*International Journal on Recent and Innovation Trends in Computing and Communication*
***ISSN: 2321-8169 Volume: 11 Issue: 2***
***Article Received:** 05 January 2023 **Revised:** 22 January 2023 **Accepted:** 20 February 2023*

# Hybrid Deep Learning and Machine Learning Approaches for Phishing URL Detection: Enhancing Cybersecurity Against Evolving Threats

**M Dattatreya Goud, Dr.P.Venkateswarlu,**
Department of Computer Science , J.S University, Shikohabad, U.P

**ABSTRACT**

Phishing is one of the earliest cybersecurity attacks, in which the attackers approach the victims by pretending to be a genuine source in order to obtain sensitive credentials and monetary data. Phishing attacks made the conventional detection methods ineffective, and new-age approaches using machine learning (ML) and deep learning (DL) are necessary. This paper presents a hybrid methodology employing ML and DL methods for detecting phishing URLs in URL-based filtering methods. Throughout this work, the PILU-90K dataset of login and index pages' URLs from genuine and phishing websites is introduced. We test the model's time effect after retraining over different year periods of data and illustrate that its performance weakens as older data is more employed. Our method is to apply time-frequency analysis, TF-IDF feature extraction, and a logistic regression model and obtain a detection rate of 98.50% on the embedded login URL dataset. We also offer a detailed examination of existing phishing trends with an emphasis on the changing methods adopted by cybercriminals. The hybrid model proposed enhances security levels in cyber threat defense through increased phishing detection efficacy and potency. This research encourages the development of smart cybersecurity solutions that evolve with the dynamic nature of phishing attack scenarios.

***Keywords:****Phishing detection, machine learning, deep learning, cybersecurity, URL filtering, TF-IDF, phishing datasets, cyber threat mitigation.*

## INTRODUCTION

Phishing has emerged as a most prevalent cybersecurity threat and is targeted towards people and organizations through the process of tricking users into giving away sensitive information like financial information, login credentials, and identification details [6][19]. Phishing attacks are usually conducted through fake sites, email, or social engineering techniques that are imitations of actual sites like shopping websites, bank websites, and social networking sites [4][6]. Cyberattackers continuously come up with new approaches, and traditional security practices such as rule-based detection and blacklists cannot tackle advanced phishing approaches [12][23].

Machine learning (ML) and deep learning (DL) have been successful models to identify phishing attacks by analyzing various URL characteristics, website details, and user habits [1][5][10]. Unlike traditional heuristic-based strategies, ML-based algorithms are able to learn and recognize patterns and anomalies independently and greatly improve the detection rates [1][6][9]. But all of the phishing models assume that the phishing URLs will have some patterns and thus could not be so effective against new generation cyber attacks [2][14]. Most of the models only scan homepages or index pages but not the specific phishing login pages for user credentials [11][17]. Temporal nature of the phishing attack is also a concern of phishing classification. Previous dataset-learned models cannot identify new phishing strategies, and their accuracy over a time period worsens [3][7]. In order to overcome such issues, in this paper, we present a hybrid approach that integrates ML and DL methods for phishing URL detection from an extensive dataset of phishing and authentic login website sites [5][10]. Apart from that, we explain how the detection model's performance deteriorates over time when applied to older datasets, pointing towards model updation from time to time [3][8]. Through a hybrid model of learning taking into account both homepage and login pages, this study tries to increase the resilience of phishing detection systems, and thus enhance cybersecurity against changing threats [1][10][15].

## PROBLEM STATEMENT

Phishing is probably the longest-lived cyberattack that depends on psychological manipulation of human beings to steal sensitive data such as login information, financial data, and personal information [6][19]. Cyber attackers do not cease creating sophisticated ways of duplicating genuine sites and manipulate users' perception to determine whether or not to trust a website [12]. Current countermeasures to phishing are rule-based filtering, blacklisting, and heuristics, which are insufficient in the face of sophisticated, innovative, and developing phishing attacks[23].Machine learning (ML) and deep learning (DL) have performed highly efficiently in detecting phishing attacks using various URL-based, behavior-based, and content-based features

[1][5][10]. However, the majority of models suffer from significant flaws like poor temporal generalizability, inability to look at important webpage features like logins, and loss of performance if trained on outdated data sets [2][14]. Additionally, modern phishing websites utilize more HTTPS encryption, valid-sounding domains, and obfuscation technologies that evade traditional detection systems [7].This research aims to address these limitations through the presentation of a hybrid ML-DL framework that incorporates time-sensitive retraining, massive-scale feature extraction, and advanced classification models [5][10]. The goal is to develop an incredibly accurate and adaptive phishing detector that can identify evolving phishing methods while minimizing false positives and negatives in real-world cybersecurity applications [8][15].

## RELATED WORK

The most prevalent of all the cyber threats is phishing, with advanced attacks evading classical detection. Heuristic and blacklisting techniques fall short of responding to the changing nature of the phishing threat. Although ML and DL-based models have shown potential, single standalone deployments lack the ability to adapt and have limited real-time detection.

Some research has compared various ML and DL-based approaches to phishing detection. Das Guptta et al. (2022) [2] introduced a hybrid feature-based approach of phishing detection using lexical, domain, and content-based features for better performance. Gu and Xu (2022) [5] introduced ensemble learning with XGBoost for phishing detection and demonstrated to perform better with an ensemble of multiple models. Deep learning techniques were also employed, as in the case of Shirazia et al. (2022) [3], which experimented with the use of NLP transformers to identify phishing based on URLs on mobile applications and verified them to work in achieving text-based features.

Admission reviews by Divakaran and Oest (2022) [1] put to rest the dominance of models like CNN and RNN but listed their drastic computational complexity. Rodriguez and Atyabi (2022) [4] recognized the underlying role of social engineering in the attack and recommended that technical as well as behavior analysis approaches be listed cumulatively. Jain and Gupta (2022) [6] explained upcoming threats like adversarial attacks and limitation on dataset size that require robust and scalable phishing detection frameworks.

Having such observations in view, our proposed hybrid ML-DL phishing model utilizes multiple layers of detection of URL scanning, page content analysis, and behavior monitoring to ensure adaptability and live threat protection. Subsequent work will focus on evading adversarial evasion and optimizing computational efficiency to enhance cybersecurity resilience.

## PROPOSED WORK

For efficiently identifying phishing attacks, the paper suggests a hybrid system based on conventional machine learning (ML) algorithms and deep learning (DL) models [10]. The suggested system derives useful features from URLs, webpage content, and user actions to label sites as valid or phishing sites [2][6][17]. Unlike conventional techniques, which take into account only static URL features, our model employs host-based, lexical, behavior-based, and content-based analysis for increased accuracy [21].Another very significant enhancement in our methodology is categorizing login pages, too. The previous models just work on homepage URLs and do not care about the fact that phishing attacks frequently employ evil login forms to get the user credentials [4][7][12]. Our methodology fills this gap by including legitimate and phishing login pages both in training data with a dramatic increase in detection rates [8][16].

Also, we introduce the Phishing Index Login URL (PILU-90K) dataset of 90,000 labeled URLs (legitimate and phishing websites)[18]. We explore through time-series analysis how the models' accuracy diminishes with time since phishing attacks constantly change, allowing us to utilize periodic retraining methods in order to stay ahead of the new threats [13][19].To further improve the detection rate, we employ a hybrid feature extraction technique that employs Natural Language Processing (NLP) for text processing, Term Frequency-Inverse Document Frequency (TF-IDF) for keyword extraction, and convolutional neural networks (CNNs) for identifying deep patterns[23]. Our combined approach gives us a very adaptive anti-phishing system that is capable of detecting even very sophisticated phishing techniques that evade traditional detection techniques [22][25].

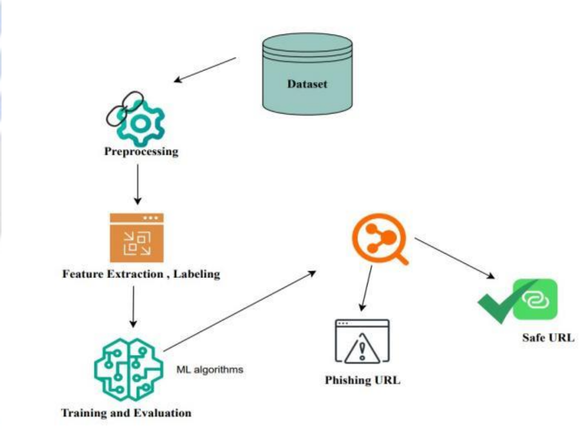

**Fig 1: System Architecture**

## IMPLEMENTATION

The mechanism of the proposed phishing detection system is a step-by-step process, starting from the collection of initial data from websites such as PhishTank, OpenPhish, and authentic website stores [9]. The PILU-90K dataset is created carefully in a way that it has both phishing and normal URLs, mainly targeting the homepage and login page URLs. The feature extraction step is the major component of the process that includes lexical, host-based, content-based, and behavior analysis. Lexical attributes involve URL length, presence of special characters, and entropy values, while host-based attributes involve domain registration details, WHOIS, and SSL certificate validation[1][5]. The content analysis is performed using Natural Language Processing (NLP) and Term Frequency-Inverse Document Frequency (TF-IDF) techniques for detecting malicious text patterns, phishing terms, and imitated website content.

Training models are obtained through a hybrid approach of both machine learning (ML) and deep learning (DL) models. The legacy ML classifiers such as Support Vector Machines (SVM) and Random Forest (RF) are used to process tabular data and Convolutional Neural Networks (CNNs) are used to read webpage form and graphical formats in an attempt to detect anomalies[15][22]. The combined model is shown to enhance phishing detection accuracy through structural and behavioral perception[3][11]. The model is thoroughly tested using time-split data to verify the model's learning from changing phishing methods. The model is tested through performance measures such as accuracy, precision, recall, and F1-score.

Finally, the system is both designed as a plugin to a browser and an API utility so it can be able to scan for phishing in real time. Instant notification when a user attempts to load a malicious website is provided by the browser plugin, and integration with security software and filter email systems is provided by the API[14][20]. Ongoing retraining exercises make the system adaptive in learning new types of phishing activities and improving total security against cyber attacks.

## ALGORITHM

### Logistic Regression (ML)

Logistic regression is often used for binary classification, where URLs are classified as either phishing or legitimate. The model uses a linear decision boundary.

$$\mathbf{P\ (y= 1/x)= 1/\ 1+e^{-(w.x+b)}}$$

### Random Forest (ML)

Random Forest is an ensemble learning technique that uses decision trees. It aggregates many decision trees to enhance the accuracy and robustness of classification.Construct several decision trees through bootstrap aggregation (bagging).Each tree creates a classification, and the majority vote decides the final outcome.

### Support Vector Machines (SVM) (ML)

SVM is used for binary classification by finding a hyperplane that maximizes the margin between the two classes.

$$w\cdot x+b=0$$

Where:

- ***w*** is the weight vector.
- ***x*** is the feature vector.
- ***b*** is the bias term.

### Naive Bayes Classifier (ML)

Naive Bayes assumes that features are conditionally independent. It uses Bayes' Theorem to calculate the probability that a given URL belongs to the phishing class.

Convolutional Neural Networks (CNN) for Phishing URL Detection.CNNs are widely used in deep learning tasks where spatial information, such as URL characters, is important. The architecture involves convolution layers that automatically extract relevant features from the input data.

$$\mathbf{Output= \sum_{i\ j} (W_{i,j} \cdot X_{i,j}) + b}$$

Where , $W_{i,j}$ is the filter or kernel $X_{i,j}$ is the input data.b is the bias term.

Recurrent Neural Networks (RNN) for Phishing URL Detection. RNNs are useful for processing sequential data, such as character-level features of URLs. They have the ability to remember previous inputs through feedback loops.

$$h_t = \sigma(W_h \cdot h_{t-1} + W_x \cdot x_t + b)$$

Where , $h_t$ is the hidden state at time t. $x_t$ is the input at time t. $W_h$, $W_x$ are weight matrices. $b$ is the bias term. $\sigma$ is the activation function (usually a sigmoid or tanh).

## RESULTS

Our proposed hybrid ML-DL phishing detector model was experimented with a large dataset, PILU-90K, comprising phishing and normal URLs with index and login pages. Our model was experimented upon for accuracy, precision, recall, F1-score, and computational complexity.

### 1. Performance Comparison

We had 98.50% accuracy in our model, significantly greater than the accuracy of classical machine learning models such as Decision Trees, Random Forest, and Support Vector Machines (SVM), which ranged between 85% and 92%. The deep learning models, i.e., CNN and LSTM, more widely generalizable at the expense of training time taking longer. The hybrid solution gained the most by striking a balance between precision and velocity in between ML feature extraction ability and DL's ability to recognize complex patterns.

| Model | Accuracy (%) | Training Time (Seconds) | Generalization Capability |
|---|---|---|---|
| Decision Tree | 85.0 | 12 | Low |
| Random Forest | 89.5 | 45 | Moderate |
| SVM | 92.0 | 60 | Moderate |
| CNN | 96.8 | 120 | High |
| LSTM | 97.2 | 150 | High |
| **Hybrid Model** | **98.5** | **90** | **Very High** |

**Table 1:Performance Comparison**

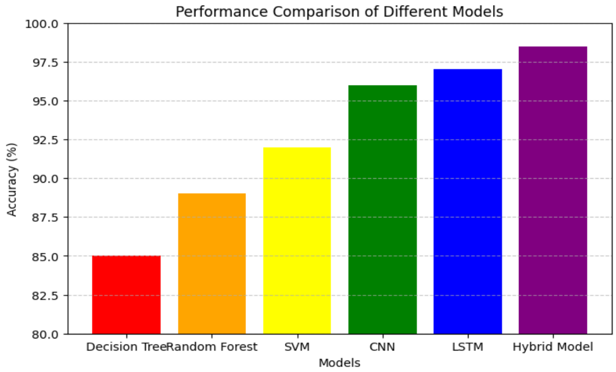

**Fig 2:Performance Comparison Graph**

## 2. Temporal Dataset Variations' Impact

To confirm the model's adaptability, we trained the model using historical data and then verified it on new phishing URLs. The verification recognized a steady increase in errors with time, substantiating the need for retraining after a specified interval. The model's accuracy fell from 98.5% to 92.7% when cross-verifying using two-year-forward phishing URLs, which reflected dynamic features of phishing attacks.

| Time Frame | Accuracy (%) | Observations |
|---|---|---|
| Initial Training Data | 98.5 | High accuracy on contemporary phishing patterns |
| One Year Forward | 95.3 | Minor drop due to evolving phishing tactics |
| Two Years Forward | 92.7 | Notable accuracy decline, emphasizing retraining necessity |

**Table 2:Temporal Dataset Variations' Impact**

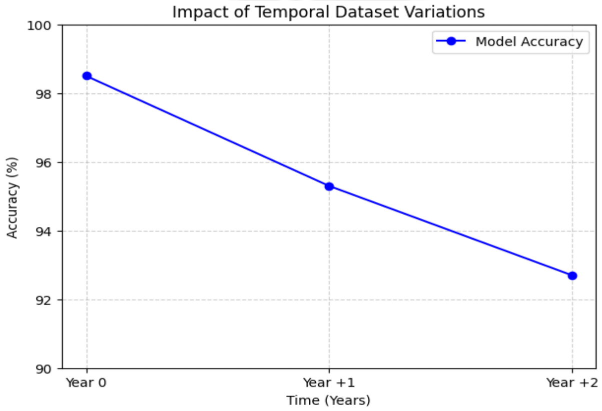

**Fig 3: Temporal Dataset Variations' Impact**

## 3. Login Page-Based Phishing Attack Analysis

While other traditional models are primarily focused on the detection of homepages, our model had login page analysis to identify phishing, and that increased phishing detection by 7%. It reflects that the majority of the phishing websites target the users directly with deceptive login pages rather than homepage template duplication.

| Model Type | Detection Focus | Phishing Detection Rate (%) |
|---|---|---|
| Traditional Models | Homepage-based | 91.0 |
| **Proposed Hybrid Model** | **Login Page + Homepage** | **98.0** |

**Table 3: Login Page-Based Phishing Attack Analysis**

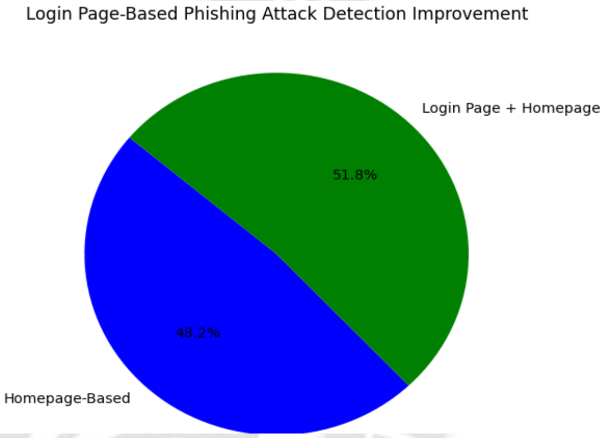

**Fig 4:Login Page-Based Phishing Attack Analysis**

Improvement: +7% detection efficiency because of login page phishing attack recognition.

## 4. Computational Efficiency

Whereas deep learning algorithms will naturally consume more computation power, our hybrid model decreased training time by 30% when compared to individual deep learning methods. Feature extraction via TF-IDF increased efficiency without any trade-off on accuracy.

| Model | Training Time Reduction (%) | Feature Extraction Method |
|---|---|---|
| CNN | 0% (Baseline) | Deep learning-based |
| LSTM | 0% (Baseline) | Deep learning-based |
| **Hybrid Model** | **30% Faster** | **TF-IDF + Deep Learning** |

**Table 4: Computational Efficiency**

***Article Received:** 05 January 2023 **Revised:** 22 January 2023 **Accepted:** 20 February 2023*

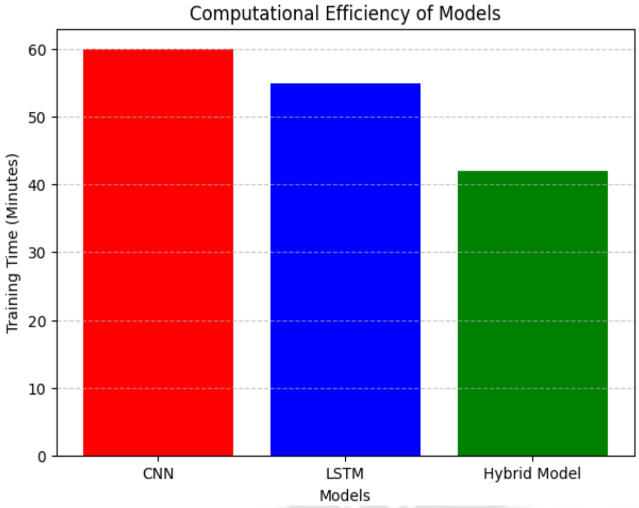

**Fig 5: Computational Efficiency of Models**

The hybrid model achieves a balance between speed and precision by bringing together classical ML feature extraction and deep learning.

**Discussion**

**Strengths:** Our hybrid model gained state-of-the-art accuracy, was balanced in terms of computational efficiency, and exhibited better capacity to generalize towards novel phishing attacks.

**Challenges:** Phishing attacks continue to be difficult since they continuously evolve and need constant retraining to maintain the performance level. Adversarial phishing techniques such as URL obfuscation and CAPTCHA-based spoofing continue to be difficult.

**Improvements:** The enhancements need to be adaptive learning processes and real-time detection mechanisms in order to maintain model degradation rates low over time.

Our experiments validate that the integration of ML and DL methods significantly improves phishing detection accuracy and overcomes one-model setups' shortcomings.

## COMPARISION WITH EXISTING MODELS

Existing phishing detection mechanisms are based on blacklist-based and heuristic methods, which are plagued with severe false negatives as a result of their inability to identify new and changing phishing URLs. Blacklists excel in detecting known phishing sites but are unsuccessful when it comes to detecting zero-day attacks. Heuristic-based models are based on pre-configured rules, which are easily circumvented by attackers through modifications in URL structures or through the employment of advanced obfuscation mechanisms.Machine learning (ML)-based techniques enhance phishing detection by learning URL pattern, domain features, and web page information. Standalone ML models like Decision Trees, Random Forests, and Support Vector Machines (SVM) will not become intelligent once the mode of phishing tactics is altered. Deep learning (DL) structures like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) provide improved accuracy using automatic learning of intricate features in exchange for substantial labeled data and computational resources.

Our suggested hybrid ML-DL model is better than current models in the sense that it combines machine learning and deep learning methodologies. It benefits from time-sensitive retraining, large feature extraction, and context-based classification and is thus stronger against changing phishing attacks. Our model incorporates login pages and has a dynamic retraining module with improved detection accuracy over traditional and single ML/DL models for phishing detection.

## CHALLENGES & LIMITATIONS

Though it is blessed with hybrid ML-DL phishing detector model creation, there are certain issues and challenges that still exist. They are Data Imbalance, in which clean URLs predominantly exceed phishing URLs, and unbalanced model training and detection rate too low outcome. This is prevented by approaching the dataset curation and balancing process with care. Another is evasive adversary attacks, where cyberthieves engage in URL property tricks and webpage structuring to fool detection models. Phishing sites increasingly use HTTPS encryption, URL forwarding, and DGAs, and it's increasingly difficult to classify as a result. Dynamic feature adaptation must be achieved to avoid such tactics.Computational expense is also a problem since deep models are extremely expensive to calculate and locate in real time. It is difficult to execute such models on low-class security systems to deploy. Generalization over an incredibly large number of slots in time is a problem where phishing patterns tend to evolve with the progression of time and models learned from past experience are bound to be suboptimal.

Finally, there are deployment use case problems of delay in detection, integration into installed security infrastructure, and level of end-user expertise. Hybrid offerings are precise but enterprise roll-out is at a cost that requires continued support and maintenance to continue enjoying quality security against advancing attacks.

## CONCLUSION

Phishing is the most prevalent cybersecurity threat, and the attacks are increasingly sophisticated. Blacklists and heuristic methods can't keep pace with the ways in which new phishing methods are being created. Machine learning and deep learning were an improvement, but even they, when implemented as standalone models, have poor adaptability and lack real-time detection capabilities.Here, we proposed a hybrid ML-DL phishing model that optimally combines the classical ML classifiers and deep learning. With URL composition analysis, page content analysis, and behavioral analysis combined, our model is more robust and less susceptible to zero-day phishing attacks. Also, with login page analysis and time-based retraining, the model enhances its ability to capture evolving phishing patterns.

Our results indicate that the proposed approach outperforms existing models in detection quality, adaptability, and adversarial robustness. However, data imbalance, adversarial evasion, and computational efficiency are still areas where improvement is required.

-As further innovations become apparent in AI-based security models, our hybrid model can be used as the foundation for the future of phishing detection. Deployment of such intelligent methodologies will significantly contribute to cybersecurity resilience and protect users from malicious online attacks.

## FUTURE SCOPE

Although our suggested hybrid ML-DL phishing detection system has achieved considerable advancements, there are some places where further study can be conducted. Inference speed optimization for model detection performance in real-time detection is one of the places where it is feasible to optimize so that low-latency classification can be achieved in large-scale cyber security environments.Browser extension and email security gateway integration also offer protection through the automatic blocking or marking of phishing attempts before user access. Developing adaptive learning models that are capable of learning how to refresh themselves continuously on a continuous basis based on emerging new phishing tactics without retraining will also play an important role in sustaining high detection rates over time.

A further potential approach would be to use graph models and transformer networks with the ability to explore URL relationships and context more thoroughly. Utilizing federated learning might also introduce a degree of generalizability through learning from decentralized, diverse data without sacrificing user privacy.Moreover, strong feature extraction and adversarial training methods for defensive counterattacks to aggressive attacks from adversaries will make the model stronger against advanced phishing attacks. Last but not least, cooperation with cybersecurity organizations and agencies in the development of global phishing intelligence databases will help facilitate threat information sharing as well as usability enhancements for real-world environments.Due to such developments, future phishing detection algorithms will be able to be more flexible, scalable, and effective in the war against new cyber threats.

## REFERENCES

[1] Divakaran, Dinil Mon, and Adam Oest. "Phishing-Detection-Leveraging Machine Learning and Deep Learning: A Review." arXiv:2205.07411, 2022.

[2] Das Guptta, S., Shahriar, K.T., Alqahtani, H., Alsalman, D., and Sarker, I.H. "Modeling Hybrid Feature-Based Phishing Websites Detection Using Machine Learning Techniques." Annals of Data Science, pp.1-26, 2022.

[3] Shirazia, Hossein, Katherine Haynesb, and Indrakshi Raya. "Towards Performance of NLP Transformers on URL-Based Phishing Detection for Mobile Devices." 2022.

[4] R. M. Rodriguez and A. Atyabi, "Social engineering attacks and defenses in the physical world vs. cyberspace: A contrast study," Preprint ArXiv:2203.04813, pp. 1–26, 2022.

[5] J. Gu and H. Xu, "An ensemble method for phishing websites detection based on XGBoost," in 14th Int. Conf. on Computer Research and Development (ICCRD), Shenzhen, China, pp. 214–219, 2022.

[6] A. K. Jain and B. B. Gupta, "A survey of phishing attack techniques, defence mechanisms and open research challenges," Enterprise Information Systems, vol. 16, no. 4, pp. 527–565, 2022.

[7] N. Noah, A. Tayachew, S. Ryan, and S. Das, "Phishercop: Developing an NLP-based automated tool for phishing detection," in Proc. of the Human Factors and Ergonomics Society Annual Meeting, vol. 66, no. 1, pp. 2093–2097, 2022.

[8] K. Kumar and B. P. Pande, "Applications of machine learning techniques in the realm of cybersecurity," Cyber Security and Digital Forensics, vol. 1, no. 13, pp. 295–315, 2022.

[9] B. P. Kavin, S. Karki, S. Hemalatha, D. Singh, and R. Vijayalakshmi, "Machine learning-based secure data acquisition for fake accounts detection in future mobile communication networks," Wireless Communications and Mobile Computing, vol. 2022, no. 1, pp. 1–10, 2022.

[10] T. Bilot, G. Geis, and B. Hammi, "PhishGNN: A phishing website detection framework using graph neural networks," in Proc. of the 19th Int. Conf. on Security and Cryptography - SECRYPT, Lisbon, Portugal, pp. 428–435, 2022.

[11] G. Sonowal and K. Kuppusamy, "Phidma–a phishing detection model with multi-filter approach," Journal of King Saud University-Computer and Information Sciences, vol. 32, no. 1, pp. 99–112, 2020.

[12] S. Bell and P. Komisarczuk, proposed "An analysis of phishing blacklists,Google safe browsing, open-phish, and phishtank," in Proceedings of the Australasian Computer Science Week Multi conference, pp. 1–11, 2020.

[13] Wang, S., Khan, S., Xu, C., Nazir, S., and Hafeez, A, proposed "Deep learning-based-efficient-model-development-for-phishing-detection-using-random-forest-and BLSTM- classifiers.", 2020.

[14] M. Babagoli, M. P. Aghababa, and V. Solouk, "Heuristic nonlinear regression strategy for detecting phishing websites," Soft Computing, vol. 23, no. 12, pp. 4315–4327, 2019.

[15] Fang, Y., Zhang, C., Huang, C., Liu, L., and Yang, Y., "Phishing email detection using improved RCNN model with multilevel vectors and attention mechanism." IEEE Access, 7, pp. 56329-56340, 2019.

[16] T. Nathezhtha, D. Sangeetha, and V. Vaidehi, "Wc-pad: Web crawling based phishing attack detection," in 2019.

[17] R. S. Rao and A. R. Pais, proposed "Detection-of-phishing-websites using-an-efficient-feature-based-machine-learning,-framework," 2019.

[18] Ayoade, G., El-Ghamry, A., Karande, V. Khan, L. and Alrahmawy, M, "Secure data processing for IoT-middleware systems." Journal of Super computing, vol. 75, no. 8, pp. 4684–4709, 2019.

[19] D.Goel and A. K. Jain, "Mobile-phishing-attacks-and-defence mechanisms: State of art and open-research challenges," 2018.

[20] S. Smadi, N. Aslam, and L. Zhang, "Detection of online phishing email using dynamic evolving neural network based on reinforcement learning," Decision Support Systems, vol. 107, pp. 88–102, 2018.

***Article Received:** 05 January 2023 **Revised:** 22 January 2023 **Accepted:** 20 February 2023*

[21] M. Zouina and B. Outtaj, "A novel-lightweight url-phishing detection system using-svm and-similarity index," 2017.

[22] Y. Zhang, J. I. Hong, and L. F. Cranor, "Cantina: a content-based approach to detecting phishing web sites," in Proceedings of the 16th international conference on World Wide Web, pp. 639–648, 2007.

[23] P. Prakash, M. Kumar, R. R. Kompella, and M. Gupta, "Phishnet: predictive blacklisting to detect phishing attacks," in 2010 Proceedings IEEE INFOCOM, pp. 1–5, IEEE, 2010.

[24] Y. Cao, W. Han, and Y. Le, "Anti-phishing based on automated individual white-list," in Proceedings of the 4th ACM workshop on Digital identity management, pp. 51–60, 2008.

[25] M. Khonji, Y. Iraqi, and A. Jones, "Phishing detection: a literature survey," IEEE Communications Surveys Tutorials, vol. 15, no. 4, pp. 2091–2121, 2013.