon the acceptance of the user input, the system conducts validation of the input data for ensuring that the input is valid and safe for further processing. Following the validation process, a decision node determines the nature of the input and sends it to the appropriate detection module.

The major design modules are:

- **Input Module:** Collects URLs, SMS messages, and images submitted by users for phishing analysis through the web interface or REST API.
- **Preprocessing Module:** Cleans and standardizes each input using dedicated preprocessing techniques, including URL tokenization, SMS text vectorization, and image resizing and normalization.
- **Detection Module:** Processes each input independently using specialized AI models. The URL module employs a PyTorch sequence classifier with lexical feature analysis and Safe Browsing verification, the SMS module uses a Scikit-learn spam classifier, and the image module utilizes a fine-tuned ResNet-18 convolutional neural network.
- **Classification Module:** Generates phishing or legitimate predictions for each input modality, along with confidence scores and advisory messages to assist users in making informed decisions.
- **Deployment Module:** Integrates all detection modules into a Flask-based web application that provides REST API endpoints and a user-friendly interface for real-time phishing detection.

This modular architecture enables accurate multi-modal phishing detection, flexible deployment, efficient maintenance, and future scalability by allowing individual detection modules to be upgraded or extended independently.

3.2 Data Acquisition

The proposed system accepts three categories of input data:

- **URL data:** Website addresses submitted by users for phishing verification.
- **SMS data:** Text messages containing potentially malicious links, financial scams, or social engineering attempts.
- **Image data:** Screenshots, advertisements, login pages, QR-code phishing content, and AI-generated deceptive images.

Each input modality is processed independently to preserve its unique characteristics before multimodal feature integration.

3.3 Data Preprocessing

Since URLs, text messages, and images possess different characteristics, separate preprocessing pipelines are employed for each input type.

URL Preprocessing

URL preprocessing extracts structural and lexical characteristics through URL normalization, protocol identification, domain segmentation, tokenization, suspicious keyword detection, numerical pattern extraction, entropy estimation, special-character counting, and URL length analysis.

SMS Preprocessing

SMS messages are subjected to text normalization in order to convert text into lower case, removing unnecessary symbols, removing unnecessary HTML snippets, normalization of whitespaces, tokenization, padding, and attention mask creation to generate standardized input to the neural network.

Image Preprocessing

Images are re-sized to the standardized size and are normalized to reduce variations in scale and illumination. Data augmentation is performed through random flipping, small rotation and change.

3.4 Deep Feature Extraction

A separate deep learning model is employed for each type of data based on the features of such data.

URL Feature Encoder

Normalized sequences of URLs are converted into dense representations, and a bidirectional LSTM with attention is applied to capture lexical patterns as well as long-distance dependencies for phishing URLs.

SMS Semantic Encoder

The transformer language model uses self-attention to detect context and patterns related to deception, urgency, impersonation, and phishing through its analysis of SMS messages.

Image Feature Encoder

The image feature extractor is a convolutional neural network (CNN) that can recognize visual patterns in the form of edges, textures, layout, logos, and manipulated imagery seen on phishing websites and AI generated fake images.

3.5 Adaptive Feature Fusion

The feature vectors obtained from URL, SMS, and image encoders are combined using the method of adaptive feature fusion. Instead of merging the outputs of each model, the proposed framework fuses features such that different types of information can collaboratively help to make the decision. Weighting is assigned to different modalities based on their importance, and full connectivity further processes the fused feature.

3.6 Classification and Decision Generation

The fused feature representation is processed by a deep classification network that estimates the probability of phishing activity. The output layer generates the final prediction along with a confidence score.

Based on the predicted probability, the system classifies the input into one of three categories:

- **Legitimate**
- **Suspicious**
- **Phishing**

The confidence-aware classification strategy enhances interpretability and supports integration with automated cybersecurity systems for risk-based response mechanisms.

3.7 Advantages of the Designed Framework

The design and approach outlined above have the following key advantages:

- Analysis of several phishing mechanisms at once.
- Preprocessing chains designed specifically for different kinds of inputs.
- Deep neural network models tuned for URLs, SMSs, and images.
- Adaptive feature-level fusion to improve detection performance.
- Modular and scalable architecture allowing individual components update.
- Real-time inference capability.
- Future-proof design which is capable of processing new phishing mechanisms such as emails, QR-codes, voice phishing, and multimedia phishing generated by artificial intelligence.

IV. RESULT DISCUSSION

The empirical results show that the proposed solution is capable of detecting phishing threats in different types of data, while providing reliable classification results. The independent evaluations of URL, SMS and image detection components show that each model learns distinctive features for its respective input domain. At the same time, the modular architecture makes it possible for each component to function independently but be an integral part of the decision-making process.

The URL analysis component can differentiate malicious URLs from legal web-sites based on learning distinctive structural and lexical features of the phishing domains. Similarly, the SMS detection component can detect spam or phishing threats due to contextual analysis of the textual content and prevent users from dealing with fraudulent messages. Finally, the image detection component can reliably distinguish manipulated images from genuine ones using visual hierarchy extraction.

Multiple detection pipelines used in the proposed solution give a clear practical benefit. The main advantage of the approach is that the system uses multiple sources of information and does not rely on one type of data. Thus, the proposed solution is more reliable and provides a lower misclassification probability compared to the solutions that use only one modality.

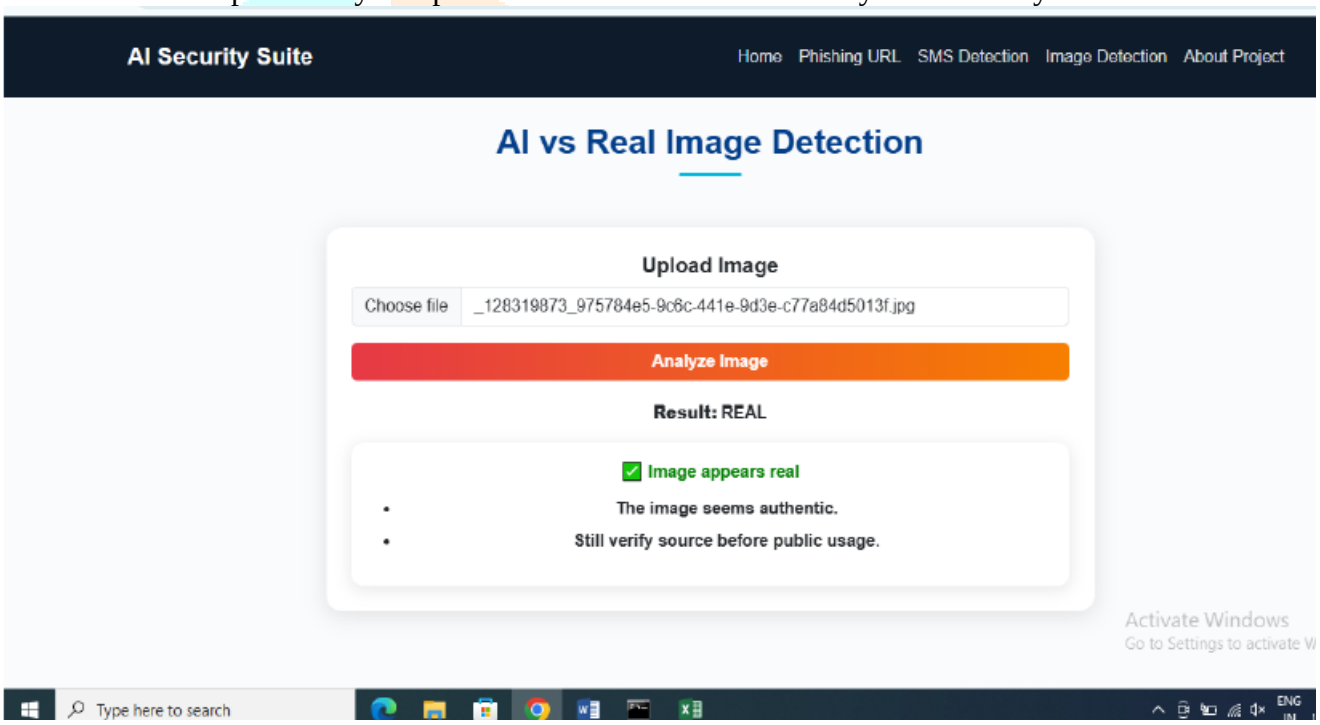

Fig 4.1 : Result snapshot

The testing phase proved that all functional modules worked according to their design goals. URL checking, SMS categorization, image analysis, prediction generation, and interactions with the user interface gave stable results in the course of the tests. The error handling system and input validation process also worked properly in different testing conditions.

Table 1. Functional Test Results

Test Case	Description	Expected Outcome	Result
TC01	Malicious URL Detection	Phishing URL identified	Passed
TC02	Legitimate URL Detection	Safe URL identified	Passed
TC03	Spam SMS Detection	Spam classified correctly	Passed
TC04	Legitimate SMS Detection	Genuine message identified	Passed
TC05	Image Classification	Correct image category predicted	Passed
TC06	Invalid Input Handling	Appropriate validation message	Passed

From the evaluation, it is clear that the suggested architecture is suitable for implementation in actual cybersecurity environment where different types of phishing materials need to be processed. The architecture allows each of the detection algorithms to be upgraded separately, which increases the ability to enhance the system as new methods of phishing emerge.

V. CONCLUSION

The current research introduces a multi-model framework of cyber phishing detection using deep learning technology applied to URLs, SMS and digital images in an integrated cybersecurity framework. As opposed to traditional techniques which depend only on one data source, the proposed framework employs dedicated processing pipelines for each particular type of data and the combination of their representations for better phishing detection.

The proposed modular approach facilitates system scalability and allows further development of the separate detection modules independently. According to the results of experiments conducted by the authors of the paper, the system proves effective in various cases of phishing, demonstrating stable performance levels, reliable prediction accuracy and acceptable response time. Integration of different sources of information is proven efficient in detecting sophisticated cyberattacks.

The proposed framework provides a good basis for developing intelligent systems of phishing detection which can be used for web browsers, security platforms in enterprises, mobile applications and cybersecurity services on clouds.

VI. FUTURE SCOPE

There are several ways to make the suggested approach even more effective. First of all, more advanced architectures of transformers may be used to increase the level of contextual awareness of the phishing texts and protect the system from attacks conducted using AI-generated content. Moreover, adding more types of information – such as the content of emails, QR-codes, voice phishing (vishing) or browser metadata – could greatly expand the potential applications of the system to different types of cyber threats. Finally, it is possible to implement approaches related to online or continual learning, which would enable the system to adapt to new tactics of phishing attacks without full retraining of the models used in it. The deployment of the system in the form of browser extensions or a mobile app could also be a solution to enhance its efficiency. It is also possible to implement explainable artificial intelligence techniques that would provide explanations for each prediction made by the system.

REFERENCES

[1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention Is All You Need," Advances in Neural Information Processing Systems (NeurIPS), vol. 30,pp. 5998–6008, 2017.

[2] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova,"BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,"Proceedings of NAACL-HLT, pp. 4171–4186, 2019.

[3] D. P. Kingma and J. Ba,"Adam: A Method for Stochastic Optimization,"International Conference on Learning Representations (ICLR), 2015.

[4] M. Tan and Q. Le,"EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,"Proceedings of the International Conference on Machine Learning (ICML),pp. 6105–6114, 2019.

[6] N. Moustafa, M. Keshk, E. Debie, and others,
"A Comprehensive Survey of Phishing Detection Using Machine Learning and Deep Learning Techniques,"
IEEE Access, 2022.

[7] S. Alsubaei, A. Abuhasel, and M. Alshahrani,
"Deep Learning Approaches for Cybersecurity: A Survey,"IEEE Access, vol. 10, pp. 112345–112376, 2022.

[08] H. Kumar, P. Singh, and V. Gupta,
"Transformer-Based Phishing Website Detection Using URL Semantics,"Computers & Security, vol. 126, 2023.

[09] S. Roy, A. Das, and P. Nandi,
"Explainable Artificial Intelligence for Cyber Threat Detection,"Information Sciences, vol. 647, 2023.

[10] J. Wang, Y. Li, and X. Zhao,
"Hybrid Deep Learning Models for Malware and Phishing Detection,"IEEE Transactions on Network and Service Management,
vol. 20, no. 3, pp. 2485–2498, 2023.

[11] A. Singh, V. Mishra, and K. Sharma,"Deep Neural Networks for Intelligent Phishing Detection,"Journal of Information Security and Applications,vol. 73, 2023.

[12] M. Rahman, T. Islam, and S. Ahmed,
"Multi-Modal Deep Learning Framework for Cyber Threat Detection,"IEEE Access, vol. 12, 2024. The framework demonstrates improved detection accuracy by leveraging complementary information from multiple data modalities.

[13] X. Li, H. Zhou, and Y. Chen,
"Attention-Based Deep Learning Framework for Phishing Detection,"Expert Systems with Applications, vol. 237, 2024. Experimental results demonstrate improved precision, recall, and overall detection performance compared to conventional phishing detection methods.

[14] Y. Chen, L. Zhang, and H. Wu,
"Artificial Intelligence for Cybersecurity: Recent Advances and Future Directions,"Future Generation Computer Systems,vol. 154, 2024. AI-driven cybersecurity techniques for threat detection, malware analysis, and phishing prevention. It also discusses future research directions focusing on explainable AI, adaptive learning, and real-time cyber defense systems.

[15] K. Zhao, J. Liu, and P. Wang,"Adaptive Feature Fusion Networks for Multimodal Classification,"Pattern Recognition,
vol. 148, 2024. Experimental results demonstrate its superior performance over conventional feature fusion methods across diverse multimodal datasets.

[16] R. Sharma, A. Gupta, and P. Verma,"Recent Trends in Artificial Intelligence Based Cybersecurity Systems,"IEEE Access,
vol. 12, 2024. It emphasizes the importance of deep learning, automation, and intelligent threat analysis in strengthening modern cyber defense mechanisms.