



PHISHING EMAIL DETECTION USING NATURAL LANGUAGE PROCESSING TECHNIQUES

1Prof. Vaishali Suryawanshi, 2Mr. Faisal Shaikh, 3Mr. Devesh Gondole, 4Mr. Aditya Khunteta,
5Mr. Vaibhav Wani

1Professor, 2,3,4,5 Students

Department of Computer Applications, JSPM's Rajarshi Shahu College of Engineering, Pune, India

Abstract : The most common cybercrime technique that coerces victims into divulging private information, such as bank account numbers, passwords, and account IDs, is phishing. Such cyber attacks are frequently initiated via phone calls, instant messages, and emails. Even though the strategies for preventing these cyberattacks are constantly being updated, the final result is still insufficient. However, the number of phishing emails has skyrocketed in recent years, indicating the need for more sophisticated and efficient ways to combat them. Although many techniques have been developed to filter phishing emails, a comprehensive solution is still required. As far as we are aware, this is the first survey that focusses on identifying phishing emails using machine learning (ML) and natural language processing (NLP) techniques. The many cutting-edge NLP techniques currently being used to detect phishing emails at different phases of the attack are analysed in this study, with a focus on machine learning techniques. These methods undergo a comparative evaluation and analysis. The problem, its immediate solution space, and the anticipated future research directions are all conveyed in this way.

Keywords – Machine Learning, Natural Language Processing, Cyber Security, Email Phishing

1. Introduction

1.1 Background: Overview of Phishing Attacks and Their Impact on Cybersecurity

Phishing attacks, which use deceit to take advantage of human psychology rather than system flaws, are a major threat in today's digital environment. In order to obtain sensitive credentials or financial information, these attacks frequently pose as reliable communications, such as emails, messages, or websites. Phishing campaigns can get past security filters by imitating trustworthy tandard, unlike conventional malware or intrusion attempts. Phishing prevention is a top priority in cybersecurity since it has become a major vector for data breaches and identity theft as people's and organisations' digital footprints expand. [1]

1.2 Importance of Phishing Detection

The consequences of phishing attacks have increased due to the growing dependence on digital platforms in all industries, from healthcare to finance. A single successful phishing email has the potential to destroy consumer trust, compromise entire networks, and result in financial harm. Blacklists and manual reviews are examples of traditional detection techniques that are unable to keep up with the volume and complexity of contemporary phishing attempts. Intelligent, adaptable systems that can identify phishing content in real time, no matter how delicately constructed the attack may be, are desperately needed. [2]

1.3 Role of Natural Language Processing (NLP)

Because NLP allows systems to interpret and analyse the content and intent of textual data, it is particularly well-suited for phishing detection. Language patterns, context, and tone—essential elements that distinguish phishing from legitimate communications—can be assessed by sophisticated natural language processing models. Phishing detection systems can process emails at a semantic level by employing models such as BERT, GPT, and RoBERTa. This allows them to identify grammatical errors, spoof intent, and manipulative

language—all of which are frequently signs of phishing. Phishing detection moves from static pattern recognition to dynamic content analysis with the integration of NLP. [3][4]

1.4 Objective of the Review

The purpose of this paper is to thoroughly examine the use of natural language processing (NLP) in phishing detection, with a focus on real-time email analysis. It will examine how NLP improves detection accuracy and adaptability, evaluate current approaches, and contrast machine learning and deep learning approaches. Additionally, the review aims to pinpoint existing gaps in the literature and suggest creative avenues for further study and advancement.

1.5 Organization of the Paper

This review is organized as follows to give a thorough understanding:

Section 2 outlines the fundamental concepts and challenges in phishing detection.

Section 3 details NLP techniques and their role in analyzing phishing content.

Section 4 surveys commonly used datasets and data handling strategies.

Section 5 presents a comparative analysis of existing approaches.

Section 6 discusses ongoing challenges and unresolved issues.

Section 7 proposes future research directions.

Section 8 concludes with key findings and implications. Section 9 lists all references used throughout the paper.

2. Fundamentals of Phishing Detection

2.1 What is Phishing?

Phishing is a dishonest tactic used by cybercriminals to obtain private data, including credit card numbers, usernames, passwords, and access credentials, by posing as a reliable source in online communications. Phishing, in contrast to other hacking techniques, takes advantage of human error as opposed to technical shortcomings. There are several types of phishing:

Spear phishing is more focused on a single person or standardized and frequently uses personal information to make the message seem reliable. Whaling, sometimes referred to as CEO fraud, is a variation of this that targets senior executives or decision-makers in an attempt to obtain sensitive company information or standardized wire transfers.

Text messages are used in smishing, also known as SMS phishing, to trick victims into clicking on dangerous links or phoney phone numbers. Vishing is the practice of attackers using voice calls to pretend to be banks or government standardized in order to obtain financial or personal information.

Clone phishing is the practice of attackers mimicking a genuine email and substituting malicious attachments or links in order to take advantage of preexisting trust. Without the user's knowledge, pharming diverts them from legitimate websites to fraudulent ones, typically by altering DNS settings.[5]

2.2 Common Characteristics of Phishing Emails

Phishing emails frequently have unique content and structure patterns that set them apart from authentic correspondence.

Important characteristics include:

Linguistic Features: odd sentence structures, misspellings, or strange grammar.

These frequently arise from non-native authorship or automated translation.

Behavioral Manipulation: Use of fear or urgency to compel users to make snap decisions, such as "Your account will be closed unless you act now."

Structural Features: Inaccurate subject lines, phoney hyperlinks, mismatched sender addresses, or attachments masquerading as authentic documents.

Visual Mimicry: stealing branding, templates, and logos from reputable businesses in order to give them a false sense of legitimacy. [6]

Understanding these traits is crucial for user education programs as well as automated systems.

2.3 Role of Natural Language Processing (NLP)

By carefully examining the content and context of messages, natural language processing (NLP) allows automated systems to do more than just identify keywords.

In contrast to conventional rule-based filters, NLP methods are able to:

Examine the message's semantic similarity to previous phishing attempts. Recognize tone and sentiment, looking for language that is deceptive or intimidating.

Extract entities such as names, dates, and organizations to detect impersonation. Sort messages according to their contextual intent, such as whether they seek to deceive, ask for credentials, or demand quick action.

Phishing detection has been transformed by contemporary NLP models like BERT and RoBERTa, which can comprehend subtleties and colloquial language that more basic models frequently overlook. [7]

2.4 Key Challenges in Phishing Detection Using NLP

NLP's effectiveness in phishing detection is still limited by a number of issues, despite its advancements:

Adversarial Attacks: Attackers alter text slightly to trick NLP models while keeping it readable by using homoglyphs or lookalikes for characters.

Evasion Techniques: Phishers get around text-based filters by using shortened URLs or embedding malicious content in images.

Data Scarcity and Imbalance: Phishing datasets are often limited in scope and heavily imbalanced, with legitimate messages far outnumbering phishing samples.

Multilingual Phishing: The generalisability of many NLP systems is limited because they are designed for English, but phishing attacks can happen in any language.

Real-Time Constraints: Accuracy must be maintained while implementing NLP at scale in real-time email systems using quick, lightweight models.[8]

Deploying NLP-driven phishing detection tools in dynamic, real-time environments requires resolving these problems.

3. Natural Language Processing Techniques in Phishing Detection

3.1 Text Preprocessing Techniques

Text needs to be cleaned and standardized before machine learning or deep learning models can analyse it. Preprocessing enhances the quality of input data and lowers noise.

Typical methods consist of:

Tokenisation is the process of breaking up unprocessed text into smaller chunks, like words or phrases.

This makes it possible for models to examine individual linguistic components.

The stop-word Removal: Eliminates commonly used but uninformative words like “is,” “and,” or “the” that could lessen the impact of analysis.

Lemmatisation is the process of applying contextual rules to change words to their base or dictionary form (for example, “running” “run”).

By eliminating suffixes, stemming reduces words to their most basic forms (for example, “fishing” “fish”), frequently at the expense of some grammatical accuracy.

By lowering dimensionality and computational overhead, these procedures assist NLP systems in concentrating on significant patterns in the data. [9]

3.2 Feature Extraction Methods

Text is transformed into numerical representations through feature extraction so that algorithms can process it. Several feature categories are used in phishing detection, such as:

Lexical Features: These represent surface-level properties of the text:

Term Frequency-Inverse Document Frequency, or TF-IDF, assesses a document word importance in relation to a corpus.

Bag of Words (BoW) ignores grammar and order in favor of creating a frequency matrix of words.

N-grams record N-word sequences in order to preserve context and word order information.

Semantic Features: These capture the meaning and relationships between words: Word embeddings, like Word2Vec and GloVe, use context to map words into dense vector spaces.

Deep comprehension of sentence structure and semantics is made possible by transformer-based embeddings from models such as BERT and GPT, which preserve contextual information.

Content-Based Features: These incorporate metadata and patterns unique to phishing:

The existence of URLs, particularly misspelled or shortened domains.

The use of dubious terms like “click to verify” or “urgent action required. Use sender analysis to find phony or deceptive email addresses. [10]

These features form the foundation of classification models for phishing detection.

3.3 Machine Learning Approaches

Phishing detection and other text classification tasks have long been handled by traditional machine learning (ML) models.

Important strategies consist of:

Support Vector Machines (SVM) are useful for high-dimensional data, such as text. **Naïve Bayes:** Suitable for baseline spam and phishing detection models, this model is based on probabilistic inference.

Random Forest: A group of decision trees that decreases overfitting and increases accuracy. [11]

Both feature selection and feature engineering are still essential. By identifying the most pertinent variables, methods like chi-square, mutual information, and recursive feature elimination enhance the performance and interpretability of the model.

3.4 Deep Learning Approaches

By automatically learning abstract features, deep learning (DL), which has access to large datasets and computational resources, provides notable improvements in phishing detection.

Neural Networks:

Because recurrent neural networks (RNNs) model sequences, they can be used to analyze language flow.

Convolutional Neural Networks (CNNs) quickly detect phishing cues by extracting localized patterns in text.

By remembering long-term dependencies, Long Short-Term Memory (LSTM) networks outperform RNNs.

Transformer Models:

Bidirectional Encoder Representations from Transformers, or BERT, allows for rich semantic understanding by capturing context from both directions in a sentence.

The Generative Pre-trained Transformer, or GPT, is very good at producing and interpreting text that looks human. O A robustly optimized version of BERT, ROBERTa frequently outperforms it in particular classification tasks.[12]

These models are excellent at identifying subtle patterns of deception and intent that more straightforward models frequently overlook in particular classification tasks.

3.5 Hybrid Approaches

To attain better performance, hybrid systems integrate the advantages of several paradigms, including deep learning, natural language processing, and traditional machine learning. For instance:

A Random Forest classifier may be used for decision-making in a hybrid model after BERT has been used for feature embedding.

A different strategy might use a multi-branch neural network to combine contentbased indicators with lexical and semantic features.

To create reliable, consensus-based predictions, ensemble learning techniques combine the outputs of several models (e.g., SVM + LSTM + BERT).

These systems are particularly useful in real-world applications where no single method provides complete coverage or resilience against novel phishing strategies. [13]

4. Datasets Used in Phishing Detection

4.1 Publicly Available Datasets

Phishing detection models are trained and evaluated using publicly available datasets. They give researchers access to benchmark corpora so they can compare performance across studies and test algorithms. Among the frequently used datasets are:

The Enron email dataset, which was initially gathered as part of a legal investigation, includes a significant number of authentic emails. Despite not being made for phishing detection, it is frequently used as a standard for classifying benign messages.

Thousands of phishing emails gathered from actual attacks are included in the Nazario Phishing Corpus, one of the first phishing datasets made publicly available.

The PhishTank Dataset is a collaborative community-maintained list of verified phishing URLs and related metadata that is updated on a regular basis.

The Anti-Phishing Working Group provides feeds of phishing reports and URLs through the APWG eCrime Exchange, though access may require affiliation.

The Canadian Institute for Cybersecurity curated the CIC-Phishing Email Dataset, a labeled dataset that includes both phishing and authentic emails. The generalizability of trained models is greatly impacted by these datasets, which differ in format, size, labeling quality, and data type (emails vs. URLs).[14]

4.2 Proprietary Datasets

Many businesses and cybersecurity companies gather proprietary data from real-world settings in addition to publicly available datasets. Usually, these datasets provide:

Higher fidelity and recency: These comprise enterprise-grade complexity, domainspecific attack vectors, and contemporary phishing techniques. Private user data: emails that have been intercepted from client-side or internal email traffic and are frequently anonymized before use.

Custom labeling provides high-quality ground truth for supervised learning by using human-verified phishing and authentic emails.

Due to privacy, legal, and competitive concerns, these datasets are typically not made publicly available, despite the fact that they are extremely useful for training productiongrade models. Reproducibility consequently becomes problematic in research comparisons.

4.3 Data Imbalance Issues

Class imbalance is a persistent issue in phishing detection research, as legitimate emails significantly outnumber phishing ones in the majority of datasets. Models may become biased in favor of the majority class as a result of this imbalance, making them less vulnerable to phishing attempts. Several strategies are used to address this:

Resampling Methods: Oversampling: Increases the quantity of phishing samples, frequently producing new examples through techniques such as SMOTE (Synthetic Minority Over-sampling Technique).

Undersampling: To balance the classes, fewer valid emails are sent, but this may result in information loss.

Cost-sensitive learning: Adjusts the training algorithm to more severely penalize instances of minority (phishing) misclassification.

Ensemble methods: By combining several weak learners, strategies like boosting or bagging can enhance minority class detection.

Evaluation Metrics: To better assess performance on imbalanced data, researchers employ Precision, recall, F1-score, and Area Under the Curve (AUC) in addition to accuracy. [15]

Deploying models that can identify phishing in real-world settings, where malicious messages are uncommon but potentially disastrous, requires skillfully managing imbalance.

5. Comparative Analysis of Existing Studies

5.1 Criteria for Comparison

Researchers usually use a set of precise criteria to assess how well phishing detection systems work. Among the most popular parameters are:

Evaluation Metrics: Model performance is evaluated using common metrics like accuracy, precision, recall, and F1-score. False Positive Rate (FPR), Detection Rate (DR), and ROC-AUC are examples of more complex analyses. **Features of the dataset:** Results are significantly impacted by the size, diversity, and balance of datasets. Exaggerated metrics that don't generalize are frequently reported in studies that use small or highly skewed datasets.

Model Architecture: The outcomes depend on the kind of algorithm used, whether it is hybrid, deep learning, or machine learning. Real-time constraints, feature engineering, and hyper parameter tuning are also taken into account. These standards facilitate meaningful comparisons between different methodologies and aid in standardizing the benchmarking process. [16]

5.2 Performance Comparison of Machine Learning Techniques

For phishing detection, traditional machine learning models have been thoroughly tested.

The following trends are revealed by a comprehensive survey:

When paired with engineered lexical features, Support Vector Machines (SVM) consistently provide high precision and F1-score. On datasets that are balanced or somewhat unbalanced, they perform especially well.

Despite being quick and easy to use, Naïve Bayes (NB) typically performs poorly in recall because of its independence assumptions. Nonetheless, it continues to function well as a lightweight classifier in limited settings. Because of their ensemble nature and capacity to manage feature interactions, Random Forests (RF) frequently perform better than other classifiers. On small to medium-sized datasets, they are particularly resistant to overfitting. [17]

In contrast to deep learning approaches, traditional models generally have trouble understanding complex language and rely significantly on the caliber of feature engineering, despite their advantages.

5.3 Performance Comparison of Deep Learning Techniques

New standards for phishing detection have been established by deep learning models, particularly those that are based on natural language processing. A number of important findings are revealed:

Both Long Short-Term Memory (LSTM) networks and Recurrent Neural Networks (RNNs) can learn from text's sequential patterns. On longer inputs where word order and context are important, they perform better than traditional models.

Convolutional Neural Networks (CNNs) have been successfully modified to capture local n-gram patterns and operate competitively with low latency, despite being less intuitive for textual data.

In terms of accuracy and semantic comprehension, transformer-based models like BERT, GPT, and RoBERTa are the best. Even in highly obfuscated content, they are adept at spotting subtle manipulation in phishing language.[18]

Despite requiring a large amount of processing power for both training and inference, studies show that transformer models routinely obtain F1-scores above 95% on benchmark phishing datasets.

5.4 Summary of Findings

The following conclusions are drawn from a synthesis of the comparative studies:

Top-Performing Methods: Because of their capacity for transfer learning and contextual understanding, transformer-based models—such as BERT and RoBERTa—perform better than any other.

The Most Useful Models: Performance and computational efficiency can be effectively balanced with Random Forests and lightweight neural networks (like LSTM).

Typical Mistakes: When accuracy is overused as the only metric, it can mask a model's actual efficacy, especially when datasets are unbalanced.

The field is still developing quickly, and as NLP models and training methods get better and more widely available, new standards are being set.

6. Challenges and Open Issues

6.1 Adversarial Attacks on NLP Models

Vulnerability to adversarial attacks is one of the most urgent issues with NLP-based phishing detection. In order to trick models without warning human readers, attackers purposefully alter email content. Methods like:

- character changes (such as changing "o" to "0"),

- Using visually similar characters from various scripts, homoglyph attacks.

Insertion of text that appears harmless but weakens phishing signals can lead to malicious emails being incorrectly classified by NLP models. Because they might overfit to patterns seen during training, even sophisticated models like BERT and GPT are vulnerable.

The creation of robust training paradigms and real-time defense mechanisms is necessary to address the significant research gap in adversarial robustness. [19]

6.2 Scalability and Real-Time Detection Issues

Several engineering and performance issues arise when phishing detection systems are implemented at scale:

Volume: Every day, businesses receive millions of emails. System bottlenecks or latency may result from processing each message using computationally demanding NLP models.

Throughput: Low-latency inference is necessary for real-time detection in order to stop malicious emails from getting to users prior to classification.

Resource management: Transformer models require a lot of memory and processing power, which frequently calls for distributed architectures or hardware accelerators.

For operational deployment, striking a balance between speed, accuracy, and resource efficiency is essential. [20]

To solve these problems, lightweight models and model compression strategies like knowledge distillation are being investigated.

6.3 Ethical and Privacy Concerns

By nature, phishing detection systems scan private communications, which raises questions about:

User privacy may be compromised if private or corporate information is revealed through email content analysis.

Data Ownership: Particularly in cloud-based detection systems, concerns are raised about who is in charge of the data used for training and inference.

Bias and Fairness: Emails written in non-standard dialects, minority languages, or from specific domains may be disproportionately misclassified by models trained on small or biased datasets.

Strict adherence to data governance guidelines, such as anonymization, consent-based use, and open auditing of AI systems, is necessary to allay these worries.

6.4 Generalization and Adaptability

Maintaining generalization—the capacity to accurately recognize phishing emails that depart from established patterns—is a major challenge in phishing detection. Attackers' tactics are always changing, and they use:

Phishing emails that dynamically alter their wording or structure are known as polymorphic emails. Localized assaults that make use of language or cultural allusions unique to a given area.

Zero-day phishing is when methods that haven't been seen before get past conventional detection systems.

In these situations, static models that were trained on historical data frequently fall short. For future systems to function well in real-world scenarios, they must include domain adaptation strategies, active feedback loops, and continuous learning. [20]

7. Future Research Directions

7.1 Advancements in NLP Models: Potential of Large Language Models (LLMs)

Large language models like GPT, BERT, and their derivatives have been instrumental in recent advances in natural language processing. The accuracy of phishing detection can be significantly improved by these models' exceptional comprehension of context, semantics, and minute linguistic nuances. In order to improve systems' ability to identify complex phishing attempts that employ more convincing or context-aware language, future research can investigate optimizing and customizing LLMs for cyber security tasks.

7.2 Improved Feature Engineering: Leveraging Contextual and Semantic Features

Syntactic and lexical features are frequently used in traditional phishing detection techniques. Nonetheless, classification robustness can be increased by integrating more profound contextual and semantic features, such as sentiment analysis, discourse structure, and intent detection. To better distinguish between legitimate and phishing emails, future research can concentrate on developing sophisticated feature extraction techniques that capture sender-receiver relationships, dynamic contextual cues, and the intent behind email content.

7.3 Robust Detection Mechanisms: Addressing Adversarial Robustness

Fishing attackers are constantly improving their methods to get around detection systems, such as creating adversarial examples that cause machine learning models to become confused. Future studies should focus on creating detection systems that are resistant to these hostile attacks by utilizing strategies like anomaly detection, model interpretability, and adversarial training. Maintaining the dependability of phishing detection systems in hostile environments requires making sure they are resilient to evasion techniques.

7.4 Real-Time Phishing Detection Systems: Deployment in Real-World Environments

Deploying these systems in real-time, large-scale environments is still difficult, even though many phishing detection models exhibit encouraging results in controlled or offline settings. The optimization of detection algorithms for low latency, scalability, and integration with current email infrastructure should be the main focus of future research. Furthermore, to increase adoption and practical efficacy in consumer and business email platforms, research on user friendly alert mechanisms and reducing false positives will be crucial. [21]

8. Conclusion

8.1 Summary of Review Findings

The importance of natural language processing methods in improving the precision of phishing detection has been emphasized in this review. The efficiency of machine learning and deep learning models in detecting phishing emails using textual and contextual characteristics, the significance of real-time analysis for prompt threat mitigation, and the difficulties presented by changing phishing strategies and adversarial attacks are some of the most important findings. The synthesis of existing approaches emphasizes the necessity of ongoing innovation to stay up with phishing tactics that are becoming more complex.

8.2 Practical Implications

The review's conclusions offer a road map for upcoming phishing detection research and development initiatives. Researchers and practitioners can create detection systems that are more effective and adaptable by concentrating on sophisticated NLP models, enhanced feature engineering, and strong adversarial defenses. The drive for real-time deployment also emphasizes how practical it is to have low-latency, scalable solutions that can be easily incorporated into current email infrastructures to proactively protect users.

8.3 Closing Remarks

Phishing is still a common and dynamic cyber threat that necessitates creative and interdisciplinary solutions. There are encouraging avenues to improve email security by utilizing developments in natural language processing and real-time analysis. Ongoing research and industry-academia collaboration will be essential to creating robust, user friendly detection systems that successfully protect digital communication as attackers continue to hone their techniques.

9. References

- [1] S. Hussain, M. Hussain, and K. Kim, "Phishing Email Detection Using Natural Language Processing and Machine Learning Techniques," *IEEE Access*, vol. 10, pp. 56545-56560, 2022.
DOI: 10.1109/ACCESS.2022.3183083
- [2] A. Aljofey and M. Basher, "A Deep Learning Approach for Phishing Email Detection Using NLP Features," in *Proc. IEEE Int. Conf. Big Data*, 2021, pp. 2234-2239.
DOI: 10.1109/BigData52589.2021.9671234
- [3] Y. Zhang, J. Ma, and H. Wang, "An NLP-Based Phishing Email Classification Framework with Attention Mechanism," *IEEE Access*, vol. 9, pp. 140985-140996, 2021.
DOI: 10.1109/ACCESS.2021.3119876
- [4] K. Gupta and P. Bansal, "Phishing Detection Model Based on NLP and Deep Learning Techniques," *IEEE Access*, vol. 10, pp. 45020-45031, 2022.
DOI: 10.1109/ACCESS.2022.3164987
- [5] P. Singh and S. Raj, "Multi-Feature Phishing Detection Using NLP and Machine Learning," *IEEE Access*, vol. 9, pp. 128746-128757, 2021.
DOI: 10.1109/ACCESS.2021.3113341
- [6] M. Amaz Uddin and I. H. Sarker, "Explainable Transformer-based Model for Phishing Email Detection," *arXiv preprint*, 2024.
- [7] S. Salloum, T. Gaber, S. Vadera, and K. Shaalan, "A Systematic Literature Review on Phishing Email Detection Using NLP Techniques," *IEEE Access*, 2022

DOI:10.1109/ACCESS.2022.3183083

- [8] J. Lee and H. Song, "NLP-Based Phishing Email Detection Using Ensemble Learning," IEEE Transactions on Information Forensics and Security, vol. 15, pp. 2372-2381, 2020.
DOI: 10.1109/TIFS.2020.2975241
- [9] T. Koide, N. Fukushi, H. Nakano, and D. Chiba, "Chat Spam Detector: ever aging Large Language Models for Phishing Email Detection," arXiv preprint, 2024.
- [10] S. G. Kim and Y. K. Lee, "Phishing Email Detection Using NLP and Feature Fusion," in Proc. IEEE Int. Conf. on Communications (ICC), 2021, pp. 1-6.
DOI: 10.1109/ICC42927.2021.9500160
- [11] M. Ahmad and A. Khan, "Deep Learning-Based Phishing Detection with NLP Features," IEEE Access, vol. 9, pp. 97281-97291, 2021.
DOI: 10.1109/ACCESS.2021.3094503
- [12] D. Roy and S. Ghosh, "NLP and Deep Learning Based Email Phishing Detection Using Transformers," IEEE Access, vol. 10, pp. 67280-67291, 2022.
DOI: 10.1109/ACCESS.2022.3180503
- [13] R. Singh and N. Sharma, "Phishing Email Detection Using NLP and Machine Learning Techniques: A Survey," in Proc. IEEE ICCCI, 2022, pp. 1-6.
DOI: 10.1109/ICCCI54376.2022.9744561
- [14] M. L. Hassan and S. A. Zaki, "Phishing Email Detection Using Natural Language Processing and Text Classification Techniques," IEEE Access, vol. 7, pp. 78432-78443, 2019.
DOI: 10.1109/ACCESS.2019.2927443
- [15] H. Patel and D. Patel, "Phishing Email Detection Using NLP and SVM Classifier," Proc. IEEE Int. Conf. on Computer Networks and Communications, 2020, pp. 242-247.
DOI: 10.1109/ICNC48265.2020.9035032
- [16] L. Chen and H. Zhang, "Phishing Email Identification via Multi-Head Attention and NLP," in Proc. IEEE Int. Conf. on Artificial Intelligence and Big Data, 2022, pp. 278-284.
DOI: 10.1109/ICAIBD.2022.9784532
- [17] P. Kumar and S. Saha, "BERT-Based NLP Model for Phishing Email Detection," IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 6, no. 5, pp. 792-801, 2022.
DOI: 10.1109/TETCI.2021.3098843
- [18] S. K. Das and R. Biswas, "Hybrid Feature-Based Phishing Email Detection Using Natural Language Processing and Ensemble Classifier," IEEE Trans. Inf. Forensics Security, vol. 16, pp. 1293-1305, 2021.
DOI: 10.1109/TIFS.2021.3051294
- [19] M. Nguyen, T. Nguyen, and T. H. Nguyen, "A Deep Learning Model with Hierarchical LSTMs and Supervised Attention for Anti-Phishing," arXiv preprint, 2018.
- [20] M. S. Farooq, M. A. Shah, and A. Alhussein, "Phishing URL Detection Using NLP and Deep Learning," IEEE Trans. on Industrial Informatics, vol. 17, no. 5, pp. 3702-3710, 2021.
DOI: 10.1109/TII.2020.3038353
- [21] S. K. Gupta, H. Agarwal, and M. Sharma, "LLM-Phish: Leveraging Large Language Models for Phishing Email Detection with Contextual Understanding," in Proc. 2024 IEEE International Conference on Artificial Intelligence and Data Science (ICAIDS), pp. 218-225, 2024.