DOI 10.52209/1609-1825_2023_2_413 UDC 004.056.5

Detection of Phishing Threats Using Neural Networks

[1]**USSATOVA Olga,** *PhD, Senior Lecturer, uoa_olga@mail.ru,*
[1]***ZHUMABEKOVA Aidana,** Doctoral Student, Lecturer, zhumabekova2702@gmail.com,*
[2]**Eric T. MATSON,** *Professor, ematson@purdue.edu,*
[1]**KARYUKIN Vladislav,** *Master, Senior Lecturer, vladislav.karyukin@gmail.com,*
[1]**ILESSOVA Bakytgul,** *Master, Lecturer, bakhytgul92@gmail.com,*
[1]*NPJSC «Al-Farabi Kazakh National University», Kazakhstan, Almaty, Al-Farabi Avenue, 71,*
[2]*Purdue University, USA, Indiana, West Lafayette, 610 Purdue Mall,*
**corresponding author.*

Abstract. *Today, the Internet is an effective channel for social interaction worldwide, but it also opens up great opportunities for cyberattacks. Recently, the number of botnets and phishing cyberattacks has increased dramatically. Phishing is a cyber attack carried out using a fake website designed to steal sensitive information. Detecting such attacks using machine learning algorithms and neural systems is an efficient and fast method. This paper discusses the types of Internet threats and neural network algorithms used to counter phishing attacks. The main goal of this paper is to build neural network models capable of effectively detecting phishing attacks. The methods of research are data analysis, data preprocesses and neural networks. In the first step, the Kaggle dataset containing data from 5000 legitimate and 5000 malicious web pages was gathered. The features of the dataset were analyzed and normalized. Then the Deep neural network, Convolutional neural network and Long short-term memory network were applied for threat classification. The Deep neural network and Long short-term memory network showed the best results of accuracy up to 0.96 and 0.99, while the Convolutional neural network reached only 0.75. Generally, the Deep neural network and Long short-term memory achieved the best results among the neural network methods.*

Keywords: *cybersecurity, neural network, machine learning, phishing, botnet, internet threats, dataset.*

Introduction. All government agencies and organizations manage various databases and own citizens' personal data, so the database, technical infrastructure, and services must always have a high level of information security to protect against internal threats and cyber attacks. In 2017, N. Nazarbayev, the head of the Republic of Kazakhstan, declared the statement «Third Modernization of Kazakhstan: Global Competitiveness», where the responsible authorities were instructed to take the necessary measures to create the «Kazakhstan Cyber Shield» system with an emphasis on information security, combating cyber attacks, and protecting against various threats. Cyber Shield ensures the security of electronic information resources from various cyber-attacks. The main goal of the cyber shield system of Kazakhstan is to increase the level of information security of the country, protect information resources, databases, information and communication infrastructure, and personal data of citizens from external and internal threats, counter potential cyber attacks, and develop the security of the national information space [1].

Materials and methods. In recent years, during the global COVID-19 pandemic, in order to protect against the epidemic, all life processes have moved to the online Internet system, and web servers have developed dramatically. Still, at this moment, cyber-attacks and various threats have spread like an epidemic and are developing day by day. In such a situation, in order to ensure information security, it is essential to get in touch with foreign terrorist organizations and protect the detective forces [2].

Recently, botnets have been the most common form of a cyber attack against users. A botnet is a network of virus programs that allow attackers to remotely control other computers without the owner's permission, which means criminals can remotely attack computers and servers and spread spam and viruses [3]. This zombie network is a good way to get rich quickly by doing massive crime missions. The computer owner does not realize this until his email is stolen, blocked or the money disappears from the account.

Another common online scam is phishing. The main purpose of phishing is to steal an Internet user's personal information (passwords, bank accounts and other confidential information). Popular phishing attack schemes are considered further. First, these are phishing sites. This attack method steals the user's personal data and offers to download or purchase malicious software [4]. Phishing sites are specially created fakes of government agencies, companies, and other popular social networks using URLs or subdomains with deliberate errors (ollx.kz, gmall.

com, gooogle.com, etc.) while maintaining the originality of the site design. The next type of phishing attack scheme is spear phishing. It is carried out by sending fake emails to users on behalf of government agencies, banks, and other organizations. The main purpose is access to the personal data of users or employees of the organization. The third type of phishing attack is whaling. Here the victim is an employee of the organization. That is, the attacker sends a fake email to employees on behalf of the head of the organization to encourage them to share sensitive data or make a payment. Another type of phishing attack is vishing. This type of threat is not limited to email; it is carried out by calling customers of second-tier banks. Voice phishing has been rampant in the country since 2020. The attacker calls the victim's mobile phone number to the technical support service of the second-tier bank and convinces them to provide the code of the sent SMS message [5].

In order to enhance the security of information and databases, various protection tools are used to protect and prevent attacks by intruders and various threats. These include cryptography, anti-virus programs, anti-phishing systems, and other methods. Neural networks (NN) [6] and machine learning (ML) algorithms effectively detect threats [7]. A neural network is a complex function that is applied to input data, and the output of this function is a number, vector, or matrix, depending on the task. This article focuses on the detection of phishing URLs using neural network methods.

Methodological and experimental department. We conducted an experimental study on the detection of phishing attacks using NN. The proposed work considers the classification of Internet phishing attacks. The development of NN for detecting Internet threats is effective in solving the problem of detecting phishing attacks. In Figure 1, we can see that the database identified 5,000 harmless (green) and 5,000 very harmful (red) symptoms (Figure 1).

Features of the dataset. We used the Kaggle dataset for this research paper. This database contained 48 features, with 5,000 legitimate web pages and 5,000 malicious phishing web pages uploaded between 2015 and 2017. The browser automation framework was used for efficient feature extraction.

Classification by classes was carried out using Deep neural networks (DNN), Convolutional neural networks (CNN), and recurrent neural networks (RNN). The results were evaluated using the accuracy, precision, recall, and F1-score measures.

Neural network algorithms. Traditional machine learning algorithms allow you to achieve good data classification results. Nevertheless, in the last decade, the Deep neural network approach has become dominant in the field of artificial intelligence, showing high accuracy in tasks such as speech and image recognition, text classifications, and natural language processing. Several types of NN are often used: DNN, CNN, and RNN.

Deep neural networks are a model of NN with

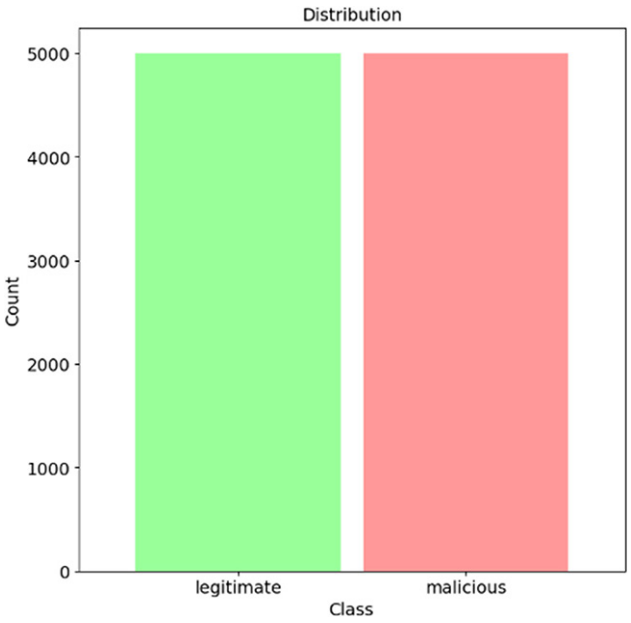

Figure 1 – Legitimate and Malicious Classes

two or more hidden layers. The neural network consists of an input layer containing input data, hidden layers including nodes called neurons, and an output layer containing one or more neurons. In this case, $x = x_1, x_2, ..., x_f$ is the input vector; $w_1, w_2, ..., w_i$ is the connection weights vector of each level; $b_1, b_2, ..., b_i$ is the displacement vector. Levels l_2 to l_{n-1} form hidden layers, and the l_n level is represented by the corresponding $y_1, y_2, ..., y_m$ output vector.

Convolutional neural networks are one of the popular and often used types of NN, which have gained great popularity due to their use in classification and image recognition tasks. They have become more popular due to their use in image classification, where the filter moves through the image. CNN are also common in speech recognition, natural language processing, and SA.

A recurrent neural network is a type of neural network that sequentially processes data and generates current conclusions based on previous calculations. In a given neural network, the latent state h_t is determined using input data x_i at the current time t and represents previous latent states h_{t-1}. The expression h_t can then be evaluated as (1):

$$h_t = \varphi(W^{t-1}h_{t-1} + W^t x_t), \quad (1)$$

where W^{t-1} and W^t are the weights of the previous hidden state h_{t-1} and the current input x_t. The output y_t is given by the following equation (2):

$$y_t = \varphi(W^y h_t). \quad (2)$$

The latent state h_t is represented as memory in the network, and the result y_t depends only on the memory h_t at a time t and the matrix of weights W^y. One of the most popular models of RNN is the Long short-term memory (LSTM) architecture. The LSTM structure includes four main components: an input

element, an output element, a forgetting element, and a memory cell.

The graphs below show the values of the graphs for correct classification and loss in test samples. DNN classification results are shown in Figures 2-3 and Table 1.

Evaluation metrics of DNN: Accuracy – 0.99; Precision – 0.99; Recall – 0.99; F1-score – 0.99.

CNN classification results are shown in Figures 4-5 and Table 2.

Evaluation metrics of CNN: Accuracy – 0.75; Precision – 0.72; Recall – 0.83; F1-score – 0.77.

LSTM classification results are shown in Figures 6-7 and Table 3.

Evaluation metrics of LSTM: Accuracy – 0.97; Precision – 0.96; Recall – 0.99; F1-score – 0.97.

Conclusion. With the help of neural networks, we achieve high results in rapidly detecting cyber attacks. Using the example of detecting phishing at-

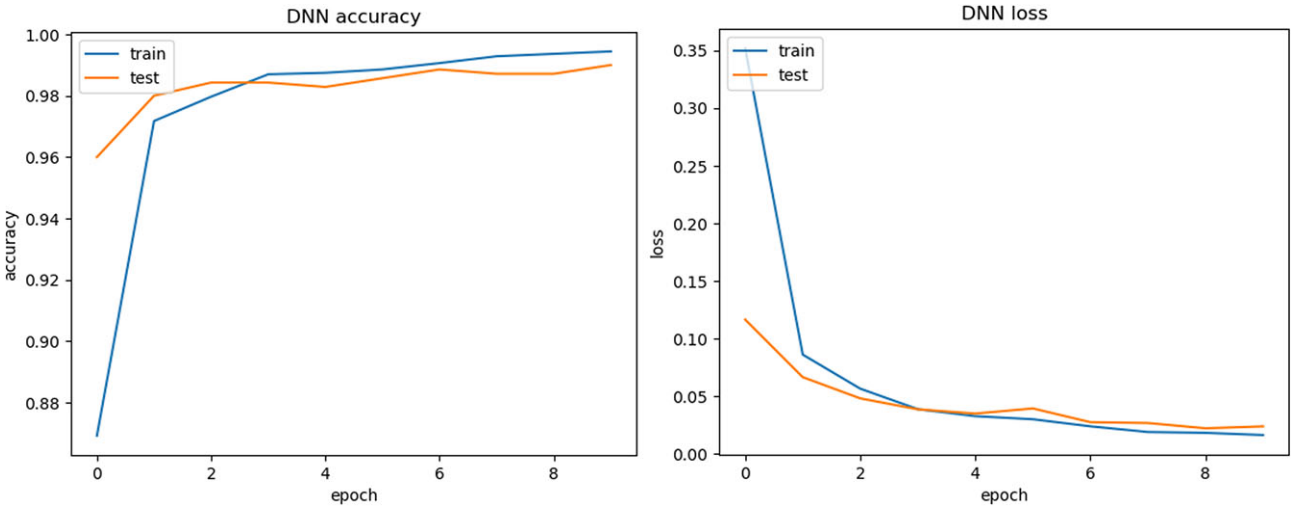

Figure 2 – DNN accuracy and loss

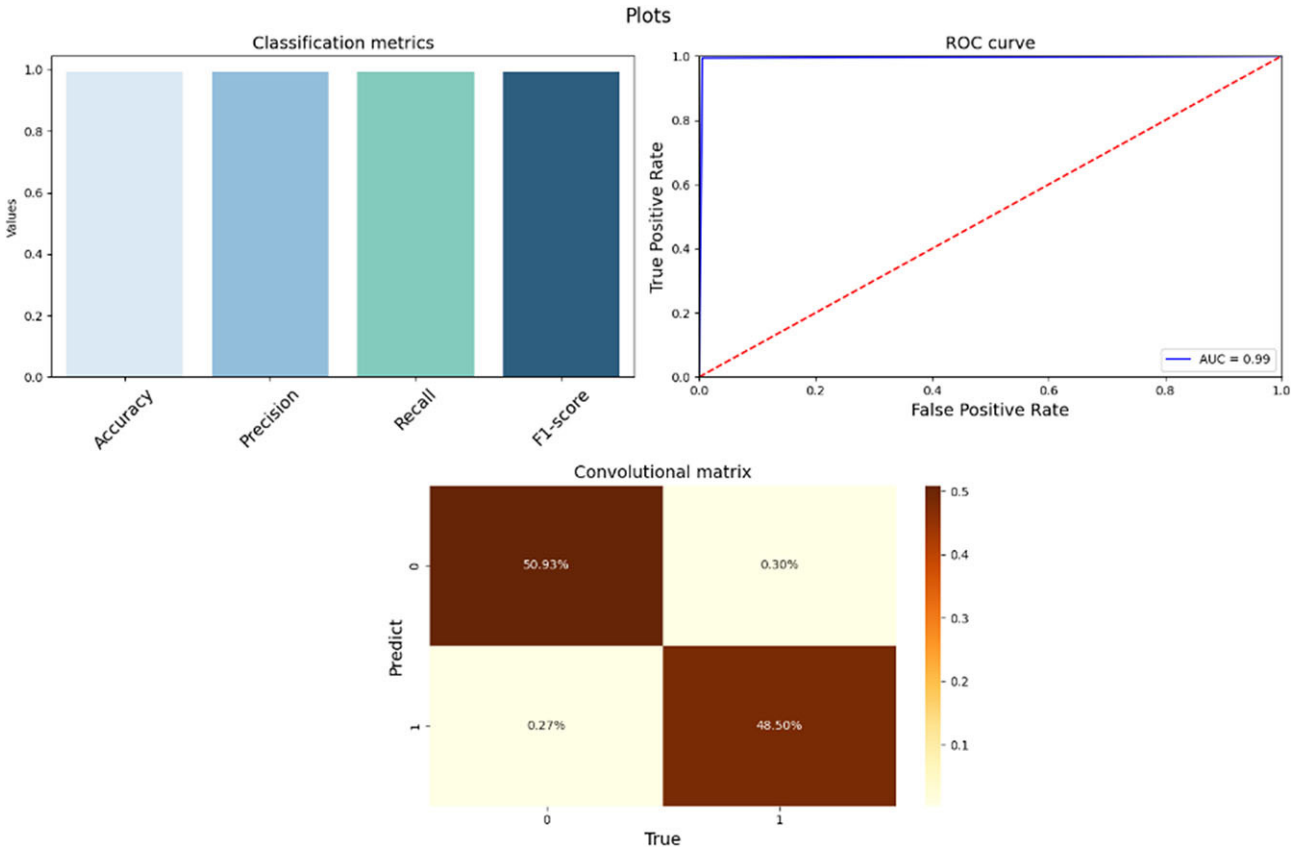

Figure 3 – DNN classification

Table 1 – Classification report

Classifier	Precision	Recall	F1-score	Support
legitimate	0.99	0.99	0.99	1463
malicious	0.99	0.99	0.99	1537
accuracy			0.99	3000
macro avg	0.99	0.99	0.99	3000
weighted avg	0.99	0.99	0.99	3000

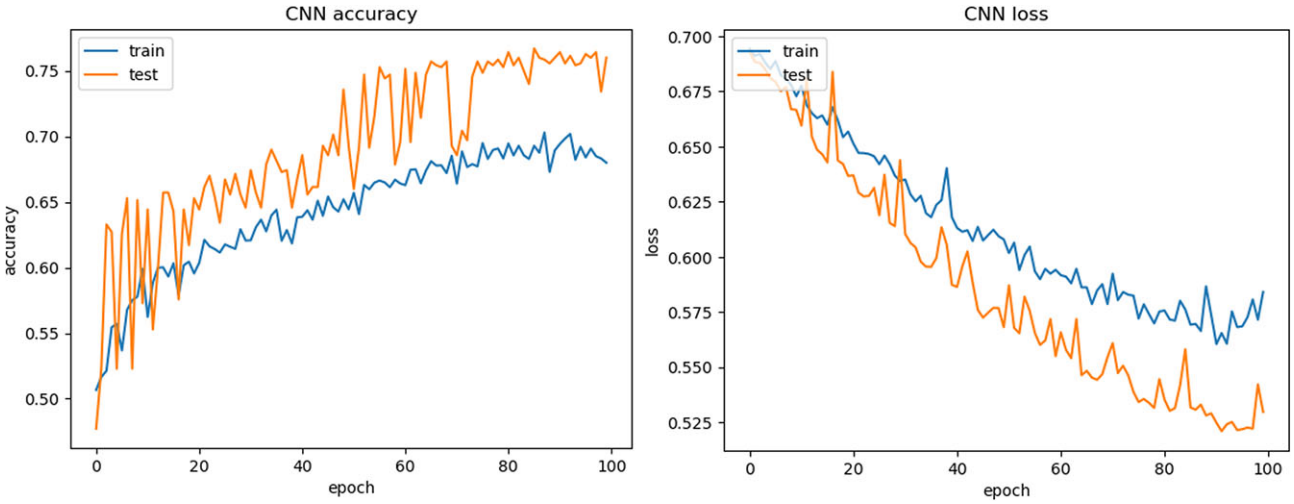

Figure 4 – CNN accuracy and loss

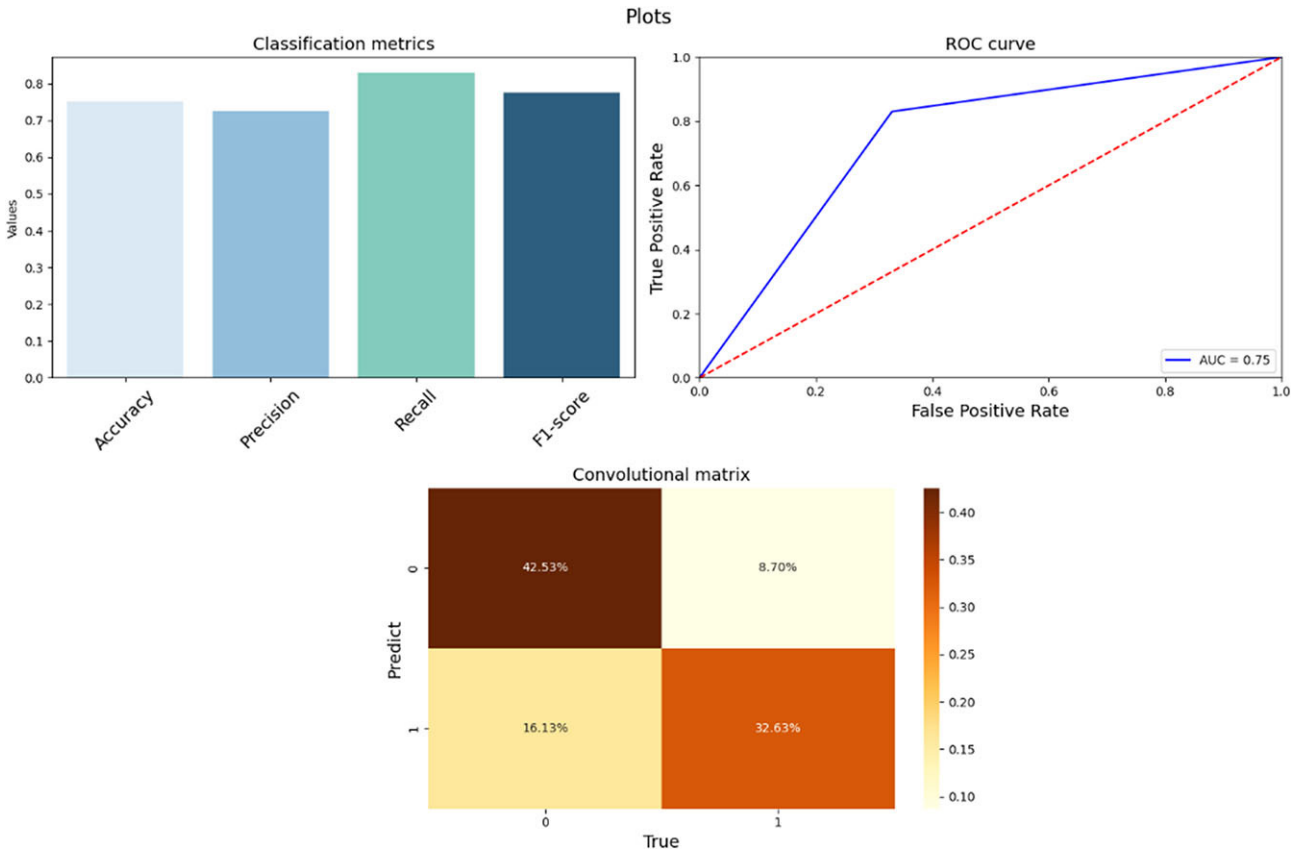

Figure 5 – CNN classification

Table 2 – Classification report

Classifier	Precision	Recall	F1-score	Support
legitimate	0.79	0.67	0.72	1463
malicious	0.72	0.83	0.77	1537
accuracy			0.75	3000
macro avg	0.76	0.75	0.75	3000
weighted avg	0.76	0.75	0.75	3000

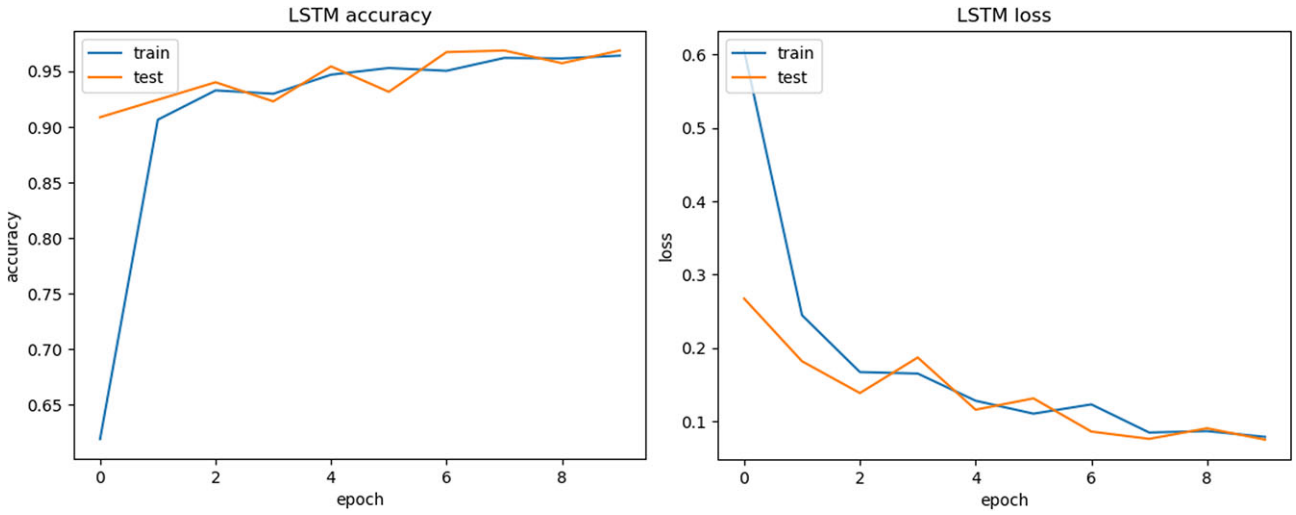

Figure 6 – LSTM accuracy and loss

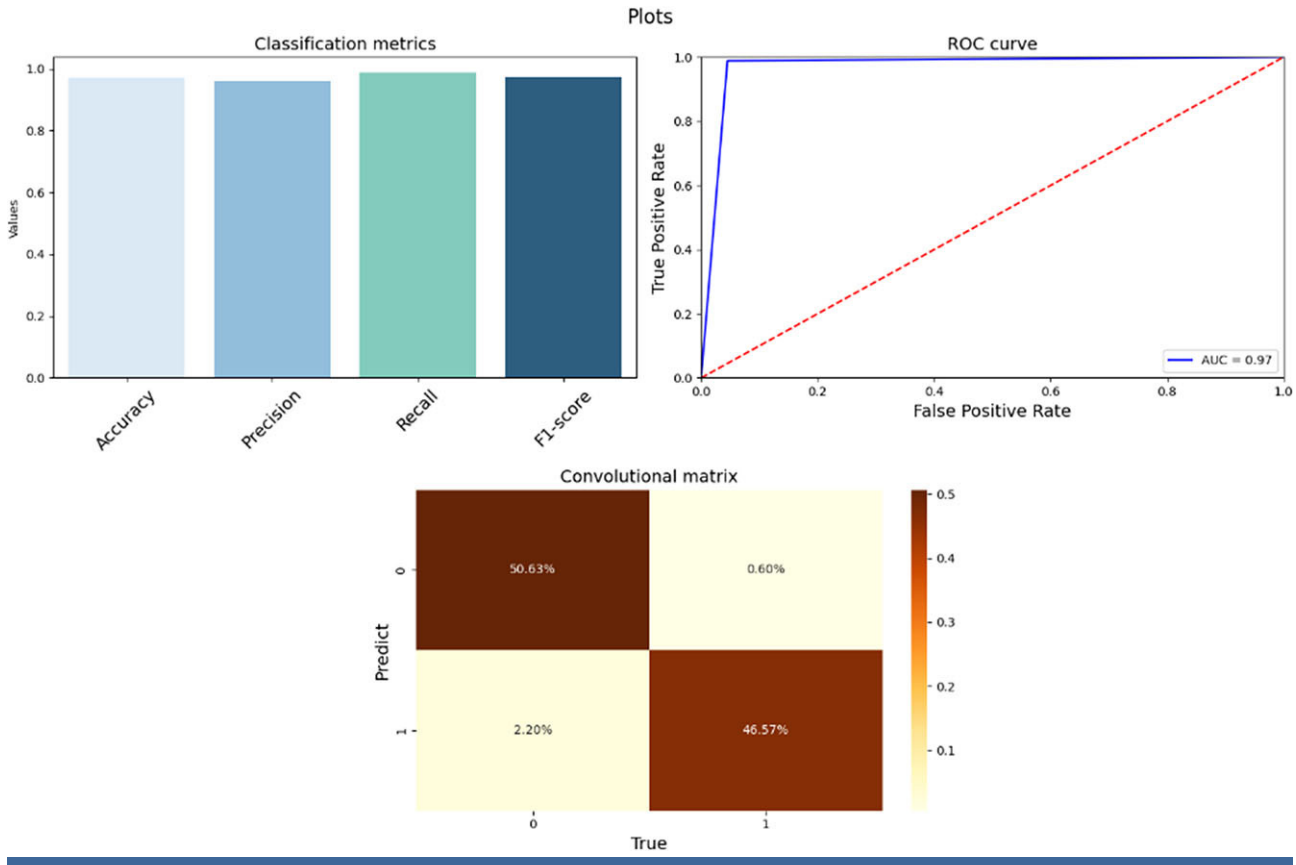

Figure 7 – LSTM classification

Table 3 – Classification report

Classifier	Precision	Recall	F1-score	Support
legitimate	0.99	0.95	0.97	1463
malicious	0.96	0.99	0.97	1537
accuracy			0.97	3000
macro avg	0.97	0.97	0.97	3000
weighted avg	0.97	0.97	0.97	3000

tacks using neural networks, according to the experiment, the Deep Neural Network showed 0.99, and Long-term memory showed 0.97, reaching a high result. The Convolutional neural networks showed 75. To summarize the study, we confirm that Deep neural networks and Long short-term memory neural networks are effective in detecting phishing attacks by intruders on sensitive user data.

REFERENCES

1. On approval of the Cyber Security Concept («Cyber Shield of Kazakhstan»). 30.09.2022. https://adilet.zan.kz/kaz/docs/P1700000407
2. The dark side of the Internet: Attacks, costs and responses. Kim, W., Jeong, O.-R., Kim, C., & So, J. (2011). Information Systems, 36 (3), 675-705. doi:10.1016/j.is.2010.11.003
3. Botnets: threats and responses. Jeong, O., Kim, C., Kim, W., & So, J. (2011). International Journal of Web Information Systems, 7 (1), 6-17. doi:10.1108/17440081111125635
4. Detecting Phishing Cyber Attack Based on Fuzzy Rules and Differential Evaluation. Rawaa Mohammed Abdul-Hussein, Ahmed H. Mohammed, Amal Abbas Kadhim. TEM Journal. Volume 11, Issue 2, pages 543-551, ISSN 2217-8309, DOI: 10.18421/TEM112-07, May 2022.
5. A phishing website detection and recognition method based on naive bayes. Xu Yingying, Chen Guangxuan, Liu Qiang, Xu Wanpeng, Zhang, Lei, Wu Jiajian,Fan Xiaoshi. IEEE 6th Information Technology and Mechatronics Engineering Conference, ITOEC 2022.
6. Internet Threat Detection Using Neural Networks. Zhumabekova A.T., Usatova O.A. International scientific journal. «Global science and innovations 2021: central asia». Nur-Sultan, Kazakhstan, December 2021.
7. Comprehensive DDoS Attack Classification Using Machine Learning Algorithms. O. Ussatova, A. Zhumabekova, Y. Begimbayeva, E.T. Matson, N. Ussatov. Computers, Materials & Continua, 2021. DOI: 10.32604/cmc.2022.026552

Нейрондық желілер арқылы фишингтік қауіптерді анықтау

[1]***УСАТОВА Ольга Александровна,*** *PhD, аға оқытушы, uoa_olga@mail.ru,*
[1]****ЖҰМАБЕКОВА Айдана Төлеуқызы,*** *докторант, аға оқытушы, zhumabekova2702@gmail.com,*
[2]***Эрик МЭТСОН,*** *профессор, ematson@purdue.edu,*
[1]***КАРЮКИН Владислав Игоревич,*** *магистр, аға оқытушы, vladislav.karyukin@gmail.com,*
[1]***ІЛЕСОВА Бақытгүл Ералықызы,*** *магистр, оқытушы, bakhytgul92@gmail.com,*
[1]*«Әл-Фараби атындағы Қазақ ұлттық университеті» КеАҚ, Қазақстан, Алматы, Әл-Фараби даңғылы, 71,*
[2]*Пердью университеті, АҚШ, Индиана, Уэст-Лафейетт, 610 Пердью Молл,*

**автор-корреспондент.*

Аңдатпа. *Бүгінгі таңда Интернет бүкіл әлем бойынша әлеуметтік өзара әрекеттесу үшін тиімді арна болып табылады, бірақ ол кибершабуылдар үшін үлкен мүмкіндіктер ашады. Соңғы уақытта ботнеттердің және фишингтік кибершабуылдардың саны күрт өсті. Фишинг – бұл құпия ақпаратты ұрлауға арналған жалған веб-сайтты пайдалану арқылы жүзеге асырылатын кибершабуыл. Мұндай шабуылдарды машиналық оқыту алгоритмдері мен нейрондық жүйелер арқылы анықтау тиімді және жылдам әдіс болып табылады. Бұл мақалада фишингтік шабуылдарға қарсы тұру үшін қолданылатын интернет-қауіптердің түрлері мен нейрондық желі алгоритмдері қарастырылады. Бұл мақаланың негізгі мақсаты – фишингтік шабуылдарды тиімді анықтауға қабілетті нейрондық желі модельдерін құру. Зерттеу әдістеріне деректерді талдау, деректерді алдын ала өңдеу және нейрондық желілер жатады. Бірінші қадам 5000 заңды және 5000 зиянды веб-беттерден деректерді қамтитын Kaggle деректер жинағын жинау болды. Мәліметтер жинағының сипаттамалары талданды және қалыпқа келтірілді. Содан кейін қауіптерді жіктеу үшін терең нейрондық желі, конволюционды нейрондық желі және ұзақ мерзімді, қысқа мерзімді жады желісі қолданылды. Терең нейрондық желі және ұзақ қысқа мерзімді жад желісі 0,96 және 0,99 дейінгі ең жақсы дәлдік ұпайларын көрсет-*

ті, ал конволюционды нейрондық желі тек 0,75-ке жетті. Жалпы алғанда, терең нейрондық желі және ұзақ мерзімді қысқа мерзімді жады нейрондық желі әдістерінің ішінде ең жақсы нәтижелерге қол жеткізді.

Кілт сөздер: *киберқауіпсіздік, нейрондық желі, машиналық оқыту, фишинг, ботнет, интернет қауіптері, деректер жинағы.*

Выявление фишинговых угроз с использованием нейронных сетей

[1]***УСАТОВА Ольга Александровна,*** *PhD, старший преподаватель, uoa_olga@mail.ru,*
[1]****ЖУМАБЕКОВА Айдана Толеуовна,*** *докторант, преподаватель, zhumabekova2702@gmail.com,*
[2]***Эрик МЭТСОН,*** *профессор, ematson@purdue.edu,*
[1]***КАРЮКИН Владислав Игоревич,*** *магистр, старший преподаватель, vladislav.karyukin@gmail.com,*
[1]***ИЛЕСОВА Бакытгуль Ералиевна,*** *магистр, преподаватель, bakhytgul92@gmail.com,*
[1]НАО «Казахский национальный университет имени Аль-Фараби», Казахстан, Алматы, пр. Аль-Фараби, 71,
[2]Университет Пердью, США, Индиана, Уэст-Лафейетт, 610 Пердью Молл,
**автор-корреспондент.*

Аннотация. *Сегодня Интернет является эффективным каналом социального взаимодействия во всем мире, но он также открывает большие возможности для кибератак. В последнее время резко возросло количество ботнетов и фишинговых кибератак. Фишинг – это кибератака, осуществляемая с использованием поддельного веб-сайта, предназначенного для кражи конфиденциальной информации. Обнаружение таких атак с помощью алгоритмов машинного обучения и нейронных систем – эффективный и быстрый метод. В данной статье рассматриваются типы интернет-угроз и алгоритмы нейронных сетей, используемые для противодействия фишинговым атакам. Основная цель этой статьи – построить модели нейронных сетей, способные эффективно обнаруживать фишинговые атаки. Методами исследования являются анализ данных, предварительная обработка данных и нейронные сети. На первом этапе был собран набор данных Kaggle, содержащий данные с 5000 законных и 5000 вредоносных веб-страниц. Характеристики набора данных были проанализированы и нормализованы. Затем для классификации угроз были применены глубокая нейронная сеть, сверточная нейронная сеть и сеть с долговременной кратковременной памятью. Глубокая нейронная сеть и сеть с долговременной кратковременной памятью показали лучшие результаты точности до 0,96 и 0,99, тогда как сверточная нейронная сеть достигла только 0,75. В целом, глубокая нейронная сеть и долговременная кратковременная память достигли наилучших результатов среди нейросетевых методов.*

Ключевые слова: *кибербезопасность, нейронная сеть, машинное обучение, фишинг, ботнет, интернет-угрозы, набор данных.*

REFERENCES

1. On approval of the Cyber Security Concept («Cyber Shield of Kazakhstan»). 30.09.2022. https://adilet.zan.kz/kaz/docs/P1700000407
2. The dark side of the Internet: Attacks, costs and responses. Kim, W., Jeong, O.-R., Kim, C., & So, J. (2011). Information Systems, 36 (3), 675-705. doi:10.1016/j.is.2010.11.003
3. Botnets: threats and responses. Jeong, O., Kim, C., Kim, W., & So, J. (2011). International Journal of Web Information Systems, 7 (1), 6-17. doi:10.1108/17440081111125635
4. Detecting Phishing Cyber Attack Based on Fuzzy Rules and Differential Evaluation. Rawaa Mohammed Abdul-Hussein, Ahmed H. Mohammed, Amal Abbas Kadhim. TEM Journal. Volume 11, Issue 2, pages 543-551, ISSN 2217-8309, DOI: 10.18421/TEM112-07, May 2022.
5. A phishing website detection and recognition method based on naive bayes. Xu Yingying, Chen Guangxuan, Liu Qiang, Xu Wanpeng, Zhang, Lei, Wu Jiajian,Fan Xiaoshi. IEEE 6th Information Technology and Mechatronics Engineering Conference, ITOEC 2022.
6. Internet Threat Detection Using Neural Networks. Zhumabekova A.T., Usatova O.A. International scientific journal. «Global science and innovations 2021: central asia». Nur-Sultan, Kazakhstan, December 2021.
7. Comprehensive DDoS Attack Classification Using Machine Learning Algorithms. O. Ussatova, A. Zhumabekova, Y. Begimbayeva, E.T. Matson, N. Ussatov. Computers, Materials & Continua, 2021. DOI: 10.32604/cmc.2022.026552