Scientific Eminence Group

Research Article

Journal of Artificial Intelligence Systems and Applications

Open Access

jais@scientificeminencegroup.com

Modern and Intelligent Methods for Detecting Phishing Sites Based on Machine Learning Algorithms

Hossein Moinzad[1*], Amin Mohajer[2] and Pegah Shayestehfar[3]

[1]*Department of Industrial Management, Faculty of Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran*

[2]*Department of Information Technology Management, Faculty of Management, Kish International Branch, Islamic Azad University, Kish, Iran*

[3]*Department of Information Technology Management, Faculty of Management, Hamedan Branch, Islamic Azad University, Hamedan, Iran*

*Corresponding Author

Hossein Moinzad, Department of Industrial Management, Faculty of Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran, E-mail: moinzad@gmail.com

Citation

Hossein Moinzad, Amin Mohajer, Pegah Shayestehfar (2025) Modern and Intelligent Methods for Detecting Phishing Sites Based on Machine Learning Algorithms. J Artif Intell Syst Appl 4: 1-18

Publication Dates

Received date: March 03, 2025
Accepted date: April 03, 2025
Published date: April 06, 2025

Abstract

Today, phishing is one of the most serious and dangerous online threats in the field of cybersecurity. Accordingly, providing a new and precise solution for detecting phishing attacks on websites and enhancing security and user authentication is considered a serious issue, highlighting the necessity of conducting this research. This study is categorized as quantitative methods and, in terms of analysis method, is of the analytical type. The necessary information for conducting the research will be collected through library sources. Additionally, the research data will be gathered using articles published in reputable scientific journals and conferences, as well as databases and computer networks, including reputable global scientific publication databases such as Springer, Elsevier, IEEE, etc., and the final implementation will be carried out using MATLAB software. In the proposed hybrid model, it first works with GWO to reduce the dimensions of the input data and selects only the most important features for phishing detection. These features are then fed into the RNN, which processes the data through its layers to classify websites as legitimate or phishing. This two-stage approach not only speeds up the identification process but also increases the accuracy and overall robustness of the model against various types of phishing attacks. This issue has been demonstrated in the application of the model on two datasets.

Keywords: phishing, RNN network, gray wolf algorithm, time series, preprocessing

Introduction

Today, phishing is one of the most serious and dangerous online threats in the field of cybersecurity. The use of social networks, e-commerce, online banking, and other online services has significantly increased due to the rapid development of internet technologies. We Are Social (2021 Global Overview Report) published data from the "Digital Report 2021," which shows that internet users have reached 4.66 billion, an increase of 7.3% (316 million new users) compared to January 2020. In 59.5% of cases, phishing has provided a platform for phishing attackers to earn money through extortion and theft of confidential information from internet users. Based on the phishing approach, the attacker creates a fake website and sends links to online platforms like Facebook, Twitter, emails, etc., with a message of panic, urgency, or a financial offer, instructing the recipient to take immediate action. When the user unwittingly clicks on the link and updates any sensitive credentials, cyber attackers gain access to user information such as financial data, personal information, usernames, passwords, etc. This stolen information is used by cybercriminals for various illegal activities, including extorting victims.

One way to identify phishing is by utilizing machine learning technology. Detecting phishing sites using machine learning involves developing algorithms that can identify fraudulent websites by analyzing patterns and common features of phishing attacks. Machine learning models, such as logistic regression, K-nearest neighbors, support vector machines, and neural networks, have been trained on datasets containing features of legitimate and phishing websites. These features may include URL structure, website content, and external information such as domain registration details. The goal is to enable these models to effectively distinguish between secure and malicious sites.

Cybercriminals obtain this information through various illegal means and impersonate these users to engage in illegal activities online. In the early days of the invention of the internet, network security issues had already emerged. With the development of the internet, network attack techniques have also rapidly changed, bringing numerous challenges to network security. Considering the methods and forms of network attacks, cybersecurity issues are primarily divided into denial of service (DoS) attacks, man-in-the-middle (MitM) attacks, SQL injection, zero-day exploits, DNS tunneling, phishing, and categories of malware.

Phishing is a network attack that combines social engineering and computer technology to steal users' sensitive personal information. Attackers request individuals to click on phishing links by sending emails, text messages, or social media messages with misleading content. Phishing has existed for over 30 years, and every year a large number of users fall victim to it, causing economic losses. Specifically, in 2020, the number of phishing attacks surged dramatically [1].

The goal of phishing campaigns is to steal confidential information using sophisticated methods, techniques, and tools through content injection, social engineering, online social networks, and mobile applications. To prevent and mitigate the risks of these attacks, various phishing detection methods have been developed, among which machine learning algorithms have shown promising results [2]. Phishing attacks primarily begin with sending an email that scares users into taking some urgent actions. In addition to email communication, phishing attacks can also target online social networks, blogs, forums, VoIP, mobile applications, and messaging platforms [3].

Recently, phishing scams have emerged with various systems, including blockchain platforms. As cryptocurrencies like Bitcoin and Ethereum reached their highest prices in the market, cybercriminals targeted these digital assets [4]. These attacks can not only lead to financial loss but also result in the loss of intellectual property (IP) and valuable confidential user information. Additionally, it may undermine trust and impact national security. Thus, detecting phishing is more important and crucial than ever.

Phishing detection systems are generally divided into two categories: list-based detection systems and machine learning-based detection systems. There is a wealth of background literature that utilizes third-party services such as web-based black and white lists, traffic size ranking, domain information, etc. Primarily, the use of these services increases the efficiency of the detection system. However, if the goal is real--time execution, then these services increase the detection time; therefore, they may not be useful [1].

Anti-phishing strategies include educating internet users and technical defenses. Identifying phishing websites is an effective method in the overall process of deceiving users' information. Many academic research projects and commercial products have been published to identify phishing websites. Traditional methods are list-based solutions that collect legitimate and legal websites in a whitelist or confirmed phishing websites in a blacklist and widely share the list to prevent attacks on other users. These approaches effectively prevent the reuse of the same phishing website URL and reduce the number of affected users and casualties. However, these methods have a significant drawback: the inability to identify new phishing URLs. As a result, some innocent users may be attacked before the link is added to the blacklist.

Some researchers proposed rule-based methods for detecting new fake websites. This method involves the expertise of security experts and the analysis of phishing website URLs. According to W3C standards, a main URL includes the protocol, subdomain, domain name, port, path, query, parameters, and fragment. Essentially, rules are generated from the components of URLs, such as the requirement that the domain name be similar to other legitimate domains. In these rules, some require requesting information like the domain registration date from third-party services. When rules are published in certain technical papers, phishers learn them and then find new phishing URLs that do not comply with the rules. Subsequently, cybersecurity experts formulated additional rules, some of which were based on the source code of web pages.

With the development of machine learning techniques, various machine learning-based methods for identifying phishing websites have emerged to enhance predictive performance. Phishing detection is a supervised classification approach that uses labeled datasets to fit models for classifying data. There are several algorithms for supervised learning processes, such as simple Naive Bayes, neural networks, linear regression, logistic regression, decision trees, support vector machines, K-nearest neighbors, and random forests.

Considering the issue and challenge raised, this research will provide various solutions and strategies to enhance security and identify phishing sites using a data science and machine learning approach. After that, the impact of this solution will be measured in comparison with other effective dimensions and components in this field. In other words, we intend to examine the effect of using machine learning methods on the identification of phishing sites and the enhancement of user security. Accordingly, the collected data related to the identification of phishing sites will be examined, and in the first stage, we will analyze it using a machine learning approach. It is worth mentioning that traditional deep learning approaches have a feature selection problem. These approaches require manual feature selection. The better the feature selection is performed, the more accurate the models become, and vice versa.

Jain & Gupta presented a comprehensive survey on analyzing phishing attack techniques, detection methods, and some existing challenges [5]. They included statistical reports and motivations for phishing attacks and presented various phishing attack techniques on personal computers and smartphones. Then, the authors introduced different defensive methods and compared existing anti-phishing approaches published from 2006 to 2017 for their advantages and limitations. After that, several significant challenges such as the selection of efficient features, the identification of small URLs, and the recognition of smartphones were presented.

[6] presented a machine learning-based approach for phishing detection using link information in 2019. This paper presents a new approach that can identify phishing attacks by analyzing the links present in the HTML source code of websites. The proposed approach includes new and prominent specific features for detecting phishing attacks. The suggested approach divides specific hyperlink features into 12 different categories and uses these features to train machine learning algorithms. The performance of the proposed phishing detection approach has been evaluated on various classification algorithms using datasets of phishing and non-phishing websites. The proposed approach is a fully client-side solution and does not require any services from third parties. Additionally, the proposed approach is language-independent and can identify websites written in any textual language. Compared to other methods, the proposed approach has a relatively high accuracy in detecting phishing websites as it achieved over 98.4% accuracy in logistic regression classification.

[7] introduced a new technique for feature selection in web phishing detection models in 2021. The proposed method in this paper consists of two stages. The first stage computes the impact of the absence of each feature by training a random forest model with a new dataset that removes one feature and specifies the accuracy. After analyzing the absence of each ele-

ment in the loop, a feature queue is ranked by accuracy from highest to lowest. The second stage involves training and testing the model starting from one feature, adding a new feature from the ranked feature list each time to form the dataset, calculating accuracy each time, and ultimately finding the subset of features with the highest accuracy. This method works for selecting the most effective subset of features. However, since each new dataset must go through the training and testing process of the algorithm, high computational complexity and long computation times are involved. For example, if the UCI dataset has 30 features, the first stage loops 30 times, and the second stage loops 30 times, requiring the tree algorithm to be trained each time. Therefore, this method is suitable for small feature sizes and single classifiers.

[8] presented an efficient approach for phishing detection using machine learning in 2021. This study examined the role of feature selection methods in the efficient and effective detection of phishing web pages. A comparative analysis of machine learning algorithms was performed based on their performance with and without feature selection. Experiments were conducted on a phishing dataset with 30 features, including 4898 phishing and 6157 benign web pages. Several machine learning algorithms were also used to obtain the best results. Following that, a feature selection method was applied to improve the performance of the models. The results indicate that the best accuracy is achieved by random forest both before and after feature selection, with a significant improvement in model build time. The experiments show that using a feature selection method alongside machine learning algorithms can enhance the build time of classification models for phishing detection without compromising their accuracy.

[9] focused on detecting phishing URLs using machine learning methods. This article aims to provide a solution for identifying phishing websites with the help of machine learning algorithms that concentrate on the behaviors and quality of proposed URLs. The web security community has created a blacklist service to identify malicious websites. Various methods such as manual reporting and site analysis discoveries have been used to generate these blacklists. Due to their novelty, lack of evaluation, or incorrect evaluation, many malicious websites accidentally evade the blacklist. To create a machine learning model to determine whether a URL is malicious or not, algorithms such as random forests, decision trees, GBM light, logistic regression, and Support Vector Machine (SVM) have been used. Feature extraction is the first step, and applying the model is the next stage. The results indicate that among these algorithms, the random forest model has higher efficiency.

[10] presented an effective and safe mechanism for phishing attacks using a machine learning approach in 2022. This paper focused on a three-phase phishing attack that accurately identifies problems in a content-based manner as a phishing attack mechanism, with three input values considered - a uniform source locator, traffic, and web content based on phishing attack features and non-attack features of phishing website techniques. To implement the proposed phishing attack mechanism, a dataset of recent phishing cases was collected. The results indicate that real phishing cases provide higher accuracy in both zero-day phishing attacks and the detection of phishing attacks. Three different classifiers were used to determine the classification accuracy in phishing detection, resulting in classification accuracies of 95.18%, 85.45%, and 78.89% respectively for the NN, SVM, and RF models.

[11] presented a predictive model for phishing detection in 2022. In this research, an advanced machine learning-based predictive model was suggested to enhance the effectiveness of anti-phishing schemes. The predictive model includes a feature selection module that is used to construct an effective feature vector. These features are extracted from the URL, web page features, and web page behavior using a system based on incremental authorship to provide the resulting feature vector to the predictive model. The proposed system utilizes Support Vector Machine and Naïve Bayes, which were trained on a 15-dimensional feature set. The experiments were based on a dataset consisting of 2,541 phishing instances and 2,500 benign samples. Using 10-fold cross-validation, the experimental results show significant performance with 0.04% false positives and 99.96% accuracy for both SVM and NB predictive models.

[12] in 2023 presented an intelligent cybersecurity phishing detection system using deep learning techniques. In this research, a detection model was proposed by leveraging machine learning techniques with the dataset divided for training the detection model and validating the results using test data to capture the inherent features of email texts and other attributes classified as phishing or non-phishing. Various machine learning algorithms were also evaluated for comparison across three phishing sites using three different datasets. The

results indicate that the proposed model achieved accuracies of 88%, 100%, and 97% on all three databases.

[13] presented an intelligent phishing detection design using deep learning algorithms in 2023. This study focused on the design and development of a deep learning-based phishing detection solution that leverages global source locators and website content such as images, text, and frames. In this study, a convolutional neural network (CNN) and a long short-term memory (LSTM) algorithm were used to build a combined classification model called the Intelligent Phishing Detection System (IPDS). To construct the proposed model, the CNN and LSTM classifiers were trained using a global source locator of 1 meter and over 10,000 images. The sensitivity of the proposed model was then determined considering various factors including feature type, the number of misclassifications, and segmentation issues. Extensive empirical analysis was conducted to evaluate and compare the effectiveness of IPDS in detecting phishing web pages and phishing attacks when applied to large datasets, with results showing that the proposed model achieved an accuracy rate of 93.28 percent and an average detection time of 25 seconds.

[14] in 2023 focused on identifying phishing Domains using machine learning. This paper develops and compares four models to examine the effectiveness of using machine learning for detecting phishing domains. It also compares the most accurate model of the four with existing solutions in the literature. These models were developed using Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees (DTs), and Random Forest (RF) techniques. Additionally, the UCI phishing domain dataset was used as a benchmark for evaluating the models. The findings indicate that the model based on random forest techniques is the most accurate model compared to the other four techniques and performs better than other solutions in the literature.

[15] in 2023 presented a deep learning-based phishing detection system using CNN, LSTM, and LSTM-CNN in their paper. The paper introduces three distinct deep learning techniques for identifying phishing websites, including Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) for comparison, and ultimately proposes an LSTM-CNN based approach. Experimental findings indicate that accuracies of 99.2%, 97.6%, and 96.8% were achieved for the CNN, LSTM-CNN, and LSTM models, respectively. It is observed that the proposed phishing detection method demonstrated by the CNN-based system is superior to other methods.

[16] in 2024 presented an effective phishing detection model that combines optimized artificial deep features and automated features. This study proposes Phishing Detection Based on Hybrid Features (PDHF), which is a new phishing detection model based on a combination of optimized artificial deep learning features. The optimized artificial phishing features are obtained by eliminating redundant features based on a new feature importance evaluation index, and an improved bidirectional search algorithm is utilized. To enhance the effective phishing detection time, deep features from URLs are learned using a one-dimensional character Convolutional Neural Network (CNN) and a quantized irregular attention mechanism. Experimental results show that PDHF outperforms many advanced methods, achieving an accuracy of 0.9965, precision of 0.9942, recall of 0.9940, and an F1 score of 0.9941.

[17] in 2024 addressed the detection of phishing websites using machine learning techniques. This article considers two main objectives. The first is to identify the best classifier that can detect phishing among twenty-four different classifiers that represent six learning strategies. The second objective is to identify the best feature selection method for phishing website datasets. The results demonstrated the superiority of the Random Forest, Filtered Classifier, and J-48 classifiers in identifying phishing websites, using two relevant phishing datasets with various features and considering eight evaluation criteria. Furthermore, the Info Gain Attribute Eval method showed the best performance among the four feature selection methods considered. Specifically, the results indicated that the proposed method has an accuracy of 92.409 percent in detecting phishing.

Deep learning approaches such as convolutional networks and recurrent neural networks have addressed the issues related to feature selection. These approaches have automatic feature selection. Considering the time series nature of the data, the aim of this research is to present an RNN-based approach for classifying phishing data. Therefore, the main question of this research, which we seek to answer, is stated as follows

How can we enhance the accuracy and speed of detecting phishing sites by designing a machine learning-based hybrid approach?

The objectives of this research include the following:

- Proposing a new model for detecting phishing sites using machine learning methods
- Increasing the accuracy of phishing site detection using an RNN-based method
- Increasing the speed of phishing site detection using an RNN-based method
- Evaluating the impact of feature selection in the preprocessing stage to enhance the performance of the phishing detection system

Literature Review

Methodologies for Detecting Phishing Websites

As phishing is a social engineering issue, effective countermeasures have been developed for various aspects in terms of education, legal oversight, and technical approaches [18]. This survey focuses on technical strategies for identifying phishing websites.

The methods for detecting phishing websites have evolved and are divided into three categories: list-based methods, exploratory methods, and machine learning methods [19].

List-based approaches include whitelists and blacklists that are manually reported and verified by systems. A whitelist is a collection of valid URLs or domains. Clearly, a blacklist is a group of confirmed phishing websites. When a user reports and confirms a website as a phishing site, the URL is added to the blacklists, which can be used to prevent disruption for other users. Exploratory strategies identify a phishing webpage based on a set of features extracted from the textual content of the website and compare the features with those of legitimate sites.

The idea of this approach is that attackers often deceive users by mimicking well-known websites. Machine learning methods also depend on the features of the website; create a model to learn from a set of data with structured features, and then predict whether a new website is a phishing site or not. In the field of machine learning, identifying phishing websites is a classification problem.

List-Based Approaches

Jain and Gupta proposed an automatic update and whitelist-based approach in 2016 to protect against phishing attacks on the client side. Experimental results show that its accuracy is 86.02 percent and the false positive rate is less than 1.48 percent, indicating its reliability. The alert for phishing attacks is another advantage of this approach, providing quick access time that guarantees a real-time environment and products [20].

Exploratory Strategies

Tan and colleagues introduced a phishing detection approach called PhishWHO, which consists of three stages. First, it obtains identity keywords through a weighted URL token system and combines the N-gram model from the HTML of the page. Secondly, it places the keywords in major search engines to find the legitimate website and legitimate domain.

In the next step, it compares the legal domain and the targeted website domain to determine whether the target website is a phishing website or not [21].

Chiu and colleagues used a logo image from the website to verify the legitimacy of the website [22]. In this paper, the authors extracted a logo from web page images using some machine learning algorithms, and then searched the domain through Google search engine using a logo as a keyword. Therefore, some researchers also referred to this category as a search engine-based approach.

Machine Learning-Based Methods

Machine learning-based interactive measures have been proposed to handle dynamic phishing attacks with higher accuracy and lower false-positive rates compared to other methods [18]. As a result, the machine learning approach consists of six components: data collection, feature extraction, model training, model testing, and prediction. Figure 1 shows the flowchart of each component. Existing machine learning-based phishing website detection solutions optimize one or more components based on this flowchart to achieve better performance.

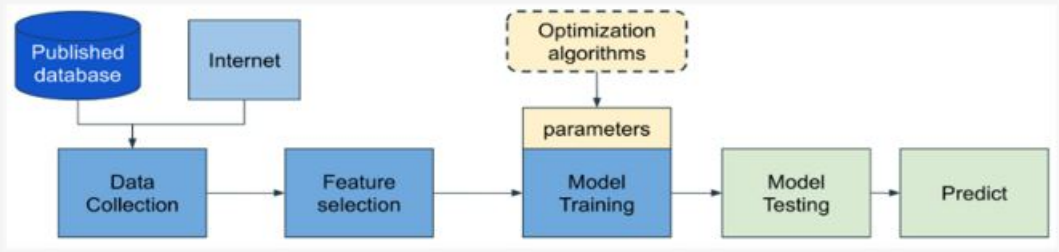

Figure 1: Flowchart of identifying phishing sites using machine learning models.

The Grey Wolf Optimization (GWO) Algorithm

The Grey Wolf Optimization algorithm simulates the hunting phases of wolves. The structure of hunting consists of three parts: chasing and encircling the prey, harassing the prey until it stops, and finally attacking the prey. Each wolf i serves as a solution to the problem in the search space with a position vector W_i = that indicates the n dimensions of the problem. The position of wolves is evaluated using a fitness function (adapted to the problem definition).

According to the values of the first best wolf with alpha (α), the second-best wolf with beta (β), and the third best wolf with delta (δ), it is shown. During the hunting process (optimization), the wolves update their positions based on the positions of the three wolves: alpha, beta, and delta. In the end, the algorithm returns the alpha wolf as the final solution [23].

Recurrent Neural Network (RNN)

A type of deep model that uses supervised learning methods, recurrent neural networks are distinct from other types of networks in terms of structure and training. The training of these networks occurs through recurrent loops. This means that the desired input data is trained in the network, and its information is sent back to the network for training subsequent data. In other words, these types of networks have a short-term memory that retains previous data and their associated information, and then processes and learns new data based on them.

For this reason, these networks are suitable for sequential and time-dependent data such as audio, video, text, etc. Data in which the current word or image is dependent on previous words or images [24].

Data Collection and Feature Extraction

Data is the source of every approach and proves to have a critical impact on performance. There are two methods for data collection: downloading published datasets and scraping URLs directly from the internet.

Research Methodology

The present research falls under the category of quantitative methods and is analytical in terms of its analysis method. The necessary information for conducting the research will be gathered through library research. Additionally, the research data will be collected using published articles in reputable scientific journals and conference proceedings as well as databases and computer networks, specifically from globally recognized scientific publishing platforms such as Springer, Elsevier, IEEE, and others. This paper will present a distributed solution based on artificial intelligence for phishing detection using the RNN model. The proposed method employs machine learning-based algorithms to identify the patterns of phishing attacks on the site. However, each data mining technique has its own specific advantages and disadvantages. Therefore, one cannot rely on a specific learning algorithm for detecting attacks. The use of a hybrid technique can enhance the accuracy of learning algorithms compared to situations where each of these algorithms is used individually. In the proposed method, we intend to provide a consolidated system for detecting phishing attacks using the RNN model. The model used in the proposed method will be deep RNN networks. In applying deep RNN networks to classification problems, we face two fundamental challenges: "determining the optimal topology of the deep network" and "determining the optimal weight vector." In the proposed method, an optimization algorithm will be used to solve these two issues simultaneously. With these explanations, the proposed method performs the detection of phishing attacks through the stages of preprocessing, feature extraction (based on the grey wolf algorithm), and classification (based on the RNN model).

In other words, the present research will be conducted in the form of the following steps:

Design

In this section, the implementation of the methods under investigation will be addressed through the equations and algorithms of each one. In this section, the flowchart for executing the method will be presented based on how they are implemented on the desired datasets.

Simulation

Preprocessing

In this section, the implementation of the hybrid gray wolf algorithm will be conducted to identify the effective features for detecting phishing sites.

Main Operations

After feature selection, we will simulate the RNN model-based method in MATLAB software and evaluate the results obtained from attack detection based on assessment criteria, and we will present the findings.

Evaluation of obtained results

In this section, after analyzing the results obtained from the proposed method, we will assess and compare its performance with other articles and methods.

The flowchart of the problem is as shown below.

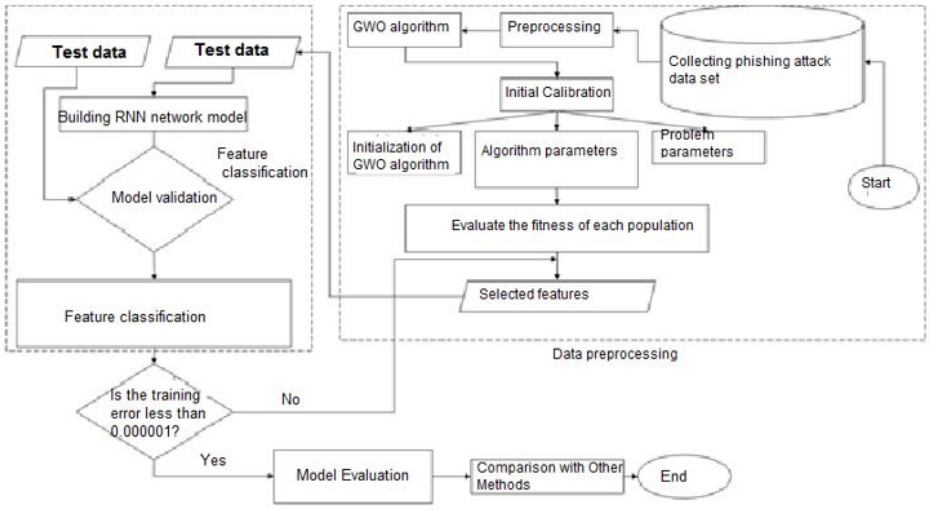

Figure 2: Problem flowchart

Implementaion

Section 1: Database

In this research, two databases related to phishing attacks are collected from the kaggle site. The specifications of these two databases are presented in Table 1.

Table 1

Dataset	Sample size	Access link
The first dataset	10,000 websites	https://www.kaggle.com/datasets/shashwatwork/phishing-dataset-for-machine-learning
The second dataset	11430 links related to types of phishing	https://www.kaggle.com/datasets/shashwatwork/web-page-phishing-detection-dataset

The first dataset contains 48 features extracted from 5000 phishing web pages and 5000 legitimate web pages that were downloaded from January to May 2015 and from May to June 2017.

The second dataset includes 11430 URLs with 87 extracted features. This dataset is designed to be used as a benchmark for machine learning-based phishing detection systems. The features of this dataset come from three different classes: 56 from the structure and syntax of the URLs, 24 from the content of their respective pages, and 7 are extracted through external service searches. The dataset is balanced and consists of exactly 50% phishing and 50% legitimate URLs.

Analysis

We present the results of this section in the form of two scenarios for both datasets. In the first scenario, the RNN algorithm is applied to all features of both datasets. In this scenario, the result obtained from the RNN for detecting phishing websites is compared with the neural network algorithms MLP and KNN without applying the gray wolf algorithm. In the second scenario, the gray wolf algorithm is applied first, and by determining the optimal features, the selected features are entered into the RNN, MLP, and KNN algorithms. Finally, the results of both scenarios are compared to determine the impact of applying the gray wolf algorithm on selecting optimal features.

As previously mentioned, the first dataset includes 48 features extracted from 5000 phishing web pages and 5000 legitimate web pages that were downloaded from January to May 2015 and from May to June 2017. However, the second dataset consists of 11430 URLs with 87 extracted features. This dataset is designed to be used as a benchmark for machine learning-based phishing detection systems. The features of this dataset are derived from three distinct classes: 56 related to the structure and syntax of the URLs, 24 from the content of their corresponding pages, and 7 through the analysis of external service searches. The dataset is balanced and consists of exactly 50% phishing and 50% legitimate URLs. Tables 4-1 and 4-2 present a portion of the first and second datasets.

First Scenario

In this scenario, two datasets undergo preprocessing, and all features aimed at identifying phishing sites are fed into Recurrent Neural Network (RNN) models, MLP neural networks, and KNN.

RNN Network

Before presenting the results, it is first necessary to provide the initial settings of the RNN algorithm.

Settings of the RNN Model

As mentioned earlier, this research uses the RNN network as the main algorithm. The initial values of the RNN parameters for classification typically include the following:

The number of hidden neurons: This is an important parameter that must be chosen carefully. The performance of the RNN can be significantly influenced by the number of neurons in the hidden layer.

- **Input weights**: These weights connect the input layer to the hidden layer and are initialized randomly.
- **Bias values**: The biases of the hidden neurons are also initialized randomly.

These parameters are set at the beginning and do not require repetitive adjustments, which is one of the reasons RNNs are well-known for their fast training times. Furthermore, in this study, the stochastic gradient descent (SGD) algorithm is used for training the RNN. The initial parameters of the RNN model are provided in Table 2. Additionally, the number of complete epochs is set to 40 during the training of the network.

Table 2: Initial Settings of the RNN Model

Parameter	The amount considered
Initial learning rate	0.00611
Learning rate decay factor	0.1
Number of iterations	50

Balance of hidden neurons	40
Weights of input layers	0.65
Bias value	0.5

Also, in order to determine the training function, different training functions were tested, the results of which are presented in Table 3.

Table 3: Comparison of correlation (regression) of different training functions of deep neural networks

Number	Function type	Function definition	Correlation coefficient
1	One-step secant	Trainoss	0.883
2	Conjugate gradient backpropagation with Powell-Beale restarts	traincgb	0.632
3	Conjugate gradient backpropagation with Polak-Ribiére updates	Traincgp	0.715
4	Scaled conjugate gradient	Trainscg	0.806
5	Bayesian regularization	Trainbr	0.936
6	BFGS quasi-Newton	Trainbfg	0.859
7	Levenberg-Marquardt	Trainlm	0.913
8	Gradient descent with adaptive learning rate	Traingda	0.678
9	Gradient descent with momentum	Traingdm	0.547
10	Gradient descent with momentum and adaptive learning rate	Traingdx	0.819

As can be seen in the table, the trainbr function is used to train the desired deep network.

In this section, the results will be presented. Accordingly, Figure 3 shows the deep architecture of the RNN layers. One note regarding the model's performance is that the essential operation of the neural network is based on training and testing data. On the other hand, since in the training system, some data are randomly selected for training and others for testing each time, the results obtained may vary slightly, but this amount of variation is negligible. It is also worth mentioning that the results obtained are based on the best results after running the neural network 15 times, which has a direct relationship with the choice of data and the system executed under it.

The convergence chart of the RNN model that led to the identification of phishing sites in the first dataset and in the second dataset can also be seen in Figures 4 and 5, respectively.

As can be seen, after 280 iterations, the RMSE obtained from solving the RNN model in the first dataset is 0.642, and for the second dataset, it is 0.8356.

Based on this, the performance of the RNN model for the first dataset can be observed using the evaluation metrics in the table, and for the second dataset in the table.

As can be seen in the tables, the RNN model achieved an accuracy of 94.83% and an MSE error rate of 0.4121 for predicting phishing sites in the first dataset, and an accuracy of 90.91% and an MSE error rate of 0.8356 for phishing sites in the second dataset.

Second Scenario

In the first step, the gray wolf algorithm is used and optimal features for both datasets are selected. Figures 6 and 7 show the convergence graphs of the GWO algorithm for the first and second datasets, respectively, for feature selection.

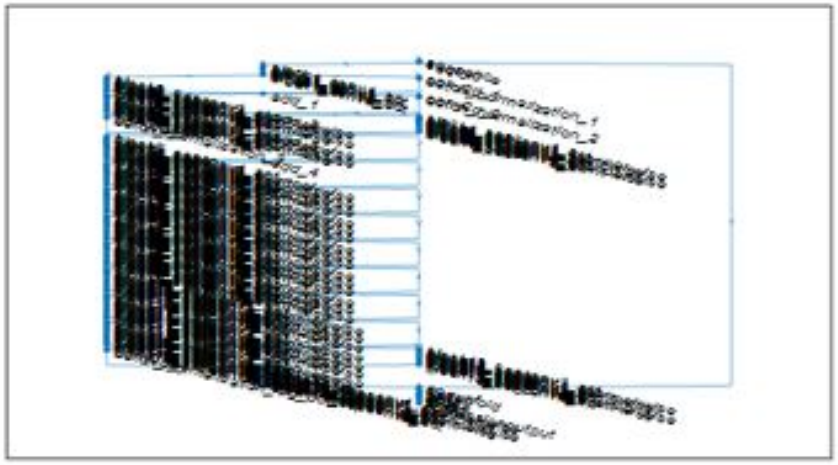

Figure 3: Architecture of Deep Layers of RNN

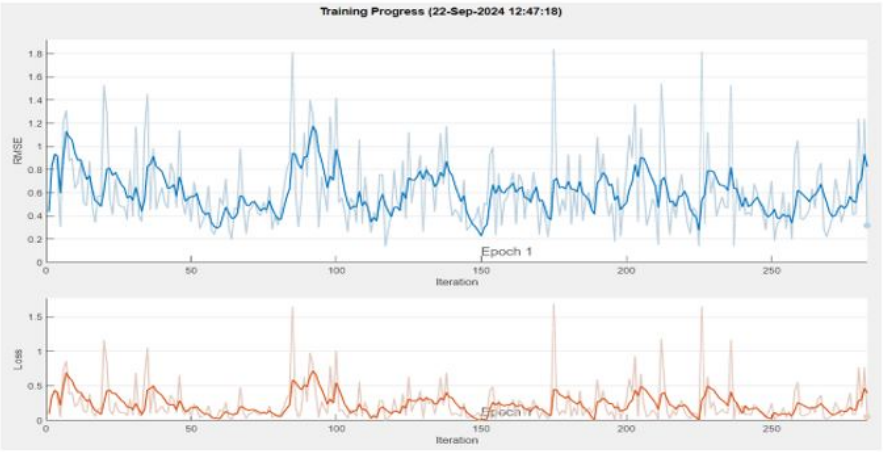

Figure 4: Convergence Chart of the RNN Model in Detecting Phishing Sites of the First Dataset

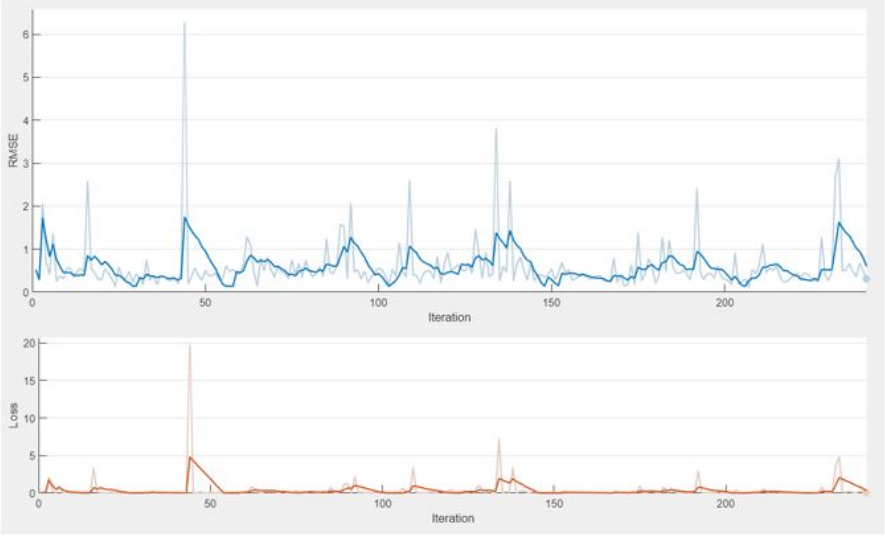

Figure 5: Convergence diagram of the RNN model in detecting phishing sites in the second dataset

Table 4: Evaluation of RNN algorithm performance in identifying phishing attacks for the first dataset in the first scenario

Criteria	Amount
Accuracy	0.9483
Precision	0.9722
Recall	0.9459
MSE error	0.4121
RMSE error	0.642

Table 5: Performance evaluation of the RNN algorithm in detecting phishing attacks for the second dataset in the first scenario.

Criteria	Amount
Accuracy	0.9091
Precision	0.9375
Recall	0.9091
MSE error	0.6982
RMSE error	0.8356

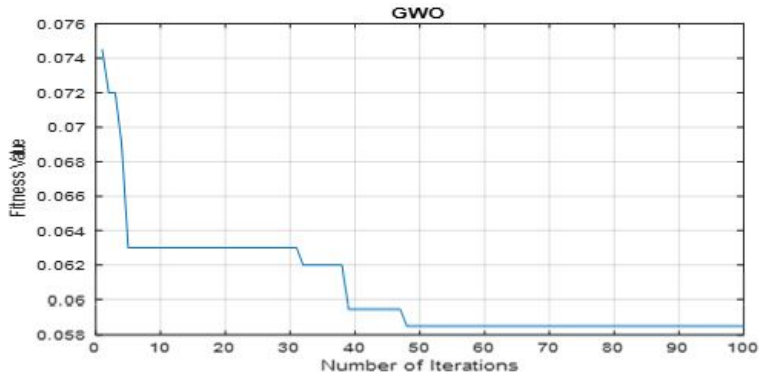

Figure 6: Convergence diagram of GWO in feature selection of the first dataset in the second scenario

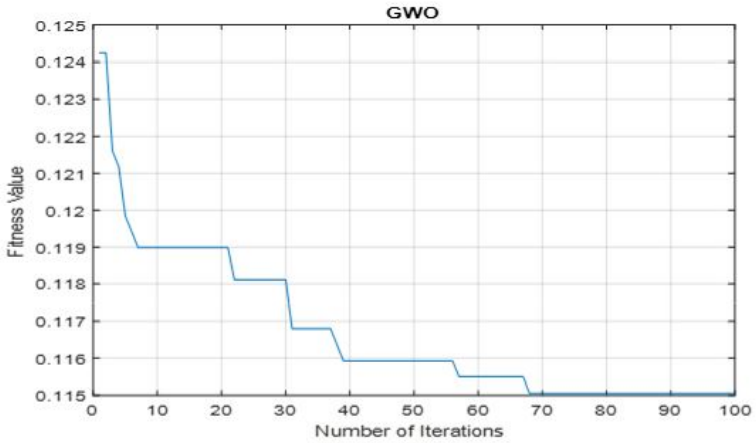

Figure 7: The convergence graph of GWO in feature selection of the second dataset in the second scenario

The result of applying the GWO optimization algorithm in selecting the optimal features for the first and second datasets is presented in Table 6.

Table 6: Result of applying the GWO optimization algorithm in feature selection of both datasets

Dataset Name	Number of Original Features	Number of Selected Features	Feature Selection Error with GWO (MSE)
First Dataset	48	37	0.0585
Second Dataset	87	65	0.1150

Accordingly, by having the selected features for both datasets, the results of all three models are presented for both datasets.

Table 7: Evaluation of the performance of the RNN algorithm in detecting phishing attacks for the first dataset in the second scenario

Criteria	Amount
Accuracy	0.9763
Precision	0.9787
Recall	0.9871
MSE error	0.0152
RMSE error	0.1233

Table 8: Evaluation of the performance of the RNN algorithm in detecting phishing attacks for the second dataset in the second scenario

Criteria	Amount
Accuracy	0.9397
Precision	0.9278
Recall	0.9709
MSE error	0.0524
RMSE error	0.2289

As can be seen in Tables 7 and 8, the RNN model was able to predict phishing sites in the first dataset with an accuracy of 97.63 percent and an MSE error rate of 0.0152, and phishing sites in the second dataset with an accuracy of 93.97 percent and an MSE error rate of 0.0524.

Table 9: Performance evaluation of the MLP algorithm in detecting phishing attacks for the first dataset in the second scenario

Criteria	Amount
Accuracy	0.95684
Precision	0.93237
Recall	0.95504
MSE error	0.0154

RMSE error	0.1242

Table 10: Performance evaluation of the MLP algorithm in detecting phishing attacks for the second dataset in the second scenario

Criteria	Amount
Accuracy	0.91238
Precision	0.91797
Recall	0.90737
MSE error	0.0282
RMSE error	0.1683

As seen in Tables 9 and 10, the MLP model has achieved an accuracy of 95.68% and an MSE error rate of 0.0154 for phishing sites in the first dataset, and an accuracy of 91.238% and an MSE error rate of 0.0282 for phishing sites in the second dataset.

Table 11: Evaluation of the KNN algorithm performance in identifying phishing attacks in the first dataset in the second scenario

Criteria	Amount
Accuracy	0.88583
Precision	0.89113
Recall	0.88108
MSE error	0.1934
RMSE error	0.4397

Table 12: Evaluation of KNN Algorithm Performance in Identifying Phishing Attacks in the Second Dataset in the Second Scenario

Criteria	Amount
Accuracy	0.88758
Precision	0.87545
Recall	0.89224
MSE error	0.2875
RMSE error	0.5361

As seen in Tables 11 and 12, the KNN model has achieved an accuracy of 88.58% and an MSE error rate of 0.1934 in predicting phishing sites in the first dataset, and an accuracy of 88.75% with an MSE error rate of 0.5361 in predicting phishing sites in the second dataset.

Result

In this section, we first compare the results of the models presented in this study. Table 13 shows the comparison of the accuracy of all models in the two scenarios considered.

Table 13: Comparison of accuracy of research models

Model	Dataset	Accuracy (%)
RNN	First	94.83
	Second	90.91
MLP	First	90.74
	Second	88.99
KNN	First	86.65
	Second	84.208
GWO+RNN	First	97.63
	Second	93.97
GWO+MLP	First	95.68
	Second	91.238
GWO+KNN	First	88.58
	Second	88.75

The results can also be viewed as a bar chart presented in Figure 8.

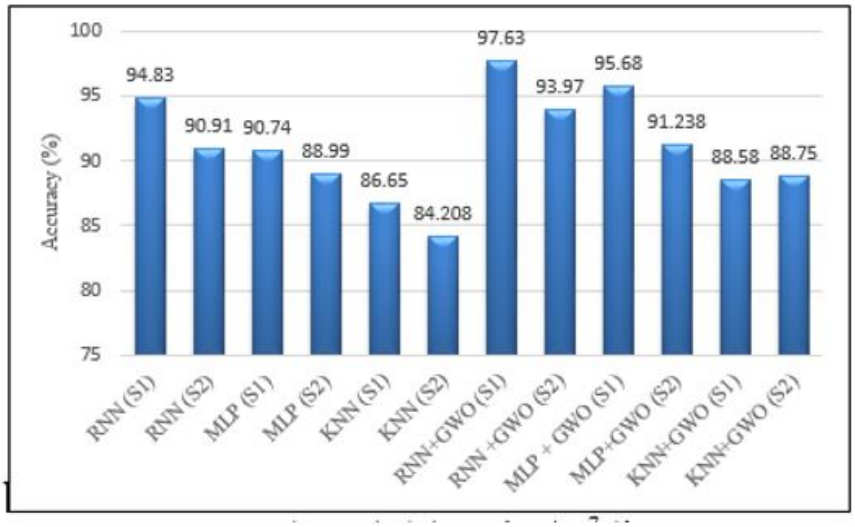

Figure 8: Comparison of the accuracy of the methods used in this research

As shown in the results of Table 13 and Figure 8, the use of the gray wolf algorithm has led to an increase in the accuracy of all models in both datasets. Additionally, among all models, the RNN model has the highest efficiency.

Now, for comparison with other articles, we will use the articles by [25,26]. In the article by [25], a large dataset of 20,000 website URLs was used, and 22 prominent features from each URL were extracted to prepare a comprehensive dataset. Along with this, another dataset containing website text was also prepared for evaluating text based on NLP. For evaluation, Support Vector Machine (SVM) models, XGBoost, random forest, multilayer perceptron, linear regression, and decision tree were used. In the article by [26], phishing attacks are also identified based on the text of suspicious web pages and not based on URLs, using natural language processing (NLP) and deep learning (DL) algorithms. In this article, Long Short-Term Memory (LSTM) models, Bidirectional LSTM (BiLSTM), Gated Recurrent Unit (GRU), and Bidirectional GRU (BiGRU) are utilized. In Table 14, a comparison between these models is presented.

Table 14

Model	Reference	Accuracy (%)
SVM	[25]	88.5
XGBoost		91.2
RF		90.6
MLP		91.2
LR		87.1
DT		90.4
LSTM	[26]	96.71
BiLSTM		97.20
GRU		97.29
BiGRU		97.39
RNN+GWO	Current research	97.63

As the table above shows, deep learning models have high efficiency in detecting phishing sites. However, the proposed model has demonstrated higher performance compared to the deep learning models presented in the study by [26].

Conclusion

Phishing website detection is a critical area of cybersecurity aimed at protecting users from malicious websites that imitate legitimate sites to steal sensitive information. Given the importance of this issue, this research employed a hybrid model based on the Grey Wolf Optimization (GWO) algorithm and the RNN deep neural network for phishing site detection using two datasets. The GWO algorithm is inspired by the social hierarchy and hunting behavior of grey wolves, which aids in optimizing the feature selection process. By selecting the most relevant features, GWO enhances the efficiency and accuracy of RNN in identifying phishing websites. The Recurrent Neural Network (RNN) is a type of deep learning model particularly effective for analyzing sequential data. In the context of phishing detection, RNNs can analyze patterns in URLs, HTML content, and other website features over time. This temporal analysis capability allows RNNs to identify subtle indicators that may signify phishing attempts. When combined with GWO, the performance of the RNN is further improved, as the optimized features lead to more accurate predictions.

In the proposed hybrid model, it first works to reduce the dimensionality of the input data using GWO and selects only the most important features for phishing detection. These features are then fed into the RNN, which processes the data through its layers to classify websites as legitimate or phishing. This two-step approach not only speeds up the identification process but also enhances the overall accuracy and robustness of the model against various types of phishing attacks. This was proven in the application of the model on two datasets. Specifically, we evaluated the research results in the form of two scenarios. In the first scenario, all features of the two datasets were used, and the performance of the RNN was compared with two models, MLP and KNN. The results showed that the RNN model had an accuracy of 94.83% on the first dataset and 90.91% on the second dataset. In contrast, the MLP had an accuracy of 90.74% and 88.99%, while the KNN had an accuracy of 86.65% and 84.208% respectively on the first and second datasets. In the second scenario, the GWO algorithm was used to select optimal features. The results indicated that this algorithm selected only 37 features from the 48 features of the first dataset with an error of 0.0585 and only 65 features from the 87 features of the second dataset with an error of 0.1150. These features were used as input features for training and testing all three models. The results showed that with the selection of optimal features by GWO, the detection accuracy of the RNN increased by 2.8% to 97.63% for the first dataset and by 3.06% to 93.97% for the second dataset. The MLP and KNN models also experienced an increase in accuracy. Subsequently, we compared the results with previous articles. For this purpose, we used two arti-

cles by [25,26], which employed various machine learning and deep learning models. The results showed that deep learning models have high efficiency in detecting phishing sites.

The results specifically showed that the BiGRU model used by [26] has an accuracy of 97.39 percent, which is the highest accuracy among the mentioned models. However, the proposed model of this research has reached an accuracy of 97.63 percent, which is an increase compared to the BiGRU.

It is suggested that the proposed algorithm in this research be used in plugins and software for fraud detection to enhance the accuracy of identification.

References

1. Suzuki YE, SAS Monroy (2022) "Prevention and mitigation measures against phishing emails: a sequential schema model." Security Journal, 35: 1162-82.

2. Catal C, Giray G, Tekinerdogan B, Kumar S, Shukla S (2022) Applications of deep learning for phishing detection: a systematic literature review. Knowledge and Information Systems, 64: 1457-500.

3. Basit A, et al. (2021) "A comprehensive survey of AI-enabled phishing attacks detection techniques." Telecommunication Systems, 76: 139-54.

4. Xia P, Wang H, Zhang B, Ji R, Gao B, Wu L, Xu G (2020) Characterizing cryptocurrency exchange scams. Computers & Security, 98: 101993.

5. Jain AK, B Gupta (2022) "A survey of phishing attack techniques, defence mechanisms and open research challenges." Enterprise Information Systems, 16: 527-65.

6. Jain AK, Gupta BB (2019) A machine learning based approach for phishing detection using hyperlinks information. Journal of Ambient Intelligence and Humanized Computing, 10: 2015-28.

7. El-Rashidy MA (2021) "A smart model for web phishing detection based on new proposed feature selection technique." Menoufia Journal of Electronic Engineering Research, 30: 97-104.

8. Gandotra E, Gupta D (2021) An efficient approach for phishing detection using machine learning. Multimedia Security: Algorithm Development, Analysis and Applications, 239-53.

9. Ahammad SH, Kale SD, Upadhye GD, Pande SD, Babu EV, et al. (2022) Phishing URL detection using machine learning methods. Advances in Engineering Software, 173: 103288.

10. Mohamed G, Visumathi J, Mahdal M, Anand J, Elangovan M (2022) An effective and secure mechanism for phishing attacks using a machine learning approach. Processes, 10: 1356.

11. Orunsolu AA, Sodiya AS, Akinwale AT (2022) A predictive model for phishing detection. Journal of King Saud University-Computer and Information Sciences, 34: 232-47.

12. Mughaid A, AlZu'bi S, Hnaif A, Taamneh S, Alnajjar A, Elsoud EA (2022) An intelligent cyber security phishing detection system using deep learning techniques. Cluster Computing, 25: 3819-28.

13. Adebowale MA, Lwin KT, Hossain MA (2023) Intelligent phishing detection scheme using deep learning algorithms. Journal of Enterprise Information Management, 36: 747-66.

14. Alnemari S, Alshammari M (2023) Detecting phishing domains using machine learning. Applied Sciences, 13: 4649.

15. Alshingiti Z, Alaqel R, Al-Muhtadi J, Haq QEU, Saleem K, Faheem MH (2023) A Deep Learning-Based Phishing Detection System Using CNN, LSTM, and LSTM-CNN. Electronics, 12: 232.

16. Zhu E, Cheng K, Zhang Z, Wang H (2024) PDHF: Effective phishing detection model combining optimal artificial and automatic deep features. Computers & Security, 136: 103561.

17. Alazaidah R, Al-Shaikh A, AL-Mousa MR, Khafajah H, Samara G, Alzyoud M, Almatarneh S (2024) M. Website Phishing Detection Using Machine Learning Techniques. Journal of Statistics Applications & Probability, 13: 119-29.

18. Alsariera YA, et al. (2020) "Ai meta-learners and extra-trees algorithm for the detection of phishing websites." IEEE Access, 8: 142532-42.

19. Zabihimayvan M, D Doran (2019) Fuzzy rough set feature

selection to enhance phishing attack detection. 2019 IEEE international conference on fuzzy systems (FUZZ-IEEE), IEEE.

20. Jain AK, BB Gupta (2016) "A novel approach to protect against phishing attacks at client side using auto-updated white-list." EURASIP Journal on Information Security, 2016: 1-11.

21. Tan CL, et al. (2016) "PhishWHO: Phishing webpage detection via identity keywords extraction and target domain name finder." Decision support systems, 88: 18-27.

22. Chiew KL, et al. (2015) "Utilisation of website logo for phishing detection." Computers & Security, 54: 16-26.

23. Mirjalili S, Mirjalili SM, Lewis A (2014) Grey wolf optimizer. Advances in engineering software, 69: 46-61.

24. Sherstinsky A (2020) Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D: Nonlinear Phenomena, 404: 132306.

25. Shaukat MW, Amin R, Muslam MMA, Alshehri AH, Xie J (2023) A hybrid approach for alluring ads phishing attack detection using machine learning. Sensors, 23: 8070.

26. Benavides-Astudillo E, Fuertes W, Sanchez-Gordon S, Nuñez-Agurto D, Rodríguez-Galán G (2023) A phishing-attack-detection model using natural language processing and deep learning. Applied Sciences, 13: 5275.