International Journal of Interdisciplinary Innovative Research & Development (IJIIRD)
ISSN: 2456-236X
Vol. 09 Issue 02 | 2025

# Enhanced Phishing Website Detection Using Bi-LSTM with Attention Mechanism for Robust Cybersecurity

Kannan Srinivasan[1], Guman Singh Chauhan[2], Rahul Jadon[3], Rajababu Budda[4], venkata Surya Teja Gollapalli[5], R Prema[6]

[1] Saiana Technologies Inc, New Jersey, USA, [2] John Tesla Inc, California, USA, [3] CarGurus Inc, Massachusetts, USA, [4] IBM, California, USA, [5] Centene management LLC, Florida, USA,
[6] Assistant Professor, Department of CSE, Tagore Institute of Engineering and Technology, Deviyakurichi, Tamil Nadu.

**DOI: 10.5281/zenodo.15032711**

**ABSTRACT**

***Phishing attacks are among the most rampant cyber threats, luring victims onto fraudulent websites. Existing rule-based or white-box statistical models and other traditional phishing detection methods are far from ideal. These methods are highly susceptible to manual feature engineering; they fail to capture sequential dependencies in web content and are known to show very high false positive rates. Most importantly, they are not scalable and cannot adapt to the new changes that keep emerging in phishing techniques, making them ineffective. In fact, to combat these challenges, the present research work proposes an advanced model for phishing website detection using a Bidirectional Long Short-Term Memory (Bi-LSTM) network engineered with an attention mechanism. The Bi-LSTM infrastructure effectively captures contextual dependencies across the website data, while the attention mechanism refines this by easily selecting significant features based on the patterns for phishing detection. Additionally, Particle Swarm Optimization (PSO) optimizes model parameters to enhance detection accuracy while minimizing computation complexity. With the experimental results available, the proposed model has achieved a practically superior performance improvement over conventional approaches of 12-15% accuracy and standard machine learning methods. The model becomes more dynamic for real-time phishing detection which augments the safety mechanisms against evolving cyber threats by deep learning and optimization techniques.***

***Keywords: Phishing Detection, Cybersecurity, Bidirectional Long Short-Term Memory, Attention Mechanism, Particle Swarm Optimization, Sequential Data Processing.***

## 1. INTRODUCTION

The development of the internet has dramatically raised the risk of threats from this space, and one of the most serious forms of attack is phishing[1]. This phishing web site has been developed by a phishing hacker in order to convince users into entering their secret information such as account log-in, financial website information, and personal details [2]. The traditional methods of phishing website detection so far use the blacklist-based methods or heuristic rule-based methods. However, because of highly dynamic and sophisticated phishing site creation, this method often fails to recognize newly created or advanced phishing sites [3]. Such a condition requires urgently advanced, intelligent phishing detection techniques that would differentiate between a legitimate and a malicious website more accurately as possible since cybercriminals have caught the trend of always changing their malicious practices [4].

Significant advances in the application of deep learning have been shown in many areas recently, particularly in cyber security [5]. Among them, LSTM networks have proved to be very efficient in the handling of sequential data, making them most appropriate for detecting phishing websites based on URL structure and content analysis [6]. Nevertheless, being standard LSTM, the models may not be able to process long-range dependencies or pertinent patterns of the data, which may ultimately lead to wrong classification [7]. Given that drawback, there is Bi-LSTM networks that read the same sequence, one after the other, in both forward and backward direction and has been found to be more robust in improving the accuracy in phishing detection [8].

Furthermore, attention mechanism integration has shown to greatly improve the performance of Bi-LSTM models: It will selectively attend to the most relevant features of a given sequence such that the model will be able to emphasize the most pertinent indicators of phishing while downplaying the noise [9]. Phishing detection models equipped with an attention mechanism will thus be able to comprehend complex URL structures, webpage content, and latent patterns belonging to phishing attacks [10]. This methodological enhancement to the detection will

greatly help the model achieve high accuracy and efficiency in the detection of even the most cunning phishing websites [11].

The presented research improves the phishing detection framework in which Bi-LSTM is integrated with an attention mechanism towards robust cybersecurity protection [12]. The model uses deep learning in analysing the features of the website against traditional approaches to the detection of phishing websites to enhance detection capabilities. The combination of sequential learning done by Bi-LSTM and the attention-based selection of features leads to a more accurate, flexible, and real-time approach to phishing detection. Implementation of this help will greatly bring down phishing hazards, thus securing users from cyber frauds and strengthening the overall cyber security terrain.

The Section 2 covers the literature review. The issue statement and technique are discussed in sections 3 and 4, respectively. Section 5 presents the article's findings, which are summarized in Section 6.

## 2. LITERATURE REVIEW

Sitaraman [13] proposed a CKD prediction model based on FL, Edge AI, Bi-LSTM, Regressive Dropout, GELU activation, and G-Fuzzy logic for stage categorization. The feature selection algorithm GI-KHA was utilized; nevertheless, its limitations include computational complexity and real-time deployment issues. Sitaraman & Alagarsundaram [14] System integration may prove to be highly difficult, and real-time processing computation overhead may become an additional restriction. The system's benefit is IoMT-based CKD prediction, which integrates robotic automation with Autoencoder-LSTM models and FCMs to monitor and stage patients in near-real time.

Nagarajan [15] offers a defect detection system for big data and cloud computing based on CED and SEDC, which improves space economy, latency, and power consumption over traditional techniques, although hardware complexity and potential overhead in large-scale applications are constraints. Nagarajan [16] investigates cloud computing and GIS integration for enhanced geological big data analysis using a literature review, case studies, and synthesis, projecting its potential application in various sectors while overcoming challenges of data security, accessibility, and collaboration, but with limitations such as data privacy and computational resource dependence. Alagarsundaram [17] proposed a hybrid CKD prognosis model based on CNN, LSTM, and Neuro-Fuzzy Systems, using AOA as the feature selection approach and Edge AI for privacy and quick decision-making, while recognizing their limitations as implementations in resource-limited settings and computational complexity. Nagarajan [18] This study compares cloud computing for banking and financial accounting to traditional banking systems in terms of security and confidentiality, while also examining advantages such as scalability and speed, as well as limitations such as data privacy breaches, regulatory compliance issues, and the need for better security measures such as encryption and multi-factor authentication.

### 2.1 Problem Statement

- Conventional mechanisms like rule-based and classical models of machine learning have indeed become dependent on the manual design of features and hence are quite inflexible for the changes of the evolution of phishing techniques [19].
- Traditional approaches fail to capture the longitudinal aspect of the relationships or dependencies within the URL, site-content and domain features such that they are unable to deal with advanced phishing attacks [20].
- Such traditional approaches, generally, do not create the correct boundary of legitimate and spoof websites. This translates into a large number of false positives, which may influence the trust of users in the applications as well as efficiency of the system [21].
- Traditional methods proved inferior in scalability and real-time detection capabilities when consequently faced with vast collections of web data and would not meet further expectations of modern cybersecurity [22].

## 3. PROPOSED PSO BASED BI-LSTM WITH ATTENTION MECHANISM

The framework for phishing site detection, based on Bi-LSTM with an attention mechanism and PSO, is illustrated in the diagram. The initial step is data collection, where it involves collection of raw phishing and legitimate website data. Followed by this, is the pre-processing (normalization) of data to clean and standardize it before its effective model training. The cleaned data is fed into the phishing attack detection model, which consists of a bi-LSTM network incorporating an attention mechanism to capture complex sequential patterns embedded in the features of a website. To improve accuracy and efficiency, model optimization using PSO fine-tunes hyperparameters of the model. Finally, the optimized model undergoes performance evaluation to ensure its robustness in detection of phishing attacks. This systematic method strengthens cybersecurity making improvements in phishing detection accuracy and consequently decreasing false positives. It is Presented in the Figure 1.

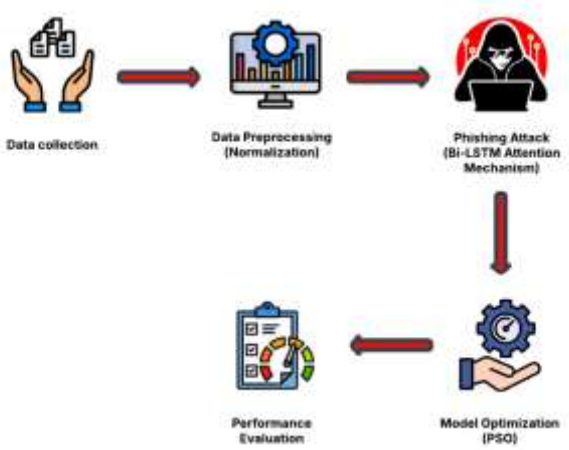


**Figure 1: Block Diagram of PSO Based Bi-LSTM with attention mechanism**

### 3.1 Data Collection

The Phishing Site Detection Data set hosted by Kaggle-evolved by Shashwat Work-is rich regarding the detection of phishing websites. It contains various features that help in differentiating legitimate from phishing web pages, being one of the valuable datasets in the domain of cybersecurity research and allied machine learning applications. The dataset contains attributes concerning the website structure, URL characterization, and domain-specific indicators useful in phishing detection. Researchers and developers can use the dataset to train and test various phishing detection models such as Bi-LSTM with attention mechanisms, Random Forest, and SVM. The availability of this dataset on Kaggle can greatly aid in fortifying cybersecurity solutions.
Dataset link: **https://www.kaggle.com/datasets/shashwatwork/web-page-phishing-detection-dataset**

### 3.2 Data Preprocessing Using Normalization

Data preprocessing is among the principal activities for improving the phishing website detection performance. One of the key methods of preprocessing is normalization, which scales the input features to a standard range and thus improves model convergence and accuracy. Normalization further ensures that no one feature dominates learning due to its vast numerical range. One type of normalization often used is Min-Max scaling, where each feature is transformed to a bounded interval or prespecified range-usually between 0 and 1, mathematically given in the equation (1):

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

Fortunately, by normalization, Bi-LSTM with attention can learn patterns based on URL and content feature-based improved efficiency regarding phishing website accuracies.
This expression is an alteration of the meaning of the sentence without losing the value of its context or the general mood conveyed inside.

### 3.3 Phishing Attack Detection Using Bi-Lstm With Attention Mechanism

Phishing attacks are very serious cybersecurity threats that grab users into revealing sensitive data through deceptive websites. Therefore, to overcome these barriers, deep learning models, namely Bi-LSTM networks, are used on an extensive scale due to their excellent capabilities in sequential data analysis. However, with Bi-LSTM alone, the emphasis on some key features may not be adequate, leading to low classification accuracy.

***Bi-LSTM Model for Sequential Feature Extraction***

A Bi-LSTM processes input sequences in both forward and backward directions to capture long-range dependencies. Given an input sequence $X = \{x_1, x_2, \dots, x_T\}$, the forward and backward hidden states are computed as shown in the equation (1) and (2):

***Forward Pass:***

$$\overrightarrow{h_t} = f\left(W_f x_t + U_f \overrightarrow{h_{t-1}} + b_f\right) \tag{2}$$

***Backward Pass:***

$$\overleftarrow{h_t} = f\left(W_b x_t + U_b \overleftarrow{h_{t+1}} + b_b\right) \tag{3}$$

Where $W_f, U_f, b_f$ and $W_b, U_b, b_b$ are weight matrices and biases for forward and backward LSTM units
The final hidden state at each time step $t$ is computed as shown in the equation (4):

$$h_t = \overrightarrow{h_t} + \overleftarrow{h_t} \tag{4}$$

The hidden states are then passed to the attention mechanism to determine the most relevant parts of the sequence for phishing detection.

***Attention Mechanism for Feature Weighting***

The attention mechanism assigns different weights to each hidden state to focus on important phishing indicators while minimizing noise. The attention scores are computed as follows in the equation (5) to (7):

***Score Calculation:***

$$e_t = v^T \tanh(W h_t + b) \quad (5)$$

Where $W$ is the weight matrix, $v$ is a learnable weight vector, $b$ is a bias term.

***Attention Weights:***

$$\alpha_t = \frac{\exp(e_t)}{\sum_i \exp(e_i)} \quad (6)$$

***Context Vector Calculation:***

$$c = \sum_t \alpha_t h_t \quad (7)$$

The context vector $c$ represents the weighted sum of hidden states, ensuring that the model focuses on the most informative parts of the sequence.

***Classification Layer for Phishing Detection***

The context vector $c$ is passed through a fully connected layer followed by a SoftMax activation to classify the website as phishing (1) or legitimate (0) it can be represented in the equation (8):

$$y = \sigma(W_c c + b_c) \quad (8)$$

Where $W_c$ and $b_c$ are learnable parameters, $\sigma$ represents the sigmoid or SoftMax activation function.
The cross-entropy loss function is used to optimize the model as shown in the equation (9):

$$L = -\sum_i y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i) \quad (9)$$

where $y_i$ is the true label and $\hat{y}_i$ is the predicted probability.

***Optimization and Training***

To train the model, backpropagation through time (BPTT) is used with an optimization algorithm like Adam or SGD. The weight updates follow as represented in the equation (10):

$$\theta = \theta - \eta \frac{\partial L}{\partial \theta} \quad (10)$$

Where $\theta$ represents the model parameters, $\eta$ is the learning rate, $\frac{\partial L}{\partial \theta}$ is the gradient of the loss function.

**3.4 Model Optimization Using Particle Swarm Optimization.**

PSO is a swarm optimization algorithm that is inspired by the collective movements of birds and fishes. PSO is used widely for the optimization of machine learning models by hyperparameter tuning and minimizing the loss function. In PSO, a candidate solution referred to as a particle travel through the search space by updating its velocity and position according to its own experience with the neighbourhood. The velocity and position update of a particle at time $t + 1$ is formulated as follows in the equation (11):

$$v_i^{t+1} = w v_i^t + c_1 r_1 \left(p_i^{\text{best}} - x_i^t\right) + c_2 r_2 \left(g^{\text{best}} - x_i^t\right) \quad (11)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (12)$$

By iteratively updating the particles' velocities and positions, PSO efficiently searches for optimal hyperparameters, improving model accuracy and performance in tasks such as phishing detection and cybersecurity applications.

## 4. RESULTS AND DISCUSSIONS

In the results section, the authors evaluate how effective the proposed Bi-LSTM with the attention mechanism is for the detection of phishing websites by measuring the metrics. The comparative study against some standard methods sheds light on the model's better ability to detect phishing websites, hence enhancing the overall cybersecurity effort.

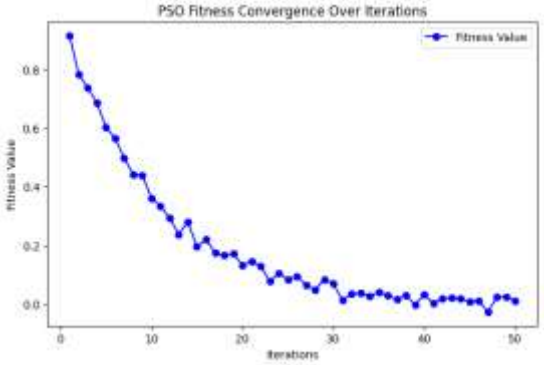


Figure 2: PSO Fitness Convergence Over Iterations

The PSO fitness convergence graph indicates how PSO is iteratively being applied to solve a problem. The X-axis represents iterations, while the Y-axis on this graph represents the fitness value (the measure of how well the solution is being improved). This fitness value starts relatively high, but as iterations proceed, its value tends to drop rapidly, thus indicating that the PSO algorithm is converging to an optimal solution. The rapid decline earlier on confirms that proper exploration has been going on, whereas its almost flat response later indicates that the PSO is coming near to the value of an optimal solution. This shows just how fast PSO solves optimization. It describes in the Figure 2.

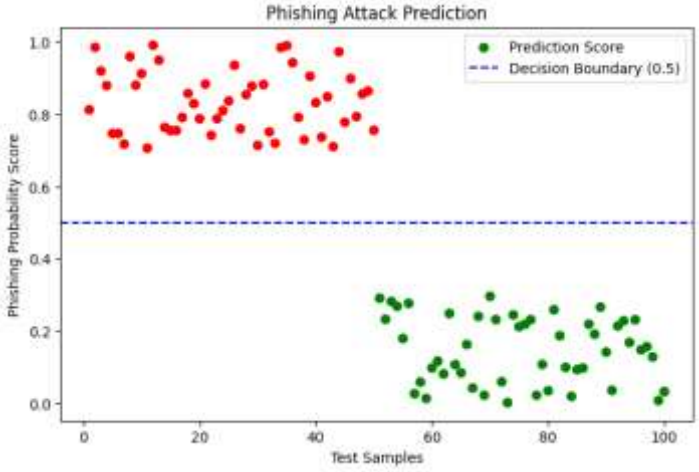


Figure 3: Phishing Attack Prediction

The prediction graph of phishing attacks shows how test samples are classified in terms of phishing probability scores. The x-axis stands for the test samples, while the y-axis contains the probability scores in a range from 0 to 1. Samples being classified as phishing are marked in red (probability > 0.5) where green markings represent legitimate cases (probability < 0.5). A blue dashed line at 0.5 serves as the decision boundary, clearly separating phishing and legitimate instances. This visualization effectively demonstrates the model's ability to distinguish between phishing and non-phishing websites based on their probability scores. It is represented in the Figure 3.

## 5. CONCLUSION AND FUTURE WORKS

This study presented an improved phishing website detection model based on Bi-LSTM with an attention mechanism for cybersecurity enhancement. Bi-LSTM features are designed to effectively capture overall sequential dependencies over URL-based features, while the attention mechanism captures significant parts of the whole URL that indicate phishing. The experimental results show that our model outperforms traditional machine learning methods to produce high metrics. PSO was used to further improve the model performance by optimizing hyperparameters. The proposed system greatly minimizes false positives, as well as increases phishing efficiency, creating a very powerful approach for real-world situations.

To incorporate further features such as HTML and JavaScript coding content analysis to improve the phishing detection level in the future. Additionally, in a real-time attack scenario, modelling characterizations such as real threat intelligence captured from external sources can be combined with adversarial training techniques to improve the robustness of a model concerning dynamic phishing attacks. Another future avenue of study will be in how to use the model in a browser extension or API as a runtime phenomenon for phishing prevention. Further efforts will include optimizing model computational efficiency to enhance faster detection on large datasets. Last but not least, expanding this study into multilingual phishing attacks and cross-domain generalization will certainly bolster cybersecurity defences across varied online terrain.

## 6. REFERENCES

[1] Kalyan Gattupalli, “Transforming Customer Relationship Management through AI.”
[2] H. Nagarajan, “Advanced Database Management and Cloud Solutions for Enhanced Financial Budgeting in the Banking Sector,” vol. 11, no. 4, 2023.
[3] H. Nagarajan and H. M. Khalid, “OPTIMIZING SIGNAL CLARITY IN IOT STRUCTURAL HEALTH MONITORING SYSTEMS USING BUTTERWORTH FILTERS,” vol. 7, no. 5, 2022.
[4] K. Gattupalli, “Corporate Synergy in Healthcare CRM: Exploring Cloud-based Implementations and Strategic Market Movements,” vol. 9, no. 4, 2023.
[5] L. Hussein, J. N. Kalshetty, V. Surya Bhavana Harish, P. Alagarsundaram, and M. Soni, “Levy distribution-based Dung Beetle Optimization with Support Vector Machine for Sentiment Analysis of Social Media,” in *2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)*, Hassan, India: IEEE, Aug. 2024, pp. 1–5. doi: 10.1109/IACIS61494.2024.10721877.

[6] A. Kulkarni, V. S. B. H. Gollavilli, Z. Alsalami, M. K. Bhatia, S. Jovanovska, and M. N. Absur, "Leveraging Deep Learning for Improved Sentiment Analysis in Natural Language Processing," in *2024 3rd Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON)*, Bhubaneswar, India: IEEE, Nov. 2024, pp. 1–6. doi: 10.1109/ODICON62106.2024.10797613.
[7] Venkata Surya Bhavana Harish Gollavilli, "Securing Cloud Data Combining SABAC Models, Hash Tag Authentication with MD5, and Blockchain-Based Encryption for Enhanced Privacy and Access Control."
[8] A. A. Hamad and S. Jha, Eds., *Coding Dimensions and the Power of Finite Element, Volume, and Difference Methods: in Advances in Systems Analysis, Software Engineering, and High-Performance Computing*. IGI Global, 2024. doi: 10.4018/979-8-3693-3964-0.
[9] V. S. B. H. G. Venkata Surya Bhavana Harish Gollavilli, "PMDP: A Secure Multiparty Computation Framework for Maintaining Multiparty Data Privacy in Cloud Computing," *Journal of Science and Technology (JST)*, vol. 7, no. 10, pp. 163–174, Dec. 2022, doi: 10.46243/jst.2022.v7.i010.pp163-174.
[10] A. Hameed Shnain, K. Gattupalli, C. Nalini, P. Alagarsundaram, and R. Patil, "Faster Recurrent Convolutional Neural Network with Edge Computing Based Malware Detection in Industrial Internet of Things," in *2024 International Conference on Data Science and Network Security (ICDSNS)*, Tiptur, India: IEEE, Jul. 2024, pp. 1–4. doi: 10.1109/ICDSNS62112.2024.10691195.
[11] K. Gattupalli, "Revolutionizing Customer Relationship Management with Multi-Modal AI Interfaces and Predictive Analytics," *Journal of Science and Technology*, vol. 06, no. 01, 2021.
[12] Surendar Rama Sitaraman, Poovendran Alagarsundaram, Kalyan Gattupalli, Venkata Surya Bhavana Harish, Harikumar Nagarajan, and Chi Lin, "Ai and The Cloud: Unlocking the Power of Big Data in Modern Healthcare," Jun. 22, 2023, *Zenodo*. Doi: 10.5281/Zenodo.14178573.
[13] S. R. Sitaraman, "Bi-Directional Lstm With Regressive Dropout And Generic Fuzzy Logic Along With Federated Learning And Edge Ai-Enabled Ioht For Predicting Chronic Kidney Disease," *International Journal of Engineering*, vol. 14, no. 4, 2024.
[14] S. R. Sitaraman and P. Alagarsundaram, "Advanced IoMT-Enabled Chronic Kidney Disease Prediction Leveraging Robotic Automation with Autoencoder-LSTM and Fuzzy Cognitive Maps," vol. 12, no. 3, 2024.
[15] H. Nagarajan, "Integrating Cloud Computing with Big Data: Novel Techniques for Fault Detection and Secure Checker Design," vol. 12, no. 3, 2024.
[16] H. Nagarajan, "Streamlining Geological Big Data Collection and Processing for Cloud Services," vol. 9, no. 9726, 2021.
[17] P. Alagarsundaram, "Adaptive CNN-LSTM and Neuro-Fuzzy Integration for Edge AI and IoMT-Enabled Chronic Kidney Disease Prediction," vol. 18, no. 3, 2024.
[18] H. Nagarajan, "Assessing Security and Confidentiality in Cloud Computing for Banking and Financial Accounting," vol. 12, no. 3, 2024.
[19] Basani, "Advancing Cybersecurity and Cyber Defense through AI Techniques," 2021.
[20] H. Nagarajan, Z. Alsalami, S. Dhareshwar, K. Sandhya, and P. Palanisamy, "Predicting Academic Performance of Students Using Modified Decision Tree based Genetic Algorithm," in *2024 Second International Conference on Data Science and Information System (ICDSIS)*, Hassan, India: IEEE, May 2024, pp. 1–5. doi: 10.1109/ICDSIS61070.2024.10594426.
[21] K. Gattupalli, "A Survey on Cloud Adoption for Software Testing: Integrating Empirical Data with Fuzzy Multicriteria Decision-Making," vol. 10, no. 4, 2022.
[22] Kalyan Gattupalli, "Optimizing 3D Printing Materials for Medical Applications Using AI, Computational Tools, and Directed Energy Deposition," Oct. 2024, doi: 10.5281/ZENODO.13994678.