Early Detection of Phishing Sites with Enhanced Neural Network Models

[1]Isa Suarti, [2]Totok Chamidy, [3]Cahyo Crysdian
[1]SekolahTinggi Informatika dan Komputer Indonesia, Indonesia
[1,2,3]Magister Informatika, Fakultas Sains dan Teknologi,
Universitas Islam Negeri Maulana Malik Ibrahim, Malang, Indonesia
Email: [1]isa.suarti@stiki.ac.id,[2]to2k2013@ti.uin-malang.ac.id,[3]cahyo@ti.uin-malang.ac.id

Article Info

Article history:
Received Apr 14th, 2024
Revised May 20th, 2024
Accepted Jul 31th, 2024

Keyword:
Early Detection
Neural Network
Phishing
URL
Web Site

ABSTRACT

Phishing is a digital crime committed with the aim of obtaining personal data by creating a link or website that resembles the original. This form of cyber attack is caused by a notification in a text message, email, or phone call. A common anti-phishing countermeasure technique is to perform early detection of potentially phishing sites, primarily according to the source code features, which are required to traverse web page content, as well as third parties that slow down the process of clarifying phishing URLs. Although the latest technology has long been used in phishing early detection, there is still a need for manual feature engineering that is important and reliable enough to detect emerging phishing offenses. One of these involves training a neural network (NN) using a dataset of known phishing URLs and legitimate URLs. The research was conducted using 200 data, Data were separated into training and testing categories. Training was done using 100 and 120 data. Training results on 100 data and 160 data had lower iterations and errors on the tanh activation function compared to the logistic activation function. The number of iterations that occur in logistic activation is as many as 400 iterations, while when using the tanh activation function only 175 iterations are needed.

Corresponding Author:
Cahyo Crysdian,
Magister Informatika, Fakultas Sains dan Teknologi,
Universitas Islam Negeri Maulana Malik Ibrahim, Malang, Indonesia
Jl. Gajayana 50, Malang, Indonesia.
Email: cahyo@ti.uin-malang.ac.id

DOI: http://dx.doi.org/10.25014/ijaidm.v7i2.30068

1. INTRODUCTION

Phishing is a form of digital attack including the creation of fake sites that look genuine, so that users feel they have entered the original site [1], [2] On this fake site, users will enter personal data such as username, password or ATM PIN. Currently, there are many incidents of web fraud or phishing. Apart from being used for data theft, phishing sites are also used to deceive internet users on behalf of legitimate sites and spread computer malware/viruses over the internet . Social engineering attacks frequently culminate in phishing sites, which are the final and most crucial stage. The majority of phishing schemes trick visitors into visiting the websites they contain [4]. Data theft through phishing sites can also be done by fraudulent actions on behalf of legitimate sites and by spreading computer malware/viruses. According to Anti-Phishing Working Group (APWG), 32% of data theft is always related to phishing activities [1], [5]. In fact, in early 2020 alone, APWG found 165,772 phishing sites ready to catch their victims. This also triggers public awareness of phishing sites to increase every year, but the number of losses caused by phishing sites is still growing rapidly.

Phishing websites look similar to legitimate websites and use similar domain names. This is called domain spoofing and is a social engineering attack that misleads users [4]. Some characteristics of phishing web sites can be seen from the Uniform Resource Locator (URL) or web site address and some other special characteristics [6]. This is what makes many researchers turn to artificial intelligence (AI) techniques, especially to detect phishing more accurately and efficiently [3], [5], [7], [8]. One such technique is the use of neural networks [4], [5], [9], [10], [11], [12], [13], [14], which have demonstrated good outcomes in detecting phishing assaults. Neural Network is a type of machine learning model inspired by the structure and function of the human brain. These networks consist of layers of interconnected artificial neurons that process and analyze input data to make predictions or classifications. By training the network with a large dataset of known phishing URLs and legitimate URLs, the neural network can learn patterns and features that distinguish between the two [15], [16]. This trained neural network can then be used to detect new and unknown phishing URLs by analyzing their characteristics and comparing them with the patterns learned during training. This approach enables real-time and proactive phishing detection, as the neural network can quickly analyze URLs and determine the likelihood that they are malicious. In addition, neural networks can also handle the complexity and variability of phishing attacks, as they can learn from diverse data sets and adapt to new attack patterns. In addition to machine learning, deep learning techniques have also been explored for phishing detection. Deep learning entails training deep neural networks with numerous layers to autonomously derive high-level features from raw data [14]. These deep learning models can capture complex patterns and relationships in data, resulting in improved classification performance for detecting phishing attacks. The accuracy of systems that use heuristic techniques depends on a set of discriminative criteria selected from the website.

Determining features and how to process them is important in classifying websites correctly, while effective and fast information retrieval is essential for making good decisions. Data mining is a technique that can be used to extract features from a website so that patterns and relationships between features are found, especially URL features [7], [17], [18], [19]. Along with the increasing number of phishing attacks targeting individuals and organizations, the need for effective phishing detection techniques has become very important [20]. In previous studies with more phishing detection cases using the randomforest algorithm, while in this study using a neural network with 6 fewer parameters than previous research used by Ojewumi. This research can also show that the neural network algorithm model is able to achieve 100% accuracy has the best performance in detecting phishing sites, and focuses on website URL link data because at this time what is very easy to modify to trap ordinary people is website URL links. in addition, in the implementation of this research, researchers really pay attention to the use of features for detection such as the importance of knowing the age of the domain and meta tags, phishing websites often have domains that are newly registered and short-lived. this is because fraudsters tend to create new domains for each phishing attack to avoid detection and blocking.

In previous studies, no one has used the meta tag feature where this meta tag is used to optimize their page to look more legitimate or official. their pages to make them look more legitimate or official so that they are easily found by search engines. In this research, the dataset can be accessed publicly, the comparison of the test data set is different from previous research. in previous research, the discussion of the activation function used was not discussed or even mentioned. but in this study it can be seen selection of the activation function used and the function can affect the results. Therefore, researchers feel the need to conduct research related to how we as individuals and organizations conduct early detection of phishing sites that are always troubling, one of which uses an enhanced neural network model.

2. RESEARCH METHOD

Early detection of phishing sites on websites by collecting web URLs and classifying them into major sections representing the specified type of classification model [10]. All classification processes follow the same methodology which all start with data set pre-processing where the phishing data set is normalized according to the chosen model. Since phishing activities are carried out through various mediums and one of the main ones is the web, the main data comes from websites. This is because, the web interface can also be included in the email body, to further convince the victim. Phishing is often done by mimicking the appearance of the web with its original appearance. In addition, phishing can also be done by providing convincing information and providing additional information that has the effect of forcing the victim to do something.

The design of the phishing detection research conducted is shown in Figure 1 below, starting from data collection in the form of URL data collecting, system design, system implementation, experiment, and evaluation.

2.1. Data Collection

The data used in this research is primary data. Primary data is a type of data that is gathered directly from primary sources like interviews, surveys, experiments, and similar methods. Primary data is typically specific as it is customized to meet the researcher's particular requirements. Primary data is usually available in unprocessed form. Collecting data from the website https://phishtank.org/ in figure 2.

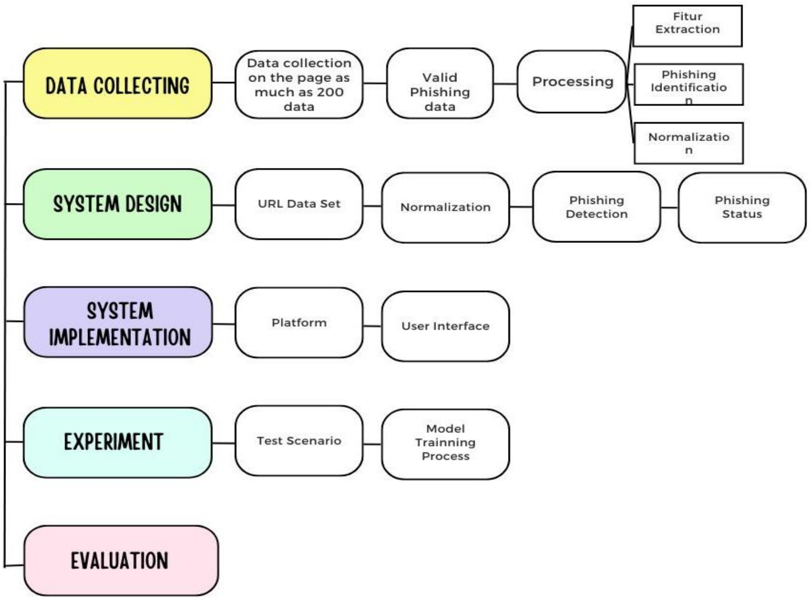
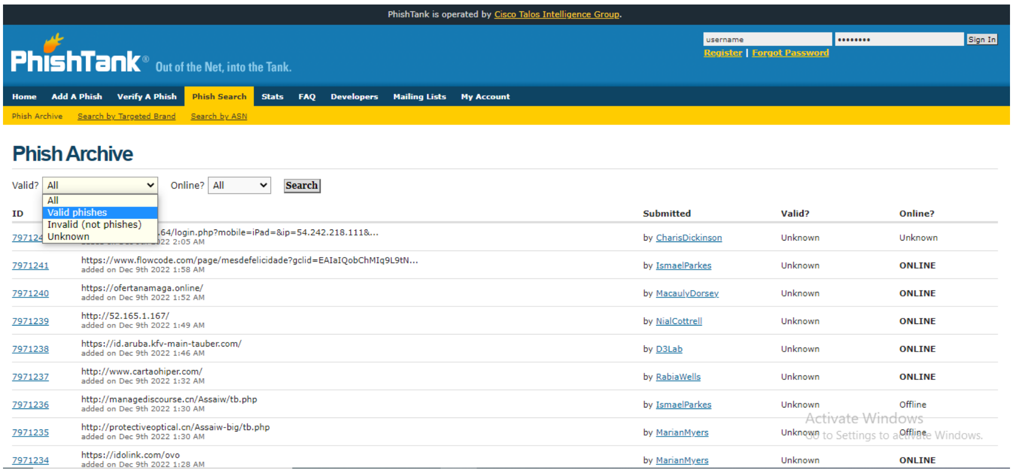


Figure 1. Design of the Phishing Detection Research



Phish Archive

Valid? All   Online? All

All
Valid phishes
Invalid (not phishes)
Unknown

ID		Submitted	Valid?	Online?
797124	.64/login.php?mobile=iPad=&ip=54.242.218.111&... 22 2:05 AM	by CharisDickinson	Unknown	Unknown
7971241	https://www.flowcode.com/page/mesdefelicidade?gclid=EAIaIQobChMIq9L9tN... added on Dec 9th 2022 1:58 AM	by IsmaelParkes	Unknown	ONLINE
7971240	https://ofertanamaga.online/ added on Dec 9th 2022 1:52 AM	by MacaulyDorsey	Unknown	ONLINE
7971239	http://52.165.1.167/ added on Dec 9th 2022 1:49 AM	by NialCottrell	Unknown	ONLINE
7971238	https://id.aruba.kfv-main-tauber.com/ added on Dec 9th 2022 1:46 AM	by D3Lab	Unknown	ONLINE
7971237	http://www.cartaohiper.com/ added on Dec 9th 2022 1:32 AM	by RabiaWells	Unknown	ONLINE
7971236	http://managediscourse.cn/Assaiw/tb.php added on Dec 9th 2022 1:30 AM	by IsmaelParkes	Unknown	Offline
7971235	http://protectiveoptical.cn/Assaiw-big/tb.php added on Dec 9th 2022 1:30 AM	by MarianMyers	Unknown	Offline
7971234	https://idolink.com/ovo added on Dec 9th 2022 1:28 AM	by MarianMyers	Unknown	ONLINE

Figure 2. PhishTank Layout

Process the initial data in the form of Excel containing the URLs that have been collected, then carry out the data checking process, namely in the form of criteria from phishing sites, checking the data includes knowing the number of dots, the presence of the @ symbol, character length, checking the age of the domain, and checking DNS (Domain name system), after checking the data. Then create a dataset. The dataset in question is a collection of data or documents containing one or more records presented in the form of a CSV file. Divide the data into 2 parts. Training data and testing data. Perform a min-max data normalization process to produce a balanced comparison value between the data before and after normalization.

2.2. Normalization

The data collected has different units, therefore a normalization process is needed, which is to make the data have the same range of values or none of the values are too large or too small (range 0 to 1). The min-max normalization method involves a linear transformation of the original data to create a balanced comparison of values between the data before and after normalization. This research uses URL parameters as input for phishing or non-phishing detection (y). The URL parameters are the number of @ ($x1$), number of dots ($x2$), domain age ($x3$), IP address ($x4$) SSL (Secure Sockets Layer) ($x5$, and Meta tags All of these variables are used as input because this variable is one of the factors of assessing a website that can be seen from the URL. Because input and output data have different units, a normalization process is needed first. The goal is to make the data have the same range of values or no value is too large or too small (range 0 to 1). The data that has been found will be analyzed in 6 attributes, namely the number of @, the number of dots (.), domain age, IP check, SSL check and meta tags. analysis of the number of @ and the number of dots will be calculated using the normalization formula so that it becomes a certain number value, while the age of the domain is calculated by calculating the difference in the creation of the domain at the whois address then compared to today, for IP check if the IP of the domain is found it will be given a value of 1 and if not found it will be given a value of 0, as well as SSL check and meta tags.

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{1}$$

With: x' : normalization result data
x : original data
x_{min} : minimum score from data
x_{max} : maximum score from data

2.3. Training with neural networks

Neural network is one of the machine learning methods that uses the basis of a calculation structure resembling neural networks in the brain. Neural networks in the brain are interconnected with each neural network to form a pattern and the core of the neural network structure is the activation function [21]. Neural Network architecture is shown in Figure 3.

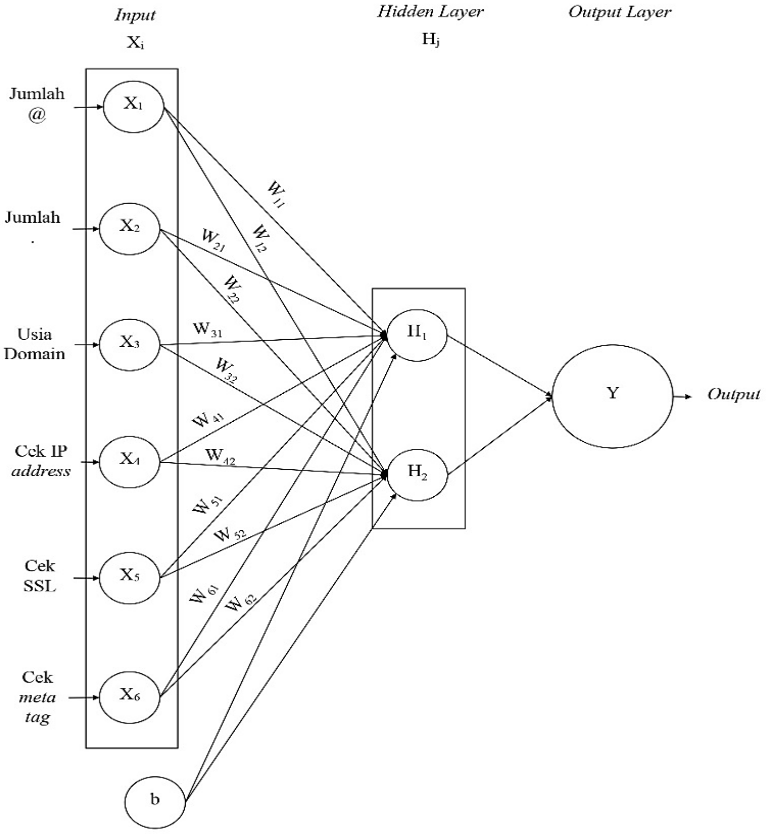


Figure 3. NN Using 1 Hidden Layer with 2 Nodes

The form of architecture used (Figure 3) in training and testing with one layer, with a hidden layer composition of 2 nodes, then the output layer. The training mechanism can be seen in Figure 1, starting with calculating the weight between the input node and the hidden layer, after that, it continues by calculating the weight between the hidden layer and the output layer. To get the weight H_1 is done by summing up all the results of multiplying the input value and the initial weight and bias as in equation 2.

$$H_j = f\left(\sum_{i=1}^{n} x_i w_{ji} + bw_j\right) \tag{2}$$

After obtaining the value of Hj, the next step is to calculate the value of Y. To get the value of Y, the calculation is carried out as before, with the input value being H_j, as in equation 3.

$$Y = f\left(\sum_{j=1}^{n} H_j w_j + bw_j\right) \quad (3)$$

These activation functions help introduce non-linearity into the model, allowing neural networks to model more complex relationships in the data. Some commonly used activation functions include ReLU, sigmoid/logistic, and tanH (tangent hyperbolic).

2.4. *tanH (tangent hyperbolic)*

Tanh is a Neural Network activation that uses an output value reference of -1 and 1 by performing a calculation process from an input. This method is based on the difference of the exponential ratio of the input value with the exponential of the depedency variable which is processed using an exponential sum. The formula of the Tanh method is in equation 4.

$$f(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}} \quad (4)$$

2.5. Sigmoid

Sigmoid is a fundamental activation in Neural Networks, sigmoid calculation produces a value between 0 and 1 as in the Logistic Regression method. The graph of the calculation results of the sigmoid method resembles the letter "S" and starts from flat with a slow increase and ends with flat as well. The sigmoid method formula is in equation 5:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

2.6. System Design

The system that will be created to perform the analysis is as shown in Figure 4 below.

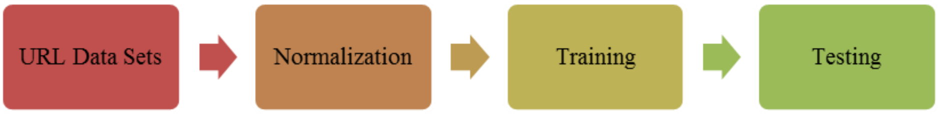


Figure 4. System Design

From the system design in Figure 4, the process that will be carried out is testing with the initial stages of initializing weights and parameters such as learning rate = 0.1, each input-output receives a signal and is forwarded to the hidden layer, calculating all outputs in the hidden layer, in the calculation process towards the output is carried out using the activation function.

3. RESULTS AND ANALYSIS

In theory, there is no standard rule in determining the optimal architecture in the Artificial Neural Network method, so the determination of the network is done by trial and error to find the maximum results. On this basis, this research tries to conduct experiments by varying the composition of training data and test data, besides that, an approach is also taken by varying the activation function on the neural network. The 200 existing data will be divided into two groups with the composition of training data and test data 50%:50%, and 80%:20%. The full details are shown in Table 1.

Table 1. Composition of the Input Data

Model	Composition Data	Number of Training Data	Number of Testing Data
A	50 : 50	100 data training	100 Data Testing
B	80 : 20	160 data training	40 Data Testing

Then in addition to the composition of the input data, this research also approaches by varying the activation function as in Table 2.

Table 2. Composition of the Input Data

Number of Layer	Activate Funtion	Number of Data
1	Logistic	200 Data
2	Tanh	200 Data

3.1. Training Outcomes

The training stage is the stage where existing data becomes learning for the system. This process is carried out continuously until the most optimal weight value is obtained. In the end, the system will get the best weight when the error is considered the smallest and does not change anymore. Then the weight obtained will be used as the basis for classifying sites. The classification results obtained are phishing or not phishing. Training will be carried out using a scenario of taking part of the existing data, and the rest will be used as testing data. The data used is 50 phishing data and 50 non-phishing data as training, and the remaining 100 data will be used as testing data. For the second experiment, 80 phishing data and 80 non-phishing data will be used, and the remaining 40 data will be used as testing data.

In the training process, a JST structure with 6 (six) inputs, a hidden layer with 2 nodes, and 1 output will be used. Then the predicted value between the classification and the real results will be calculated so that it can be analyzed for the confusion matrix. The figure above displays the weights found, then the weight value is used as a reference for testing. In this testing stage, 50% phishing data is used, and 50% non-phishing data. When viewed from the figure, the classification column shows the results of the artificial neural network process. Based on the value in the column, it can be seen that the target and process results provide the same value, so it can be said that the weight value generated from the learning stage can be used to determine whether a web is phishing or not phishing.

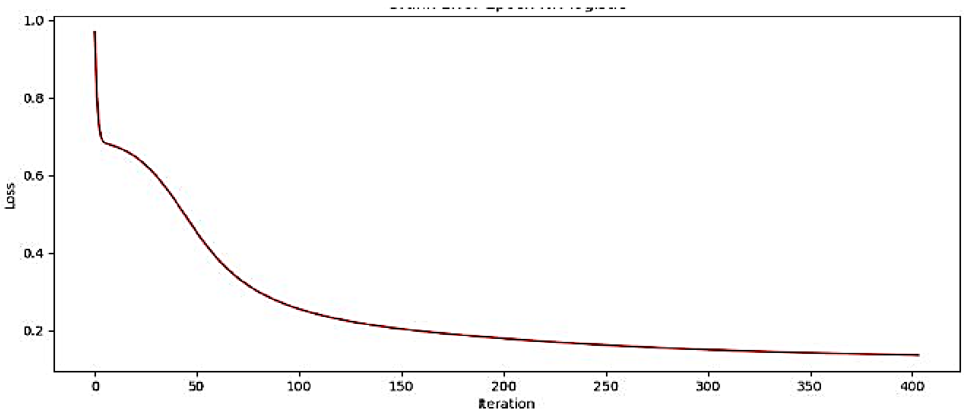


Figure 5. Graph of Error & Epoch neural network logistic 50:50

Training is done using 1 hidden layer, with 2 neurons and the results are shown in the graph in Figure 6. The graph shows that in trials at iterations 0 - 150th the error value is still unstable and high, the error starts to be low when iterations more than 200 and starts to stabilize at iterations 300, Iterations above 300, the error shows a difference that is not too significant so, optimal results are shown at iterations 300 and above.

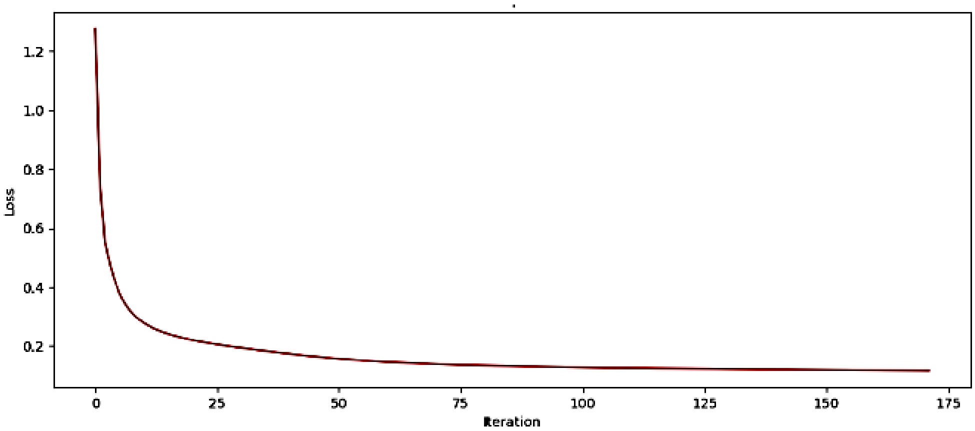


Figure 6. Graph of Error & Epoch neural network tanh 50:50

The graph shows that in trials at iterations 0-25th the error value is still unstable and high, the error starts to be low when iterations are above 50 and starts to stabilize at iteration 75, at iterations above 100 the error shows a difference that is not too significant so, optimal results are shown at iteration 100 and above.

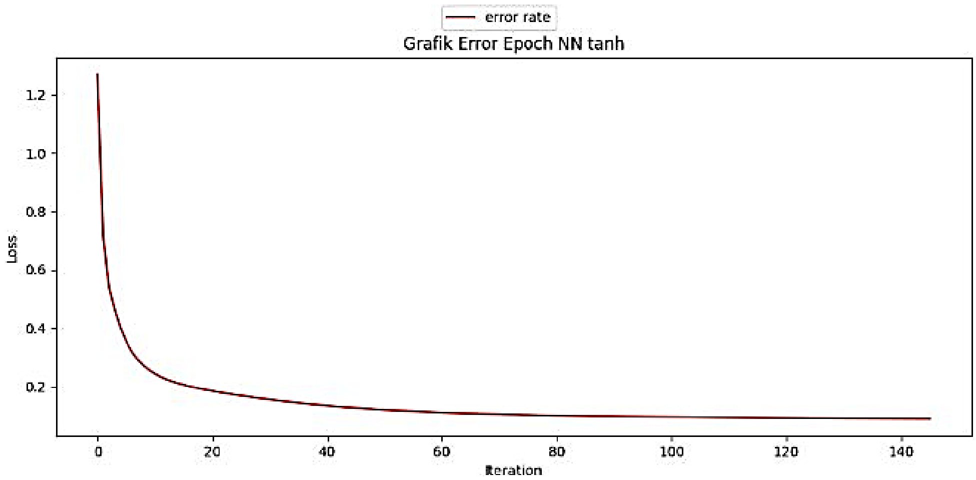


Figure 7. Graph of Error & Epoch neural network Tanh 80:20

The results of the graph in Figure 7 shows that in the trial at iterations 0-40th the error value is still unstable, the error starts to be low when the iteration is above 60, at the iterations above 80 the error shows a difference that is not too significant so, the optimal results are shown at iterations 80 and above.

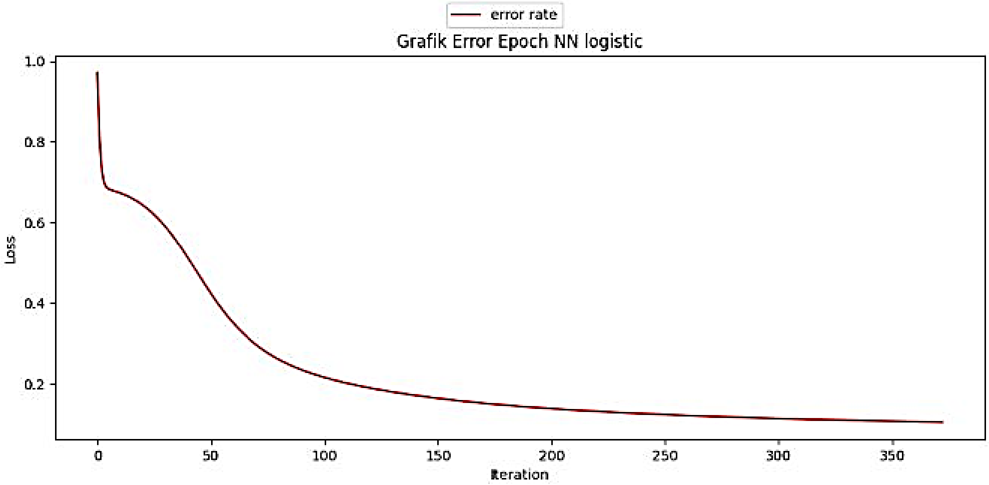


Figure 8. Graph of Error & Epoch neural network logistic 80:20

The results of the graph in Figure 8 shows that in the trial at iterations 0-40th the error value is still unstable, the error starts to be low when the iteration is above 60, at the iterations above 80 the error shows a difference that is not too significant so, the optimal results are shown at iterations 80 and above.

3.2. Evaluation

The evaluation of this research uses confusion matrices with the results of scenario 1 50: 50 obtained as follows:

True Positive (TP)	: 50 phishing data are correctly predicted as phishing.
True Negative (TN)	: 50 non-phishing data is correctly predicted.
False Positive (FP)	: 0, no non-phishing data is predicted as phishing.
False Negative (FN)	: 0, no phishing data was predicted as non-phishing.

Table 3. Confussion Matrics

Actual / Prediction	Phishing	Non-Phishing
Phishing	TP : 50	TN : 0
Non-Phishing	FP : 0	FN : 50
Accuracy	1.0	
Precison	1.0	
Recall	0.5	

The table consists of several columns, as bellow.

1. Actual: Shows the actual class (ground truth), i.e. Phishing and Not Phishing.
2. Predicted: Shows the predicted results of the model, also consisting of Phishing and Not Phishing.
3. The number of values in each cell indicates the amount of data classified into that category.
4. Accuracy: 1.0 or 100%, meaning the model predicts all data correctly.
5. Precision: 1.0 or 100%, meaning there are no false positive predictions.
6. Recall: 0.5 or 50%, meaning the model can only identify half of the actual phishing data. In the confusion matrix analysis above, it shows that true positives get a value of 50% and false negatives get a value of 50% so that the precision reaches 1.0 or 100%.

The following table is the result of calculating the weights in the neural network.

Table 4. Calculation of Weights

	URL	Number of (@)	Number of (.)	Domain ages	IP	SSL	Meta Tags	Result	Classifications	Predictions	TP	TN	FP	FN
1	htttp://web.iib.mizuhebanki-japan.com/client/index/php	0.0	0.375	0.036827	0.0	0.0	0.0	1	1	0.0184	1	0	0	0
2	https://web.ob.miizuhabunk-japan.armstronw.com/client/index_sp.php	0.0	0.500	0.036827	0.0	0.0	0.0	1	1	0.0113	1	0	0	0
3	https://web.ob.miizuhabunk-japan.armstronw.com/client/index.php	0.0	0.500	0.036827	0.0	0.0	0.0	1	1	0.0113	1	0	0	0
4	https://aazodeabyz.duckdns.org	0.0	0.125	0.000000	1.0	0.0	0.0	1	1	0.0142	1	0	0	0
5	https://acreqgfbzh.duckdns.org	0.0	0.125	0.000000	1.0	0.0	0.0	1	1	0.0142	1	0	0	0
6	https://adcqgyzhye.duckdns.org	0.0	0.125	0.000000	0.0	0.0	0.0	1	1	0.0464	1	0	0	0
7	https://adkovgdcmn.duckdns.org	0.0	0.125	0.000000	1.0	0.0	0.0	1	1	0.0142	1	0	0	0
8	https://afedgyybrf.duckdns.org	0.0	0.125	0.000000	1.0	0.0	0.0	1	1	0.0142	1	0	0	0
9	http://rakuten.xonreungfop.com/pc	0.0	0.125	0.036827	0.0	0.0	0.0	1	1	0.0649	1	0	0	0
10	http://rakuten.xonreungfop.com/	0.0	0.125	0.036827	0.0	0.0	0.0	1	1	0.0649	1	0	0	0

Table 4 contains a data explanation of the web URL feature value weight table. This table analyzes and classifies whether the URL is a phishing site or not by using 6 features such as number of @, number of dots, domain age, IP address, SSL, and Meta tags html to assess the performance of the classification model.

3.3 Discussion

Phishing is one of the most common and harmful cybersecurity threats. One method that is quite effective for phishing detection is neural networks. In the implementation of neural networks, activation functions play an important role in the training process and performance of the model. The activation function is responsible for introducing nonlinearity into the neural network, which allows the model to learn complex patterns in the data. The two activation functions used are the tanh (hyperbolic tangent) function and the logistic (sigmoid) function. In this study, the experimental results show that the tanh activation function activation function provides the best performance in phishing site detection using neural networks. This may be due to the ability of the tanh function to speed up the training process and maintain the stability of the weight values. In previous research on phishing detection using neural networks, most of them used a logistic activation function. However, the results from this study show that the tanh activation function provides better performance in phishing site detection. One of the previous studies used a neural network with a logistic activation function for phishing detection. They reported an accuracy of 92.37% in identifying phishing sites. While in this study, by using tanh activation function with 1 hidden layer, the accuracy of phishing site detection increased to 99% with faster learning time and fewer iterations.

4. CONCLUSION

Phishing is a crime committed by creating something similar to the original, or inviting people to believe that the information provided is true information. Several features or variables will be used to identify phishing. These include web address or URL, such as @ character, number of dots, SSL, domain age, IP address, meta tags. These features are then used to analyze 200 URL data from phishing and non-phishing websites. In the above trial, it is found that training and testing have been carried out with different amounts of data. In the 50:50 composition, 100 data were used for training and 100 data for testing. The training data contains 50 phishing data, and 50 non-phishing data, as well as the testing data. While in the 80:20 composition, 160 data were trained, namely 80 phishing data and 80 non-phishing data, and 40 data were used for testing, which contained 20 phishing data and 20 non-phishing data. Training results on 100 data and 160 data have lower iterations and errors on the tanh activation function compared to the logistic activation function. The number of iterations that occur in logistic activation is as many as 400 iterations, while when using the tanh activation function only 175 iterations are needed. In this case, the implementation of the activation function in artificial neural networks is very important and can have an influence on the prediction results. Choosing the right activation function in a neural network can affect the performance of the model in phishing site detection. The results of this study show that the tanh activation function provides better performance than the logistic function in the neural network training process for phishing site detection. However, keep in mind that the performance of the model is also affected by other factors such as network architecture, optimization, and training data quality. Further research can be conducted to explore the optimal combination of activation functions and other techniques to improve the accuracy of phishing site detection.

REFERENCES

[1] S. S. Roy, A. I. Awad, L. A. Amare, M. T. Erkihun, and M. Anas, “Multimodel Phishing URL Detection Using LSTM, Bidirectional LSTM, and GRU Models,” *Futur. Internet*, vol. 14, no. 11, 2022, doi: 10.3390/fi14110340.

[2] M. K. Moussavou Boussougou and D. J. Park, “Attention-Based 1D CNN-BiLSTM Hybrid Model Enhanced with FastText Word Embedding for Korean Voice Phishing Detection †,” *Mathematics*, vol. 11, no. 14, pp. 1–25, 2023, doi: 10.3390/math11143217.

[3] R. Yang, K. Zheng, B. Wu, C. Wu, and X. Wang, “Phishing website detection based on deep convolutional neural network and random forest ensemble learning,” *Sensors*, vol. 21, no. 24, pp. 1–18, 2021, doi: 10.3390/s21248281.

[4] D. J. Liu, G. G. Geng, X. B. Jin, and W. Wang, “An efficient multistage phishing website detection model based on the CASE feature framework: Aiming at the real web environment,” *Comput. Secur.*, vol. 110, p. 102421, 2021, doi: 10.1016/j.cose.2021.102421.

[5] A. Basit, M. Zafar, X. Liu, A. R. Javed, Z. Jalil, and K. Kifayat, “A comprehensive survey of AI-enabled phishing attacks detection techniques,” *Telecommun. Syst.*, vol. 76, no. 1, pp. 139–154, 2021, doi: 10.1007/s11235-020-00733-2.

[6] F. E. Purwiantono and A. Tjahyanto, “Classification model based on url and content feature approach for detection phishing website in Indonesia,” *J. Theor. Appl. Inf. Technol.*, vol. 95, no. 17, pp. 4181–4191, 2017.

[7] E. A. Aldakheel, M. Zakariah, G. A. Gashgari, F. A. Almarshad, and A. I. A. Alzahrani, “A Deep Learning-Based Innovative Technique for Phishing Detection in Modern Security with Uniform Resource Locators,” *Sensors*, vol. 23, no. 9, 2023, doi: 10.3390/s23094403.

[8] A. Aljofey, Q. Jiang, Q. Qu, M. Huang, and J. P. Niyigena, “An effective phishing detection model based on character level convolutional neural network from URL,” *Electron.*, vol. 9, no. 9, pp. 1–24, 2020, doi: 10.3390/electronics9091514.

[9] Y. Kim and Y. K. Kim, “Time-Frequency Multi-Domain 1D Convolutional Neural Network with Channel-Spatial Attention for Noise-Robust Bearing Fault Diagnosis,” *Sensors*, vol. 23, no. 23, pp. 1–20, 2023, doi: 10.3390/s23239311.

[10] P. Secchi, S. Vantini, and V. Vitelli, “Bagging voronoi classifiers for clustering spatial functional data,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 22, no. 1, pp. 53–64, 2013, doi: 10.1016/j.jag.2012.03.006.

[11] H. Salah and H. Zuhair, “Deep learning in phishing mitigation: a uniform resource locator-based predictive model,” *Int. J. Electr. Comput. Eng.*, vol. 13, no. 3, pp. 3227–3243, 2023, doi: 10.11591/ijece.v13i3.pp3227-3243.

[12] L. N. Smith, “Cyclical learning rates for training neural networks,” *Proc. - 2017 IEEE Winter Conf. Appl. Comput. Vision, WACV 2017*, no. April, pp. 464–472, 2017, doi: 10.1109/WACV.2017.58.

[13] Y. Ida, Y. Fujiwara, and S. Iwamura, “Adaptive learning rate via covariance matrix based preconditioning for deep neural networks,” *IJCAI Int. Jt. Conf. Artif. Intell.*, vol. 0, pp. 1923–1929, 2017, doi: 10.24963/ijcai.2017/267.

[14] Z. Wang, S. Li, B. Wang, X. Ren, and T. Yang, “A malicious url detection model based on convolutional neural network,” *Commun. Comput. Inf. Sci.*, vol. 1298 CCIS, pp. 34–40, 2020, doi: 10.1007/978-981-15-9031-3_3.

[15] N. H. Hassan and A. S. Fakharudin, “Web Phishing Classification Model using Artificial Neural Network and Deep Learning Neural Network,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 7, pp. 535–542, 2023, doi: 10.14569/IJACSA.2023.0140759.

[16] E. Zhu, Y. Chen, C. Ye, X. Li, and F. Liu, “OFS-NN: An Effective Phishing Websites Detection Model Based on Optimal Feature Selection and Neural Network,” *IEEE Access*, vol. 7, pp. 73271–73284, 2019, doi: 10.1109/ACCESS.2019.2920655.

[17] D. Zhang, Z. Yan, H. Jiang, and T. Kim, “A domain-feature enhanced classification model for the detection of Chinese phishing e-Business websites,” *Inf. Manag.*, vol. 51, no. 7, pp. 845–853, 2014, doi: 10.1016/j.im.2014.08.003.

[18] D. J. Liu, G. G. Geng, and X. C. Zhang, “Multi-scale semantic deep fusion models for phishing website detection,” *Expert Syst. Appl.*, vol. 209, no. August, p. 118305, 2022, doi: 10.1016/j.eswa.2022.118305.

[19] S. Ruder, "An overview of gradient descent optimization algorithms," pp. 1–14, 2016, [Online]. Available: http://arxiv.org/abs/1609.04747

[20] J. Wei *et al.*, "Machine learning in materials science," *InfoMat*, vol. 1, no. 3, pp. 338–358, 2019, doi: 10.1002/inf2.12028.

[21] Y. Wang, Y. Li, Y. Song, and X. Rong, "The influence of the activation function in a convolution neural network model of facial expression recognition," *Appl. Sci.*, vol. 10, no. 5, 2020, doi: 10.3390/app10051897.

BIBLIOGRAPHY OF AUTHORS

Isa Suarti received a Bachelor's degree from the Indonesian College of Informatics & Computers (STIKI), in 2016, and is currently completing a master's program in informatics at the science and technology faculty of Maualana Malik Ibrahim State Islamic University, Malang. She can be contacted via email: isa.suarti@stiki.ac.id

Totok Chamidy Received a Bachelor's degree from Brawijaya University, in 1994, and a Master's degree from the Sepuluh Nopember Institute of Technology, in 2002. He obtained a Doctoral degree in 2021 from Malang State University. Since 2006 until now he has been active as a permanent lecturer at the science and technology faculty of Maualana Malik Ibrahim State Islamic University, Malang. He can be contacted via email: to2k2013@ti.uin-malang.ac.id

Cahyo Crysdian Received a Bachelor's degree from Brawijaya University, in 1997, and a Master's degree from Universiti Teknologi Malaysia, in 2003. He received a Doctoral degree in 2006 from Universiti Teknologi Malaysia. Since 2014 until now he has been active as a permanent lecturer at the science and technology faculty of Maualana Malik Ibrahim State Islamic University, Malang. He can be contacted via email: cahyo@ti.uin-malang.ac.id