POLİTEKNİK DERGİSİ

JOURNAL of POLYTECHNIC

ISSN: 1302-0900 (PRINT), ISSN: 2147-9429 (ONLINE)
URL: http://dergipark.org.tr/politeknik

Phishing detection system using extreme learning machines with different activation function based on majority voting

Çoğunluk oylamasına dayalı farklı etkinleştirme işlevine sahip aşırı öğrenme makinelerini kullanan kimlik avı tespit sistemi

***Yazar (Author)**: Murat UÇAR[1]*

ORCID[1]: 0000-0001-9997-4267

To cite to this article: Uçar M., "Phishing detection system using extreme learning machines with different activation function based on majority voting", ***Journal of Polytechnic***, 26(1): 401-414, (2023).

Bu makaleye şu şekilde atıfta bulunabilirsiniz: Uçar M., "Phishing detection system using extreme learning machines with different activation function based on majority voting", ***Politeknik Dergisi***, 26(1): 401-414, (2023).

Erişim linki (To link to this article): http://dergipark.org.tr/politeknik/archive

DOI: 10.2339/politeknik.1098037

Phishing Detection System Using Extreme Learning Machines with Different Activation Function based on Majority Voting

Çoğunluk Oylamasına Dayalı Farklı Etkinleştirme İşlevine Sahip Aşırı Öğrenme Makinelerini Kullanan Kimlik Avı Tespit Sistemi

Highlights

- *ELM model, which provides a faster and generalizable performance was used for phishing detection.*
- *Performances of ELM models with different activation functions were evaluated.*
- *This study provides a fast, low cost, high performance and generalization capacity system.*

Graphical Abstract

In the proposed system, the individual performances of each of the ELM classifiers with different activation functions were evaluated, and then the results of the first three ELM models with the best performance were majority voted and the final result was reached.

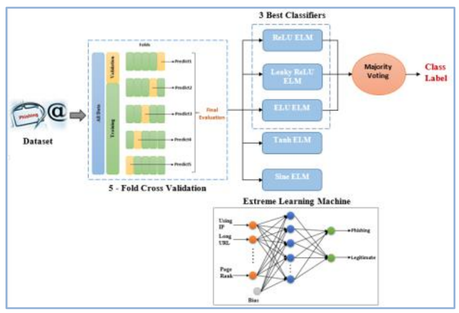

Figure. Structure of the proposed phishing detection model

Aim

Phishing is a type of software-based cyber-attack carried out to steal private information such as login credentials, user passwords, and credit card information. When the security reports published in recent years are examined, it is seen that there are millions of phishing spoofing web pages. Therefore, in this study, it is aimed to develop an effective phishing detection model.

Design & Methodology

In this study, an extreme learning machine based model using different activation functions such as sine, hyperbolic tangent function, rectified linear unit, leaky rectified linear unit and exponential linear unit was proposed and comparative analyses were made. In addition, the performances of the models when combined with the majority vote were also evaluated.

Originality

An overview is presented based on the studies developed for phishing detection in the literature, and a novel and effective model is proposed by combining extreme learning machine models using different activation functions with majority voting.

Findings

In the study, the highest accuracy value of 97.123% was obtained when the three most successful activation functions were combined with the majority vote.

Conclusion

Experimental results show the effectiveness and applicability of the model proposed in the study.

Declaration of Ethical Standards

The author of this article declares that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission

Phishing Detection System Using Extreme Learning Machines with Different Activation Function based on Majority Voting

Araştırma Makalesi / Research Article

Murat UÇAR[1*]

[1]Faculty of Business and Management Science, Department of Management Information Systems, Iskenderun Technical University, Türkiye

(Geliş/Received : 04.04.2022 ; Kabul/Accepted : 03.07.2022 ; Erken Görünüm/Early View : 24.08.2022)

ABSTRACT

Phishing is a type of software-based cyber-attack carried out to steal private information such as login credentials, user passwords, and credit card information. When the security reports published in recent years are examined, it is seen that there are millions of phishing spoofing web pages. Therefore, in this study, it is aimed to develop an effective phishing detection model. In the study, an extreme learning machine based model using different activation functions such as sine, hyperbolic tangent function, rectified linear unit, leaky rectified linear unit and exponential linear unit was proposed and comparative analyses were made. In addition, the performances of the models when combined with the majority vote were also evaluated and it was seen that the highest accuracy value of 97.123% was obtained when the three most successful activation functions were combined with the majority vote. Experimental results show the effectiveness and applicability of the model proposed in the study.

Keywords: Phishing detection, extreme machine learning, majority voting.

Çoğunluk Oylamasına Dayalı Farklı Etkinleştirme İşlevine Sahip Aşırı Öğrenme Makinelerini Kullanan Kimlik Avı Tespit Sistemi

ÖZ

Kimlik avı, oturum açma kimlik bilgileri, kullanıcı şifreleri, kredi kartı bilgileri gibi özel bilgileri çalmak amacıyla gerçekleştirilen yazılım tabanlı bir siber saldırı türüdür. Son yıllarda yayınlanan güvenlik raporları incelendiğinde milyonlarca kimlik avı sahteciliği yapan web sayfasının olduğu görülmektedir. Bu nedenle bu çalışmada etkili bir kimlik avı tespit modelinin geliştirilmesi amaçlanmıştır. Çalışmada sinüs, hiperbolik tanjant fonksiyonu, doğrultulmuş doğrusal birim, sızıntılı doğrultulmuş doğrusal birim ve üstel doğrusal birim gibi farklı aktivasyon fonksiyonlarının kullanıldığı aşırı öğrenme makineleri tabanlı bir model önerilmiş ve karşılaştırmalı analizler yapılmıştır. Ayrıca modellerin çoğunluk oyu ile birleştirildiğindeki performansları da değerlendirilmiş ve en yüksek doğruluk değerinin %97.123 ile en başarılı üç aktivasyon fonksiyonun çoğunluk oyu ile birleştirildiğinde elde edildiği görülmüştür. Deneysel sonuçlar, çalışmada önerilen modelin etkinliğini ve uygulanabilirliğini göstermektedir.

Anahtar Kelimeler: Kimlik avı tespiti, aşırı makine öğrenimi, çoğunluk oylaması.

1. INTRODUCTION

Phishing is a cybercrime aimed at obtaining usernames, passwords and personal financial information using social engineering methods and technological tricks. [1]. In order to obtain this information, fake emails or websites that are very similar to the original are generally used. According to the report of the AntiPhishing Working Group (APWG), the number of phishing attacks has doubled since the beginning of 2020. In addition, 260,642 phishing attacks were seen in July 2021, the highest monthly level compared to previous years [2]. These statistics show that anti-phishing solutions and work need to be improved. One of the most used methods for detecting phishing websites is phishing URL tanks. [3]. However, in order to keep phishing URL tanks up to date, individuals or organizations must manually report phishing websites. This situation can cause problems such as more human effort and not detecting phishing URLs in a timely manner [4].

To tackle these disadvantages of phishing URL tanks, researchers primarily focused on traditional machine learning methodologies that can provide a more intelligent phishing detection [5-12]. In the traditional machine learning approach, feature selection is made with the help of cyber security experts, and then phishing detection is performed by using traditional machine learning algorithms. Deep learning methods, which have come to the forefront with their rapid development and successful results in many different fields in recent years, have also started to be used for phishing detection. [13-17]. In deep learning algorithms, data can be used

*Sorumlu Yazar (Corresponding Author)
e-posta : murat.ucar@iste.edu.tr

directly without the need for a manual feature selection step.

In this study, an extreme learning machine (ELM) based approach is proposed for phishing detection. In the proposed approach, the effect of different activation functions on the prediction accuracy of ELM models was also investigated. In the study, five different activation functions, namely sine, hyperbolic tangent function (Tanh), rectified linear unit (ReLU), leaky RELU and exponential linear unit (ELU), were used and the results obtained from each ELM model were analysed. Then three ELM models with the best performance were determined and the final result was reached by majority voting of these three ELM models. The main contribution of this study are:

- In this study, the ELM model, which provides a faster and generalizable performance and does not require parameters such as learning rate and momentum in classical artificial neural network architectures, was used for phishing detection.
- Performances of ELM models with different activation functions were evaluated. As a result of the experimental tests, it was seen that the three best activation functions in ELM models were ELU, leaky ReLU and RELU, respectively.
- The proposed model that focused majority voting of the ELM models with the three best activation functions reached a high accuracy value of 97.123%.
- In addition, this study provides a fast, low cost, high performance and high generalization capacity system for phishing detection.

The remainder of the article is organized as follows. In Section 2, a brief review of the studies performed for phishing detection is presented. In Section 3, the model and methodology proposed in the study are presented in detail. Section 4 describes the dataset and the experimental considerations and results for the selection of the best parameters for the ELM models used in the study. Section 5 provides detailed performance comparisons of the proposed model and previous work in this area. Finally, the paper concluded in Section 6.

2. RELATED WORK

Researchers have proposed various approaches for phishing detection, including traditional machine learning methods and deep learning-based methods.

Zhu et al. proposed an approach based on optimal feature selection and neural networks for the detection of phishing attacks. The feature selection algorithm designed in the study reduces the time cost as it does not take into account many useless and small-impact features by determining a threshold value. They reported that the proposed approach was successful in detecting many types of phishing websites [1]. Xiang et al. proposed a feature-based model for phishing detection, which they called Cantina+. In the study, in which they evaluated the performance of six different machine learning methods as classifiers, they reported that the best algorithm was the Bayesian network and it performed quite well in catching the ever-evolving new phishing attacks [5]. Şahingöz et al. created and shared a rather large dataset containing 36,400 legitimate and 37,175 phishing records. They utilized seven different machine learning algorithms for real-time phishing detection. They reported that the Random Forest method obtained the highest accuracy with 97.98%, using the features extracted based on natural language processing (NLP) [8]. In another study Rao and Pais used eight different traditional machine learning methods in their study by extracting the heuristic features of phishing sites. Among these models, the RF model achieved the best performance with 99.31% accuracy. In addition, in this study, tests were carried out with all RF types to obtain the best result, and they reported that the highest accuracy value was obtained with 99.55% with the Principal Component Analysis-RF classifier [10]. Priya et al. proposed an approach to detect drive-by download attacks using useful information they extracted by analysing web pages. They achieved 92% accuracy with the KNN algorithm and reported that better performance could be achieved with more HTML and JavaScript features [18]. Toğaçar used support vector machine (SVM), k-nearest neighbor (KNN), decision tree (DT) and random forest (RF) methods from traditional machine learning methods for phishing detection, and obtained the highest accuracy value of 96.73% with the RF method [19]. Similarly, when Koşan et al. compared the performances using C4.5, ID3, PRISM, RIPPER, NB, KNN and RF methods for the detection of phishing web pages, they reported that the best accuracy value was obtained with the RF method with 97.3%. Although the RF method has the best accuracy value, the model creation and estimation time takes a little longer than other methods [20]. Ali and Malebary proposed an approach for phishing detection using feature weighting based on particle swarm optimization (PSO). They indicated that the PSO-based feature weighting proposed in the study had a positive effect on success and reached 96.83% accuracy performance [21]. Minocha and Singh utilized the KNN method as a classifier in their study where they designed a new transfer function for phishing detection. As a result of the performance evaluations of the proposed method, they reported that it produced better results compared to the state-of-the-art techniques [22]. Kaytan and Hanbay used the ELM method to detect phishing websites. The average classification accuracy of the proposed method was 95.05% when the 10-fold cross validation test was applied [23]. Li et al. performed phishing detection using the features they extracted by analysing URL addresses and HTML codes of web pages. In the study, they proposed a stacking model approach by combining various boosting algorithms. They stated that the proposed approach achieved 97.30% and 98.60% accuracy values as a result of the tests performed on two different data sets. The study stands out as a real-time phishing detection system which can be utilized for protecting users from phishing attacks [24]. In another study, Yang et al. noted that they achieved 97.5% accuracy in phishing detection with the improved ELM approach [25]. Savaş and Savaş utilized 8 different machine learning algorithms such as SVM, RF, KNN, DT, Gaussian Naive Bayes, logistic regression, multilayer perceptron and XGBoost to classify the URL addresses whether they are phishing or not. They have reached a high accuracy of 99.8% in many models they tested on the data obtained from USOM, Alexa and Phishtank. [26].

Wei et al. utilized convolutional neural networks (CNN) in the study that they designed a light-weight phishing detection sensor. They reported that the proposed method reached 86.63% accuracy and reduced execution time by 30% [4]. Yang et al. proposed a deep learning-based approach using multidimensional features. As a result of experimental tests, they indicated that the proposed approach provides high accuracy performance quite quickly [16]. Feng et al. proposed a hybrid deep model approach by using a new method called Web2Vec for feature extraction. As a result of the experimental tests, the proposed model reached quite high accuracy performance [17]. Somesha et al. used deep learning methods. They reported that the best performance was obtained with the long short-term memory (LSTM) method with 99.57% in the study, where they minimized the number of features and diminished the dependency on third-party services [27]. Özcan et al. proposed hybrid models called DNN-LSTM and DNN-BiLSTM based on LSTM and deep neural network (DNN) for the detection of phishing attacks. They tested proposed models on two different datasets and reported that the DNN-BiLSTM model achieved a very high performance with 98.79% and 99.21% accuracy rates. They stated that hybrid architectural models give better results thanks to using both NLP features and character embedding features at the same time. [28]. Al-Ahmadi et al. proposed a generative adversarial network-based approach, which they called PDGAN, for the detection of phishing attacks. They tested the proposed approach on a very large dataset created by PhishTank and DomCop and reported that the model achieved an accuracy of 97.58% [29].

3. METHODS

3.1. Proposed Model

The aim of this study is to develop a new ELM based system for phishing detection using the features of a data set obtained from Kaggle, a public data science platform. The architecture of the proposed system is illustrated in Figure 1. In the proposed system, the individual performances of each of the ELM classifiers with different activation functions were evaluated, and then the results of the first three ELM models with the best performance were majority voted and the final result was reached.

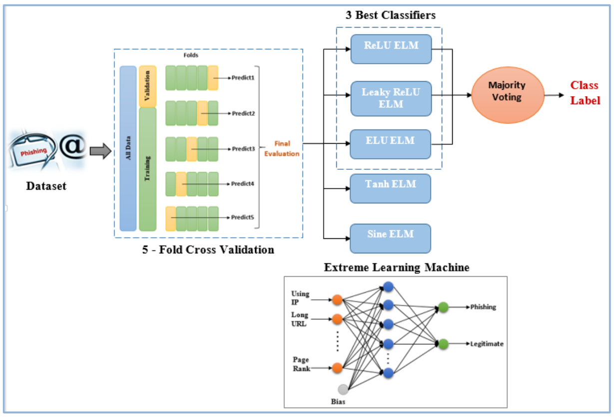

Figure 1. Structure of the proposed phishing detection model

3.2. ELM for Phishing Detection

ELM is a method developed to train single hidden layer feedforward neural networks proposed by Huang et al. in 2006 [30]. In traditional feedforward neural networks, weights and threshold values are adjusted by choosing the most appropriate system to be modelled. In gradient-based learning approaches such as the back propagation learning algorithm, all weights and threshold values are changed iteratively until the training error is minimized. However, the learning process takes a lot of time to achieve the best performance and sometimes the error can be stuck in a local point. Changing the momentum value may prevent the error from getting stuck at a local point, but it will not be useful in shortening the learning process [31]. In ELM, input weights and threshold values are randomly assigned and output weights are calculated accordingly. Therefore, ELM provides faster and better performance in some tasks compared to traditional methods [30, 31]. The structure of the ELM is presented in Figure 2.

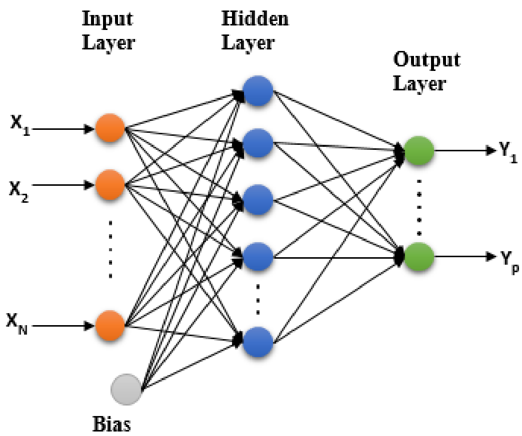

Figure 2. Structure of an ELM network with a single hidden layer

The artificial network shown in the figure $X_1, X_2, X_3, \dots, X_N$ denotes input vectors and Y indicates output vectors. The mathematical representation of this network, where the number of neurons in the hidden layer is M, is as in equation 1.

$$\sum_{i=1}^{M} \beta_i g(W_i X_k + b_i) = Y_k, \;\; k = 1,2,\dots,N \qquad (1)$$

Here, $W_{i1}, W_{i2}, W_{i3}, \dots, W_{iN}$ represent the connection weights between the input layer and hidden layer, while $\beta_{i1}, \beta_{i2}, \beta_{i3}, \dots, \beta_{im}$ indicate the threshold values, b_i hidden layer neurons, Y_k output values and $g(.)$ activation function in the output layer [32].

3.3. ELM Models with Different Activation Functions for Phishing Detection

ELM is a type of algorithm that tends to perform well in extremely fast learning speed, and choosing the right activation function is very important for the prediction performance of ELM. Non-differentiable or discrete activation functions can be used in ELM [31]. In this study, sine, Tanh, ReLU, leaky ReLU and ELU, which are frequently utilized in the literature, were selected.

The sine activation function is sinusoidal in nature. Although the training time is short in this activation function, it causes overfitting problems as it adjusts the weights easily and quickly [33]. The sine activation function has the following form:

$$f(x) = \sin(x) \qquad (2)$$

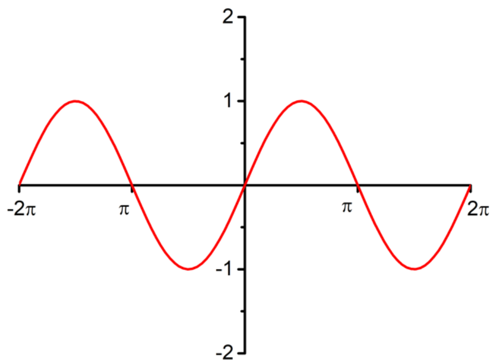

Figure 3. Sine activation function

The Tanh activation function is very similar to the sigmoid activation function, but unlike the sigmoid, it converts inputs to outputs between -1 and +1. This means that its derivative is steeper, that is, it can take more values, and it means that it will be more efficient for the classification process. However, gradient vanishing problem is also a disadvantage of this activation function [34]. The Tanh function is defined as in equation 3.

$$f(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \qquad (3)$$

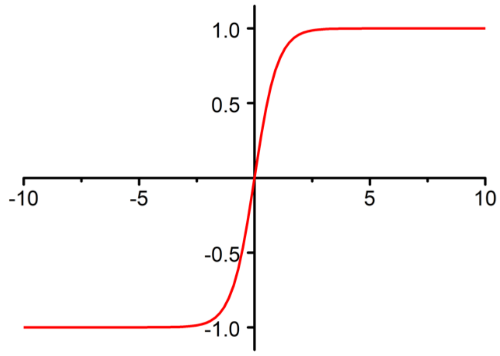

Figure 4. Tanh activation function

The ReLU activation function converts inputs to outputs between 0 and +∞. For this reason, ReLU is called an unsaturated function. The biggest advantage of this function is that the computational load is low and it does not activate all neurons at the same time. It is also resistant to ReLU gradient vanishing problems [35, 36].

$$f(x) = \begin{cases} 0 & x < 0 \\ x & x \geq 0 \end{cases} \qquad (4)$$

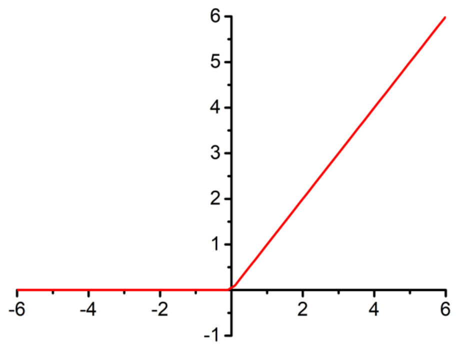

Figure 5. ReLU activation function

Leaky ReLU is one of the solutions developed against the dying ReLU problem, which occurs when the ReLU activation function directly equals negative values to zero. In Leaky ReLU, negative values are very close to zero, but not exactly zero. Thus, its derivative is prevented from being zero, and learning takes place on the negative side as well [36].

$$f(x) = \begin{cases} 0.01x & x < 0 \\ x & x \geq 0 \end{cases} \qquad (5)$$

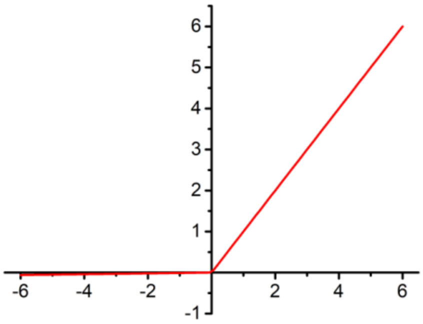

Figure 6. Leaky ReLU activation function

ELU is a more advanced activation function compared to ReLU and has further reduced the gradient vanishing effect. The ELU hyperparameter α controls the value ELU saturates for negative net inputs and has negative values that bring the mean of ELU activations closer to zero. These near-zero activations result in faster learning and higher classification accuracies as the slope approaches the natural gradient [37].

$$f(x) = \begin{cases} \alpha\,(e^{x} - 1) & x \le 0 \\ x & x > 0 \end{cases} \tag{6}$$

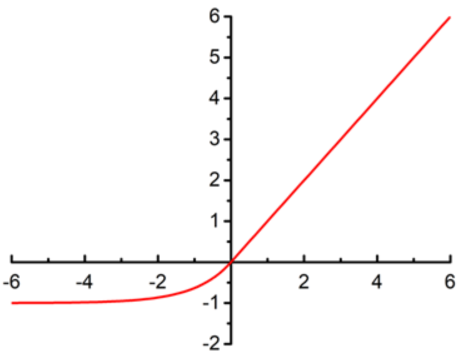

Figure 7. ELU activation function

4. EXPERIMENTAL STUDY

4.1. Data Description

In this study, experiments were carried out on a phishing dataset obtained from Kaggle platform [38]. This dataset were mostly obtained from Phishtank and MillerSmiles archives. It consists of two files, the text-based file containing 11055 website content and "csv" file extension containing 11054 website content. In this study, 11054 examples and 30 features in the csv file were used. The dataset contains 4897 examples classified as phishing and 6157 examples classified as legitimate and is balanced in terms of the distribution of the classes. The dataset is categorized under four main headings: address bar-based features, abnormality-based features, HTML and JavaScript-based features, and domain-based features. These properties contain values between {-1, 1} and {-1, 0, 1}. Among these values, {1} is Legitimate, {0} is Suspicious, and {-1} is Phishing. The 30 features used in the study are presented in Figure 8 [21].

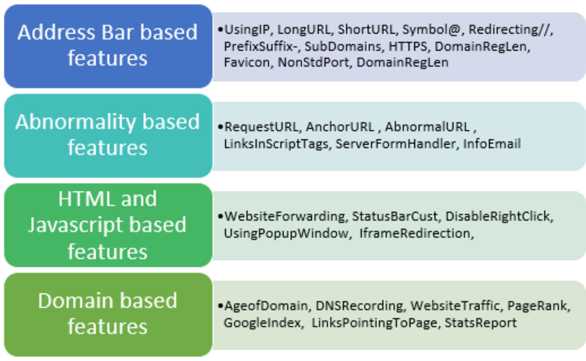

Figure 8. Features in the dataset

4.2. Experimental Evaluation

The proposed model was run on a computer which has Intel Core i5 8250U, 1.60 GHz processor, 12GB RAM and Windows 10 64 bit operating system and it was written with the python programming language. For ELM algorithms with different activation functions used in the study, the number of hidden layer neurons was used as 512, 1024, 2048, 4096 and 6144, respectively. In addition, classification algorithms were applied on the dataset using cross-validation technique. Cross validation is utilized based on the generally accepted and highly reliable 5-fold cross validation techniques. To evaluate ELM models, accuracy (Acc), sensitivity (Sen), precision (Pre), specificity (Spe) and F1 score, which are widely used metrics in many studies, were used. These metrics given in Equation 7-11 are calculated using values such as True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) obtained in the confusion matrix. Here TP occurs when the model correctly predicts an instance belonging to the phishing class. FP occurs when an exemplary model belonging to the legitimate class is mistakenly predicted as phishing. TN occurs when the model correctly predicts an instance of the legitimate class. Finally, an FN occurs when the model incorrectly classifies an instance of the phishing class as legitimate. Accuracy assesses the ability of the proposed model to distinguish between phishing and legitimate examples.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{7}$$

$$Sensitivity = Recall = \frac{TP}{TP+FN} \tag{8}$$

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

$$Specificity = \frac{TN}{TN+FP} \tag{10}$$

$$F1-score = 2\,\frac{Precision \times Recall}{Precision + Recall} \tag{11}$$

4.3. Results

In this section, the results obtained from ELM models with different activation functions and hidden layer neuron numbers are presented in detail. The binary classification performances of the models were evaluated separately for each fold (Appendix A). In addition, an overlapped confusion matrix was created for the general evaluation of the models and performance criteria representing the model in general were calculated using this matrix (Table 1).

Table 1. Performance results of each ELM models.

Number of hidden neurons	Models	Total TP	Total FN	Total FP	Total TN	Performance Results % Spe	Sen	Pre	F1 score	Acc
512	ELU-ELM	4511	386	236	5921	96.167	92.118	95.036	93.551	94.373
	Leaky ReLU-ELM	4517	380	242	5915	96.069	92.240	94.925	93.561	94.373
	ReLU-ELM	4512	385	241	5916	96.086	92.138	94.947	93.517	94.337
	Sine -ELM	4338	559	435	5722	92.935	88.584	90.891	89.719	91.007
	Tanh-ELM	4485	412	254	5903	95.874	91.586	94.646	93.087	93.975
	Overlapped					**95.426**	**91.333**	**94.089**	**92.687**	**93.613**
1024	ELU-ELM	4569	328	185	5972	96.995	93.302	96.134	94.685	95.359
	Leaky ReLU-ELM	4575	322	189	5968	96.930	93.424	96.046	94.711	95.377
	ReLU-ELM	4554	343	197	5960	96.800	92.995	95.863	94.403	95.115
	Sine -ELM	4457	440	313	5844	94.916	91.015	93.454	92.211	93.188
	Tanh-ELM	4561	336	236	5921	96.167	93.138	95.094	94.101	94.825
	Overlapped					**96.362**	**92.775**	**95.318**	**94.022**	**94.773**
2048	ELU-ELM	4625	272	165	5992	97.320	94.445	96.561	95.488	96.047
	Leaky ReLU-ELM	4611	286	171	5986	97.223	94.159	96.430	95.277	95.866
	ReLU-ELM	4630	267	164	5993	97.336	94.547	96.589	95.552	96.101
	Sine -ELM	4551	346	259	5898	95.793	92.934	94.623	93.767	94.527
	Tanh-ELM	4605	292	188	5969	96.946	94.037	96.084	95.047	95.658
	Overlapped					**96.924**	**94.025**	**96.057**	**95.026**	**95.640**
4096	ELU-ELM	4647	250	142	6015	97.694	94.894	97.041	95.952	96.454
	Leaky ReLU-ELM	4629	268	174	5983	97.174	94.527	96.388	95.445	96.001
	ReLU-ELM	4618	279	175	5982	97.158	94.302	96.354	95.315	95.893
	Sine -ELM	4494	403	312	5845	94.933	91.770	93.510	92.628	93.532
	Tanh-ELM	4598	299	193	5964	96.865	93.894	95.980	94.921	95.549
	Overlapped					**96.765**	**93.878**	**95.855**	**94.852**	**95.486**
6144	ELU-ELM	4663	234	132	6025	97.856	95.221	97.252	96.223	96.689
	Leaky ReLU-ELM	4630	267	174	5983	97.174	94.547	96.383	95.454	96.010
	ReLU-ELM	4632	265	177	5980	97.125	94.588	96.324	95.446	96.001
	Sine -ELM	4461	436	308	5849	94.998	91.097	93.559	92.307	93.269
	Tanh-ELM	4609	288	174	5983	97.174	94.119	96.365	95.227	95.820
	Overlapped					**96.865**	**93.915**	**95.976**	**94.931**	**95.558**

When the performances of ELM models with different numbers of hidden layer neurons are examined, it can be seen from Table 1 that the highest accuracy values were obtained by ELM models using the ELU, Leaky ReLU and ReLU activation functions, with accuracy values very close to each other. On the other hand, the ELM model, in which the sine activation function is used, has the lowest accuracy value. In the study, in addition to the individual performance of each classifier, their performance when combined with the majority vote was also evaluated. The values obtained by combining the five classifiers with the majority vote are presented in Table 2.

Table 2. The performance results of majority voting with all ELM model

Number of hidden neurons	Model	Fold	TP	FN	FP	TN	Performance Results %				
							Acc	Sen	Pre	Spe	F1 Score
512	Majority voting with all ELM models	1	911	69	33	1198	95.387	92.959	96.504	97.319	94.699
		2	913	67	58	1173	94.346	93.163	94.027	95.288	93.593
		3	918	61	38	1194	95.522	93.769	96.025	96.916	94.884
		4	898	81	38	1194	94.618	91.726	95.940	96.916	93.786
		5	897	82	57	1174	93.710	91.624	94.025	95.370	92.809
	Overlapped		**4537**	**360**	**224**	**5933**	**94.716**	**92.648**	**95.304**	**96.362**	**93.954**
1024	Majority voting with all ELM models	1	924	56	22	1209	96.472	94.286	97.674	98.213	95.950
		2	927	53	47	1184	95.477	94.592	95.175	96.182	94.882
		3	928	51	35	1197	96.110	94.791	96.366	97.159	95.572
		4	906	73	23	1209	95.658	92.543	97.524	98.133	94.969
		5	911	68	40	1191	95.113	93.054	95.794	96.751	94.404
	Overlapped		**4596**	**301**	**167**	**5990**	**95.766**	**93.853**	**96.507**	**97.288**	**95.155**
2048	Majority voting with all ELM models	1	939	41	22	1209	97.151	95.816	97.711	98.213	96.754
		2	936	44	42	1189	96.110	95.510	95.706	96.588	95.608
		3	939	40	25	1207	97.060	95.914	97.407	97.971	96.655
		4	924	55	18	1214	96.698	94.382	98.089	98.539	96.200
		5	921	58	30	1201	96.018	94.076	96.845	97.563	95.440
	Overlapped		**4659**	**238**	**137**	**6020**	**96.608**	**95.140**	**97.151**	**97.775**	**96.131**
4096	Majority voting with all ELM models	1	942	38	14	1217	97.648	96.122	98.536	98.863	97.314
		2	944	36	27	1204	97.151	96.327	97.219	97.807	96.771
		3	941	38	26	1206	97.105	96.118	97.311	97.890	96.711
		4	928	51	18	1214	96.879	94.791	98.097	98.539	96.416
		5	932	47	31	1200	96.471	95.199	96.781	97.482	95.984
	Overlapped		**4687**	**210**	**116**	**6041**	**97.051**	**95.711**	**97.589**	**98.116**	**96.639**
6144	Majority voting with all ELM models	1	938	42	25	1206	96.970	95.714	97.404	97.969	96.552
		2	948	32	25	1206	97.422	96.735	97.431	97.969	97.081
		3	944	35	25	1207	97.286	96.425	97.420	97.971	96.920
		4	929	50	12	1220	97.196	94.893	98.725	99.026	96.771
		5	928	51	31	1200	96.290	94.791	96.767	97.482	95.769
	Overlapped		**4687**	**210**	**118**	**6039**	**97.033**	**95.711**	**97.549**	**98.083**	**96.619**

In addition, the results obtained by combining the three ELM models which have the highest accuracy with the majority vote are also evaluated and presented in Table 3. When Table 2 and Table 3 are compared, it is seen that the performance in the case of combining the three models which have the highest accuracy values with the majority vote is higher than the performance in the case of combining all the models with the majority vote.

Table 3. The performance results of majority voting with best three ELM models

Number of hidden neurons	Model	Fold	TP	FN	FP	TN	Performance Results %				
							Acc	Sen	Pre	Spe	F1 Score
512	Majority voting with best three ELM models	1	913	67	30	1201	95.613	93.163	96.819	97.563	94.956
		2	911	69	56	1175	94.346	92.959	94.209	95.451	93.580
		3	918	61	38	1194	95.522	93.769	96.025	96.916	94.884
		4	897	82	40	1192	94.482	91.624	95.731	96.753	93.633
		5	897	82	57	1174	93.710	91.624	94.025	95.370	92.809
	Overlapped		**4536**	**361**	**221**	**5936**	**94.735**	**92.628**	**95.362**	**96.410**	**93.972**
1024	Majority voting with best three ELM models	1	922	58	23	1208	96.336	94.082	97.566	98.132	95.792
		2	924	56	47	1184	95.341	94.286	95.160	96.182	94.721
		3	924	55	36	1196	95.884	94.382	96.250	97.078	95.307
		4	906	73	19	1213	95.839	92.543	97.946	98.458	95.168
		5	908	71	39	1192	95.023	92.748	95.882	96.832	94.289
	Overlapped		**4584**	**313**	**164**	**5993**	**95.685**	**93.608**	**96.561**	**97.336**	**95.055**
2048	Majority voting with best three ELM models	1	941	39	23	1208	97.196	96.020	97.614	98.132	96.811
		2	942	38	40	1191	96.472	96.122	95.927	96.751	96.024
		3	936	43	24	1208	96.970	95.608	97.500	98.052	96.545
		4	918	61	18	1214	96.427	93.769	98.077	98.539	95.875
		5	920	59	36	1195	95.701	93.973	96.234	97.076	95.090
	Overlapped		**4657**	**240**	**141**	**6016**	**96.553**	**95.099**	**97.070**	**97.710**	**96.069**
4096	Majority voting with best three ELM models	1	939	41	16	1215	97.422	95.816	98.325	98.700	97.054
		2	948	32	31	1200	97.151	96.735	96.834	97.482	96.784
		3	939	40	25	1207	97.060	95.914	97.407	97.971	96.655
		4	926	53	16	1216	96.879	94.586	98.301	98.701	96.408
		5	928	51	31	1200	96.290	94.791	96.767	97.482	95.769
	Overlapped		**4680**	**217**	**119**	**6038**	**96.960**	**95.568**	**97.527**	**98.067**	**96.534**
6144	Majority voting with best three ELM models	1	941	39	25	1206	97.105	96.020	97.412	97.969	96.711
		2	944	36	27	1204	97.151	96.327	97.219	97.807	96.771
		3	945	34	25	1207	97.332	96.527	97.423	97.971	96.973
		4	934	45	8	1224	97.603	95.403	99.151	99.351	97.241
		5	929	50	29	1202	96.425	94.893	96.973	97.644	95.922
	Overlapped		**4693**	**204**	**114**	**6043**	**97.123**	**95.834**	**97.636**	**98.148**	**96.723**

Individually and overlapped confusion matrices for each fold in the case of combining the three best ELM models with 6144 hidden neurons, where the most successful accuracy value was obtained, are presented in Figure 9.

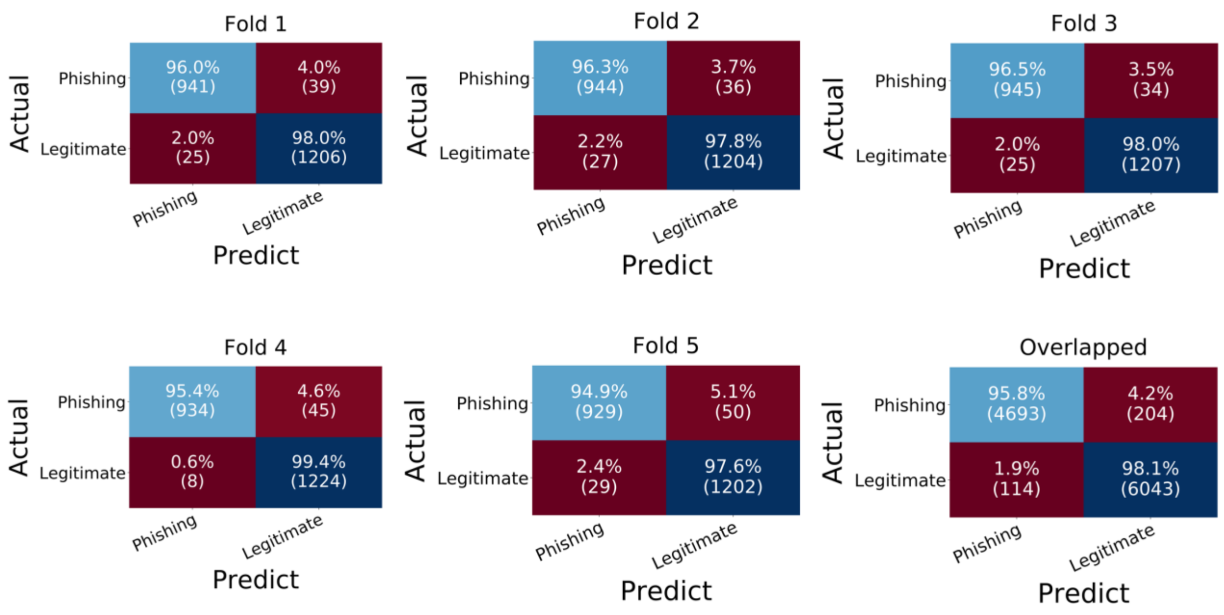

Figure 9. Confusion matrices of majority voting with best three ELM models

In addition, the performance of the model obtained as a consequence of combining the best three ELM models with the majority vote was also evaluated according to the ROC curve metric and presented in Figure 10.

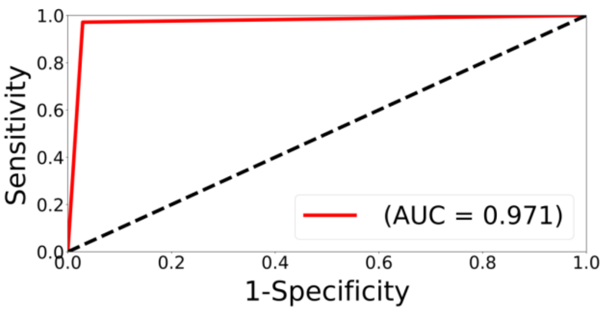

Figure 10. The ROC curve of majority voting with best three ELM models

5. DISCUSSION

Especially in recent years, it has been seen that researchers have carried out studies on the detection of web pages related to phishing fraud, which has increased with the rise in web applications. While traditional machine learning methods are used in many studies, it is noteworthy that deep learning methods have also been used, especially in recent years. In studies using traditional machine learning methods, it was observed that the best performance was mostly obtained with the Random Forest algorithm [8, 10, 19, 20, 21, 26]. When the studies using deep learning methods were examined, it was seen that the LSTM model came to the fore and achieved high accuracy values [17, 27, 28, 29]. In the study, the performance of the proposed method was compared directly with only studies using the same dataset for a fair comparison, and these studies were summarized in Table 4.

Table 4. Comparison of the results of ELM model with related studies

Author	Method	Acc (%)	Sen (%)	Spe (%)
Toğaçar [19]	SVM, KNN, DT, RF	RF: 96.53	RF: 97.88	RF: 94.86
Koşan et al [20]	C4.5, ID3, PRISM, RIPPER, NB, KNN, RF	RF: 97.3	-	-
Ali and Malebary [21]	ML models with PSO based feature weighting	RF-PSO: 96.83	RF-PSO: 95.37	RF-PSO: 98.00
Kaytan and Hanbay [23]	ELM	ELM: 95.93	-	-
Proposed Model	**Majority voting of ELM models with different activation functions**	**ELM: 97.12**	**95.83**	**98.15**

As can be seen from Table 4, Toğaçar [19], Koşan et al. [20] and Ali and Malebary[21] used various traditional machine learning methods to detect phishing websites, and when they evaluated the performances of these models, all three of them achieved the best results with RF machine learning. Another study using this dataset belongs to Kaytan and Hanbay [23]. Kaytan and Hanbay achieved 95.93% accuracy performance with the ELM model they analysed using 10-fold cross-validation technique. In this study, the ELM method was used similarly to Kaytan and Hanbay. However, in this study, the individual achievements of five ELM models using different activation functions and then the success of these models by combining them with the majority vote were evaluated. In this study, the highest accuracy value was obtained as 97.12% by combining the three ELM models with the best individual accuracy with the majority vote. It has been observed that this result is very close to Koşan et al [20], which has the highest accuracy value in Table 4, and also that combining ELM models with different activation functions with majority vote positively affects the classification performance.

6. CONCLUSION

In this paper, ELM models using different activation functions are proposed for effective and efficient phishing detection. Then, the most successful three of these ELM models were combined with the majority vote and the final result was reached. The 5-fold cross-validation technique was used to evaluate the performance of the proposed model in the study. In consequence of comprehensive evaluations, it has been observed that the highest accuracy value of the proposed method is 97.123%. It is thought that the proposed ELM model in the study will contribute to the literature in terms of having a faster and effective performance compared to classical artificial neural networks and providing a high performance at a lower cost.

In future studies, it is planned to observe the performance of the proposed method by evaluating it on larger and different datasets.

DECLARATION OF ETHICAL STANDARDS

The author of this article declares that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

AUTHORS' CONTRIBUTIONS

Murat UÇAR: Performed the study, analysed the results and wrote the manuscript.

CONFLICT OF INTEREST

There is no conflict of interest in this study.

REFERENCES

[1] Zhu, E., Chen, Y., Ye, C., Li, X., & Liu, F., "OFS-NN: an effective phishing websites detection model based on optimal feature selection and neural network", ***IEEE Access***, *7*, 73271-73284, (2019).

[2] Anti-Phishing Working Group, "Phishing Activity Trends Report 3rd Quarter 2021," https://apwg.org/trendsreports/#:~:text=APWG%20saw%20260%2C642%20phishing%20attacks,monthly%20in%20APWG's%20reporting%20history.&text=The%20number%20of%20brands%20being,Q2%20to%207%2C741%20in%20Q3 Erişim Tarihi: 03.01.2022

[3] Phishtank, https://www.phishtank.com/ Erişim Tarihi 10.01.2022.

[4] Wei, B., Hamad, R. A., Yang, L., He, X., Wang, H., Gao, B., & Woo, W. L., "A deep-learning-driven light-weight phishing detection sensor", ***Sensors***, *19*(19): 4258, (2019).

[5] Xiang, G., Hong, J., Rose, C. P., & Cranor, L., "Cantina+ a feature-rich machine learning framework for detecting phishing web sites", ***ACM Transactions on Information and System Security (TISSEC)***, *14*(2): 1-28, (2011).

[6] El-Alfy, E. S. M., "Detection of phishing websites based on probabilistic neural networks and K-medoids clustering", ***The Computer Journal***, *60*(12): 1745-1759, (2017).

[7] Jain, A. K., & Gupta, B. B., "Towards detection of phishing websites on client-side using machine learning based approach". ***Telecommunication Systems***, *68*(4): 687-700, (2018).

[8] Sahingoz, O. K., Buber, E., Demir, O., & Diri, B.,"Machine learning based phishing detection from URLs", ***Expert Systems with Applications***, *117*, 345-357, (2019).

[9] Chiew, K. L., Tan, C. L., Wong, K., Yong, K. S., & Tiong, W. K., "A new hybrid ensemble feature selection framework for machine learning-based phishing detection system", ***Information Sciences***, *484*, 153-166, (2019).

[10] Rao, R. S., & Pais, A. R., "Detection of phishing websites using an efficient feature-based machine learning framework", ***Neural Computing and Applications***, *31*(8): 3851-3873, (2019).

[11] Kasım Ö., "Malicious XSS code detection with decision tree", ***Politeknik Dergisi***, 23(1): 67-72, (2020).

[12] Çıtlak, O., Dörterler, M. & Dogru, İ. "A Hybrid Spam Detection Framework for Social Networks", ***Politeknik Dergisi,*** 1-1. (2022).

[13] Uçar, E., Ucar, M., and İncetaş, M. O., "A Deep learning approach for detection of malicious URLs", ***In 6th International Management Information Systems Conference,*** pp.10-17, (2019).

[14] Bahnsen, A. C., Bohorquez, E. C., Villegas, S., Vargas, J., & González, F. "Classifying phishing URLs using recurrent neural networks", **In *2017 APWG symposium on electronic crime research (eCrime)***, IEEE, pp.1-8, (2017).

[15] Yi, P., Guan, Y., Zou, F., Yao, Y., Wang, W., & Zhu, T., "Web phishing detection using a deep learning framework", ***Wireless Communications and Mobile Computing***, (*2018)*.

[16] Yang, P., Zhao, G., & Zeng, P., "Phishing website detection based on multidimensional features driven by deep learning", ***IEEE Access***, *7*, 15196-15209, (2019).

[17] Feng, J., Zou, L., Ye, O., & Han, J., "Web2Vec: Phishing Webpage Detection Method Based on Multidimensional Features Driven by Deep Learning", ***IEEE Access***, *8*, 221214-221224, (2020).

[18] Priya, M., Sandhya, L., & Thomas, C., "A static approach to detect drive-by-download attacks on webpages", **In *2013 International Conference on Control Communication and Computing (ICCC),*** IEEE, pp. 298-303, (2013).

[19] Toğaçar, M., "Web Sitelerinde Gerçekleştirilen Oltalama Saldırılarının Yapay Zekâ Yaklaşımı ile Tespiti. ***Bitlis***

Eren Üniversitesi Fen Bilimleri Dergisi, 10(4): 1603-1614, (2021).

[20] Koşan, M. A., YILDIZ, O., & Karacan, H., "Kimlik avı web sitelerinin tespitinde makine öğrenmesi algoritmalarının karşılaştırmalı analizi", ***Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi***, *24*(2): 276-282, (2018).

[21] Ali, W., & Malebary, S., "Particle swarm optimization-based feature weighting for improving intelligent phishing website detection", ***IEEE Access***, *8,* 116766-116780, (2020).

[22] Minocha, S., & Singh, B., "A novel phishing detection system using binary modified equilibrium optimizer for feature selection", ***Computers & Electrical Engineering***, *98*, 107689, (2022).

[23] Kaytan, M., & Hanbay, D., "Effective classification of phishing web pages based on new rules by using extreme learning machines", ***Computer Science***, *2*(1): 15-36, (2017).

[24] Li, Y., Yang, Z., Chen, X., Yuan, H., & Liu, W., "A stacking model using URL and HTML features for phishing webpage detection", ***Future Generation Computer Systems***, *94*, 27-39, (2019).

[25] Yang, L., Zhang, J., Wang, X., Li, Z., Li, Z., & He, Y., "An improved ELM-based and data preprocessing integrated approach for phishing detection considering comprehensive features", ***Expert Systems with Applications***, *165*, 113863, (2021).

[26] Savaş, T. & Savaş, S. "Tekdüzen Kaynak Bulucu Yoluyla Kimlik Avı Tespiti için Makine Öğrenmesi Algoritmalarının Özellik Tabanlı Performans Karşılaştırması", ***Politeknik Dergisi,*** 1-1. (2021).

[27] Somesha, M., Pais, A. R., Rao, R. S., & Rathour, V. S., "Efficient deep learning techniques for the detection of phishing websites", ***Sādhanā***, *45*(1): 1-18, (2020).

[28] Ozcan, A., Catal, C., Donmez, E., & Senturk, B. "A hybrid DNN–LSTM model for detecting phishing URLs", ***Neural Computing and Applications***, 1-17. (2021).

[29] Al-Ahmadi, S., Alotaibi, A., & Alsaleh, O. "PDGAN: Phishing Detection with Generative Adversarial Networks", ***IEEE Access,*** (2022).

[30] Huang, G. B., Zhu, Q. Y., & Siew, C. K., "Extreme learning machine: theory and applications", ***Neurocomputing***, *70*(1-3): 489-501, (2006).

[31] Suresh, S., Saraswathi, S., & Sundararajan, N., "Performance enhancement of extreme learning machine for multi-category sparse data classification problems", ***Engineering Applications of Artificial Intelligence***, *23*(7): 1149-1157, (2010).

[32] Kaya, Y., & Tekin, R., "Epileptik nöbetlerin tespiti için aşırı öğrenme makinesi tabanlı uzman bir system", ***Bilişim Teknolojileri Dergisi***, *5*(2): 33-40, (2012).

[33] Sopena, J. M., Romero, E., & Alquezar, R., "Neural networks with periodic and monotonic activation functions: a comparative study in classification problems", In ***9th International Conference on Artificial Neural Networks: ICANN '99***, (1999).

[34] Sharma, S., Sharma, S., & Athaiya, A., "Activation functions in neural networks", ***towards data science***, *6*(12): 310-316, (2017).

[35] Nair, V., & Hinton, G. E., "Rectified linear units improve restricted boltzmann machines", **In *Icml***, (2010).

[36] Pedamonti, D., "Comparison of non-linear activation functions for deep neural networks on MNIST classification task", ***arXiv preprint arXiv:1804.02763***, (2018).

[37] Clevert, D. A., Unterthiner, T., & Hochreiter, S., "Fast and accurate deep network learning by exponential linear units (elus)", ***arXiv preprint arXiv:1511.07289***, (2015).

[38] Dataset, Chand E. 2021. Phishing website Detector. Kaggle. https://www.kaggle.com/datasets/eswarchandt/phishing-website-detector Erişim Tarihi: 05.12.2021

APPENDIX A

Number of Hidden Neurons	Model	Fold	TP	FN	FP	TN	Performance Results %				
							Acc	Sen	Pre	Spe	F1 Score
512	ELU-ELM	1	901	79	34	1197	94.889	91.939	96.364	97.238	94.099
		2	907	73	58	1173	94.075	92.551	93.990	95.288	93.265
		3	914	65	45	1187	95.025	93.361	95.308	96.347	94.324
		4	898	81	43	1189	94.392	91.726	95.430	96.510	93.542
		5	891	88	56	1175	93.484	91.011	94.087	95.451	92.523
	Leaky ReLU-ELM	1	903	77	34	1197	94.980	92.143	96.371	97.238	94.210
		2	908	72	58	1173	94.120	92.653	93.996	95.288	93.320
		3	910	69	45	1187	94.844	92.952	95.288	96.347	94.105
		4	905	74	40	1192	94.844	92.441	95.767	96.753	94.075
		5	891	88	65	1166	93.077	91.011	93.201	94.720	92.093
	ReLU-ELM	1	908	72	34	1197	95.206	92.653	96.391	97.238	94.485
		2	905	75	62	1169	93.804	92.347	93.588	94.963	92.964
		3	908	71	42	1190	94.889	92.748	95.579	96.591	94.142
		4	894	85	36	1196	94.527	91.318	96.129	97.078	93.662
		5	897	82	67	1164	93.258	91.624	93.050	94.557	92.331
	Sine - ELM	1	869	111	73	1158	91.678	88.673	92.251	94.070	90.427
		2	885	95	90	1141	91.633	90.306	90.769	92.689	90.537
		3	871	108	93	1139	90.909	88.968	90.353	92.451	89.655
		4	863	116	81	1151	91.090	88.151	91.419	93.425	89.756
		5	850	129	98	1133	89.729	86.823	89.662	92.039	88.220
	Tanh-ELM	1	900	80	43	1188	94.437	91.837	95.440	96.507	93.604
		2	906	74	62	1169	93.849	92.449	93.595	94.963	93.018
		3	911	68	43	1189	94.980	93.054	95.493	96.510	94.258
		4	888	91	42	1190	93.985	90.705	95.484	96.591	93.033
		5	880	99	64	1167	92.624	89.888	93.220	94.801	91.524
1024	ELU-ELM	1	910	70	20	1211	95.929	92.857	97.849	98.375	95.288
		2	934	46	55	1176	95.432	95.306	94.439	95.532	94.870
		3	918	61	37	1195	95.568	93.769	96.126	96.997	94.933
		4	907	72	25	1207	95.613	92.646	97.318	97.971	94.924
		5	900	79	48	1183	94.253	91.931	94.937	96.101	93.409
	Leaky ReLU-ELM	1	919	61	28	1203	95.975	93.776	97.043	97.725	95.381
		2	920	60	55	1176	94.799	93.878	94.359	95.532	94.118
		3	930	49	32	1200	96.336	94.995	96.674	97.403	95.827
		4	902	77	26	1206	95.341	92.135	97.198	97.890	94.599
		5	904	75	48	1183	94.434	92.339	94.958	96.101	93.630
	ReLU-ELM	1	917	63	29	1202	95.839	93.571	96.934	97.644	95.223
		2	920	60	49	1182	95.070	93.878	94.943	96.019	94.407
		3	915	64	45	1187	95.070	93.463	95.313	96.347	94.379
		4	900	79	29	1203	95.115	91.931	96.878	97.646	94.340
		5	902	77	45	1186	94.480	92.135	95.248	96.344	93.666
	Sine - ELM	1	893	87	45	1186	94.030	91.122	95.203	96.344	93.118
		2	891	89	71	1160	92.763	90.918	92.620	94.232	91.761
		3	906	73	69	1163	93.578	92.543	92.923	94.399	92.733
		4	876	103	54	1178	92.899	89.479	94.194	95.617	91.776
		5	891	88	74	1157	92.670	91.011	92.332	93.989	91.667
	Tanh-ELM	1	915	65	32	1199	95.613	93.367	96.621	97.400	94.966
		2	921	59	56	1175	94.799	93.980	94.268	95.451	94.124
		3	918	61	57	1175	94.663	93.769	94.154	95.373	93.961
		4	906	73	35	1197	95.115	92.543	96.281	97.159	94.375
		5	901	78	56	1175	93.937	92.033	94.148	95.451	93.079
2048	ELU-ELM	1	936	44	29	1202	96.698	95.510	96.995	97.644	96.247
		2	931	49	43	1188	95.839	95.000	95.585	96.507	95.292
		3	929	50	27	1205	96.517	94.893	97.176	97.808	96.021
		4	911	68	26	1206	95.749	93.054	97.225	97.890	95.094
		5	918	61	40	1191	95.430	93.769	95.825	96.751	94.786

	Leaky ReLU-ELM	1	928	52	27	1204	96.427	94.694	97.173	97.807	95.917
		2	931	49	40	1191	95.975	95.000	95.881	96.751	95.438
		3	932	47	36	1196	96.246	95.199	96.281	97.078	95.737
		4	910	69	25	1207	95.749	92.952	97.326	97.971	95.089
		5	910	69	43	1188	94.932	92.952	95.488	96.507	94.203
	ReLU-ELM	1	939	41	30	1201	96.789	95.816	96.904	97.563	96.357
		2	937	43	45	1186	96.020	95.612	95.418	96.344	95.515
		3	923	56	28	1204	96.201	94.280	97.056	97.727	95.648
		4	918	61	20	1212	96.336	93.769	97.868	98.377	95.775
		5	913	66	41	1190	95.158	93.258	95.702	96.669	94.465
	Sine - ELM	1	922	58	45	1186	95.341	94.082	95.346	96.344	94.710
		2	912	68	68	1163	93.849	93.061	93.061	94.476	93.061
		3	920	59	46	1186	95.251	93.973	95.238	96.266	94.602
		4	903	76	40	1192	94.754	92.237	95.758	96.753	93.965
		5	894	85	60	1171	93.439	91.318	93.711	95.126	92.499
	Tanh-ELM	1	924	56	30	1201	96.110	94.286	96.855	97.563	95.553
		2	923	57	46	1185	95.341	94.184	95.253	96.263	94.715
		3	928	51	42	1190	95.794	94.791	95.670	96.591	95.228
		4	921	58	25	1207	96.246	94.076	97.357	97.971	95.688
		5	909	70	45	1186	94.796	92.850	95.283	96.344	94.051
4096	ELU-ELM	1	936	44	16	1215	97.286	95.510	98.319	98.700	96.894
		2	944	36	36	1195	96.744	96.327	96.327	97.076	96.327
		3	931	48	27	1205	96.608	95.097	97.182	97.808	96.128
		4	919	60	24	1208	96.201	93.871	97.455	98.052	95.630
		5	917	62	39	1192	95.430	93.667	95.921	96.832	94.780
	Leaky ReLU-ELM	1	932	48	33	1198	96.336	95.102	96.580	97.319	95.835
		2	922	58	33	1198	95.884	94.082	96.545	97.319	95.297
		3	932	47	46	1186	95.794	95.199	95.297	96.266	95.248
		4	922	57	19	1213	96.563	94.178	97.981	98.458	96.042
		5	921	58	43	1188	95.430	94.076	95.539	96.507	94.802
	ReLU-ELM	1	929	51	25	1206	96.563	94.796	97.379	97.969	96.070
		2	931	49	40	1191	95.975	95.000	95.881	96.751	95.438
		3	925	54	39	1193	95.794	94.484	95.954	96.834	95.214
		4	915	64	28	1204	95.839	93.463	97.031	97.727	95.213
		5	918	61	43	1188	95.294	93.769	95.525	96.507	94.639
	Sine - ELM	1	909	71	56	1175	94.256	92.755	94.197	95.451	93.470
		2	907	73	69	1162	93.578	92.551	92.930	94.395	92.740
		3	908	71	65	1167	93.849	92.748	93.320	94.724	93.033
		4	890	89	55	1177	93.487	90.909	94.180	95.536	92.516
		5	880	99	67	1164	92.489	89.888	92.925	94.557	91.381
	Tanh-ELM	1	925	55	28	1203	96.246	94.388	97.062	97.725	95.706
		2	929	51	47	1184	95.568	94.796	95.184	96.182	94.990
		3	923	56	39	1193	95.703	94.280	95.946	96.834	95.106
		4	906	73	31	1201	95.296	92.543	96.692	97.484	94.572
		5	915	64	48	1183	94.932	93.463	95.016	96.101	94.233
6144	ELU-ELM	1	931	49	21	1210	96.834	95.000	97.794	98.294	96.377
		2	944	36	27	1204	97.151	96.327	97.219	97.807	96.771
		3	936	43	36	1196	96.427	95.608	96.296	97.078	95.951
		4	929	50	19	1213	96.879	94.893	97.996	98.458	96.419
		5	923	56	29	1202	96.154	94.280	96.954	97.644	95.598
	Leaky ReLU-ELM	1	935	45	40	1191	96.156	95.408	95.897	96.751	95.652
		2	927	53	32	1199	96.156	94.592	96.663	97.400	95.616
		3	934	45	32	1200	96.517	95.403	96.687	97.403	96.041
		4	921	58	23	1209	96.336	94.076	97.564	98.133	95.788
		5	913	66	47	1184	94.887	93.258	95.104	96.182	94.172
	ReLU-ELM	1	937	43	33	1198	96.563	95.612	96.598	97.319	96.103
		2	927	53	32	1199	96.156	94.592	96.663	97.400	95.616
		3	928	51	40	1192	95.884	94.791	95.868	96.753	95.326
		4	920	59	26	1206	96.156	93.973	97.252	97.890	95.584
		5	920	59	46	1185	95.249	93.973	95.238	96.263	94.602

	Sine -ELM	1	895	85	46	1185	94.075	91.327	95.112	96.263	93.181
		2	891	89	55	1176	93.487	90.918	94.186	95.532	92.523
		3	890	89	61	1171	93.216	90.909	93.586	95.049	92.228
		4	897	82	58	1174	93.668	91.624	93.927	95.292	92.761
		5	888	91	88	1143	91.900	90.705	90.984	92.851	90.844
	Tanh-ELM	1	919	61	35	1196	95.658	93.776	96.331	97.157	95.036
		2	931	49	38	1193	96.065	95.000	96.078	96.913	95.536
		3	925	54	30	1202	96.201	94.484	96.859	97.565	95.657
		4	913	66	30	1202	95.658	93.258	96.819	97.565	95.005
		5	921	58	41	1190	95.520	94.076	95.738	96.669	94.900