¹, * Peng Wang

¹Yong Li

² Jing Liu

Research on Classification and Classification Model of Data Security Based on Graph Neural Network



Abstract: - This paper proposes a data security classification and classification model based on graph neural network, aiming to realize the fine management of data assets through advanced neural network technology and expert knowledge. Firstly, a multi-level graph neural network structure is constructed, which can capture the complex dependency relationship between data and learn the deep features of data through the message passing mechanism between nodes. The model introduces an expert knowledge base to embed the experience and rules of domain experts into the learning process of the network, so as to enhance the interpretability and accuracy of the model. In this way, the model can not only automatically identify the security level of the data, but also classify the data in a fine granularity, thus providing strong support for the security management of the data. In order to verify the validity of the model, a series of simulation analyses are carried out in this paper. The experiment draws on real data sets from different industries, including finance, healthcare and education. The results show that compared with the traditional data classification methods, the accuracy and recall rate of this model are significantly improved. Especially when dealing with high dimensional and nonlinear data, the advantages of the model are more obvious. In addition, with the help of expert knowledge, the model can better adapt to the safety norms of specific industries, showing good generalization ability and practicability.

Keywords: Neural Network; Data Analysis; Security Classification; Expert Knowledge.

I. INTRODUCTION

In the wave of digital transformation, data has become the oil of the new era, driving economic development and social progress, but also spawning unprecedented security challenges. Data security is no longer an isolated technical issue, but has risen to the level of national strategy. Traditional security protection measures, such as firewalls, intrusion detection systems, etc., have been difficult to cope with increasingly covert and intelligent network attacks [1]. Therefore, how to build an efficient and intelligent data security classification system has become the hot and difficult point of current research.

The current research status of data security processing shows a trend of diversification and depth. Researchers have begun to focus on data lifecycle management, from the generation of data, storage, transmission to the destruction of every link of strict control [2]. Neural networks, especially deep neural networks, are considered to be an effective way to improve data security processing ability because of their excellent self-learning and adaptive capabilities.

The application of neural network in data security processing is mainly concentrated in two aspects: First, the use of neural network for anomaly detection, by learning the pattern of normal data, to identify abnormal behavior that deviates from the normal pattern; The second is the use of neural networks for threat intelligence analysis, by analyzing a large number of network traffic data, to find potential security threats. However, although neural networks have achieved good results for some specific tasks, they still face some challenges in practical applications [3]. For example, the black-box nature of neural networks makes their decision-making process difficult to interpret, which is particularly important in the security world, where security personnel need to understand how an attack is detected. In addition, problems such as data imbalance, noise, and adversarial attacks also pose challenges to the performance of neural networks.

The research content of this paper focuses on constructing a data security classification and classification model based on graph neural network. In the field of data security, data often exists in the form of graphs, such as social networks, communication networks, etc. Therefore, graph neural networks are very suitable for data security analysis and processing [4]. The model proposed in this paper will combine the powerful representation ability of graph neural networks with the expertise in the field of data security to achieve efficient classification and classification of data. The model can not only improve the accuracy and efficiency of data security processing, but also provide some interpretability to help security personnel better understand and trust the decision of the model.

¹ Inner Mongolia Power Research Institute, Hohhot City, 010020, China

² Inner Mongolia Power (Group)CO, Hohhot City, 010010, China.

^{*}Corresponding author: Peng Wang.

Copyright © JES 2024 on-line : journal.esrgroups.org

The structure of this paper is as follows: First, this paper will review the research status of data security processing, analyze the current challenges and opportunities; Secondly, this paper will introduce the application status of neural network in data security processing, focusing on the characteristics and advantages of graph neural network; Then, the design principle and implementation method of data security classification and classification model based on graph neural network are described in detail [5]. Finally, this paper will verify the validity of the model through a series of experiments, and analyze and discuss the experimental results, and look forward to the future development direction of the model [6]. The research content of this paper aims to provide a new perspective and tool for the field of data security, promote the development of data security processing technology by integrating graph neural network and data security expertise, and contribute to the construction of a more secure and reliable digital environment.

II. NETWORK DATA SECURITY MODEL

A. PPDR model

An American information technology company (PPDR), which first appeared on the Internet, is an initial dynamic information security model composed of protection, detection, response and policy [7]. The Model Framework is shown in Figure 1(image cited in Beyond PS-LTE: Security Model Design Framework for PPDR Operational Environment).



Figure 1. PPDR security mode

Aiming at the problem of data security in PPDR system, a dynamic and time-dependent security mechanism is proposed. The components of the PPDR mode have the following roles:

In the security model of PPDR system, policy plays a core role, and all behaviors of the system are based on it, including system protection, anomaly detection and reaction protection. However, the formulation of security strategy is not random, and a safe and effective strategy is gradually formed based on the understanding of the system through the formulation, evaluation and implementation steps [8]. The specific security that is implemented, and the certain results that are achieved, depend on this security strategy.

The process of protection is a kind of counter process to the threat, attack and other harm to the operation of the system, which is generally carried out through the firewall, net gate and other security devices [9]. This includes the development of safety facilities regulations, configuration and installation of various safety facilities.

Detection is an important part in PPDR model. All operational feedback responses are based on detection, which provides the basis for real-time protection, dynamic response and decision execution [10]. Through the real-time monitoring of the information in the network and the system, when there is an emergency such as danger, abnormal and so on, you can respond quickly through periodic information feedback.

Response as the most critical part of PPDR mode, its implementation is mainly for network security issues, such as network security issues [11]. In order to better solve the problem of information security, people must start from the two aspects of emergency response and emergency handling.

B. WPDRRC Mode

WPDRRC model is a construction mode of information system security assurance system established by some domestic researchers according to the unique network data security form of the country [12]. This model includes all kinds of security elements. Compared with other data security modes, it fully reflects the important role of people in data security. The model's construction is shown in Figure 2.



Figure 2. WPDRRC security model

In the data security mode of WPDRRC, except for the addition of "early warning", "recovery" and "counter" three modules, the rest are completely consistent with the traditional PPDR system. Various methods are used to detect abnormal behavior, collect and analyze whether there is abnormal data and attack behavior in the data, and evaluate the characteristics of attack behavior and possible damage [13]. In addition to technical factors, WPDRRC research also includes personnel participation and strategy, among which the security problem is the synergistic effect of personnel, technology and strategy. In this process, human resources are the most fundamental, and the middle level is the link between human resources and science and technology; Technology is only one layer, and its operation process is supervised by employees and monitored by policies. These three elements play their respective functions in these six parts, and the technology itself is not a simple technology, it provides support for the security of information systems.

C. PDRR Model

PDRR is proposed in the Information Assurance Technology Framework based on the PPDR pattern and is an evolution of the PPDR pattern (Figure 3 is quoted in the Security Model).



Figure 3. PDRR security model

PDRR's protection, detection, and response are basically the same as PPDR's protection, detection, and response, but the difference between PDRR's "recovery" function and PPDR's function is that once it is compromised, the function can repair it to the initial state through the repair function, and add new security, thereby improving the security of the entire system [14].

III. IMPROVED BP NEURAL NETWORK CLASSIFIER

A. Abnormal network data detection classifier model

Because of the serious harm to the security of data, it is very necessary to carry out effective analysis. For new unknown, new and unknown data, conventional security inspection technology cannot effectively detect new and unknown abnormal data. For example, WannaCry ransomware occurred in May 2017, which can spread rapidly on clients and the Internet that have installed conventional anti-virus software, bringing serious harm to users [15]. The root cause is that the existing defense system is unable to actively, efficiently and in advance to prevent new viruses.

This project intends to adopt the fusion method of PCA and BP neural network. On the premise of retaining the original features, the principal element information is obtained through dimensionality reduction processing, and then the principal metadata obtained after dimensionality reduction is used as the input layer of BP neural network. This method uses the multi-layer, multi-layer and multi-level neuron composition of artificial neural network to train the training sample so that it has a certain weight and threshold value, so that it can distinguish the existing samples and effectively identify the similar samples [16]. This project intends to introduce this feature into data security research, so as to solve the deficiency of existing security equipment detection technology that cannot actively identify unknown information, and improve detection efficiency and accuracy. the topology of the network is shown in Figure 4 (the image is referenced in Soft Computing, 2020, 24: 13219-13237).



Figure 4. Network topology relationship diagram between PCA and BP neural network classifier

In this study, the algorithm is combined with the classification algorithm of neural network to perform PCA on the data processed by PCA, which not only maintains all the original information characteristics, but also reduces a large number of samples. BP neural network is used to deal with nonlinear problems, and the generalization and robustness of BP network are enhanced [17]. The accuracy and detection effect have been greatly improved.

B. BP neural network model

Among them, BP neural network is the most representative one. BP neural network is a kind of feedforward neural network, by comparing the real input of each neuron with the predetermined expected output, when there is a certain deviation between the two, the BP neural network will feedback it back until the end of training. BP

neural network is a hierarchical neural network in structure, it has a level between the input, output and hidden layers [18]. Since there are connections between neurons of different levels, while neurons of the same level are not connected, this paper believes that neurons of the same level receive signals from upper neurons, and neurons of other levels only affect the output of lower neurons (Figure 5).



Figure 5. BP neural network

When BP neural network performs information security, it must be trained first, and its performance will affect the discovery of anomalies [19]. Therefore, how to train neural network effectively is very critical. Figure 6 shows the schematic diagram of the BP neural network and the specific training process.



Figure 6. BP neural network training diagram

C. Data security evaluation method of fuzzy neural network

When evaluating the security of data, some indicators have strong subjectivity, and it is difficult to judge their values, while BP is only suitable for processing partially quantified data, and the analysis and processing ability

of qualitative indicators is poor. A data security evaluation method based on fuzzy neural network is proposed. The following methods are used to blur the data security risk assessment index:

1) Identify the main risk factors that may pose threats to system assets, threats, vulnerabilities, threats and vulnerabilities through correlation research.

2) Use fuzzy mathematics method to establish the set of risk factors in the data.

$$A = \{a_1, a_2, \mathcal{L}, a_n\} \tag{1}$$

3) Establish a comprehensive evaluation system based on fuzzy evaluation. The research includes the privacy, integrity, availability value of the system, the severity of vulnerability, the ease with which vulnerability can be transformed into benefits, the availability of vulnerability and the technicality of vulnerability. According to the score given by each expert, it is divided into m levels, among which fuzzy evaluation is as follows:

$$B = \{b_1, b_2, \mathcal{L}, b_m\}$$
⁽²⁾

4) Each index was evaluated by experts and fuzzy mapping was established.

$$g: A = G(B) \tag{3}$$

$$a_{i} = g(a_{i}) = (s_{i1}, s_{i2}, L, s_{im}) \in G(B)$$
(4)

G(B) represents all the fuzzy sets on B. g represents the likelihood of a data risk factor index a_i associated with a particular review. Assuming that S_i represents the membership vector of the risk factor a_i in the evaluation set B, the membership matrix S can be calculated from the membership degree of the factor.

$$S_{i} = \{s_{i1}, s_{i2}, L, s_{im}\} (i = 1, 2, L, n)$$
(5)

5) Each evaluation index in the set of judgments is weighted with reference to expert reviews. Assign weights to set Q, assigning weights according to the value of the evaluation, and the sum of each weight is 1. According to the fuzzy transformation formula, O:

$$Q = \{q_1, q_2, \mathcal{L}, q_m\}$$
(6)

$$O = QS^{W} \tag{7}$$

O represents the weighting of the risk factors in each data, with a value between (0,1). Finally, the data security risk assessment index after fuzzy processing is obtained, which is used as the input of BP neural network algorithm.

A BP neural network is set as layer L, the first layer is the input layer, the second layer is the hidden layer, and the L layer is the output layer. If the l layer has $n_l (l = 1, 2, L, L)$ neuron, then the weight factor of the ineuron input to the l layer is $\lambda_{ij}^l (i = 1, 2, L, n_l; j = 1, 2, L, n_{l-1})$, then the input/output conversion relationship of this BP neural network can be expressed as:

$$D_{i}^{(l)} = \sum_{j=1}^{n_{l-1}} \lambda_{ij}^{l} x_{j}^{l-1} \beta_{i}^{(l)}$$
(8)

$$X_i^{(l)} = g(D_i^{(l)}) = \frac{1}{1 + e^{-\eta D_i^{(l)}}}$$
(9)

Let the input and output sample expressions for group P be:

 $X_{k}^{(0)} = [X_{k1}^{(0)}, X_{k2}^{(0)}, L, X_{kn1}^{(0)}]^{T}$ (10)

$$f_k = [f_{k1}, f_{k2}, \mathcal{L}, f_{kn_l}]^T (l = 1, 2, \mathcal{L}, m)$$
(11)

From the above, it can be seen that the learning algorithm of BP neural network is as follows:

$$\lambda_{ij}^{(l)}(t+1) = \lambda_{ij}^{(l)}(t) + \delta F_{ij}^{(l)}(t+1)$$
(12)

$$F_{ij}^{(l)} = \sum_{k=1}^{P} \varepsilon_{ki}^{(l)} X_{kj}^{(l)}$$
(13)

$$\varepsilon_{ki}^{(l)} = \left(\sum_{t=1}^{l+1} \varepsilon_{kt}^{(l+1)} \lambda_{yi}^{(l+1)}\right) \eta X_{ki}^{(l)} (1 - X_{ki}^{(l)})$$
(14)

$$\varepsilon_{ki}^{(L)} = (f_{ki} - X_{ki}^{(L)})\eta X_{ki}^{(L)} (1 - X_{ki}^{(L)})$$
(15)

D. Data security evaluation process of fuzzy neural networks

Firstly, by judging the evaluation factors and grading criteria, and then using the fuzzy set transformation principle, using the membership degree method to describe the fuzzy boundary between each element, a fuzzy judgment matrix is constructed, and this matrix is the input of BP neural network. Then BP neural network is used to calculate and obtain the level of the target to be evaluated [20]. The detailed implementation steps are:

1) Establish a set of security risk factors and set $Q = \{q_1, q_2, L, q_n\}$, where n represents the number of elements in the factors;

2) Construct the decision set. Different evaluation sets of resources, threats and vulnerabilities are established and represented by $O = \{o_1, o_2, L, o_m\}$, where m is the number of elements in the evaluation set.

3) Each factor in factor Q is evaluated by reference to evaluation set O, and A fuzzy graph $g: Q = G(O), q_i \rightarrow g(q_i) = (k_{i1}, k_{i2}, L, k_{im}) \in G(O)$ is constructed by expert evaluation of each factor. Here g is the correspondence between the support degree of security risk factor q_i for each comment in the evaluation set. Calculated from the membership vector $P_i = (k_{i1}, k_{i2}, L, k_{im})$ of the judgment set O, the following membership matrix is obtained:

$$P = \begin{bmatrix} k_{11} & k_{12} & L & k_{1m} \\ k_{21} & k_{22} & L & k_{2m} \\ M & M & M & M \\ k_{n1} & k_{n2} & L & k_{nm} \end{bmatrix}$$
(16)

4) k_c, k_t and k_f represent the membership degree of each factor related to asset, threat and vulnerability level respectively, and $\phi = (\phi_1, \phi_2, L_{-}, \phi_n)$ is the weight vector corresponding to each factor. After the calculation, the index weight vector $A = (a_1, a_2, L_{-}, a_{n1})$ of asset evaluation set, the index weight vector of threat judgment set is $B = (b_1, b_2, L_{-}, b_{n2})$, and the index weight of vulnerability judgment set is $W = (w_1, w_2, L_{-}, w_{n3})$.

5) Based on the weight vector of asset, threat and weakness judgment set, BP neural network is constructed, and the security level of data is obtained through network learning and training.

IV. CASE ANALYSIS

Security risk assessment of an information system This section considers the security risks of data facilities. It can be seen from Figure 7 that data facilities involve five types of risk factors, namely computer operating system A, network operating system B, network communication protocol C, general application platform D and network management data E, as shown in Figure 7. The following takes building BP network model of computer operating system for security risk assessment as an example. Firstly, fuzzy theory is used to evaluate the system and the target estimate of the system is obtained. Then the trained network is used to complete the evaluation.

The security of a particular information system is evaluated. As can be seen from Figure 7, the data device includes the following five types of risk factors: computer operating system R_1 , network operating system R_2 , network communication protocol R_3 , common application platform R_4 , and network management data R_5 (figure cited in Sustainability 2020, 12(10), 415). In the following, the security risk assessment will be carried out by establishing a computer operating system model based on BP network [21]. Firstly, fuzzy theory is used to evaluate the system, the index of the system is obtained, and it is used as the sample set of the network, and then the network is used to evaluate.



Figure 7. Data facility security risk factors

1) Construct factor set and judgment set, obtain membership matrix Q_c , Q_t and Q_f , and calculate entropy weight coefficients of each factor and weight vectors of each index [22]. The set of factors represents $V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$. $v_i (i = 1, 2, ..., 7)$ stands for "defect", "backdoor", "corruption", "password acquisition", "Trojan Horse", "virus", "upgrade defect" and other dangerous factors]. The structure decision set is represented by $U_c = \{u_{c1}, u_{c2}, u_{c3}, u_{c4}, u_{c5}\}, U_t = \{u_{t1}, u_{t2}, u_{t3}, u_{t4}, u_{t5}\}, U_f = \{u_{f1}, u_{f2}, u_{f3}, u_{f4}, u_{f5}\}$. The membership matrix Q_c , Q_t and Q_f were obtained by integrating the probability of each risk factor associated with each index with the evaluation results of each risk factor by experts. According to the entropy weight coefficient method, the weight vector corresponding to each factor is obtained: $\Theta_{c1} = (\theta_{c1}, \theta_{c2}, \theta_{c3}, \theta_{c4}, \theta_{c3}) = (0.111, 0.096, 0.219, 0.170, 0.108, 0.188, 0.108)$

$$\Theta_{c} = (\theta_{c1}, \theta_{c2}, \theta_{c3}, \theta_{c4}, \theta_{c5}, \theta_{c6}, \theta_{c7}) = (0.111, 0.096, 0.219, 0.170, 0.108, 0.188, 0.108),$$

$$\Theta_t = (\theta_{t1}, \theta_{t2}, \theta_{t3}, \theta_{t4}, \theta_{t5}, \theta_{t6}, \theta_{t7}) = (0.101, 0.120, 0.109, 0.115, 0.120, 0.240, 0.195),$$

$$\Theta_f = (\theta_{f1}, \theta_{f2}, \theta_{f3}, \theta_{f4}, \theta_{f5}, \theta_{f6}, \theta_{f7}) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.138, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.128, 0.128, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.128, 0.128, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.128, 0.128, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.128, 0.128, 0.127, 0.116, 0.221, 0.235, 0.045) = (0.118, 0.128,$$

The weight of each index in the evaluation set is

A = (1/15, 2/15, 1/5, 4/15, 1/3),

$$B = (1/15, 2/15, 1/5, 4/15, 1/3),$$

$$C = (1/15, 2/15, 1/5, 4/15, 1/3)_{\circ}$$

2) Construct three-layer BP neural network for computer operating system R_1 . The neural network of the neural network consists of 35 neurons. The fuzzy dominance matrix composed of seven factors accepts the values of 35 factors each. The result of operation $l = \log_2 35 \approx 6$ shows that the hidden layer consists of 6 elements. The output layer contains a risk assessment corresponding to each element of the system.

Log-sigmoid function was used to analyze log-consistency. In order to enhance the generalization performance of the model, a strategy of ending first is proposed, which divides the samples into two groups: training and testing [23]. The function gradient of the network performance and the modified network weight and threshold are obtained by using the training samples. This test case is used to calculate and verify the artificial

Table 1. Results of fuzzy algorithm and neural network training of samples								
	1	2	3	4	5	6	7	8
FT evaluation	0.1745	0.1748	0.1565	0.1597	0.1679	0.1654	0.1546	0.1592
FNN evaluation	0.1852	0.1812	0.1726	0.1723	0.1699	0.1666	0.1608	0.1573
	9	10	11	12	13	14	15	16
FT evaluation	0.1506	0.1545	0.1501	0.1514	0.1405	0.1271	0.1124	0.1654
FNN evaluation	0.1573	0.1557	0.1545	0.1531	0.1485	0.1212	0.1168	0.1590

neural network. Fifteen samples were selected for this article, one of which was for testing. The mean variance e of the sample is set to 10-4, and the learning rate α is set to 0.05(Table 1).

The number of output errors and cycles for the neural network sampling set is shown in Figure 8. The results of the artificial neural network and the evaluation method based on fuzzy logic are shown in Figure 9. The results show that the model is close to the risk degree evaluated by the fuzzy evaluation method, and the adaptability of the method is good.



Figure 8. Network output error and number of cycles



Figure 9. Comparison of theoretical values with network output

3) Risk frequency and risk factor of vulnerability are 0.1686 and 0.232; By establishing BP neural network model, people can get the corresponding security risk level from the aspects of network operating system, network communication protocol, general application platform, network management data and so on. According to the importance of each component in data equipment and the idea of system comprehensive evaluation, a weight-based method is proposed to evaluate the security risk of data system, which will not be introduced in detail here. Through comparative analysis, a small level of security risk is obtained, indicating that the system is safe and reliable.

Three representative data security models such as K-mean, fuzzy C-mean and support vector set are compared with existing data security models to verify the advantages of this model in terms of recognition rate and false positive rate. K-means is a common method used to find abnormal data in clustering [24]. It divides a data set into several categories or categories, and makes use of clustering to make the data in the same category or category have the greatest similarity while having the greatest difference, so as to achieve effective detection and identification of abnormal data. FCM is an improvement of the traditional C-means clustering method, which can maximize the similarity of objects divided into a class and apply it to the detection of abnormal data. A new algorithm based on cloud modeling and semi-supervised clustering is proposed to detect and identify the abnormal data effectively by differentiating the weights among the features.

Through comparative experiments, some indexes such as detection rate and false alarm rate are selected to evaluate the performance of the method. Three different test samples are set up, three test samples are tested, and then the average of the test results of the three samples is calculated. Compared with K-mean, fuzzy C-mean, support vector machine, and feature extraction methods in BP network, the recognition rate and false alarm rate of this method are shown in Figure 10.







As can be seen from the test results in Figure 10, compared with K-means, FCM and support vector set, this method has higher accuracy and lower false positive rate.

V. CONCLUSION.

After the in-depth research and experimental verification in this paper, people can confidently conclude that the data security classification and classification model based on graph neural network shows significant advantages in dealing with complex and changeable data security problems. Compared to traditional clustering algorithms such as K-means and fuzzy C-means (FCM), and classification algorithms such as support vector machines (SVM), graph neural networks (GNN) provide more powerful and flexible analytical tools for data security applications. First, GNN is able to naturally deal with data structures in non-Euclidean space, such as social networks, communication networks, etc. These network structures usually exist in the form of graphs, with nodes and edges representing data entities and their interrelationships. This ability allows GNN to capture deep

dependencies between data that vector-based clustering algorithms such as K-means and FCM do not. This deep relationship capture is particularly critical in data security classification, as it can reveal potential risks and threats hidden beneath the surface of the data. Second, GNN shows greater efficiency and accuracy when dealing with high-dimensional data and large-scale datasets. In contrast to SVM, GNN does not need to manually select features and is not affected by the "dimensional disaster", it can automatically learn useful feature representations from the data. This is particularly important in the field of data security, where secure data tends to be highdimensional and noisy, and manual selection of features is time-consuming and prone to missing important information. In addition, GNN has shown its unique advantages in dealing with unbalanced data sets and adversarial attacks. Traditional algorithms tend to suffer performance degradation in these cases, while GNN, through its distributed presentation learning ability, is better able to adapt to the unbalance of data and to a certain extent resist adversarial attacks, which is essential to ensure data security. When dealing with data security problems, the data security classification and classification model based on graph neural network can not only provide higher classification accuracy and stronger robustness, but also capture the complex structural relationship between data, so as to provide more abundant and in-depth information for security analysts. Therefore, GNN has a broad application prospect in the field of data security and is expected to become an important tool for data security processing in the future. Future research will further optimize the GNN model to improve its performance and interpretability in real-world applications to meet changing data security needs.

VI. REFERENCES

- Liu, X., Zhao, J., Li, J., Cao, B., & Lv, Z. Federated neural architecture search for medical data security. IEEE transactions on industrial informatics, Vol.18(2022) No.8, p. 5628-5636.
- [2] Gupta, R., Gupta, I., Saxena, D., & Singh, A. K. A differential approach and deep neural network based data privacypreserving model in cloud environment. Journal of Ambient Intelligence and Humanized Computing, Vol.14(2023) No.5, p. 4659-4674.
- [3] Chen, Z. Research on internet security situation awareness prediction technology based on improved RBF neural network algorithm. Journal of Computational and Cognitive Engineering, Vol. 1(2022) No.3, p. 103-108.
- [4] Yu, F., Shen, H., Yu, Q., Kong, X., Sharma, P. K., & Cai, S. Privacy protection of medical data based on multi-scroll memristive Hopfield neural network. IEEE Transactions on Network Science and Engineering, Vol.10(2022) No.2, p. 845-858.
- [5] Pankova, M., Kwilinski, A., Dalevska, N., & Khobta, V. Modelling the Level of the Enterprise'Resource Security Using Artificial Neural Networks. Virtual Economics, Vol. 6(2023) No.1, p. 71-91.
- [6] Yang, Y. G., Niu, M. X., Zhou, Y. H., Shi, W. M., Jiang, D. H., & Liao, X. A visually secure image encryption algorithm based on block compressive sensing and deep neural networks. Multimedia Tools and Applications, Vol. 83(2024) No.10, p. 29777-29803.
- [7] Pan, X., Zhao, T., Chen, M., & Zhang, S. Deepopf: A deep neural network approach for security-constrained dc optimal power flow. IEEE Transactions on Power Systems, Vol.36(2020) No.3, p. 1725-1735.
- [8] Balan, N., & Ila, V. A Novel Biometric Key Security System with Clustering and Convolutional Neural Network for WSN. Tehnički vjesnik, Vol.29(2022) No.5, p. 1483-1490.
- [9] Song, Y., Fu, Y., Yu, F. R., & Zhou, L. Blockchain-enabled Internet of Vehicles with cooperative positioning: A deep neural network approach. IEEE Internet of Things Journal, Vol.7(2020) No.4, p. 3485-3498.
- [10] Kumar, M., Mukherjee, P., Verma, K., Verma, S., & Rawat, D. B. Improved deep convolutional neural network based malicious node detection and energy-efficient data transmission in wireless sensor networks. IEEE Transactions on Network Science and Engineering, Vol.9(2021) No.5, p. 3272-3281.
- [11] Wang, H., Cao, Z., & Hong, B. A network intrusion detection system based on convolutional neural network. Journal of Intelligent & Fuzzy Systems, Vol. 38(2020) No.6, p. 7623-7637.
- [12] Drewek-Ossowicka, A., Pietrołaj, M., & Rumiński, J. A survey of neural networks usage for intrusion detection systems. Journal of Ambient Intelligence and Humanized Computing, Vol.12(2021) No.1, p. 497-514.
- [13] Guo, Z., Shen, Y., Bashir, A. K., Imran, M., Kumar, N., Zhang, D., & Yu, K. Robust spammer detection using collaborative neural network in Internet-of-Things applications. IEEE Internet of Things Journal, Vol.8(2020) No.12, p. 9549-9558.
- [14] Guo, C., Chen, C. H., Chang, C. C., Hwang, F. J., & Chang, C. C. De-correlation neural network for synchronous implementation of estimation and secrecy. IEEE Communications Letters, Vol.27(2022) No.1, p. 165-169.
- [15] Wong, L. W., Tan, G. W. H., Lee, V. H., Ooi, K. B., & Sohal, A. Psychological and system-related barriers to adopting blockchain for operations management: an artificial neural network approach. IEEE Transactions on Engineering Management, Vol. 70(2021) No.1, p. 67-81.
- [16] Xu, R., Joshi, J., & Li, C. Nn-emd: Efficiently training neural networks using encrypted multi-sourced datasets. IEEE Transactions on Dependable and Secure Computing, Vol.19(2021) No.4, p. 2807-2820.
- [17] Zhou, K., Wang, W., Wu, C., & Hu, T. Practical evaluation of encrypted traffic classification based on a combined method of entropy estimation and neural networks. Etri Journal, Vol. 42(2020) No.3, p. 311-323.

- [18] Zitar, R. A., Al-Dmour, N., Nachouki, M., Hussain, H., & Alzboun, F. Hashing generation using recurrent neural networks for text documents. ICIC Express Letters Part B: Applications, Vol.12(2021) No.3, p. 231-241.
- [19] Rathee, G., Jaglan, N., Garg, S., Choi, B. J., & Jayakody, D. N. K. Handoff security using artificial neural networks in cognitive radio networks. IEEE Internet of Things Magazine, Vol.3(2020) No.4, p. 20-28.
- [20] Cheng, H., Liu, X., Wang, H., Fang, Y., Wang, M., & Zhao, X. SecureAD: A secure video anomaly detection framework on convolutional neural network in edge computing environment. IEEE Transactions on Cloud Computing, Vol. 10(2020) No.2, p. 1413-1427.
- [21] Lin, G., Wen, S., Han, Q. L., Zhang, J., & Xiang, Y. Software vulnerability detection using deep neural networks: a survey. Proceedings of the IEEE, Vol. 108(2020) No.10, p. 1825-1848.
- [22] Dastres, R., & Soori, M. Artificial neural network systems. International Journal of Imaging and Robotics (IJIR), Vol. 21(2021) No.2, p. 13-25.
- [23] Alwan, A. H., & Kashmar, A. H. Block Ciphers Analysis Based on a Fully Connected Neural Network. Ibn AL-Haitham Journal for Pure and Applied Sciences, Vol. 36 (2023) No.1, p. 415-427.
- [24] Dhawan, S., & Gupta, R. Analysis of various data security techniques of steganography: A survey. Information Security Journal: A Global Perspective, Vol.30(2021) No.2, p. 63-87.