International Journal of Innovative Computer Science and IT Research

Volume 01, Issue 04 – Jul-Aug 2025 • E-ISSN: 3067-1108 | P-ISSN: 3067-1094

Phishing Detection Using Natural Language Processing and Behavioral Analytics

Prof. Ronan Gallagher

Associate Professor, University College Cork, Ireland

Abstract

Phishing remains one of the most persistent cybersecurity threats because attackers continuously modify communication patterns, social engineering techniques, website structures, and message content to deceive users. Traditional phishing detection approaches commonly depend on predefined signatures, blacklists, URLs, or manually engineered rules, which may become ineffective against newly generated and rapidly changing attacks. This paper examines an intelligent phishing detection approach that combines Natural Language Processing (NLP) with behavioral analytics to identify suspicious communication and user interaction patterns. NLP provides mechanisms for analyzing textual characteristics in emails, messages, web content, and other digital communications, while behavioral analytics examines deviations in interaction patterns, access behavior, communication timing, navigation sequences, and contextual user activities. The proposed conceptual framework integrates these two analytical dimensions through a multi-stage detection architecture involving data collection, preprocessing, linguistic feature extraction, behavioral profiling, feature fusion, classification, risk scoring, and adaptive response.

The analysis demonstrates how textual and behavioral evidence can complement each other to improve phishing identification, particularly when individual indicators are insufficient for reliable classification. NLP-based analysis can identify suspicious linguistic structures, urgency cues, impersonation patterns, unusual requests, and semantic inconsistencies, whereas behavioral analytics can identify abnormal interactions that may indicate credential theft or malicious redirection. The paper further examines major application scenarios, implementation challenges, and future research directions associated with intelligent phishing detection. Particular attention is given to contextual modeling, explainability, privacy preservation, adversarial adaptation, and real-time detection. The study concludes that integrating NLP with behavioral analytics can provide a more adaptive and context-sensitive approach to phishing detection than isolated content-based or rule-based techniques. Such integration can support intelligent cybersecurity systems capable of continuously evaluating communication and interaction evidence while adapting to evolving phishing strategies.

Keywords: Phishing Detection, Natural Language Processing, Behavioral Analytics, Cybersecurity, Machine Learning, Social Engineering, Text Classification, Anomaly Detection, Risk Assessment, Intelligent Security Systems

International Journal of Innovative Computer Science and IT Research

Volume 01, Issue 04 – Jul-Aug 2025 • E-ISSN: 3067-1108 | P-ISSN: 3067-1094

1. Introduction

Phishing is a major cybersecurity challenge in which attackers manipulate users into revealing sensitive information, opening malicious resources, transferring funds, or performing unauthorized actions. Modern phishing attacks increasingly imitate legitimate organizations, services, colleagues, and digital platforms, making detection difficult through simple rule-based mechanisms. Attackers can modify domains, messages, layouts, sender identities, and communication strategies quickly. Consequently, effective phishing detection requires analytical techniques capable of evaluating multiple indicators rather than depending exclusively on static signatures or previously identified malicious resources.

Natural Language Processing has become an important component of intelligent cybersecurity because phishing communications frequently contain linguistic signals associated with deception and social engineering. Email messages, instant messages, notifications, and web content can be processed to identify lexical patterns, semantic relationships, urgency expressions, suspicious requests, impersonation language, and inconsistencies in communication style. NLP-based classification can therefore transform unstructured communication content into meaningful features for machine learning models. However, textual analysis alone may produce limitations when attackers generate highly convincing messages or use legitimate-looking language.

Behavioral analytics provides a complementary perspective by analyzing how users, accounts, devices, and digital sessions behave over time. Instead of evaluating only the content of a message, behavioral models can examine unusual login patterns, abnormal navigation sequences, unexpected access locations, atypical communication activity, and deviations from established user profiles. Such behavioral deviations can provide important contextual evidence when a phishing attempt results in suspicious user interaction. Behavioral analysis is particularly valuable for identifying attacks that cannot be recognized solely from message content.

Combining NLP and behavioral analytics creates an opportunity for developing more comprehensive phishing detection architectures. Textual analysis can determine whether communication contains suspicious linguistic and semantic characteristics, while behavioral analytics can determine whether associated actions differ from established patterns. The combination of these evidence sources can enable risk-based classification rather than binary decisions based on individual indicators. Feature fusion can also improve the contextual understanding of phishing scenarios across email, web, messaging, and enterprise environments.

The objective of this paper is to develop a conceptual and analytical framework for phishing detection using NLP and behavioral analytics. The study reviews relevant intelligent detection approaches, defines an integrated detection architecture, examines important analytical components, and evaluates major application areas and implementation considerations. The paper also discusses challenges associated with evolving phishing strategies, data quality, privacy, explainability, adversarial manipulation, and real-time processing.

2. Objectives

2.1 Primary Objective

The primary objective of this study is to examine an integrated phishing detection framework that combines Natural Language Processing and behavioral analytics for identifying suspicious digital communications and interaction patterns.

The framework seeks to improve detection contextuality by combining linguistic evidence with behavioral deviations and risk indicators rather than depending on isolated signatures or predefined rules.

2.2 Development of an NLP-Based Detection Model

The first objective is to examine NLP techniques capable of extracting meaningful information from phishing-related communications. The analysis considers lexical, syntactic, semantic, contextual, and intent-related characteristics that can differentiate suspicious messages from legitimate communications.

Machine learning and language representation methods can transform textual information into structured indicators for phishing classification and risk assessment.

2.3 Integration of Behavioral Analytics

The second objective is to investigate behavioral analytics as a complementary detection mechanism. User and system behavior can provide contextual information about whether an interaction is consistent with established activity patterns.

The study therefore considers behavioral profiling, anomaly identification, contextual deviation, session analysis, and risk scoring as important components of integrated phishing detection.

2.4 Evaluation of Integrated Detection Challenges

The third objective is to identify technical and operational challenges associated with combining NLP and behavioral analytics. These include data quality, privacy, adversarial adaptation, false positives, explainability, computational requirements, domain changes, and real-time processing.

Understanding these challenges is essential for designing reliable and scalable intelligent phishing detection systems.

3. Review of Literature

3.1 Natural Language Processing for Phishing Detection

NLP-based cybersecurity research has demonstrated the value of analyzing textual content for identifying malicious or deceptive communications. Traditional approaches commonly use lexical features, word frequencies, n-grams, stylometric indicators, and statistical representations for text classification. Machine learning algorithms can subsequently classify messages based on learned patterns. More recent approaches use distributed representations and transformer-based language models to capture contextual and semantic relationships. These approaches provide greater flexibility for analyzing sophisticated phishing messages where suspicious characteristics may depend on the relationship between multiple words or sentences.

Phishing messages frequently exploit psychological mechanisms such as urgency, authority, fear, curiosity, and reward. NLP systems can identify these characteristics by analyzing linguistic patterns and communication intent. Named-entity recognition can also help identify references to organizations, financial services, individuals, and digital platforms. Semantic similarity can support comparison between suspicious communication and known legitimate communication styles. However, attackers can deliberately produce grammatically correct and contextually convincing messages, creating a need for models that consider deeper semantic and contextual evidence.

3.2 Behavioral Analytics in Cybersecurity

Behavioral analytics focuses on identifying deviations from established patterns of user, device, application, or network activity. In cybersecurity, behavioral models can establish baselines representing normal activity and subsequently detect anomalous behavior. Indicators may include login timing, access frequency, navigation behavior, device characteristics, session duration, geographic inconsistencies, and resource access patterns. Such approaches are particularly useful when attacks generate behavioral consequences that cannot be identified through static content inspection.

Behavioral analytics can strengthen phishing detection by identifying suspicious activities following interaction with potentially malicious communications. For example, an unusual authentication attempt after a suspicious message or abnormal navigation toward unfamiliar resources may increase the risk associated with an event. Nevertheless, behavioral systems can face difficulties when legitimate users naturally change their activity patterns. Effective implementation therefore requires adaptive baselines and contextual modeling capable of distinguishing genuine anomalies from normal behavioral variation.

3.3 Integrated Intelligent Phishing Detection

Integrated detection approaches seek to combine multiple evidence sources to improve cybersecurity decision-making. Combining content-based analysis with behavioral information allows systems to evaluate both what a communication contains and how associated users or systems respond to it. This creates a richer representation of potential phishing events. Machine learning models can use fused features to produce classification results or continuous risk scores.

Multimodal and ensemble detection approaches can also improve resilience against individual weaknesses. If attackers manipulate message content to avoid linguistic detection, behavioral deviations may still provide evidence of malicious activity. Similarly, if behavioral patterns are ambiguous, NLP indicators can contribute additional evidence. The major research challenge is therefore not simply combining datasets but developing appropriate feature representations, fusion mechanisms, model architectures, and decision policies that maintain accuracy while controlling false positives.

4. Methodology

4.1 Research Design

This study adopts a conceptual and analytical research design for examining phishing detection through integrated NLP and behavioral analytics. The methodology organizes phishing detection into multiple analytical stages, beginning with the collection of communication and behavioral information and continuing through preprocessing, feature extraction, representation, fusion, classification, and risk assessment. The approach does not claim experimental performance from an original dataset; instead, it provides a structured framework that can guide future implementation and empirical evaluation.

The proposed design considers phishing detection as a contextual classification problem in which multiple evidence sources contribute to a final security decision. Textual characteristics are processed using NLP techniques, while behavioral information is modeled using anomaly detection and profiling approaches. These outputs are subsequently combined through a feature-fusion layer. The resulting representation can be evaluated using machine learning or deep learning classifiers and transformed into an interpretable risk score for security response.

4.2 NLP and Behavioral Feature Extraction

The NLP component processes emails, messages, web text, and other communication content. Preprocessing may include normalization, tokenization, removal of irrelevant elements, linguistic analysis, entity extraction, and semantic representation. Relevant features can include suspicious keywords, urgency expressions, request patterns, sender-message inconsistencies, semantic similarity, linguistic complexity, and impersonation indicators. Modern language representations can additionally capture contextual relationships that traditional keyword-based methods may overlook.

The behavioral component establishes contextual profiles of users, devices, accounts, and sessions. Features can include access frequency, authentication behavior, session timing, navigation sequences, device changes, interaction patterns, and deviations from historical activity. Anomaly scores can be calculated by comparing current behavior with established baselines. Behavioral features should be interpreted contextually because unusual activity does not automatically indicate malicious behavior. Combining temporal and contextual information can therefore improve the reliability of behavioral risk assessment.

4.3 Integrated Detection Architecture

The proposed architecture integrates NLP and behavioral evidence through a feature-fusion and risk-assessment layer.

The major methodological components include:

- **Data Collection:** Collection of communication content and relevant behavioral events.
- **Data Preprocessing:** Cleaning, normalization, tokenization, and event preparation.
- **NLP Feature Extraction:** Identification of linguistic, semantic, contextual, and intent-related indicators.
- **Behavioral Profiling:** Construction of normal user and system activity profiles.
- **Anomaly Detection:** Identification of deviations from expected behavior.
- **Feature Fusion:** Combination of textual and behavioral representations.
- **Classification:** Application of machine learning or deep learning models for phishing classification.
- **Risk Scoring:** Generation of continuous risk estimates based on multiple evidence sources.
- **Security Response:** Warning, blocking, investigation, or escalation according to risk level.

The integrated architecture is designed to support both content-level and context-level analysis. This structure enables the detection system to use NLP evidence when suspicious communication characteristics are apparent and behavioral evidence when user interaction provides additional contextual information. The architecture can therefore support adaptive decision-making across multiple digital environments.

5. Analysis

5.1 Architecture of NLP and Behavioral Phishing Detection

An integrated phishing detection architecture can be divided into several interconnected layers. The data acquisition layer collects communication and behavioral information, while preprocessing prepares the information for analytical processing.

International Journal of Innovative Computer Science and IT Research

Volume 01, Issue 04 – Jul-Aug 2025 • E-ISSN: 3067-1108 | P-ISSN: 3067-1094

NLP models analyze communication content, whereas behavioral analytics evaluates activity patterns. The feature-fusion layer combines both representations before classification and risk assessment. Finally, the response layer converts analytical outcomes into appropriate security actions.

Table 1: Core Components of an Integrated Phishing Detection Architecture

Component	Primary Function	Relevant Indicators
Data Collection	Gather communication and activity information	Emails, messages, sessions, access events
Preprocessing	Prepare raw information	Normalization, tokenization, event cleaning
NLP Analysis	Analyze communication content	Keywords, semantics, intent, entities
Behavioral Profiling	Model expected activity	User, device, session patterns
Anomaly Detection	Identify deviations	Unusual access, timing, navigation
Feature Fusion	Combine analytical evidence	Textual and behavioral representations
Classification	Determine phishing likelihood	ML/DL classification
Risk Scoring	Quantify security risk	Confidence, severity, contextual evidence
Response Layer	Initiate security action	Alert, block, investigate, monitor

Interpretation: Table 1 demonstrates that effective phishing detection requires coordination between content intelligence and behavioral intelligence. NLP provides detailed information about communication characteristics, whereas behavioral analytics provides information about contextual activity.

Feature fusion acts as the central mechanism for combining these signals. The architecture also supports risk-based decisions, allowing systems to differentiate between low-risk anomalies and highly suspicious events rather than treating every unusual communication as malicious.

5.2 Applications of Integrated Phishing Detection

The combination of NLP and behavioral analytics can support phishing detection across multiple communication and computing environments. Email security is an important application because phishing messages frequently rely on linguistic manipulation and impersonation.

International Journal of Innovative Computer Science and IT Research

Volume 01, Issue 04 – Jul-Aug 2025 • E-ISSN: 3067-1108 | P-ISSN: 3067-1094

Web security systems can analyze page content and subsequent navigation behavior. Enterprise systems can monitor suspicious interactions involving organizational accounts. Messaging platforms can similarly analyze communication content while considering unusual user activity.

Table 2: Application Areas of NLP and Behavioral Phishing Detection

Application Area	NLP Contribution	Behavioral Analytics Contribution
Email Security	Message and intent analysis	Unusual sender or recipient interactions
Web Phishing	Page text and semantic analysis	Navigation and session anomalies
Enterprise Security	Organizational communication analysis	Account and access deviations
Financial Platforms	Suspicious requests and language	Transaction-related behavioral anomalies
Instant Messaging	Message classification	Abnormal communication behavior
Cloud Applications	Content and notification analysis	Login and session profiling
Customer Support	Social-engineering language detection	Unusual account interaction patterns
Mobile Security	Message and notification analysis	Device and application behavior

Interpretation: Table 2 indicates that the integrated approach can be applied across diverse digital environments. NLP is especially valuable where communication content contains social-engineering signals, while behavioral analytics adds contextual evidence associated with user or system actions.

The complementary nature of the two approaches can support more comprehensive detection than either component operating independently. Application-specific feature engineering and privacy controls remain necessary because behavioral characteristics differ substantially between environments.

5.3 Factors Influencing Detection Effectiveness

Detection effectiveness depends on several technical and contextual factors. NLP performance can be influenced by language variation, message length, domain-specific terminology, multilingual communication, and adversarial text manipulation. Behavioral analytics is affected by the quality of historical profiles and the natural variability of user activity. Systems with overly rigid behavioral baselines may generate excessive false positives, whereas overly flexible models may fail to identify meaningful deviations.

Model adaptation is another important factor. Phishing campaigns change rapidly, and previously reliable indicators may lose effectiveness. Detection systems therefore require continuous monitoring and periodic model updates. Explainability is also essential because security analysts need to understand

Volume 01, Issue 04 – Jul-Aug 2025 • E-ISSN: 3067-1108 | P-ISSN: 3067-1094

why a message or interaction was assigned a particular risk level. A combined model should ideally identify the primary linguistic and behavioral evidence contributing to a decision.

6. Discussion

6.1 Combining Linguistic and Behavioral Intelligence

The integration of NLP and behavioral analytics provides a broader perspective on phishing detection. Content analysis identifies characteristics within the communication itself, while behavioral analysis evaluates the context surrounding the interaction. This combination is important because phishing is fundamentally a social-engineering problem involving both deceptive communication and manipulated user behavior. A message that appears legitimate may become suspicious when associated with abnormal account activity or an unexpected access pattern.

Feature fusion can therefore transform phishing detection from isolated message classification into contextual security analysis. Instead of asking only whether a message resembles known phishing content, an integrated system can consider whether its language, sender characteristics, user interaction, and subsequent activity collectively indicate elevated risk. Such contextualization may improve detection of previously unseen attacks while reducing dependence on static indicators.

6.2 Machine Learning and Adaptive Detection

Machine learning provides the analytical foundation for learning relationships between phishing indicators and security outcomes. NLP representations can be processed using supervised, semi-supervised, or deep learning models, while behavioral information can be modeled through anomaly detection, clustering, sequential analysis, or representation learning. Fusion models can subsequently combine these outputs to generate classification or risk scores.

Adaptive learning is particularly important because phishing campaigns evolve continuously. Attackers can modify vocabulary, sentence structures, communication channels, timing, and interaction strategies. A static model may therefore degrade as the threat environment changes. Continuous learning, domain adaptation, and drift monitoring can help maintain detection relevance. However, adaptation must be controlled carefully to avoid learning malicious behavior as normal activity.

6.3 Explainability, Privacy, and Security

Integrated phishing detection systems must balance analytical power with privacy and operational transparency. Behavioral analytics may involve sensitive information about user activity, authentication patterns, and interaction history. Excessive collection or retention of such information can create privacy risks. Privacy-preserving processing, access controls, data minimization, and appropriate governance are therefore important components of deployment.

Explainability is similarly important for operational adoption. Security analysts need understandable evidence supporting automated alerts. Instead of providing only a phishing probability, the system should ideally identify relevant linguistic indicators and behavioral deviations contributing to the decision. Explainable risk assessment can improve analyst trust, facilitate incident investigation, and support corrective action when models generate incorrect classifications.

7. Challenges

Phishing detection using NLP faces substantial challenges associated with language variability and adversarial manipulation. Attackers can use sophisticated language, multilingual content, intentional misspellings, obfuscation, indirect requests, and contextually appropriate communication. Generative AI can further facilitate the creation of highly polished phishing messages. Consequently, detection systems cannot rely solely on obvious keywords or grammatical errors and must incorporate semantic and contextual representations.

Behavioral analytics introduces another set of challenges because legitimate behavior is naturally variable. Users may travel, change devices, work at unusual times, or access unfamiliar applications without malicious intent. Establishing appropriate behavioral baselines is therefore difficult. Excessively sensitive systems may generate false positives, while overly tolerant systems may miss genuine attacks. Context-aware anomaly detection is required to distinguish legitimate deviations from suspicious activity.

Scalability, privacy, and real-time processing also remain important concerns. Enterprise environments can generate very large volumes of communication and behavioral events, requiring efficient data processing and model inference. At the same time, behavioral information may contain sensitive user-related data. Organizations must therefore balance detection capabilities with privacy requirements, secure data handling, computational efficiency, and explainable decision-making. These requirements become more complex when integrated systems operate across heterogeneous platforms.

8. Future Research

Future research should investigate transformer-based and foundation-model approaches for phishing detection while maintaining computational efficiency and security. Language models can provide stronger contextual understanding than traditional lexical representations, particularly for sophisticated social-engineering messages. Research should also explore multilingual and cross-domain phishing detection because attackers increasingly operate across languages and communication platforms. Robust evaluation against AI-generated phishing content will be particularly important.

Another promising direction is the development of temporal behavioral models capable of understanding how suspicious activity evolves over time. Rather than analyzing individual events independently, future systems could model sequences involving message receipt, link interaction, authentication, navigation, and subsequent account activity. Temporal modeling could provide stronger contextual evidence and help distinguish isolated anomalies from coordinated attack patterns. Privacy-preserving learning methods should also be explored to enable behavioral intelligence without unnecessary exposure of sensitive information.

Future systems should additionally emphasize explainable, collaborative, and adaptive security architectures. Federated learning could support distributed model development while limiting centralized access to sensitive behavioral data. Online learning and concept-drift detection could enable continuous adaptation to emerging phishing strategies. Explainable AI mechanisms could provide security analysts with interpretable evidence for automated decisions. Integration with Security Information and Event Management platforms and automated response systems could further transform phishing detection into a broader intelligent cybersecurity capability.

9. Conclusion

Phishing continues to evolve as attackers increasingly combine social engineering, impersonation, deceptive communication, and sophisticated digital infrastructure. Traditional detection methods based primarily on signatures, blacklists, or static rules may struggle against rapidly changing and previously unseen attacks. This study examined an integrated approach based on Natural Language Processing and behavioral analytics to address this limitation. NLP provides mechanisms for understanding communication content, including lexical, semantic, contextual, and intent-related characteristics, while behavioral analytics evaluates deviations in user, account, device, and session activity.

The proposed framework combines data collection, preprocessing, NLP feature extraction, behavioral profiling, anomaly detection, feature fusion, classification, risk scoring, and security response. This architecture enables phishing detection to consider both the characteristics of a communication and the context in which it is accessed or acted upon. The analysis demonstrates that the complementary nature of textual and behavioral evidence can support more contextualized detection than isolated content analysis. However, practical implementation must address evolving attack strategies, false positives, privacy requirements, computational scalability, adversarial manipulation, and explainability.

Future phishing detection systems should therefore move toward adaptive, context-aware, and trustworthy architectures. Transformer-based language models, temporal behavioral analysis, privacy-preserving learning, federated approaches, and explainable AI can provide important foundations for this development. The integration of NLP and behavioral analytics should not be viewed simply as the combination of two independent detection techniques but as a broader strategy for contextual cyber-threat intelligence. By continuously evaluating communication content and associated behavioral evidence, intelligent security systems can become more capable of identifying sophisticated phishing attempts and supporting timely, informed cybersecurity responses.

References

1. Alhogail, A., & Alsabih, A. (2021). Applying machine learning and natural language processing to detect phishing email. *Computers & Security, 110*, 102414. https://doi.org/10.1016/j.cose.2021.102414
2. Alani, M. M., & Tawfik, H. (2022). PhishNot: A cloud-based machine-learning approach to phishing URL detection. *Computer Networks, 218*, 109407. https://doi.org/10.1016/j.comnet.2022.109407
3. Atawneh, S., & Aljehani, H. (2023). Phishing email detection model using deep learning. *Electronics, 12*(20), 4261. https://doi.org/10.3390/electronics12204261
4. Bergholz, A., Chang, J. H., Paass, G., Reichartz, F., & Strobel, S. (2008). Improved phishing detection using model-based features. In *CEAS 2008: Fifth Conference on Email and Anti-Spam*.
5. Bergholz, A., De Beer, J., Glahn, S., Moens, M.-F., Paass, G., & Strobel, S. (2010). New filtering approaches for phishing email. *Journal of Computer Security, 18*(1), 7–35.
6. Chandrasekaran, M., Narayanan, K., & Upadhyaya, S. (2006). Phishing email detection based on structural properties. In *New York State Cyber Security Conference*.
7. Fette, I., Sadeh, N., & Tomasic, A. (2007). Learning to detect phishing emails. In *Proceedings of the 16th International Conference on World Wide Web* (pp. 649–656). ACM.

8. Gupta, B. B., Gaurav, A., Arya, V., Attar, R. W., Bansal, S., Alhomoud, A., & others. (2024). Advanced BERT and CNN-based computational model for phishing detection in enterprise systems. *Computer Modeling in Engineering & Sciences, 141*(3), 2165–2183. https://doi.org/10.32604/cmes.2024.056473
9. Lai, C., Selamat, A., Ibrahim, R., Quang, D. N., Fujita, H., Krejcar, O., & others. (2024). URL phishing detection by using natural language processing and deep learning model. In *New Trends in Intelligent Software Methodologies, Tools and Techniques* (pp. 1–15). IOS Press. https://doi.org/10.3233/FAIA240360
10. Marzuq, A., & others. (2021). Phishing email detection using natural language processing techniques: A literature survey. *Procedia Computer Science, 189*, 19–28. https://doi.org/10.1016/j.procs.2021.05.077
11. Rashid, F., Doyle, B., Han, S. C., & Seneviratne, S. (2024). Phishing URL detection generalisation using unsupervised domain adaptation. *Computer Networks, 245*, 110398. https://doi.org/10.1016/j.comnet.2024.110398
12. Sahingoz, O. K., Buber, E., Demir, O., & Diri, B. (2019). Machine learning based phishing detection from URLs. *Expert Systems with Applications, 117*, 345–357. https://doi.org/10.1016/j.eswa.2018.09.029
13. Tawil, A. A., Almazaydeh, L., Qawasmeh, D., Qawasmeh, B., Alshinwan, M., & Elleithy, K. (2024). Comparative analysis of machine learning algorithms for email phishing detection using TF-IDF, Word2Vec, and BERT. *Computers, Materials & Continua, 81*(2), 3395–3412. https://doi.org/10.32604/cmc.2024.057279
14. Yang, L., Li, J., & others. (2022). Predicting user susceptibility to phishing based on multidimensional features. *Computational Intelligence and Neuroscience, 2022*, Article 7058972. https://doi.org/10.1155/2022/7058972
15. *A systematic literature review on phishing email detection using natural language processing techniques*. (2022). *IEEE Access*. https://doi.org/10.1109/ACCESS.2022.3183083
16. *An intelligent cyber security phishing detection system using deep learning techniques*. (2022). *Cluster Computing, 25*, 3819–3828. https://doi.org/10.1007/s10586-022-03604-4
17. *Comparative investigation of traditional machine-learning models and transformer models for phishing email detection*. (2024). *Electronics, 13*(24), 4877.
18. *A systematic review of deep learning techniques for phishing email detection*. (2024). *Electronics, 13*(19), 3823. https://doi.org/10.3390/electronics13193823
19. *Novel interpretable and robust web-based AI platform for phishing email detection*. (2024). *Computers & Electrical Engineering, 120*, 109625. https://doi.org/10.1016/j.compeleceng.2024.109625
20. *Why people keep falling for phishing scams: The effects of time pressure and deception cues on the detection of phishing emails*. (2022). *Computers & Security, 123*, 102937. https://doi.org/10.1016/j.cose.2022.102937