Research on Data Mining Algorithm Based on BP Neural Network

Jingyou Zhang^{1*}, Haiping Zhong²

¹School of Mathmatics and Statistics, Chongqing Three Gorges University, Wanzhou 404100, China ²Mathematics and Computer Science College, Yuzhang Normal University, Nanchang 330103, China

> Received: July 9, 2021. Revised: January 7, 2022. Accepted: January 26, 2022. Published: February 25, 2022.

Abstract—The current data mining algorithm has the problem of imperfect data mining function, which leads to the algorithm taking too long time. This paper designs a data mining algorithm based on BP neural network. Analyze the basic structure of the data mining algorithm, obtain the data characteristics of the multi-objective decision-making, adjust the convergence speed with the distributed computing technology to keep the inertia factor state unchanged, construct the local minimal discrete model, measure the interest of the model, calculate the optimal output value of the network using the BP (Back Propagation) neural network model, and complete the improved design of the data mining function. Experimental results: The average computational time consumption of the designed data mining algorithm is 559.827 seconds, which saves 145.975 seconds and 174.237 seconds respectively than other traditional algorithms. It is proved that the data mining algorithm based on BP neural network reduces the computational time consumption, improves the performance of data mining, and has high application value.

Keywords—BP neural network, convergence rate, dispersion model, data mining, data mining algorithm, hidden layer structure.

I. INTRODUCTIONS

WITH the development of information technology, the information data explodes, and the amount of data is also increasing in [1]-[2]. In order to effectively use and discover a large amount of knowledge in the huge amount of information, data mining technology has emerged and developed vigorously under the joint efforts of scholars and enterprises at home and abroad. Data mining is intended for extracting useful and interesting knowledge and information [3] from large amounts of noisy and incomplete data. It can be widely used in various applications, including business management, production control, market analysis, engineering design, and scientific exploration. Data mining is able to reveal implied, previously unknown, and potentially valuable information from a large amount of data from the database. Data mining is a kind of decision support process, it is mainly based on artificial intelligence, machine learning, mode recognition,

statistics, database, visualization technology, etc., highly automated analysis of enterprise data, make inductive reasoning, dig out the potential mode, help decision makers adjust market strategy, reduce risk, make the right decision. Classification is the most important step in data mining. The traditional data mining classification algorithm shows great limitations in the growth of data volume, and the mining function is imperfect. Classification is the most important step in data mining classification algorithm shows great limitations in the growth of data volume, and the mining function is imperfect.

Li et al. [4] presented an integrated algorithm based on interval weights for the distortion problem existing in traditional data mining time series classification models. The interval weights were used to assign different weights to the different intervals of the time series to solve the problem that the subsequence characteristics are not obvious. The base classifier is then determined to ensure the performance of the integrated classifier. Then, the integrated classifier is trained on the training set, the improved integrated classifier training is parallelized, classifying the more time-consuming parts, and realizes the improvement of the data mining classification algorithm.

Chi [5] improved on the traditional classification algorithm, the decision tree ID3 (Iterative Dichotomiser 3) algorithm, for data mining. On the basis of the full analysis of the decision tree ID3 algorithm, the decision tree ID3 algorithm improves the decision tree I D 3 big data classification algorithm. Some sample set of college students' football training was selected to refer to the attribute weights given by sports experts according to experience, and quickly and reasonably.

However, although the above data mining improvement method improves the mining function of the traditional mining classification algorithm, due to the complex computing process and the long computing time, so it cannot meet the actual needs. To solve this problem, this paper proposes data mining algorithm based on BP neural network.

BP neural networks belong essentially to feedforward neural networks and are widely used in artificial neural networks. It has a perfect theoretical system and learning mechanism, and establishes a multi-layer perceptron model by simulating the response process of human brain neurons to external stimuli. BP neural networks work based on error backpropagation methods. Through repeated iterative learning, the mathematical expressions of this mapping relationship can achieve the accuracy of multiple mapping relationships. Its most prominent advantage is that its basic ideas are easy to understand, and they have a strong non-linear ability. Therefore, the BP neural network can use the prior information to obtain the mapping relationship between the data, and improve the mining performance of the data mining classification algorithms when the amount of data grows [6].

The innovation point of this paper is to analyze the basic structure of the decision level of data mining algorithm, obtain the data characteristics of multi-objective data decision, adjust the convergence rate of data mining algorithm, keep the inertia factor state unchanged, construct the local minimal discrete model, measure the model interest, use the positive calculation process of BP network, calculate the output value of the network, compare with the actual value, continuously reduce the error and complete the improved design of data mining function. The experimental results show that the data mining improvement algorithm designed here greatly reduces the computational time and has high application value compared with other data mining improvements.

II. RESEARCH ON THE DATA MINING ALGORITHM BASED ON THE BP NEURAL NETWORK MODEL

A. Obtaining Data Features of Multi-objective Decision Making

Feature extraction is very important in data mining algorithm, which directly affects the classification results and subsequent classification work. In many cases, the decision-maker does not make a decision according to the size of the objective vector in the natural partial order [7], [8]. Even under this principle, because the natural partial order is not a total order, many target vectors cannot be compared. Therefore, it is necessary to study the general description of decision-maker's judgment. Generally, the attitude given by the decision-maker to judge whether a decision is good or bad is called the decision-maker's preference for decision-making problems. The basic structure of decision hierarchy is shown in Figure 1.



Figure 1. Decision hierarchy

It can be seen from Figure 1 that the elements of the same level are independent of each other, and the elements of the upper level dominate some or all of the elements of the next level, thus forming a multi-level analysis structure dominated by layers from top to bottom. According to the relative distance from the alternative to the positive ideal solution and the negative ideal solution, the expression formula of the standardization matrix is as follows:

$$P_{ij} = \frac{a_{ij}}{\sqrt{\sum_{i=1}^{e} (a_{ij})^2}}$$
(1)

In Formula (1), *a* is the positive ideal solution, *i* is the decision coefficient, *j* is the maximum eigenvector, *e* is the negative ideal solution, and satisfies the $(1 \le i \le a, 1 \le j \le e)$ condition. Preference is the decision maker's judgment on whether any two vectors in the decision space have some properties. Therefore, in mathematical sense, preference is a binary relation in decision space, and the expression formula of multi criteria preference index is as follows:

$$\beta(A,B) = \frac{\sum_{j=1}^{a} r_j(A,B)}{\sum_{j=1}^{a} r_j}$$
(2)

In Formula (2), the meanings of a and j are the same as those in Formula (1). A, B respectively represent any two decision schemes in the decision-making level, and rrepresents the net flow. According to the size of the net flow to sort the scheme, the larger the net flow, the better the scheme. After screening, enter the data preparation stage. In the data preparation stage, it is the beginning stage of the whole data mining process, which is very important and occupies more than 60% of the time. The quality of the data preparation stage directly affects the feasibility of the subsequent steps and the quality and effectiveness of the whole mining. Data from different data sources are integrated to facilitate the unified storage and processing of subsequent calculation and analysis, and realize data integration. The process of data cleaning is to remove the data that contains noise or has nothing to do with the mining topic in the data source, so as not to affect the accuracy of the results. The main purpose of processing the integrated data is to transform the data into a storage format that is easy to be mined and calculated, so that the data can adapt to the input format and mode of mining algorithm. Data mining is the real process of mining and analysis. The stage of evaluation and representation is to evaluate the mined models, remove the patterns that do not meet the evaluation criteria, and express the results through various visualization tools, which makes it easier for people to understand and accept, and show users the relevant information mined. A feature item set must have the following two characteristics: integrity, feature items should be able to summarize all document contents; Distinguishability: feature itemsets should be able to distinguish documents from each other [9]. In addition to strong generalization ability and profound knowledge, it is better to construct feature sets manually with linguists according to the general principles of human feature extraction. In addition, we should increase the frequency of documents for the feature items with high specificity, and reduce the frequency of documents for the feature items with low specificity. Based on the above description, the multi-objective decision data feature acquisition step is completed.

B. Distributed Computing Adjusts Convergence Rate

The emergence of distributed computing just provides favorable conditions for massive data mining. In the data mining technology, the distributed computing technology is used to divide the computing tasks of large-scale data into many independent small tasks, and these small tasks are assigned to multiple computing servers for parallel computing. Finally, the computing results are merged to complete the whole mining computing task. This method can realize the mining task under the condition of large-scale data, and has high efficiency and good scalability. Applying gradient descent method to search unconstrained optimization problem is the essence of BP neural network training process. The weight matrix is modified in the direction of the first negative gradient vector of error pair weight and closed value, so as to minimize the network output error. If you want to adjust the convergence rate, you can't choose too large learning rate [10]. In the general operation process, the learning rate of the network is usually selected by people's previous experience, and the learning rate and inertia factor remain fixed in the operation process. When the network operation process is close to the local minimum or the best point, we can use quadratic function to approximate the objective function. The equipotential surface of the objective function is close to an ellipsoid. The direction of the eigenvector determined by the maximum eigenvalue and the minimum eigenvalue of the objective function represents the major axis and the minor axis of the ellipsoid respectively. The calculation formula of the objective function is as follows:

$$f(x) = \sum_{d \in Q\{x\}} L \tag{3}$$

In Formula (3), Q is the local minimum, L is the maximum eigenvalue, and d is the data dimension. When the calculated curve is flat, the zigzag phenomenon may appear in the iterative process. At this time, if the learning rate can not be adjusted in time, the whole operation process will become slow. When dealing with different problems, there are significant differences in the rules that the network needs to meet. If we only rely on experience to determine the parameters, it is difficult to widely apply the network to other fields, which seriously restricts the development and promotion of data mining algorithm. Distributed network storage technology usually refers to the large-scale data is divided into multiple blocks, the different blocks are stored on a number of independent storage devices. Distributed network storage uses easy to expand storage structure, uses multiple storage devices to reduce storage burden, and uses central server to locate and manage storage information. The search step size of gradient descent method is generally obtained by

calculation, but due to the high-dimensional complexity of BP neural network model, the corresponding learning rate is fixed by artificial selection. On the basis of the above, the goal of adjusting the convergence rate is achieved.

C. Construction of Local Minimum Point Dispersion Model

If the error function is strictly convex, then it has a unique minimum (at this time, the local minimum is also the global minimum), so the network may converge to obtain its optimal weight and closed value. However, for the algorithm, its training error function is a nonlinear function of independent variables; the weight space constructed by such a function is not a paraboloid with only one minimal point, but a hypersurface with multiple local minimal points (that is, multiple local minimal points, saddle points or flexible stationary points). Because of the existence of these local minima, the data mining algorithm is easy to fall into these local minima or saddle points when searching for the optimal connection weights and thresholds, and stop the iteration and update, so that the network can not converge to get the optimal solution of the problem, and can not get the optimal connection weights and thresholds and other parameters of the network. The basic trend of parameter development is described by smooth curve, the specific expression formula is as follows:

$$y_z = h(z, \phi) + \gamma_z \tag{4}$$

In Formula (4), h is the prediction object, z is the prediction error, ϕ is the trend rate, and γ is the trend index. According to the calculation results of formula (4), the parameter variation law is obtained, and the data discrete state is obtained:

$$S_{g+m} = u_m + v_m \times K + \frac{1}{2} \cdot K^2$$
(5)

In Formula (5), u is fluctuation coefficient, v is time series, K is error correction series, g is smoothing absolute error, and m is data fitting degree. This characteristic determines that whether in theoretical research or engineering application, it is meaningless to discuss data mining separately from closely related disciplines. For specific tasks, it is often necessary to skillfully combine different technologies. Data mining is the integration and application of various disciplines aiming at knowledge discovery. By integrating new methods from multiple disciplines, the ability of data mining can be significantly enhanced. It needs to be oriented to specific requirements and applications. This leads to different scenarios and tasks need to deal with different types of data, using different analysis and processing techniques to obtain specific models and discovery knowledge. On the other hand, in the face of specific user needs, different mining objectives will produce different results, usually need to combine the background knowledge of the task, constraints and rules, other knowledge of the research field into the mining process. Therefore, data mining does not have the best algorithm, only the most suitable algorithm for specific tasks. Large volume and type rich data. The information age gives birth to data mining, and the existence of data must be the primary premise of analysis and mining. At the same time, data generally need to reach a certain scale in order to reflect the universal characteristics in the real world and to explore the discreteness of data [11], [12]. In order to find valuable patterns for a given user effectively, the measurement of model interest is indispensable. Model interest can be divided into subjective and objective measurement. Objective measurement is usually processed by probability in statistics, such as the support or confidence of association rules; the subjective measurement is based on the user's belief in the data. For example, the pattern obtained is similar to the user's premonition or the discovery pattern is unexpected, which leads to the difference of the same knowledge discovery for different users. Generally, these two measurement methods are complementary and combined in practical application. On the basis of the above description, the construction of local minimum point dispersion model is completed. Based on the above description, the steps of constructing local minimum point dispersion model are completed.

D. BP Neural Network Design Data Mining Algorithm Function

The training process of the network is actually the calculation process of the network, and its principle is: use the forward calculation process of BP network to calculate the output value of the network, compare the output value with the actual value, if there is a large error, then carry out the reverse calculation process [13]-[15]. By calculating the nearest distance between the data to be detected and each cluster center, comparing the nearest distance and threshold size, judging whether the data is normal or abnormal, outputting the normal optimal solution, completing the improved design, and returning the anomalous solution to repeated calculation until the optimal normal solution is obtained. The specific steps are:

First of all, we should analyze the problems and determine the level objectives. By comparing the decision matrix, according to the scaling method, through expert consultation, combined with domain knowledge and expert opinions, the relative importance of each pair of indicators in each hierarchy is scored, and the pairwise comparison decision matrix is established. The maximum eigenvalue and eigenvector of pairwise comparison decision matrix are calculated, and the hierarchical single ranking is carried out. The consistency was tested according to the consistency ratio. The expression formula of decision level network is obtained as follows:

$$H_{\eta} = q\left(net_{\eta}\right) \tag{6}$$

In Formula (6), q is the number of neurons and η is the number of nodes. On this basis, the expression formula of hidden layer nonlinear mapping is obtained:

$$\Delta \delta = -\varphi \frac{Cw_s}{Cw_{\sigma k}} \tag{7}$$

In Formula (7), φ is the learning objective function, C is the input signal, w is the number of samples, k is the output node, s is the input node, σ is the network error. According to the calculation results, the signal propagation mode is set. The input signal is transmitted forward, and the initial input comes in from the input layer, which is processed by the neurons in the hidden layer, and the final result is produced by the output layer. In this process, the connection weight of the network does not change. The weights of each layer of neurons are only associated with the adjacent layer of neurons. The definition of a schema is an abstract description of a dataset. There are two types of patterns in data mining, including predictive pattern and descriptive pattern [16]-[18]. Predictive model is based on the current data set to predict the value of unknown data [19]. Typical predictive models include component model and sequence model. Descriptive pattern can not be directly used for prediction. It mainly gives a specific description of the rules and patterns mined from the existing data sets and divides similar data into the same group. Descriptive pattern mainly includes clustering pattern and association pattern. Based on the back propagation method of error gradient descent, the signal propagation formula of connecting weights and thresholds is as follows:

$$D = \tau \left(1 - E_{\zeta} \right) \left(R - E_{\zeta} \right) \tag{8}$$

In Formula (8), τ is the connection weight, *E* is the number of training samples, ζ is the initialization threshold, and *R* is the topology parameter. According to the provided sample set, the input training sample pair is determined. The basic flow of data mining is shown in Figure 2.



Figure 2. Data mining process

As can be seen from Figure 2, each link of data mining has cross parts and corresponding order. The functions of data mining algorithm are: according to the differences of the characteristics and attributes of the records, the records in the data are divided into different categories, and different things are described by different class labels. Association rules and sequential pattern discovery, that is, association rules are the probability that one event may occur and other events may also occur, so there is an association between the two events. Analyze and extract the inherent laws of the data, and classify the data according to these laws. According to the analysis of things, extract the rules of things, and predict the nature of things according to the rules. Some special cases in the data are described and analyzed, and the internal reasons are pointed out. These functions of data mining are intrinsically linked, interrelated, and interact with each other. In the process of data mining, they cooperate with each other and work together [20], [21]. At the same time, sequence pattern is an extension of association pattern, which not only represents the relationship between data, but also connects the relationship between data with time. It can be seen that in order to discover the sequence pattern in the data, it is necessary not only to judge the possibility of the event, but also to know the time of the event [22], [23]. Firstly, the data set containing only normal data is used as the training set, and the above algorithm for automatically determining the number of clusters is used for simulation. Then, the anomaly judgment is made for the data to be detected, and the nearest distance and the corresponding nearest class are found by calculating the distance between the data to be detected and each cluster center, and the size of the nearest distance and the threshold is compared, to determine whether the data is normal or abnormal. On the basis of the above, the steps of designing data mining function are completed.

III. EXPERIMENTAL TEST

A. Setting up Experimental Environment

In order to verify the effectiveness of the design algorithm, the experimental test is carried out, and the experimental environment is built according to the needs of the experimental test. This experiment chooses Java as the main programming language, eclipse as the integrated development environment, as the main development environment of Java, is also a development tool. At the same time, Spring + Struts + Hibernate framework is selected for hierarchical development of the system, which conforms to the typical J2EE (Java 2 Platform Enterprise Edition) specification. JSP (JavaServer Pages)+ Javascript + j Query is selected for the front-end portal. In the resource management layer, all kinds of distributed open source frameworks need to be integrated and encapsulated. These frameworks are deployed on a number of nodes. These frameworks are encapsulated to provide the upper level of distributed storage and distributed computing capabilities [24], [25]. In the resource management layer, it also provides the monitoring information of the cluster, so that users can observe the utilization of CPU (Central Processing Unit), memory and network of each node in the

cluster, as well as the running health of each node and task, which is convenient for the maintenance of the cluster. In the data display layer, the results of data analysis are displayed by the user in the form of charts and reports, or the result data is directly provided by the interface. In the data analysis layer and data presentation layer, the rest API (Application Programming Interface) is used to interact with each other, so that the modules can be connected in a loose way. In the data integration layer, the original relational database and common log file data can be extracted, and then the data can be stored in the big data management platform. In addition, the obtained data are preprocessed according to the configuration parameters of the task, so as to lay a good foundation for subsequent mining and analysis. In the above experimental environment, the experimental test is carried out and the experimental results are obtained [26]-[28].

B. Experimental Result

Select Apriori based data mining algorithm and Hadoop based data mining algorithm, and compare them with the designed data mining algorithm. Test the time-consuming of the three algorithms under different data set proportions. The less time-consuming, the better the performance of the algorithm. The experimental results are shown in Table 1.

Table 1 Experimental results of algorithm time consuming (s)

Dataset scale (%)	The algorithm	The algorithm	Design of
	proposed by	proposed by	data mining
	Li et al. [4]	Chi [5]	algorithm
5	163.54	171.28	75.31
15	235.10	229.34	145.19
25	306.79	313.28	203.16
35	428.22	456.33	334.58
45	693.16	641.39	426.79
55	721.06	716.48	557.02
65	774.88	802.03	620.39
75	993.12	1013.76	780.49
85	1208.51	1359.26	1130.23
95	1533.64	1637.49	1325.11

It can be concluded from Table 1 that under the condition of different data set proportion, the average value of the data mining algorithm designed in this paper and the other two data mining algorithms is 559.827 s, 705.802 s and 734.064 s. This method saves 145.975 seconds and 174.237 seconds more time than other conventional algorithms, respectively, which shows that the data mining algorithm designed in this paper has better performance in practical application. This method can save more computing time than other methods because we first analyze the basic structure of data mining algorithm, master the data characteristics of multi-objective decision, then adjust the convergence speed of data mining algorithm and improve the computing speed of the BP network model to reduce the error and improve the optimal computing speed.

IV. CONCLUSIONS

In order to further improve the mining performance of the data mining algorithm and reduce its time-consuming, this paper designed a data mining algorithm based on the BP neural network. Analyze the basic structure of the data mining algorithm, adjust the convergence rate using the distributed computing technology, construct the local minimal discrete model, obtain the model interest; Based on the reverse computation training of the BP neural network model, calculate the optimal output value of the network, and complete the improved design of the data mining function. The experimental results prove that the paper design algorithm reduces the computational time consumption, significantly improves the performance of data mining, but also enriches the academic literature, broadens the application field of data mining algorithm, and has practical application value. Due to the limitations of the research conditions, the paper lacks the algorithm efficiency under different configuration platforms. Future research can propose more adaptive data mining algorithms for different configuration platforms.

REFERENCES

- [1] J. Hu, J. Fang, Y. Du, Z. Liu and P. Ji, "Application of PLS algorithm in discriminant analysis in multidimensional data mining", Journal of Supercomputing, vol. 75, no. 9, pp. 6004-6020, 2019.
- [2] G. Casola, C. Siegmund, M. Mattern, and H. Sugiyama, "Data mining algorithm for pre-processing biopharmaceutical drug product manufacturing records", Computers & Chemical Engineering, vol. 124, pp. 253-269, 2019.
- [3] C. Yuan, X. Yu, D. Li and Y. Xi, "Overall traffic mode prediction by VOMM approach and AR mining algorithm with large-scale data", IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 4, pp. 1508-1516, 2019.
- [4] J. P. Li, X. W. Wang, L. B. Ma and M. Huang, "Research on interval time series classification algorithm", Cyberspace Security, vol. 10, no. 8, pp. 84-92, 101, 2019.
- [5] D. W. Chi, "Application of an improved decision tree ID3 algorithm". Modern Computer, no. 17, pp. 3, 2019.
- [6] L. Zhang and X. Zuo, "Credit of small and medium sized scientific and technological enterprises based on BP neural network evaluation research", MATEC Web of Conferences, 336, no. 5, pp. 09010, 2021.
- [7] T. Yatsuka, A. Ishigaki, Y. Kinoshita, T. Yamada and M. Inoue, "Control method of effect of robust optimization in multi-player multi-objective decision-making", American Journal of Operations Research, vol. 9, no. 4, pp. 175-191, 2019.
- [8] I. Stipanovic, Z. A. Bukhsh, C. Reale and K. Gavin, "A multiobjective decision-making model for risk-based maintenance scheduling of railway earthworks", Applied Sciences, vol. 11, no. 3, pp. 965, 2021.
- [9] M. Baldan, A. Nikanorov and B. Nacke, "A novel multi-surrogate multi- objective decision-making optimization algorithm in induction heating", Compel, vol. 39, no. 1, pp. 144-157, 2020.
- [10] A. V. Kitaeva, M. V. Kolupaev, N. V. Stepanova and O. I. Zhukovskiy, "Convergence rate of mean-square error of kernel estimators of non-homogeneous poisson process intensity function", Journal of Physics: Conference Series, vol. 1680, no. 1, pp. 012021, 2020.

- [11] E. Martín-Hernández, M. Martín, M. Mohammadi and I. Harjunkoski, "Modeling and analysis of organic waste management systems in centralized and decentralized supply chains using generalized disjunctive programming", Industrial & Engineering Chemistry Research, vol. 60, no. 4, pp. 1719-1745, 2021.
- [12] M. Jayachandran and G. Ravi, "Decentralized model predictive hierarchical control strategy for islanded ac microgrids", Electric Power Systems Research, vol. 170, pp. 92-100, 2019.
- [13] J. Q. James, Y. H. Hou, Y. S. Lam and O. K. Li, "Intelligent fault detection scheme for microgrids with wavelet-based deep neural networks", IEEE Transactions on Smart Grid, vol. 10, no. 2, pp. 1694-1703, 2019.
- [14] S. Lahmiri, C. Tadj and C. Gargour, "Biomedical diagnosis of infant cry signal based on analysis of cepstrum by deep feedforward artificial neural networks", IEEE Instrumentation and Measurement Magazine, vol. 24, no. 2, pp. 24-29, 2021.
- [15] X. J. Fei and G. Z. Tian, "Research on data mining algorithm based on neural network and particle swarm optimization", Journal of Intelligent and Fuzzy Systems, vol. 35, no. 3, pp. 2921-2926, 2018.
- [16] I. Parvez, "Generation of hydro energy by using data mining algorithm for cascaded hydropower plant", Energies, vol. 14, no. 2, pp. 298, 2021.
- [17] Y. P. Singh and P. Gautam, "Development of data mining algorithm for giving the loan in banks introduction", High Technology Letters, vol. 27, no. 1, pp. 284-295, 2021.
- [18] Z. Qian and H. X. Xia, "Data mining algorithm based on feature weighting", Journal of Computational Methods in Sciences and Engineering, vol. 19, no. S1, pp. S269-S276, 2019.
- [19] M. B. Ozkan and P. Karagoz, "Data mining-based upscaling approach for regional wind power forecasting: regional statistical hybrid wind power forecast technique (RegionalSHWIP)", IEEE Access, vol. 7, pp. 171790-171800, 2019.
- [20] S. Hasheminejad and M. Khorrami, "Clustering of bank customers based on lifetime value using data mining methods", Intelligent Decision Technologies, vol. 14, no. 4, pp. 1-9, 2020.
- [21] R. S. Faradonbeh and A. Taheri, "Long-term prediction of rockburst hazard in deep underground openings using three robust data mining techniques", Engineering with Computers, vol. 35, no. 2, pp. 659-675, 2019.
- [22] Q. X. Yang, G. N. Wang and T. Wang, "Simulation of multi-source log security data mining based on time series", Computer Simulation, vol. 36, no. 2, pp. 297-301, 2019.
- [23] Y. Negara, and A. F. Doni, "Comparison of data mining algorithm performance on student savings dataset", Journal of Physics: Conference Series, vol. 1569, no. 2, pp. 022081, 2020.
- [24] Y. Liu, "Incomplete big data imputation mining algorithm based on BP neural network", Journal of Intelligent and Fuzzy Systems, vol. 37, no. 4, pp. 4457-4466, 2019.
- [25] Y. Ibrahim, J. Y. Liu, X. X. Yang, H. W. Sha, P. Li and H. B. Wang, "Analyzing the impact of soft errors

in deep neural networks on GPUs from instruction level", WSEAS Transactions on Systems and Control, vol. 15, pp. 699-708, 2020.

- [26] L. Kanaan, J. Haydar, M. Samaha, A. Mokdad and W. Fahs, "Intelligent bus application for smart city based on LoRa technology and RBF neural network", WSEAS Transactions on Systems and Control, vol. 15, pp. 725-732, 2020.
- [27] S. Gochhait, Y. Rimal and S. Pageni, "The comparison of forward and backward neural network model – a study on the prediction of student grade", WSEAS Transactions on Systems and Control, vol. 16, pp. 422-429, 2021.
- [28] Q.-J. Smith and R. Valverde, "A perceptron based neural network data analytics architecture for the detection of fraud in credit card transactions in financial legacy systems", WSEAS Transactions on Systems and Control, vol. 16, pp. 358-374, 2021.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US