Computers, Materials & Continua
DOI: 10.32604/cmc.2023.030784
Article

Intelligent Deep Learning Based Cybersecurity Phishing Email Detection and Classification

R. Brindha[1], S. Nandagopal[2], H. Azath[3], V. Sathana[4], Gyanendra Prasad Joshi[5] and Sung Won Kim[6,*]

[1]Department of Computing Technologies, SRM Institute of Science and Technology, Kattankulathur, 603203, India
[2]Department of Computing Science and Engineering, Nandha College of Technology, Erode, 638052, India
[3]School of Computing Science and Engineering, VIT Bhopal University, Bhopal, 466114, India
[4]Department of Computer Science and Engineering, K.Ramakrishnan College of Engineering, Tiruchirappalli, 621112, India
[5]Department of Computer Science and Engineering, Sejong University, Seoul, 05006, Korea
[6]Department of Information and Communication Engineering, Yeungnam University, Gyeongsan-si, 38541, Gyeongbuk-do, Korea
*Corresponding Author: Sung Won Kim. Email: swon@yu.ac.kr
Received: 01 April 2022; Accepted: 12 October 2022

Abstract: Phishing is a type of cybercrime in which cyber-attackers pose themselves as authorized persons or entities and hack the victims' sensitive data. E-mails, instant messages and phone calls are some of the common modes used in cyberattacks. Though the security models are continuously upgraded to prevent cyberattacks, hackers find innovative ways to target the victims. In this background, there is a drastic increase observed in the number of phishing emails sent to potential targets. This scenario necessitates the importance of designing an effective classification model. Though numerous conventional models are available in the literature for proficient classification of phishing emails, the Machine Learning (ML) techniques and the Deep Learning (DL) models have been employed in the literature. The current study presents an Intelligent Cuckoo Search (CS) Optimization Algorithm with a Deep Learning-based Phishing Email Detection and Classification (ICSOA-DLPEC) model. The aim of the proposed ICSOA-DLPEC model is to effectually distinguish the emails as either legitimate or phishing ones. At the initial stage, the pre-processing is performed through three stages such as email cleaning, tokenization and stop-word elimination. Then, the N-gram approach is; moreover, the CS algorithm is applied to extract the useful feature vectors. Moreover, the CS algorithm is employed with the Gated Recurrent Unit (GRU) model to detect and classify phishing emails. Furthermore, the CS algorithm is used to fine-tune the parameters involved in the GRU model. The performance of the proposed ICSOA-DLPEC model was experimentally validated using a benchmark dataset, and the results were assessed under several dimensions. Extensive comparative studies were conducted, and the results confirmed the superior performance of the proposed ICSOA-DLPEC model over other existing approaches. The proposed model achieved a maximum accuracy of 99.72%.

Keywords: Phishing email; data classification; natural language processing; deep learning; cybersecurity

1 Introduction

Nowadays, phishing has become a more lucrative style of committing fraudulent activities than ever before [1]. According to criminal law, fraud can be defined as a dishonest action of an individual with an intention to gain personal interests or reveal an individual's image. The online-based fraudulent activities include misleading individuals into smearing their private data to achieve personal or financial gains. Likewise, Phishing is an act in which a cyber-attacker attempts to automatically gain confidential or delicate data from the users for the purpose of stealing. This is done by mimicking the original websites and projecting them as the real ones [2]. Usually, the phishing attack is executed through electronic devices (namely, computers and ipads) and computer networks; the cyber-attackers wait for the right opportunity, find the vulnerable places in the recognition system, bypass the security features and steal the valuable data of the end-users. Such vulnerable places are taken into account as weak components in a security chain [3,4]. As described earlier, the Phishing attacker interacts well with the users through socially-engineered messages and persuades them to disclose their private data. The collected data is then used by fraudsters to obtain illegal access to a user's private and secure data. Phishing extracts complex data from the incautious victims and remains a social engineering-based threat. Some of the transmission networks that are frequently utilized to make such types of attacks include instant messaging, emails and so on. The attacker appears as a credible and legitimate individual. Since email is commonly used by such attackers, they emphasize email messages in their work [5].

As per the literature [6], it is challenging to detect phishing messages and emails automatically. This study discussed various techniques for the detection of phishing emails. It discussed about the Deep Learning (DL) method, blacklisting and the Machine Learning (ML)-based classification algorithms for the detection of phishing emails [7]. The existing blacklist method primarily depends on personal reports, whereas phishing emails are identified after spending too much time and workforce. On the other hand, the ML classification method-based phishing detection techniques utilize Artificial Intelligence (AI) methods to detect phishing attacks. Feature engineering is important for the automatic recognition of the representation features, whereas it is not possible when the information migration scenario is applied [8]. Moreover, the recognition methods established using the DL techniques are constrained by word embedding in email content representations. In such case, the output remains opaque since the technique falls short in recognizing the significance of the phishing emails. So, both DL and the Natural Language Processing (NLP) technologies must be interchanged [9]. Following a dramatic expansion in the DL method-applications, considerable attention has been paid by the researchers to investigative phishing [10]. In contrast to the conventional ML methods, the DL techniques integrally collect the hand-engineered aspects in such a way that the ML specialists add the information without any need to obtain information on cyber security.

The authors in literature [11] validated the performance of the Convolutional Neural Network (CNN) techniques in recognizing phishing attacks through content analysis of email messages. These techniques take the embedded texts from the email body as input, and the outcome is as follows; probable demonstration of a message, whether it is malicious or not. Pan et al. [12] presented a model termed Semantic Graph Neural Network (SGNN) to overcome the challenges involved in email classifiers. This approach converted the email classification problems into graph classification problems. In this method, the emails were presented as graphs, and the SGNN method was used for

classification. The email feature was created as a semantic graph. So, there was no requirement to embed a word as a numerical vector representation.

Nayak et al. [13] presented a data science technique for Spam Email Detection (SMD) utilizing an ML approach. A hybrid bagging method was utilized in this study for the recognition of spam emails. This approach used two methods such as the Naïve Bayes (NB) method and the J48 (viz. Decision tree (DT) method. When a dataset was fed as an input in these approaches, the methods categorized the data into distinct sets with the help of data science. Hossain et al. [14] presented a method in which e-mails were classified either as spam or harmful. Both Density-Based Clustering (DBSCAN) and Isolation Forest techniques were utilized in the literature to identify the maximum values outside a particular range. Recursive Feature Elimination, Heatmap, and the Chi-Square Feature Selection (FS) approaches were utilized for the selection of effectual features. The presented method was executed using both ML and DL techniques to conduct a comparative analysis.

The current study presents an Intelligent Cuckoo Search (CS) Optimization Algorithm with Deep Learning-based Phishing Email Detection and Classification (ICSOA-DLPEC) model. The aim of the proposed ICSOA-DLPEC model is to effectually distinguish the emails as legitimate and the phishing ones. The N-gram approach is applied for the extraction of useful feature vectors. Moreover, the CS algorithm with a Gated Recurrent Unit (GRU) model is employed in this study for the detection and classification of phishing emails. Furthermore, the CS algorithm is also involved in this study for fine-tuning the parameters of the GRU model. The performance of the proposed ICSOA-DLPEC model was experimentally validated using a benchmark dataset, and the results were assessed under several dimensions. In short, the contributions of the current research study are as follows.

- A new ICSOA-DLPEC model is proposed for detection and the classification of the phishing emails.
- Data preprocessing and feature extraction processes are employed to derive the features.
- The GRU model is applied for detection and the classification of the phishing emails.
- The CS algorithm is employed for optimal fine-tuning of the hyperparameters involved in the GRU model.

2 Materials and Methods

In this study, a new ICSOA-DLPEC technique has been developed to effectually distinguish emails as legitimate and phishing ones. Primarily, the data pre-processing is performed through three phases such as email cleaning, tokenization and stop-word elimination. Besides, the N-gram approach is applied for the extraction of useful feature vectors. Further, the CS algorithm is employed with the GRU model to detect and classify phishing emails. Fig. 1 shows the block diagram of the proposed ICSOA-DLPEC technique.

2.1 Feature Engineering

In this case, the N-gram method is applied in the extraction of valuable feature vectors. The pattern is generated by concatenating the neighboring demonstrations to n-grams, in which $n = \{1,\ 2,\ 3 \ldots\}$. In this simple approach, n signifies a factor that is related to one and is also termed a unigram. Here, the n-gram approach is used to decide a word that should be considered to develop a visible word in the text with respect to another word. In general, $n - grams$ are certainly not superior to $n = 3$. The maximal values denote the probability of creating difficult patterns that are rarely equivalent.

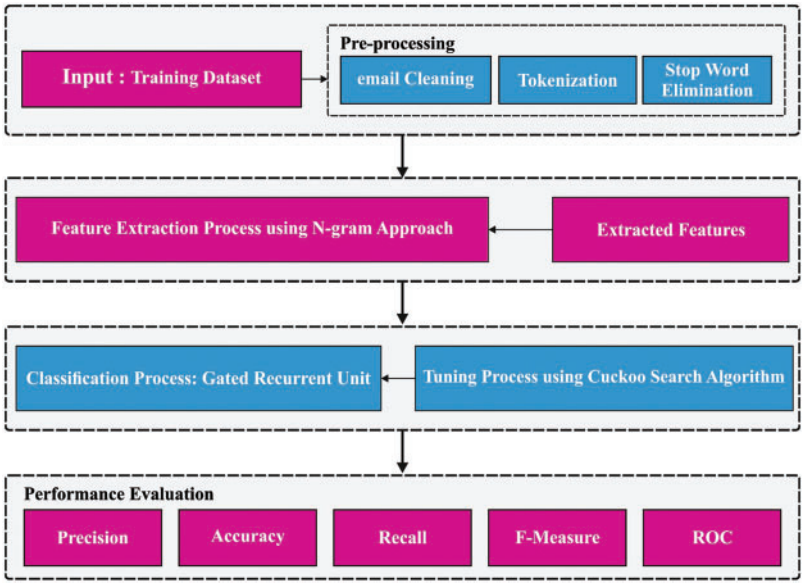

Figure 1: Block diagram of the ICSOA-DLPEC model

2.2 GRU Classification

The GRU model is exploited for the detection and classification of phishing emails.

Elman Recurrent Neural Network (RNN) is a significant example of simple RNNs in which a context layer feature is used to function as a memory layer. It is integrated with the existing state to propagate the data to the future states so as to handle the provided future input. The context layer is utilized to store the resultant neurons. The preceding time steps are constructed appropriately to suit the time-varying pattern of the data. The context layer continues as a memory of the previously hidden layer outcomes. The vectorized equation for a simple RNN is provided below.

$$h_t = \sigma h\left(W_h x_f + U_h h_{t-1} + b_h\right)$$
$$y_t = \sigma_y\left(W_y h_t + b_y\right) \tag{1}$$

Here, x_t refers to the input vector, h_t signifies the hidden layer vector, y_t represents the resultant vector, W denotes the weight of the hidden and the resultant layers, U implies the weight of the context state, b indicates the bias and σ_h and σ_y denote the respective activation functions. The Backpropagation Through Time (BPTT) approach is a prominent approach used to train simple RNNs. In comparison with the simple Neural Network (NN), the BPTT in RNNs propagate the error to in-depth network infrastructure, and the feature states are determined based on time.

The GRU model is a simplified model of the Long Short-Term Memory (LSTM) model in RNN. Unlike LSTM, the GRU combines both inputs as well as forgetting gates to update the gates. Considering the quantity of the hidden layers represented by h, a small-batch input that uses a time step t indicates $X_t \in \mathbb{R}^{n*d}$. The quantity of the samples is denoted by n, whereas the quantity of the input is denoted by d. The hidden neuron at a prior time step t1 is denoted by $H_{t-1} \in \mathbb{R}^{n*h}$. The resultant hidden layer h of the individual GRU at the existing time step t is given below [15]:

$$R_t = \sigma\left(X_t W_{xr} + H_{t-1} W_{hr} + b_r\right) \tag{2}$$

$$Z_t = \sigma\left(X_t W_{xz} + H_{t-1} W_{hz} + b_z\right) \tag{3}$$

$$\overline{H} = \tan h\left(X_t W_{xh} + \left(R_t E \odot H_{t-1}\right) W_{hh} + b_h\right) \tag{4}$$

$$H_t = \left(1 - Z_t\right) \odot H_{t-1} + Z_t E \odot \widetilde{H}_t \tag{5}$$

whereas σ denotes the sigmoid activation function, for example, $\sigma(x) = 1/1 + e^{-x}$, $\cdot W_{hz}$ W_{xr}, W_{hr} and W_{XZ} demonstrate the weights of the connecting hidden layer and the update gate, input layer and the reset gate, hidden layer and the reset gate and input layer and the update gate respectively; b_r and b_z indicate the bias values of the reset and update gates; H_t indicates the candidate hidden layer of the existing time step, t; $\odot$ denotes the matrix multiplication of two components. *Tanh* indicates the hyperbolic tangent activation function as given below.

$$\tanh(x) = 1 - \frac{2}{1 + e^{-2x}} \tag{6}$$

When a variable is estimated, the value of the present-day time is closely linked with the value of previous time and the value of the following time [16]. Fig. 2 depicts the architecture of the GRU method.

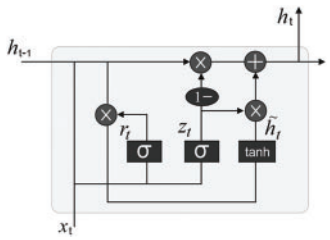

Figure 2: Structure of the GRU model

2.3 Hyperparameter Optimization

In order to optimally adjust the hyperparameters [17–20] involved in the GRU model, the CS algorithm is applied. Like other evolutionary models, the CS algorithm starts with an initial population. In general, the Cuckoos lay eggs in other birds' nests. Some of the host eggs are expected to be turned and raised by the cuckoos. However, other eggs are recognized by the host bird, whereas the host bird still raises the cuckoos' eggs. The rate of the raised eggs shows the correctness of the area. When additional eggs need to be added, it shows that the region gains no profit [21]. Consequently, the situation turns different in that the additional eggs endure for a period of time. This is a parameter to improve the performance of the cuckoos. The cuckoos seek the best area to enhance the lifetime of an egg. After hatching and turning into matured cuckoos, they generate their own society and community. All the communities have their own habitation. The optimum habitation of all the communities would be the forthcoming terminated place of the cuckoos in other groups.

All the groups immigrate towards an optimum currency area. Each group may be an inhabitant in the region near the optimum area. Egg Laying Radius (ELR) is assessed according to the number of eggs laid by the cuckoos and the distance from the existing enhanced area. Next, the cuckoos start laying their eggs arbitrarily in the nest within an egg-laying radii. This method iterates that until the

optimal position is obtained, the eggs are to be laid in the area with maximal returns. This improved area is the location where the highest number of cuckoos gather collectively. It is important to make parameters as an array so that the optimization issues are solved. In Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) techniques, the array is known as a 'chromosome' and the location of the 'particles'. However, in the CS algorithm, the array is recognized as a 'habitat'. In one-D *Nvar* optimization issue, habitation is a $1 \times Nvar$ array that expects the existing position of the cuckoo life. This is defined as follows.

$$Habitat = [x_1, x_2, \ldots, x_{Nvar}] \tag{7}$$

The quantity of the suitability or profit rate for the existing habitation is accomplished as a profit function. Thus, the following equation is applied.

$$Porofit = f_p\ (habitat) = f_p\ (x_1, x_2, \ldots, x_{Nvar}) \tag{8}$$

The CS algorithm is a procedure that exploits the profit of an entity. To exploit the CS algorithm, the cost function needs to be multiplied with a minus symbol so that the problem gets resolved. In order to initiate the optimization process, a habitation matrix of size $N_{pop} \times N_{var}$ is generated. Then, the quantity of eggs is characterized by each habitation matrix. Naturally, a cuckoo lays about 5–20 eggs. This number can be exploited by a minimal limit and a maximal limit in describing the eggs. Each cuckoo lays eggs in a certain range. So, ELR is the highest array of eggs laid upon. In optimization problems, the lower and the upper limits are denoted by var_{low} 0027 & var_{hi}, respectively. A cuckoo has an ELR, which is related to the complete amount of eggs, existing amount of eggs and lower or upper limits of the parameter. Henceforth, the ELR is described as given below.

$$ELR = \alpha \times \frac{Number\ of\ current\ cuckoo\ eggs}{Total\ number\ of\ eggs} \times (var_{hi} - var_{low}) \tag{9}$$

Consider that α represents a variable, whereas the highest ELR remains static. Each cuckoo subjectively lays its eggs in a host bird's nest with an ELR value. At last, an egg which is approximately smaller than the host eggs is recognized and thrown away. Later, each egg lays its own process, p% of eggs (usually 10%). If the profit value is lesser, it gets discarded. The residual chick in the host nests is then fed and raised.

3 Results and Discussion

In this section, the performance of the proposed ICSOA-DLPEC model was experimentally validated using the Enron email dataset [22], which can be accessed at https://www.cs.cmu.edu/~enron/. It includes 7,781 legitimate emails and 999 phishing emails.

Fig. 3 exhibits the four confusion matrices generated by the proposed ICSOA-DLPEC model on phishing email classification with distinct training (TR) and testing (TS) datasets. With a TR/TS of 90:10, the proposed ICSOA-DLPEC model recognized 773 and 99 samples as legitimate and phishing classes, respectively. In line with this, with a TR/TS of 80:40, the proposed ICSOA-DLPEC model classified 1,543 and 208 samples under legitimate and phishing classes, respectively. Moreover, on 80:20 TR/TS, the ICSOA-DLPEC model categorized 2,333 and 284 samples under legitimate and phishing classes respectively.

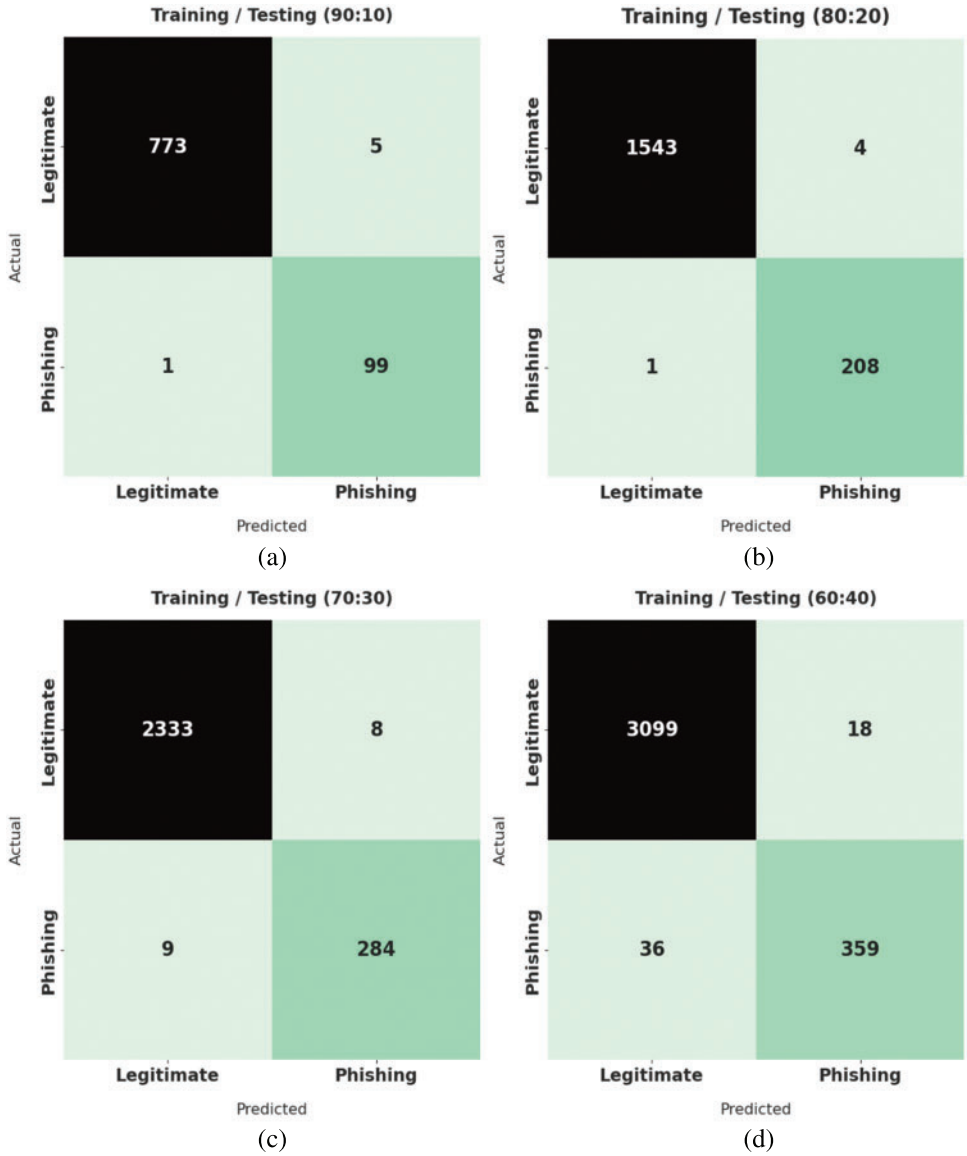

Figure 3: Confusion matrices of the ICSOA-DLPEC model

Table 1 and Fig. 4 show the overall classification outcomes accomplished by the proposed ICSOA-DLPEC model on phishing emails. With 90:10 TR/TS data, the proposed ICSOA-DLPEC model gained average $accu_y$, $prec_n$, $reca_l$ and F_{score} values such as 99.32%, 97.53%, 99.18% and 98.34%, respectively. Besides, with 80:20 TR/TS data, the proposed ICSOA-DLPEC model gained average $accu_y$, $prec_n$, $reca_l$ and F_{score} values such as 99.72%, 99.02%, 99.63% and 99.33%, respectively. Moreover, with 70:30 TR/TS data, the presented ICSOA-DLPEC model achieved average $accu_y$, $prec_n$, $reca_l$, and F_{score} values such as 99.352%, 98.44%, 98.29% and 98.37%, respectively. Furthermore, with 60:40 TR/TS data, the proposed ICSOA-DLPEC model attained average $accu_y$, $prec_n$, $reca_l$ and F_{score} values such as 98.46%, 97.04%, 95.15% and 96.07% respectively.

Table 1: Overall classification results of the ICSOA-DLPEC model

Class labels	Accuracy	Precision	Recall	F-score
Training/Testing (90:10)				
Legitimate	99.32	99.87	99.36	99.61
Phishing	99.32	95.19	99.00	97.06
Average	99.32	97.53	99.18	98.34
Training/Testing (80:20)				
Legitimate	99.72	99.94	99.74	99.84
Phishing	99.72	98.11	99.52	98.81
Average	99.72	99.02	99.63	99.33
Training/Testing (70:30)				
Legitimate	99.35	99.62	99.66	99.64
Phishing	99.35	97.26	96.93	97.09
Average	99.35	98.44	98.29	98.37
Training/Testing (60:40)				
Legitimate	98.46	98.85	99.42	99.14
Phishing	98.46	95.23	90.89	93.01
Average	98.46	97.04	95.15	96.07

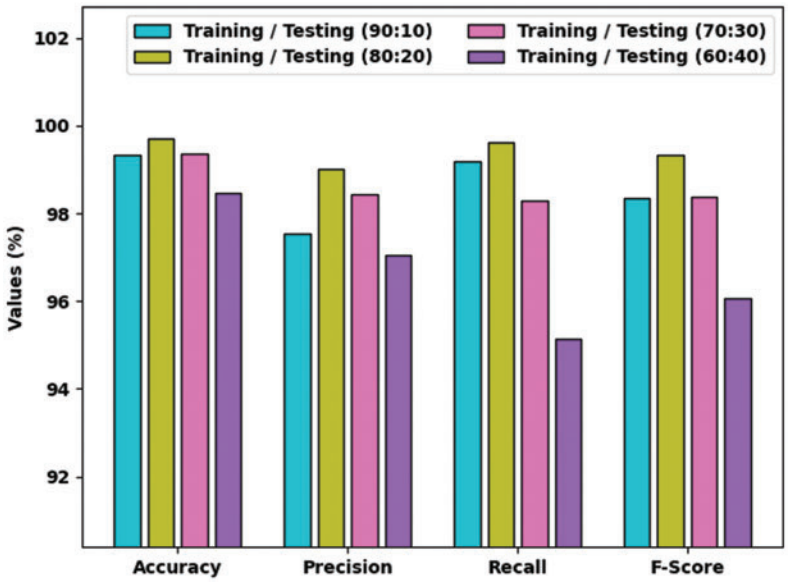

Figure 4: Phishing email classification outcomes of the ICSOA-DLPEC model

A Receiver Operating Characteristic (ROC) analysis was conducted upon the ICSOA-DLPEC model on phishing email classification and the results are revealed in Fig. 5. The figure portrays that the proposed ICSOA-DLPEC model accomplished the maximum ROC values to identify the phishing and legitimate class labels.

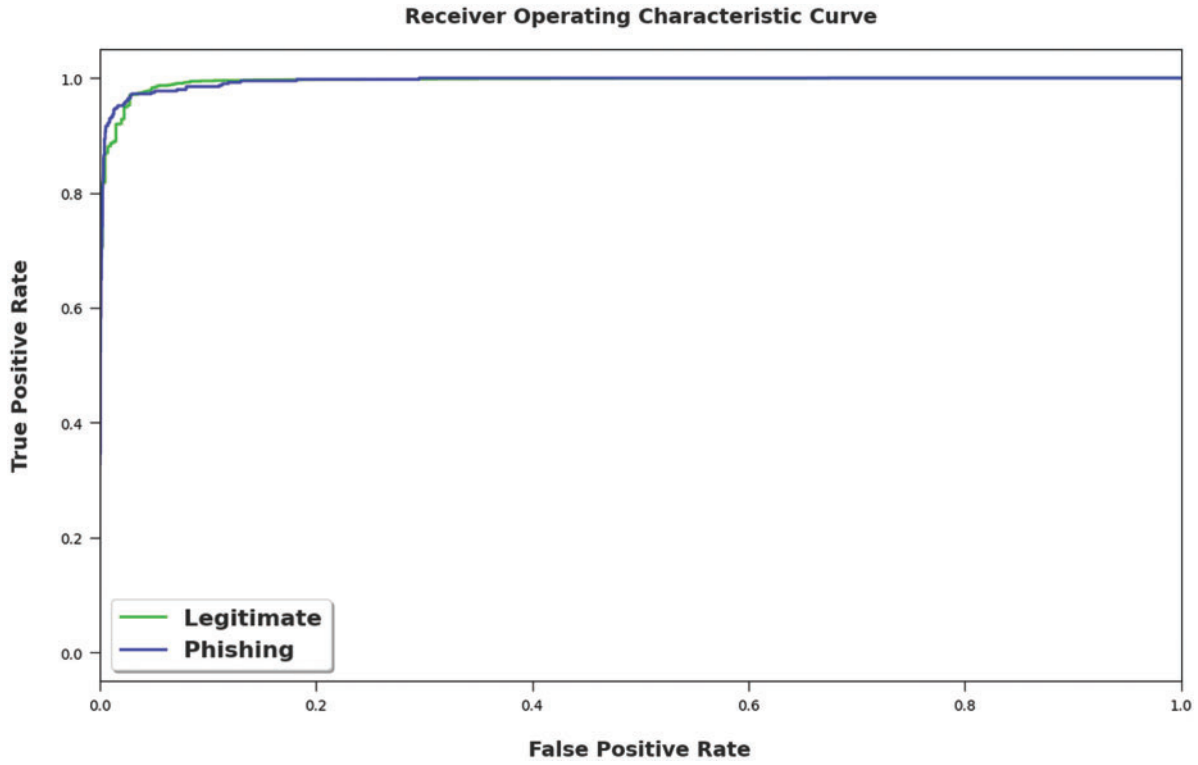

Figure 5: ROC examination results of the ICSOA-DLPEC model

Both Training Accuracy (TA) and Validation Accuracy (VA) values, attained by the proposed ICSOA-DLPEC model on phishing email classification, are demonstrated in Fig. 6. The experimental outcomes imply that the proposed ICSOA-DLPEC model gained the maximum TA and VA values. To be specific, the VA values were higher than the TA values. Then, the Training Loss (TL) and the Validation Loss (VL) values, achieved by the proposed ICSOA-DLPEC model on phishing email classification, are shown in Fig. 7. The experimental outcomes infer that the proposed ICSOA-DLPEC model accomplished the least TL and VL values whereas the VL values were lower than the TL values.

Fig. 8 shows the precision-recall curve obtained by the ICSOA-DLPEC model on phishing email classification. The results indicate that the proposed ICSOA-DLPEC model gained the maximum precision-recall values in both classes.

To validate the betterment of the proposed ICSOA-DLPEC model on phishing email classification, a comparative examination was conducted, and the results are shown in Table 2.

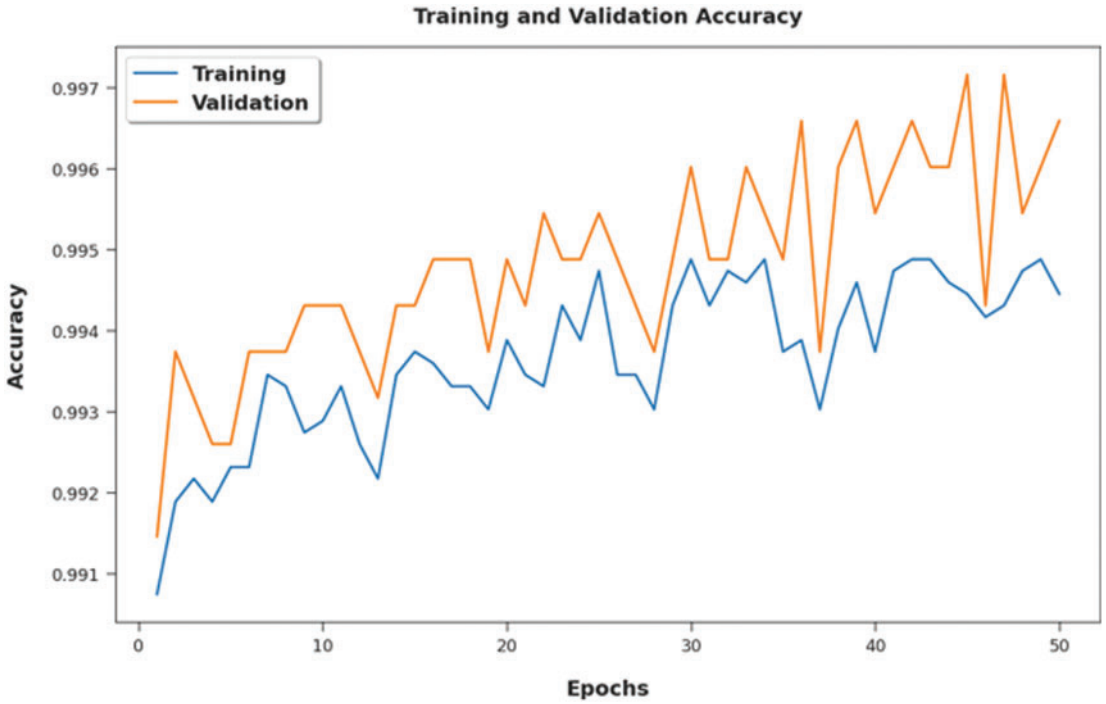

Figure 6: TA/VA examination results of the ICSOA-DLPEC model

Figure 7: TL/VL examination results of the ICSOA-DLPEC model

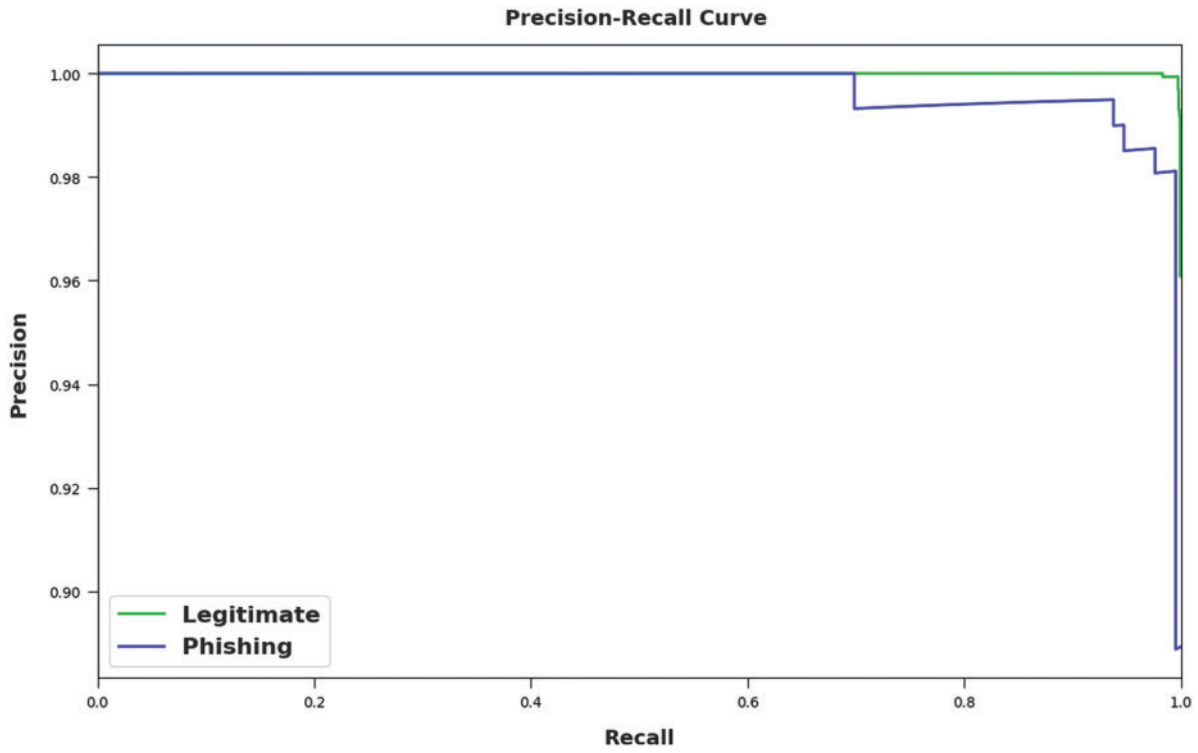

Figure 8: Precision recall examination results of the ICSOA-DLPEC model

Table 2: Comparative study results of the ICSOA-DLPEC model and other recent models

Methods	Accuracy	Precision	Recall	F-score
LSTM model	98.16	93.85	83.21	88.18
CNN model	97.58	86.27	87.53	85.19
THEMIS	99.34	98.96	99.12	99.03
ICSOA-DLPEC	99.72	99.02	99.63	99.33

Fig. 9 shows the comprehensive $accu_y$ and $prec_n$ investigation results of the ICSOA-DLPEC model and other existing models. The figure reveals that the CNN model achieved the least $accu_y$ and $prec_n$ values, such as 97.58% and 86.27%, respectively. In line with this, the LSTM model gained slightly enhanced $accu_y$ and $prec_n$ values such as 98.76% and 93.85%, respectively. Though the THEMIS model gained reasonable $accu_y$ and $prec_n$ values such as 99.34% and 98.96%, respectively, the proposed ICSOA-DLPEC model gained effectual outcomes with maximum $accu_y$ and $prec_n$ values such as 99.72% and 99.02%, respectively.

Fig. 10 portrays the extensive $reca_l$ and F_{score} examination results accomplished by the proposed ICSOA-DLPEC model and other existing models. The figure exposes that the CNN model delivered the least $reca_l$ and F_{score} values, such as 87.53% and 85.19%, respectively. In line with this, the LSTM model gained slightly enhanced $reca_l$ and F_{score} values, such as 83.21% and 88.18%, respectively. Though the THEMIS model gained reasonable $reca_l$ and F_{score} values such as 99.12% and 99.03%, the proposed

ICSOA-DLPEC model gained effectual outcomes with maximum $reca_l$ and F_{score} values such as 99.63% and 99.33%, respectively. Based on the results and the discussion, it is evident that the proposed ICSOA-DLPEC model is effective in the detection and classification of phishing emails.

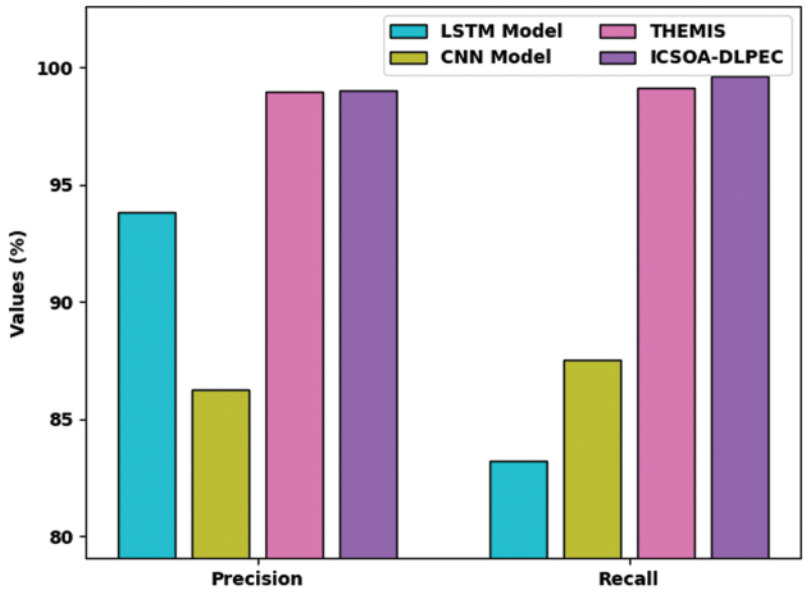

Figure 9: Comparative $accu_y$ and $prec_n$ investigation results of the ICSOA-DLPEC model

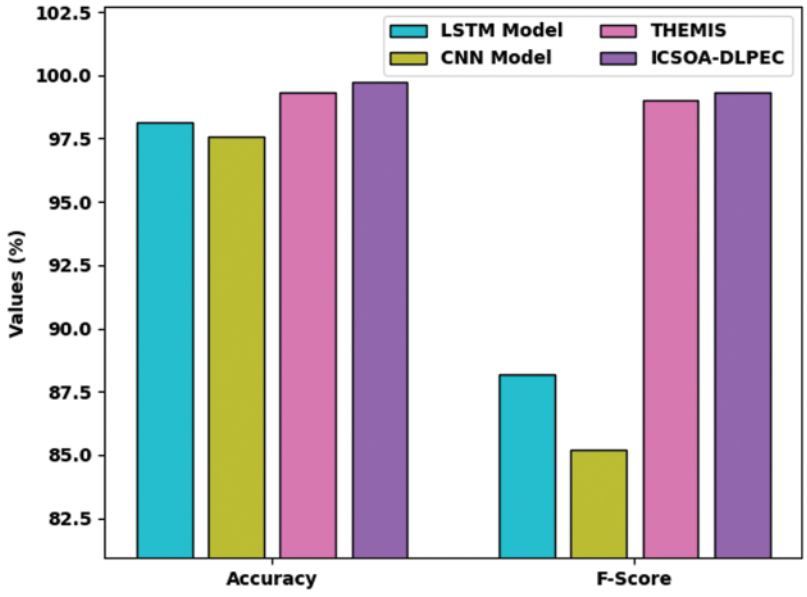

Figure 10: Comparative $reca_l$ and F_{score} investigation results of the ICSOA-DLPEC model

4 Conclusion

In this study, a new ICSOA-DLPEC technique has been developed to effectually distinguish emails as legitimate and phishing. Primarily, the data pre-processing is performed through three steps such as email cleaning, tokenization and stop-word elimination. Following this, the N-gram approach is applied for the extraction of the useful feature vectors. In addition, the CS algorithm is employed with the GRU model to detect and classify phishing emails. Finally, the CS algorithm is applied to fine-tune the parameters involved in the GRU model. The performance of the proposed ICSOA-DLPEC model was experimentally validated using a benchmark dataset, and the results were assessed under several dimensions. Extensive comparative analysis results confirmed the superiority of the ICSOA-DLPEC model over recent approaches. In the future, the hybrid DL models can be exploited to increase the detection rate further.

Funding Statement: This research was supported in part by Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (NRF-2021R1A6A1A03039493), and in part by the NRF grant funded by the Korea government (MSIT) (NRF-2022R1A2C1004401).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] S. Salloum, T. Gaber, S. Vadera and K. Shaalan, "Phishing email detection using natural language processing techniques: A literature survey," *Procedia Computer Science*, vol. 189, pp. 19–28, 2021.

[2] T. Gangavarapu, C. D. Jaidhar and B. Chanduka, "Applicability of machine learning in spam and phishing email filtering: Review and approaches," *Artificial Intelligence Review*, vol. 53, no. 7, pp. 5019–5081, 2020.

[3] Y. Fang, C. Zhang, C. Huang, L. Liu and Y. Yang, "Phishing email detection using improved rcnn model with multilevel vectors and attention mechanism," *IEEE Access*, vol. 7, pp. 56329–56340, 2019.

[4] A. Karim, S. Azam, B. Shanmugam, K. Kannoorpatti and M. Alazab, "A comprehensive survey for intelligent spam email detection," *IEEE Access*, vol. 7, pp. 168261–168295, 2019.

[5] R. Vinayakumar, B. Ganesh, A. Kumar, K. P. Soman and P. Poornachandran, "DeepAnti-PhishNet: Applying deep neural networks for phishing email detection," in *Proc. of the 1st AntiPhishing Shared Pilot at 4th ACM Int. Workshop on Security and Privacy Analytics (IWSPA 2018)*, Tempe, Arizona, USA, pp. 1–11, 2018.

[6] A. Vazhayil, N. B. Harikrishnan, R. Vinayakumar, K. P. Soman and A. D. R. Verma, "PED-ML: Phishing email detection using classical machine learning techniques," in *Proc. of the 1st AntiPhishing Shared Pilot at 4th ACM Int. Workshop on Security and Privacy Analytics (IWSPA 2018)*, Tempe, Arizona, USA, pp. 1–8, 2018.

[7] S. Kumar Birthriya and A. K. Jain, "A comprehensive survey of phishing email detection and protection techniques," *Information Security Journal: A Global Perspective*, pp. 1–30, 2021. https://doi.org/10.1080/19393555.2021.1959678.

[8] J. Lee, F. Tang, P. Ye, F. Abbasi, P. Hay *et al.,* "D-fence: A flexible, efficient, and comprehensive phishing email detection system," in *2021 IEEE European Symp. on Security and Privacy (EuroS&P)*, Vienna, Austria, pp. 578–597, 2021.

[9] J. Rastenis, S. Ramanauskaitė, I. Suzdalev, K. Tunaitytė, J. Janulevičius *et al.,* "Multi-language spam/phishing classification by email body text: Toward automated security incident investigation," *Electronics*, vol. 10, no. 6, pp. 1–10, 2021.

[10] A. Sundararaj and G. Kul, "Impact analysis of training data characteristics for phishing email classification," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 12, no. 2, pp. 85–98, 2021.
[11] C. McGinley and S. A. S. Monroy, "Convolutional neural network optimization for phishing email classification," in *2021 IEEE Int. Conf. on Big Data (Big Data)*, Orlando, FL, USA, pp. 5609–5613, 2021.
[12] W. Pan, J. Li, L. Gao, L. Yue, Y. Yang *et al.,* "Semantic graph neural network: A conversion from spam email classification to graph classification," *Scientific Programming*, vol. 2022, pp. 1–8, 2022.
[13] R. Nayak, S. A. Jiwani and B. Rajitha, "Spam email detection using machine learning algorithm," *Materials Today: Proceedings*, pp. 1–5, 2021. https://doi.org/10.1016/j.matpr.2021.03.147.
[14] F. Hossain, M. N. Uddin and R. K. Halder, "Analysis of optimized machine learning and deep learning techniques for spam detection," in *2021 IEEE Int. IOT, Electronics and Mechatronics Conf. (IEMTRONICS)*, Toronto, ON, Canada, pp. 1–7, 2021.
[15] Y. Zhang, J. Tang, Z. He, J. Tan and C. Li, "A novel displacement prediction method using gated recurrent unit model with time series analysis in the erdaohe landslide," *Natural Hazards*, vol. 105, no. 1, pp. 783–813, 2021.
[16] C. Xu, J. Shen, X. Du and F. Zhang, "An intrusion detection system using a deep neural network with gated recurrent units," *IEEE Access*, vol. 6, pp. 48697–48707, 2018.
[17] K. Shankar, E. Perumal, M. Elhoseny, F. Taher, B. B. Gupta *et al.,* "Synergic deep learning for smart health diagnosis of covid-19 for connected living and smart cities," *ACM Transactions on Internet Technology*, vol. 22, no. 3, pp. 1–14, 2022.
[18] N. Metawa, I. V. Pustokhina, D. A. Pustokhin, K. Shankar and M. Elhoseny, "Computational intelligence-based financial crisis prediction model using feature subset selection with optimal deep belief network," *Big Data*, vol. 9, no. 2, pp. 100–115, 2021.
[19] R. Saravanakumar, N. Krishnaraj, S. Venkatraman, B. Sivakumar, S. Prasanna *et al.,* "Hierarchical symbolic analysis and particle swarm optimization based fault diagnosis model for rotating machineries with deep neural networks," *Measurement*, vol. 171, pp. 1–25, 2021.
[20] M. Elhoseny, M. M. Selim and K. Shankar, "Optimal deep learning based convolution neural network for digital forensics face sketch synthesis in internet of things (IoT)," *International Journal of Machine Learning and Cybernetics*, vol. 12, no. 11, pp. 3249–3260, 2021.
[21] Y. Xu, N. Chen, X. Shen, L. Xu, Z. Pan *et al.,* "Proposal and experimental case study on building ventilating fan fault diagnosis based on cuckoo search algorithm optimized extreme learning machine," *Sustainable Energy Technologies and Assessments*, vol. 45, pp. 1–15, 2021.
[22] The First Security and Privacy Analytics Anti-Phishing Shared Task. [Online]. Available: https://dasavisha.github.io/IWSPA-sharedtask/?tdsourcetag=s_pctim_aiomsg.