EMERGING ADVANCES IN INTEGRATED TECHNOLOGY
e-ISSN: 2773-5540

EmAIT

Vol. 5 No. 1 (2024) 59-64
https://publisher.uthm.edu.my/ojs/index.php/emait

Classifying Phishing Websites Using Multilayer Perceptron

Muhammad Ikram Mohsin[1]*, Nor Hazlyna Harun[1,2]

[1] *School of Computing*
Universiti Utara Malaysia, Sintok, Kedah, MALAYSIA
[2] *Data Science Research Lab (DSRL), School of Computing*
Universiti Utara Malaysia, Sintok, Kedah, MALAYSIA

*Corresponding Author: ikram4@soc.uum.edu.my

DOI: https://doi.org/10.30880/emait.2024.05.01.008

Article Info

Received: 4 January 2024
Accepted: 13 June 2024
Available online: 30 June 2024

Keywords

Phishing, classification, multilayer perceptron, python, orange

Abstract

The prevalence of phishing as a cybercrime continues to escalate, posing significant threats to individuals' sensitive information. This paper addresses the urgent need for effective phishing detection methods, considering the limitations of existing approaches. The study employsArtificial Neural Networks, specifically Multilayer Perceptrons (MLP), trained using the backpropogationalgorithm. The study also highlights MLP's advantages in handling complex and noisy data. Through a comprehensivereview of related works, the paper identifies gaps in current research and establishes the groundwork for an innovative phishing website classification framework. The proposed solution utilizes MLPs, offering a detailed explanation of the methodology, dataset, model architecture, and trainingprocesses. The research concludes by summarizing key findings, emphasizing the solution's contributions to cybersecurity, and outlining potential avenues for future research.

1. Introduction

Phishing is a widespread form of cybercrime where attackers use deceitful methods to trick people into unintentionally, or without their consent, providing private information such as usernames, passwords, or financial details. These attackers often pretend to be trustworthy entities, such as banks or government agencies to exploit the trust of their victims. Instead of just using emails, phishing can also happen through text messages, instant messaging, or social media. In these fraudulent messages, there are usually links to fake websites or harmfulattachments. Clicking on these links or opening attachments can lead users to deceptive websites that look real, with the goal of tricking users into sharing sensitive information for identity theft or financial fraud.

Detecting phishing is crucial to stop these malicious activities. Various methods and tools are used for this purpose. Email filtering tools, for instance, analyze incoming emails using certainmethods to identify and flag suspicious ones. Web browsing protection in browsers andsecurity software helps detect and block access to known phishing websites. There is also specialized anti-phishing software that analyzes website links, email content, and user behavior to identify phishing patterns. Educating users about the signs of phishing through training programs is also important.

Addressing phishing is urgent because these attacks are becoming more frequent and sophisticated. As technology advances, so do the tactics used by cybercriminals. Successful phishing attacks can have severe consequences, and the need to adapt and improve security measures is constant. However, the current methods to detect phishing have their limitations.Some methods may not be effective against new attacks, and others might generate false alarms or miss sophisticated attacks. It's crucial to find a balanced approach that combines technical solutions, user education, and regular updates to security measures to effectively reduce the risks of phishing attacks.

Artificial Neural Networks (ANN), and Multilayer Perceptrons (MLP) in particular offer distinct advantages in the task of classifying phishing websites compared to existing methods.These include the ability to adapt to complex problems, reliable handling of noisy data, and better scalability. These advantages contribute to the effectiveness of MLPs in comparison to conventional methods, making them a promising approach in order to address the issue at hand.

In the review of related works, this research systematically examines existing literature and research on phishing website classification. The goal is to establish a contextual foundation for the proposed solution by evaluating the methodologies used in prior studies and identifying their strengths and limitations. Through this survey, the paper aims to pinpoint gaps in the current research landscape to pave the way for an innovative framework for classifying phishing websites.

At the core of this research is the proposed solution, which introduces a novel approach to classifying phishing websites using Multilayer Perceptron (MLP) and trained with the backpropogation (BP) algorithm. The section will provide an in-depth view of the proposed methodology, utilizing advanced machine and deep learning techniques to improve the precision and effectiveness of phishing website detection. The presentation includes a thorough overview of the dataset, the architecture of the MLP model, and the processes involved in training and evaluation. By delineating these technical aspects, the paper aims to ensure a clear understanding of the proposed solution.

The conclusion section serves as the endpoint of the research, summarizing key findings and insights gained throughout the study. It revisits the primary objectives outlined in the introduction, emphasizing how the proposed solution addresses identified gaps in existing research. Additionally, the conclusion discusses the broader implications of the findings for the field of cybersecurity and outlines potential directions for future research. This section provides a concise yet comprehensive summary, bringing closure to the research article by synthesizing the study's contributions and their broader significance within academic and practical domains.

2. Related Works

Al-Ahmadi was able to develop PDMLP, which was a model for detecting phishing websites using Multilayer Perceptron [1]. The MLP model outperformed other Machine Learning Models such as kNN, SVM, and Decision Tree, in terms of results in accuracy, precision, recall and F1-Score. Other research from Kalaharsha & Mehtre performed an evaluation of various Machine Learning models in the domain of detecting phishing websites, and concluded that MLP performed the best, at 98.4% accuracy [2]. Shabudin et al, opted to look into the feature selection process to determine if a proper choice of inputs could influence the accuracy of the model [3]. The two techniques tested were Feature Selection by Omitting Redundant Features (FSOR) and Feature Selection by Filtering Method (FSFM). The findings indicated that FSOR generally produced higher accuracy as opposed to FSFM.

Chang further supports MLP-based solutions by providing another case of MLP outperforming other models in terms of detecting phishing links [4]. They also recommended tuning the weights of the nodes in order to determine what specific features are significant for detecting phishing links. Paulius Vaitkevicius & Virginijus Marcinkevičius also able to come to the same conclusions [5]. Research conducted by Rendall et al., explored different Multi-Layered solutions, and concluded that among MLP, SVM, Naive Bayes and Decision Tree, MLPs consistently had showed better results [6]. Manoj P et al., also compared the results of several Machine Learning algorithms and independently came to the conclusion that MLPs would perform better in their specific case, coming up with an 86% accuracy score [7].

Rayalla et al., identified that Deep Learning methods, which MLPs are apart of, tend to overshadow Machine Learning methods in the domain of phishing detection [8]. Odehet al., have also developed their own model for the classification of phishing websites utilizing MLP, and were able to achieve similar, promising results as previous research has shown [9]. The list of related works is shown in Table 1.

Table 1 *Literature review matrix*

Author(s)	Aim	Datasets	Model/Techniques	Results
Al-Ahmadi, 2020	To propose a new method to perform phishingdetection tasks	UCI and Kaggle datasets	MLP, kNN, SVM, Decision Tree, Random Forest, RoF	96% accuracy of MLP as opposed tothe other models

Kalaharsha & Mehtre, n.d.	Compare performance of 18 different detection models	Alexa, Common-Crawl, Phish-tank, Open-Fish, UCI, Majestic, Ebbu2017, Kaggle, 5000 Best Websites	Naive Bayes, CNN, Random Forest, XCS, TWSVM, MLP	Testing on multiple datasets Produced 96% accuracy om average.
Shabudin et al, 2020	Explore two types of feature selection	UCI, "Phishing Websites"	Random Forest, MLP, Naive Bayes	FSOR method produces higher accuracy, at 96%
Chang, 2022	Prove that MLP can produce better results than conventional ML models	Kaggle dataset	MLP, Logistic Regression, Naive Bayes	MLP accuracy of 96%
Paulius Vaitkevicius & Virginijus Marcinkevičius, 2020	Compare accuracies of traditional supervised models	Two UCI and three MDP datasets	AdaBoost, Classification Tree, Regression Tree, Gradient Tree Boosting, kNN, MLP, Naive Bayes, Random Forest, SVM	MLP produced 97% accuracy
Rendall et al., 2020	Explore the usage of multi layered detection methods	Alexa, PhishTank, OpenPhish	MLP and SVM	MLP produced 89% accuracy
Manoj P et al., 2021	Develop and intelligent system that can detect phishing links	Alexa, dmoz, PhishTank	Logistic Regression, SVM, XGBoost, MLP, AutoEncoders	85% accuracy for MLP, comparable with SVM
Pavan Sai Rayallaet al., 2023	Compare performance of ML and DL models	Self-developed, labeled legitimate or phishing	Decision Trees, Random Forest, MLP	85% accuracy by MLP
Odeh et al., 2020	Use MLP to reduce runtime of detection models	PhishTank, MillerSmiles, Google search	Feed forward NN, Rule based, Logistic Regression, Naive Bayes, MLP	MLP produced an accuracy of 99%

3. Proposed Solution – Classifying Phishing Websites using Multilayer Perceptron

3.1 Data Preprocessing

The dataset (Web Page Phishing Dataset) was obtained from Kaggle, and comprised of 10505 instances of 87 features and 1 target variable. The target was to determine whether or not a website link was labeled as "legitimate" or "phishing". No missing values or outliers were found in the original dataset; therefore, all 10505 instances of data were used. The only pre- processing that was conducted on the dataset was normalization, to prepare for the dataset to be inserted into the ANN. The dataset was normalized to scale all values to range between 0 and 1.

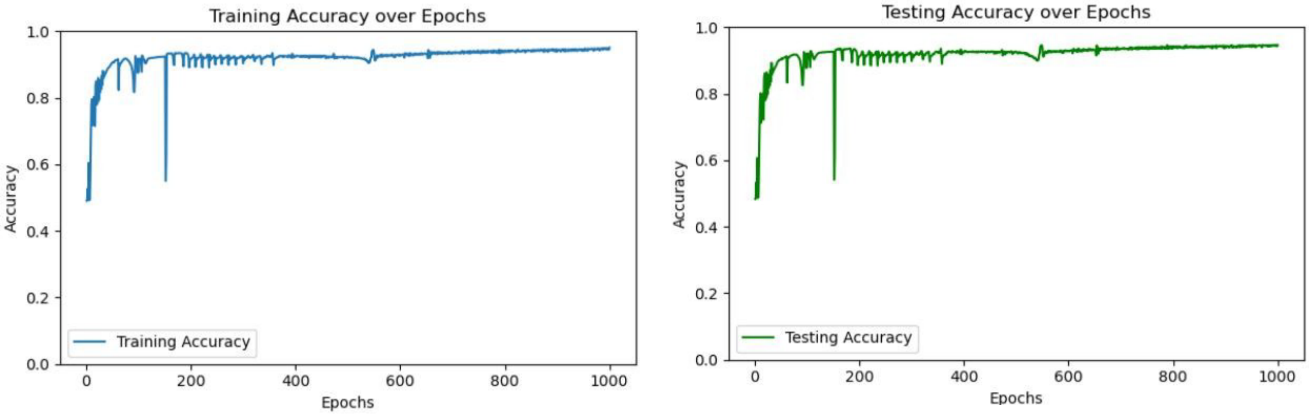
3.2 Model Building

The next step taken was building the Machine Learning model. The model was built using the Python language, utilizing the Jupyter Notebook environment. The main library used was scikit-learn, Python's free open-source machine learning library. The specific details of the model are:

- 3 hidden layers of 64, 32, and 16 neurons respectively,
- the activation function used was the rectified Linear Unit (reLU) function,
- 1000 epochs, and
- 0.001 learning rate.

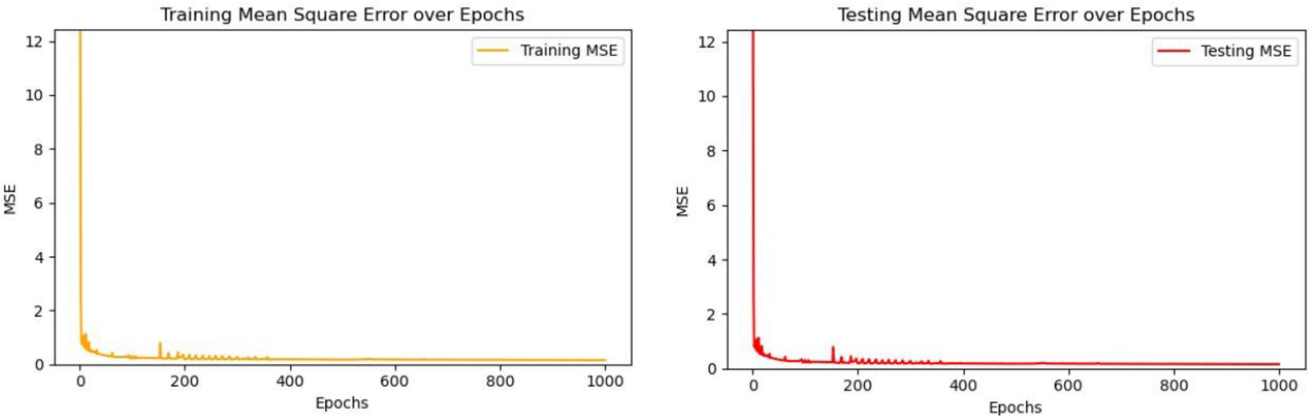
3.3 Training and Testing Procedure

Lastly, the MLP was model subjected to training and testing on the normalized dataset (Fig.1 and 2). The dataset (total of 10505 instances) was split into a 7:3 ratio of the Training set (7354 instances) and Testing set (3151 instances). The Testing set was then used to train the model, and then the results were validated using the Testing set. 4 evaluations, or graphs, were produced from the Training and Testing phase which are the Training and Testing Accuracy over Epochs, and the Training and Testing Mean Square Error (MSE) over Epochs:



(a) **(b)**

Fig. 1 *(a) Training and (b) Testing accuracy over epochs*



(a) **(b)**

Fig. 2 *(a) Training and (b) Testing MSE over epochs*

It should be noted that further changes to the values of epochs and learning rates did not positively influence the accuracy of the model.

3.4 Evaluation Metrics

A confusion matrix is a form of evaluation metric that can be produced by analyzing the performance of a given model. It provides a detailed breakdown of the model's predictions and the actual outcomes across different classes. The confusion matrices for the Training and Testing sets are as follows:

Table 2 *Confusion matrix for testing set*

	Predicted Positive (legitimate)	Predicted Negative (Phishing)
Actual Positive (legitimate)	1073	543
Actual Negative (Phishing)	19	1518

Table 3 *Confusion matrix for training set*

	Predicted Positive (legitimate)	Predicted Negative (Phishing)
Actual Positive (legitimate)	2415	1312
Actual Negative (Phishing)	75	3551

As can be seen from Table 2 and 3, both the testing and training sets are able to classify instances of legitimate links with high accuracy. In terms of phishing links, the accuracy is also high, albeit slightly lower than in legitimate cases. From the confusion matrix, four (4) new evaluation metrics can be discerned, which are:

- Accuracy, a general measure of how often the model makes correct predictions,
- Precision, the accuracy of the model when minimizing false positives,
- Recall, the accuracy of the model when accounting for false negatives, and
- F1 Score, considers both Precision and Recall in order to achieve proper balanced results

Table 4 below shows the four metrics of the Training and Testing set, calculated from thefindings of the confusion matrix. The evaluation of other Machine Learning models was also tested to compare their performance to the target model (MLP). Note that the other (non-MLP) models were tested only in the Orange Data Mining tool.

Table 4 *Performance metrics comparison*

Dataset	Performances			
	Accuracy	F1 Score	Precision	Recall
MLP (Training)	0.94	0.94	0.94	0.94
MLP (Testing)	0.95	0.95	0.95	0.95
Decission Tree	0.94	0.94	0.94	0.94
Naïve Bayes	0.90	0.90	0.90	0.90
SVM	0.65	0.64	0.67	0.65
Random Forest	0.95	0.95	0.95	0.95
kNN	0.90	0.90	0.90	0.90

As can be observed from the above table, the MLP model showed comparatively high-performance metrics, tied with the Random Forest model at 95% accuracy. The Decision Treemodel and Naive Bayes also produced high results, at 94% and 90% respectively.

4. Conclusion

In conclusion, by utilizing a dataset of 10505 cases and 87 features, the MLP-based classification model obtained favorable results and was able to accurately discern between phishing and legitimate links. With an accuracy of 95% on the Python implementation, these results show that MLP-based solutions show high promise in solving the issue of phishing detection.

Acknowledgement

We express our sincere gratitude to our lecturer for the A231 STIN2054 (Neural Network) course, for their invaluable guidance, unwavering support, and insightful feedback throughout this research. Thanks to our colleagues and fellow researchers at Universiti Utara Malaysia (UUM), for their collaborative spirit and constructive contributions. Lastly, heartfeltthanks to our friends and family for their unwavering support. This work is the culmination of the collective efforts of all mentioned, and we are grateful for your integral roles in our academic journey.

Conflict of Interest

Authors declare that there is no conflict of interests regarding the publication of the paper.

Author Contribution

The authors confirm contribution to the paper as follows: ***study conception and design:*** *Muhammad Ikram bin Mohsin;* ***data collection:*** *Muhammad Ikram bin Mohsin;* ***analysis and interpretation of results:*** *Muhammad Ikram bin Mohsin, Nor Hazlyna Harun;* ***draft manuscript preparation:*** *Muhammad Ikram bin Mohsin, Nor Hazlyna Harun. All authors reviewed the results and approved the final version of the manuscript.*

References

[1] Al-Ahmadi, S., & Lasloum, T. (2020). PDMLP: Phishing Detection using Multilayer Perceptron. *International Journal of Network Security & Its Applications*, *12*(3), 59–72. https://doi.org/10.5121/ijnsa.2020.12304

[2] Kalaharsha, P., & Mehtre, B. M. (2021). Detecting Phishing Sites -- An Overview. *ArXiv.Org*. http://arxiv.org/abs/2103.12739

[3] Shabudin, S., Sani, N. S., Ariffin, K. A. Z., & Aliff, M. (2020). Feature selection for phishing website classification. *International Journal of Advanced Computer Science and Applications*, *11*(4), 587–595. https://doi.org/10.14569/IJACSA.2020.0110477

[4] Chang, P. (2022). Multi-Layer Perceptron Neural Network for Improving Detection Performance of Malicious Phishing URLs Without Affecting Other Attack Types Classification. *ArXiv Preprint ArXiv:2203.00774*, 1–3. https://arxiv.org/abs/2203.00774%0Ahttps://arxiv.org/pdf/2203.00774

[5] Vaitkevicius, P., & Marcinkevicius, V. (2020). Comparison of Classification Algorithms for Detection of Phishing Websites. *Informatica*, *31*(1), 143–160. https://doi.org/10.15388/20-INFOR404

[6] Rendall, K., Nisioti, A., & Mylonas, A. (2020). Towards a multi-layered phishing detection. *Sensors (Switzerland)*, *20*(16), 1–18. https://doi.org/10.3390/s20164540

[7] Manoj, P., Bhuvan Kumar, Y., Rakshitha, D., & Megha, G. (2021). Detection and classification of phishing websites. *Trends in Computer Science and Information Technology*, *6*(2), 053–059. https://doi.org/10.17352/tcsit.000040

[8] Rayalla, P. S., Vivek, S. V., Golla, H., Katakam, G., & A N, M. Z. (2023). Detecting Phishing Websites using Deep Learning. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4527759

[9] Odeh, A., Keshta, I., & Abdelfattah, E. (2020). Efficient detection of phishing websites using multilayer perceptron. *International Journal of Interactive Mobile Technologies*, *14*(11), 22–31. https://doi.org/10.3991/ijim.v14i11.13903