Retraction

Retracted: Prediction of Enterprise Economy by Network Communication Business Based on BP Neural Network in Big Data Security Environment

Security and Communication Networks

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] S. Kong, "Prediction of Enterprise Economy by Network Communication Business Based on BP Neural Network in Big Data Security Environment," *Security and Communication Networks*, vol. 2022, Article ID 1132270, 9 pages, 2022.

WILEY | Hindawi

Research Article

Prediction of Enterprise Economy by Network Communication Business Based on BP Neural Network in Big Data Security Environment

Sibo Kong [1,2]

[1]*Chongqing College of Mobile Communication, Chongqing 401520, China*
[2]*Chongqing Key Laboratory of Public Big Data Security Technology, Chongqing 401420, China*

Correspondence should be addressed to Sibo Kong; woc180010@siswa.um.edu.my

From the environmental and security status quo faced by big data, although big data is unstructured or more difficult to filter and analyze, it does not mean that big data is necessarily more secure. In order to be invincible in the fierce market competition, an enterprise must conduct an in-depth understanding and investigation of the rapid changes in the market and economic development. The current era is already the Internet era. The arrival of the era of network communications has greatly facilitated people's lives. However, how to seize the opportunities in the era of network communication and make economic predictions based on actual situations are particularly important for enterprises. Therefore, we are required to make reasonable economic predictions of network communication services in order to seize opportunities and meet challenges. By analyzing the development and application of artificial neural network (NN), this article briefly introduces its development and principles. Based on the reality of a network communication industrial enterprise, it uses data modeling and comparative analysis to establish a logistic regression model, a decision tree model, and a BP NN. The model compares the customer types and data predictions of the enterprise's data traffic under the three models. From the results of the model analysis, it can be seen that in the analysis of the three models, the ROC curve analysis, the BP NN predicts that the cumulative hit rate in the ROC curve is wider. More users who can handle traffic services will be covered. In the cumulative customer lift analysis, under the first set of data (that is, when the depth is 20), the cumulative depths of the decision tree model, logic analysis model, and BP neural system prediction model are 2.5, 3.4, and 3.6, respectively. The BP neural system prediction model has the highest value of cumulative depth. In the cumulative capture response percentage analysis, the cumulative capture response percentage from low to low is the decision tree model and logic analysis. From BP NN prediction model, we can draw the conclusions. The network system can play a good role in forecasting and help enterprise managers make economic decisions.

1. Introduction

With the maturity, application and promotion of big data technology, and enterprises or government departments with important data assets increasingly recognize the concept of big data development, and data has become another core value asset after cash and technology. Data security is the core of the entire big data era, including personal privacy, business secrets, and even important national data. Once these sensitive data are tampered with or leaked, it will affect business operations at a light level, and even directly affect social security and national security steady development.

One of the key purposes of accounting management is the economy, which is an important indication for measuring an enterprise's operating conditions. Its goal is to give the leadership a solid and stable foundation on which to make decisions, to make better use of limited investment dollars, and to maximize the economic construction effect. Understanding the future cost level and its shifting pattern will help eliminate decision-making blindness through economic forecasting and make it simple for managers to select the best plan and make the best judgments possible. The important difficulties for businesses to achieve large profits include increasing economic forecasting and

management, enhancing the quality of cost management, and lowering operating costs. Forecasting the economy also aids in improving management predictability and cost control, hence supporting the enhancement of enterprise economic benefits.

Economic forecasting is a very important and complex task. Economic forecasting needs to infer the development trend of things based on historical data and current conditions, and based on certain theories and methods. Among the existing economic forecasting methods, time series forecasting and regression forecasting are the two most commonly used statistical methods. Because the economic system is affected by multiple influencing factors, it is a highly uncertain nonlinear system. The collection of historical data required for economic modeling is difficult, and the information is incomplete. To use traditional prediction methods to solve such problems.

Saraswathy found that the development of Internet technology has promoted the rapid development of the IT industry [1]. Kapil found that with the exponential growth of big data, it became more and more vulnerable to malicious attacks. These attacks can compromise the privacy, integrity, and availability of information systems. To counter these malicious intents, it is necessary to develop effective security mechanisms. He presents some important aspects of big data Hadoop security and privacy to increase the security of enterprise data [2]. Shunquan used a NN model to forecast how many tourists would visit a province. The number of tourists is an essential determinant of tourism's economic advantages and long-term development. As a result, forecasting the number of tourists has become a crucial part of tourism development planning. A forecast model of the number of tourists based on BP NN was built based on the number of tourists in a province for more than two decades, and the principles and methods of BP NN were applied [3]. Also, estimate the number of tourists expected to visit the province in the future. The BP NN model's Matlab simulation results show that the tourist number prediction model based on BP NN can accurately estimate tourist numbers.

This time based on the study of the application of the BP NN system in the economic prediction of network communication, taking a domestic network communication industrial enterprise as an example, this article briefly introduces the artificial NN, mainly introduces its concepts, principles, features, etc. Then, according to the actual needs of the enterprise, the economic forecast of the enterprise's traffic data is carried out, and the logistic regression model, the decision tree model, and the BP neural network model are used to compare the data traffic customer types and the data forecast results. According to the analysis results of the model, the BP neural system has a good application effect on the economic forecast of the enterprise.

2. Proposed Method

2.1. Overview of Artificial NNs. Artificial NN is a new information processing discipline, referred to as NN for short. It originates from the human brain nervous system and is an important way to simulate human intelligence. Although artificial NNs are not as complex as the human brain, there are two key similarities between them. First of all, the composition of the two networks is highly interconnected by the computable units; second, the connection between the processing units determines the function of the network [4]. The artificial NN is composed of simple information processing units (artificial neurons, referred to as neurons) interconnected to form a network, that receives and processes information, and its information processing is realized by the interaction between the processing units. Current NNs are mainly used in pattern recognition, image processing, nonlinear optimization, intelligent robots, language processing, predictive analysis, adaptive control, knowledge processing, cognitive science, etc. [5, 6].

The artificial NN processing flow is shown in Figure 1.

The main purpose of image preprocessing is to remove unnecessary information and retain useful information. Each artificial neuron has synapses, as shown in Figure 1. The artificial neuron gets the output of all neurons connected to it, and the signal to be generated is amplified connection strength [7]. The weighted total is compared to the net value of the neuron, and the fake neuron is triggered if it is bigger than the threshold [8, 9]. The signal is transferred to the higher-level neurons attached to it when it is triggered. The artificial NN's operation is primarily controlled by two factors: first, the network's structure, or how the artificial neurons are connected; and second, the artificial neurons' function. The second set of rules is the network learning and operation rules, which are the rules for adjusting the network's connection weights. The artificial NN research approach can handle large-scale parallel data processing, and has high fault tolerance, perception, memory, thinking, and reasoning abilities, as well as strong self-learning and adaptive abilities. It excels at learning from a wide variety of sources. Analyze statistical data and obtain macroscopic statistical laws. Therefore, the use of artificial NNs for economic forecasting can play an important role in correctly assessing the economic development level of a region, accurately predicting future economic development trends, and timely reflecting the effects of macroeconomic regulation and control [10].

The characteristics of the artificial NN mainly have the following four points:

(1) High-speed information processing and powerful information storage capacity. Artificial NNs have a large number of neurons related to each other, and neurons store and process information in parallel.

(2) It has a strong ability to deal with fuzzy data. Results are valid even if input data is missing, wrong, or unclear.

(3) Due to the robustness advantage of artificial NNs, when a single neuron is lost, biological NNs do not lose the memory of the original pattern. The strongest proof is that when the human brain is slightly damaged by accident, it does not lose all memory of the original thing. The same is true for
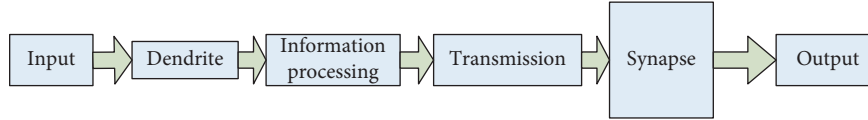
FIGURE 1: Flow chart of artificial NN.

NNs. For some reason, no matter which part of the neuron fails, it does not affect the operation of the overall network. The basic algorithm of NN has low convergence precision, slow convergence, no convergence, and difficulty in determining the network structure.

(4) Artificial NNs with highly nonlinear systems are different from current computers, breaking through the limitations of traditional linear processing computers [11, 12].

2.2. BP NN and Algorithm. The BP algorithm is still the most important and most widely used effective algorithm in automatic control. The BP NN is the error backpropagation network, which belongs to the forward multi-layer propagation and guided learning NN.

The BP artificial network model is part of the forward multi-layer artificial NN back propagation learning technique, which has three layers: input, output, and a learning layer. It's made up of multiple layers that are not visible. There are numerous neurons in each stratum. Connection weights and thresholds connect the neurons in each layer. In the same layer, there is no connectivity between neurons. A sort of supervised network is the BP network model. When data is entered into the network, it is first transferred from the input layer to the hidden layer node, then sent to the next hidden layer following the characteristic function, and ultimately passed to the output layer [13]. The neural unit's function is usually an s-type function:

$$f(x) = 1/1 + e^{-x}. \tag{1}$$

After identifying the structure of the BP network, train it by modifying the connection weight in the BP network and the network size (including n, m, and the number of hidden layer nodes) using the input and output sample sets. Realize the given input-output mapping connection and accurately approximate any nonlinear function. The BP network calculating method is based on the error backpropagation algorithm, which is made up of two processes: forward information propagation and back error propagation [14]. The movement of events in Figure 2:

It can be seen from Figure 2 that the neurons in the input layer are used as the starting point for information transmission. Hidden layer refers to the layers other than the input layer and output layer in the multi-level feedforward NN. The hidden layer does not directly receive external signals, nor does it directly send signals to the outside world. It is only required when the data is nonlinearly separated. The input layer receives the information, passes it to the middle layer, the middle layer converts it, and designs a multi-layer structure as needed to build a hidden layer, The

information is subsequently passed to the output layer by the concealed layer. It will enter the erroneous back propagation phase if the actual output does not match the predicted output. The error flows through the output layer, correcting the weight of each layer via error gradient descent, and then returning to the hidden layer and input layer by layer. Constantly changing the weights and learning and training of the NN [15] is the process of information propagation. This process should be repeated until the error has been reduced to a bare minimum. To apply the network system to the practice of solving practical problems, we must first train the network system. The BP network system has its own unique training method, that is, the algorithm of the BP system, which is also the standard BP algorithm. The main idea is: for a set of input samples, the actual output is calculated by the BP NN, and the actual output and output of the BP network have used. The error between samples is corrected until the error of both reaches the set value of the network connection weight [16, 17]. This smaller value is called the fitting error, also known as the error function. Generally, the sum of squared errors between the actual output and the output samples is expressed as:

$$E = \frac{1}{n_3} \sum_{k=1}^{n_3} (t_k - z_k)^2. \tag{2}$$

In the formula, t_k is the sample output value and z_k is the actual output value. The so-called activation function is the function that runs on the neurons of the artificial NN and is responsible for mapping the input of the neuron to the output. The steps to train the BP network with the BP algorithm are as follows:

(1) Initialization of connection weight

At the beginning of the network training, the connection weight is unknown, and generally, a smaller random number is used as the initial value of the connection weight of each layer.

(2) Calculate the output value of neurons in each layer

$$y_j = f_1 \left(\sum_{i=0}^{n_1} w_{1ij} x_i \right), (j = 1, 2, ..., n_2),$$

$$z_k = f_2 \left(\sum_{j=0}^{n_2} w_{2kj} y_j \right), (k = 1, 2, ..., n_3). \tag{3}$$

In the formula: f_1, f_2 for activation function, generally use the sigmoid function or linear function.

(3) Correction of connection weight

Gradient descent is a first-order optimization algorithm. The correction of the connection weight adopts the
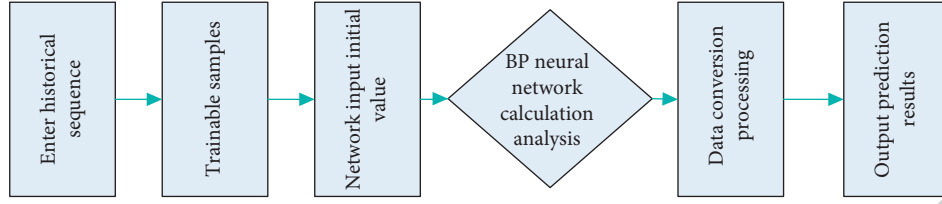
Figure 2: BP NN system process.

gradient descent method, each time the correction amount of the connection weight and the gradient of the error function.

It is proportional to the reverse transmission from the input layer to each layer. The connection weight correction amount of each layer is;

$$\Delta w_{2kj} = \eta(t_k - z_k)f_{2,}y_j, (k = 1, 2, ..., n_3; j = 0, 1, 2, n_2),$$

$$\Delta w_{1ji} = \eta(t_k - z_k)f_{2,}w_{2kj}f_{1,}x_i, (j = 1, 2, n_2; i = 0, 1, 2, ..., n_1).$$

$$(4)$$

In the formula, η is the learning rate; f_1, f_2 is the derivative of $f_{1,}, f_{2,}$ activation function The initial weight is added to the corresponding adjustment amount to calculate a new weight, and so on until the sum of squares of the output layer error reaches the set value [18].

The main purpose of parallel processing is to save time for solving large and complex problems. In order to use parallel processing, the program needs to be parallelized first. Parallel processing of information, data fusion, and self-adaptation. The standard BP algorithm is widely used, but it also has its shortcomings. The standard BP algorithm has the tendency to form a local minimum, the convergence speed is very slow, and the training period is long; it is easy to produce overtraining, etc [19, 20]. In order to make up for these shortcomings, people have made many beneficial improvements on the basis of the standard BP algorithm, such as the momentum method, conjugate gradient learning algorithm, LiebenbergMarquardt optimization method, and Bayesian regularized BP NN algorithm. Bayes' theorem is a theorem about the conditional probability (or marginal probability) of random events A and B.

The network communication business income is mainly concentrated in the Internet business income and mobile traffic business income With the continuous improvement of 4G networks and the introduction of 5G networks, along with the expansion of 4G and 5G mobile users, mobile Internet output traffic is close to 10 billion, and online consumer traffic through mobile terminals accounts for about 90%. According to the characteristics of daily traffic usage of users in the communications industry in the network communications business, this study takes the customer's demand for traffic as the index of this economic forecast study. Since the current 4G network has the widest distribution and the most perfect development, the flow business referred to in this article refers to the flow business based on the 4G network, and continues to analyze the customer flow business of a network communication company.

Many factors affect customers' handling of traffic services. Due to the differences in personal consumption needs and differences in consumption habits of users, plus the impact of whether home broadband, Wi-Fi and terminals are smart machines, although users must use some kind of main package business. The standard traffic package business is also decided based on the user's traffic demand and consumption capacity. The corresponding recommended traffic combinations are package A, package B, and package C. The month of the extracted data is a total of 9 months from January 20×9 to September 20×9. The data from January to August was analyzed as a sample, and the September data was predicted and compared with the actual results. In order to make the model prediction effect more intuitive, respectively, select the optimal model output results, and compare and evaluate the model effects.

3. Experiment Design and Simulation Analysis

3.1. Confirmation of Requirements. Solve the problem of forecasting the willingness to handle the traffic business portfolio, and determine the probability of whether the customer will handle the traffic business portfolio in different customer segments. This part is the preliminary preparation work and has been completed in the process of field investigation.

3.2. Data Preprocessing. Generate an input data source for the flow business portfolio willingness prediction, and select the non-downtime active communication customer of the enterprise as the customer group in the modeled sample data. The number of observations obtained for the total sample in September 2015 is 1637895, and the data types of all indicators have been converted to numeric types.

3.3. Build a Communication Traffic Business Combination Classification Prediction Model

3.3.1. Build a Logistic Regression Model. In the logical analysis model, whether the customer will handle the flow business combination, for the binary classification prediction problem, the value of the willingness to predict the value of P (x) is (0, 1). To solve P (x), you can use the sigmoid function, whose definition domain is $(-\infty, +\infty)$, and the value range is (0, 1). It is difficult to directly solve the sigmoid function, which is converted into a logistic function. Can be expressed as:

$$P(y = 1|x) = \pi(x) = 1/1 + e^{-g(x)}. \tag{5}$$

In the formula $\pi(x)$ describes the possibility of handling business combinations on the basis of each attribute value of customers. Use the EM module in SAS statistical analysis software to build the model. The parameter that the variable in SAS will be selected to enter is SLE, the parameter that the variable in SAS will remain in the selected variable combination is SLS, the entry of the new variable leads to the insufficient contribution of the old variable to the entire model, the SLS threshold will be This variable is eliminated. The default statistical significance level of parameters SLE and SLS is $P < 0.05$.

3.3.2. Build a Decision Tree Model. A decision tree is generally composed of block nodes, circular nodes, program branches, probability branches, etc. The block nodes are called decision nodes, and several thin branches are drawn from the nodes. The decision tree is a kind of classification process based on representative rules of representative examples. Although different customers have different combinations of attribute values, they ultimately correspond to the result of whether they will handle the flow business combination. The combination of different attribute values of customers can also be described as different judgment rules. The final judgment result is whether or not the business will be handled. The number of different attribute values of customers is not the same. When judging each attribute value, the final number of judgment rules is large. The decision tree algorithm determines the feature selection metric technology adopted at each internal node as the information gain rate. Information gain represents the difference between the experience entropy $H(D)$ of the training data set D and the experience condition entropy $H(D \mid A)$ of D under the given conditions of the variable value A, that is: first calculate the experience entropy $H(D)$:

$$H(D) = -\sum_{k=1}^{k} \frac{|C_k|}{|D|} \log_2 \frac{|C_k|}{|D|}. \tag{6}$$

Among them, $|D|$ is the number of observations in the training sample data set, $|C_k|$ is the number of the Kth classification result, where K is 2, Second, calculate the empirical conditional entropy $H(D|A)$:

$$H(D|A) = \sum_{i=1}^{n} \frac{|D_i|}{|D|} H(D_i). \tag{7}$$

Among them, because $|D_i|$ refers to the number of samples of attribute A in i values, $|D_i|/|D|$ refers to the probability of attribute A taking i values, and finally, the information gain is calculated, the expression is given as:

$$g(D, A) = H(D) - H(D|A). \tag{8}$$

The information gain method is used for feature extraction in text classification, and words with relatively large information gain for a certain category are selected as the features of this category. The principle of feature selection using the information gain method: Calculate the variable with the largest information gain of all variables in the training data set. A large information gain indicates that the variable has a stronger classification capability. It can be seen that the information gain represents the degree to which the uncertainty of the information of class Y is reduced by learning the information of the variable A. The smaller the empirical conditions, the higher the purity of the partition. Use the EM module in SAS statistical analysis software to build a decision tree model. The input training data set does not use evidence weight values to replace each group of each variable, but directly uses the original discretized data results. Before the model is fitted, the EM module can optimally divide according to the type of attribute value. In order to compare the model output with the results of the logistic regression model, the decision tree model is built based on the data set of variables with information values less than 0.5.

3.3.3. Build a BP NN Model. Due to a large number of input attributes in this study, first adopt the principal component analysis data. In order to reduce the number of model input variables, the principal component analysis method is used to reduce the dimensionality to speed up the model fitting time. The role of principal component analysis is to generate a number of dominant variable combinations from the existing variable set, and can determine the size of these variable combinations. The obtained result is used as input, that is, the normalized sample data value is converted into the main component sample data, and then the data set is used as the input training data set. The principal component analysis is a statistical method. According to kolmogorov's theorem, when the input layer node of the BP NN is n_1 and the number of output layers is m, the number of hidden layer nodes k is selected as

$$\sqrt{(n_1 + m + 1)} + a, \tag{9}$$

where a can be selected from 1 to 10, and the value of m in this article is 2. In this study, the BP NN with a single hidden layer has a strong nonlinear conversion ability, which can improve the efficiency of model fitting. The EM module in SAS statistical analysis software is used to build an algorithm model based on BP NN. Since the number of input layer variables n_1 is 9, the number of output layer variables m is 2, so by

$$sqrt(n_1 + m + 1) + a, \tag{10}$$

where a is set to 1, the number of hidden layer units is 5, the learning rate is set to 0.1, and the maximum number of iterations is 50.

3.4. Performance Evaluation and Performance Analysis Indicators. There are many indicators to evaluate the effectiveness of a model. In this study, we choose ROC linear analysis, lift, and response percentage. We use test data to evaluate the performance stability of the model, and select two indicators, K-S value and lift, for analysis. The ROC curve is composed of the predicted hit rate (TPR) and the false prediction rate (FPR). TPR represents how many of the

people who actually handle traffic business are correctly predicted to handle business. FPR represents how many of the people who have not actually handled traffic business are wrongly predicted to handle business. The area under the ROC curve is named using the AUC value and is used to evaluate the actual effect of the model. The range of AUC values is [0.5,1]. The larger the value, the more the ROC curve is toward the upper left corner, indicating that under the same threshold, the cumulative hit rate is greater than the cumulative error prediction rate, indicating that more will be handled Users of traffic services will be covered. This article uses the K-S value to test the user's ability to distinguish between traffic service portfolios and user non-service portfolios. If the K-S value is larger, it indicates that the ability to distinguish users from handling traffic services and users from not handling services is stronger, and the predictive model is better. The degree of promotion refers to that users are ranked in descending order according to the model probability and divided into 10 groups on average. The ratio of the event rate of each group to the overall event rate is the degree of promotion. The higher the lift of the first group, the better the model effect.

3.5. Linear Analysis Effect of Three Models ROC.
The results of the ROC linear graphs of the three models are shown in Figure 3.

Figure 3 shows that the BP NN prediction model is positioned above the linear model analysis and decision tree models under the same cardinality, indicating that the BP NN model has a higher cumulative hit rate in the ROC curve under the same conditions. Because there will be more users handling the traffic business, the BP NN model's prediction effect will be better.

3.6. Analysis of the Effect of the Cumulative Lift of the Three Models.
The effect of the cumulative lift of the three models is shown in Figure 4.

It can be seen from Figure 4 that under the first set of data (that is, when the depth is 20), the cumulative depth of the decision tree model, logical analysis model, and BP neural system prediction model are 2.5, 3.4, and 3.6, respectively. The cumulative depth of the model is the highest, which is better than the decision tree model. The logical analysis model is 44% and 8% higher, respectively, indicating that under the same prediction conditions, the BP model has the best prediction effect, followed by the logical analysis model, and the decision tree model is the worst.

3.7. Analysis of the Cumulative Effect of the Percentage Captured by the Three Models.
The effect of the cumulative lift of the three models is shown in Figure 5.

It can be seen from Figure 5 that under the same cardinality conditions, the cumulative capture response percentages range from low to decision tree model, logic analysis model, and BP NN prediction model. When the cardinality is large enough, the logic analysis model cumulative capture response percentage trend When the cumulative capture
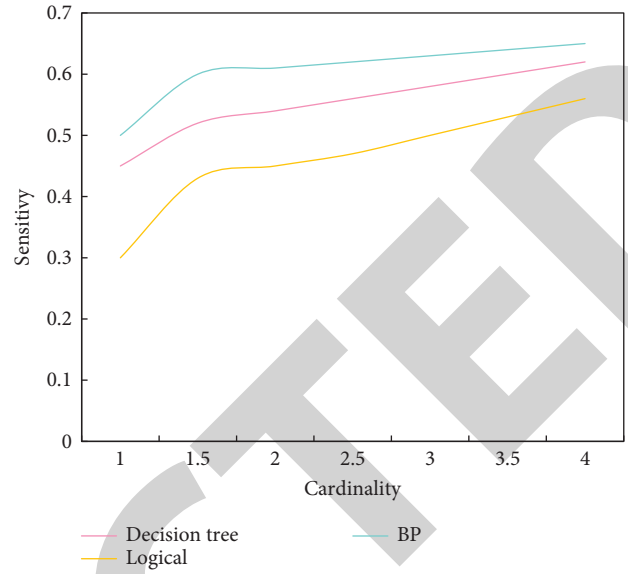
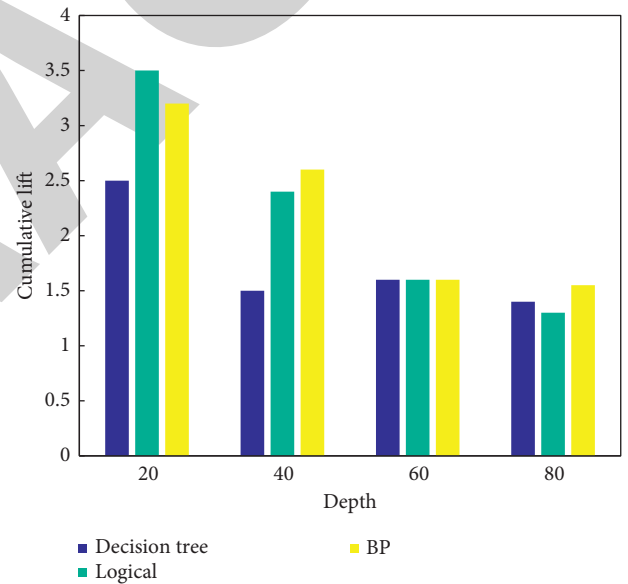

Figure 3: ROC linear analysis of three models.

Figure 4: Analysis of the cumulative lifting graph of the three models.

response percentage of the near BP NN model is still lower than the cumulative capture response percentage under the BP prediction model, it can be concluded that the cumulative capture response percentage value under the BP NN model is the highest and the model prediction effect is the best.

The results of the three models of different value customer groups are shown in Figure 6.

It can be seen from Figure 6 that for the modeling samples of three subdivided customer groups, Choose the optimal model output results respectively for the model effect evaluation ratio. From the perspective of cumulative lift and cumulative response percentage, BP is used in high-value customer groups, medium-value customer groups, or low-value customer groups. The effect of the NN model is better.
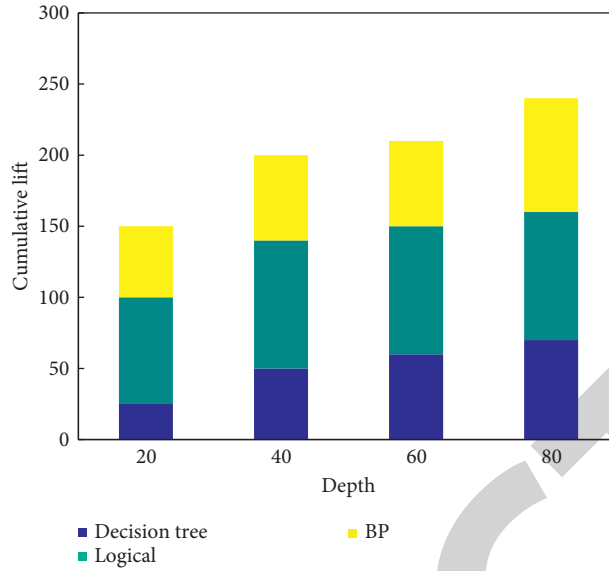
Figure 5: Analysis of the cumulative capture response percentage of the three models.
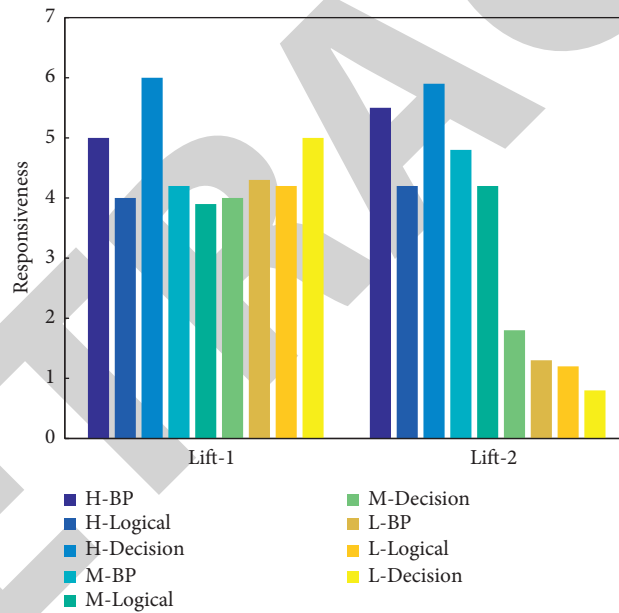

Figure 6: Results of three models of different value customer groups.

3.8. Analysis of the Performance Stability of the Three Models. The test data set was used to analyze the effects of the three model results. With the help of the K-S value, we can intuitively find a segmented interval with the largest difference in the prediction model and output the lift and cumulative lift of the test data set as an index to verify the performance stability of the model. The effect of the cumulative lift of the three models is shown in Table 1.

It can be seen from Table 1 that in low-value customer groups, the BP NN model is expected to have higher values, followed by logistic regression analysis, and the decision tree model is the worst. BP NN is expected to have the highest value. In the prediction of high-value groups, the predicted value of BP NN reaches 0.506988.

The lifting degree and cumulative lifting degree of the verification data output in the three models are shown in Table 2 (Table 2 selects the representative data with a depth of 10%)

It can be seen from Table 2 that the actual lift and cumulative lift corresponding to the top 10% of the users are the best in the BP model. Among the low-value customer groups, the BP NN model has a higher value from the actual lift and actual cumulative lift corresponding to the top 10% of users, followed by the logical review of the NN, and the decision tree model is the worst. And the cumulative improvement of the verification results is lower than the training results, and similar conclusions are obtained in the analysis of the middle and high customer groups. It can be

TABLE 1: Verify the maximum output K-S value in the three models.

Serial number	Segment name	Algorithm type	Maximum K-S value
1	High-value customer groups	Model 1	0.483745
		Model 2	0.322588
		BP	0.506988
2	Medium-value customer groups	Model 1	0.352912
		Decision tree	0.319108
		BP	0.378465
3	Low-value customer groups	Model 1	0.440798
		Model 2	0.381677
		BP	0.448482

TABLE 2: Verify the lift and cumulative lift of the output of the data in the three models.

Serial number	Segment name	Depth (%)	Algorithm type	Actual lift of each layer	Actual cumulative lift
1	High-value customer groups	10	Model 1	4.7565	4.7565
			Model 2	4.2728	4.2728
			BP	4.9730	4.9730
2	Medium-value customer groups	10	Model 1	4.3004	4.3004
			Model 2	3.8201	3.8201
			BP	1.3097	3.1407
3	Low-value customer groups	10	Logistic regression model	4.0064	4.0064
			Model 2	3.4012	3.4012
			BP	4.1220	4.1220

seen that from the two aspects of interpretability and model performance stability, the model results fitted with the BP NN model can be used to classify and predict mobile communication traffic services and marketing applications.

4. Conclusions

In recent years, with the rapid development of information technology, especially big data, it has become easier to collect, store, publish and analyze massive data. From the perspective of data security and personal privacy protection, big data applications also bring great hidden dangers to data security. Among the many security problems faced by big data, how to analyze and mine more value from big data and well protect the privacy and security of data is particularly important.

In the era of big data, data security has become the lifeblood of countries, governments, and enterprises. After entering the information age, data not only has the characteristics of diversification, but also the amount of information gradually increases which has become two issues that China's communication industry needs to pay attention to correct decision-making is the cornerstone to successful business management, and accurate prediction is the foundation of both decision-making and scientificity. We can only achieve "marketing based on demand and production based on sales" if we estimate the market scientifically and precisely. We can only exploit potential and cut costs by anticipating profits and costs in advance, allowing businesses to make appropriate arrangements in their company activities and support growth. In the strong market battle, businesses have a position. It must be stated that network communication has played a pivotal part in the development of today's period, as well as an essential role in the overall economic condition. Economic forecasting has an unparalleled impact on the development of businesses due to its foresight and practicality. Therefore, this article is based on the economic analysis and economic prediction of the network communication business industry, and its role cannot be underestimated. Based on the application of the BP network system in the network communication economic prediction, the comparative analysis of the application of logic analysis systems and decision tree models in this field, Demonstrated the role of the BP NN system in network communication economic forecast.

This article is based on the application research of the BP NN system in the economic forecast of network communication. It focuses on the reality of a network communication enterprise. According to its business type characteristics and enterprise reality, it first briefly introduces the concept and characteristics of artificial NN. Then, based on the characteristics of the enterprise, using its network communication flow index as the economic forecast index of this enterprise, respectively, established three sets of data prediction models, namely logistic regression model, decision tree model, and BP NN model, and compared the data traffic customers of the enterprise under the three models Type and data prediction.

According to the model test results in this study, the analysis results show that in the ROC linear graph analysis, under the same conditions, the cumulative hit rate of the ROC curve of the BP NN model is wider, and more users will handle traffic services will be covered. . In the analysis of the cumulative lift effect diagram, the analysis results show that at the same depth, the cumulative lift from high to low is BP NN analysis, decision tree model analysis, and logistic

regression model analysis. Under the same cardinality conditions, the cumulative capture response percentages range from low to low, respectively: decision tree model, logic analysis model, and BP NN prediction model. The cumulative capture response percentage value under the BP NN model is the highest and the model prediction effect is the best. In the model stability analysis, the BP NN model is expected to have a higher value in the three customer groups of the enterprise's high, middle and low, followed by logistic regression analysis, and the decision tree model is the worst. In the enterprise, the actual lift and cumulative lift corresponding to the top 10% of users are the best in the BP model. Among the low-value customer groups, the BP NN model has a higher value from the actual lift and actual cumulative lift corresponding to the top 10% of users, followed by the logical review of the NN, and the decision tree model is the worst. And the cumulative improvement of the verification results is lower than the training results, and similar conclusions are obtained in the analysis of the middle and high customer groups. From this, we can conclude that the model results fitted with the BP NN model have a good prediction effect on the economic prediction of network communication business, and it is feasible to apply the BP NN system to the economic prediction of network communication business.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare no potential conflicts of interest in this study.

References

[1] K. S. Saraswathy and S. S. Sujatha, "Internet of things big data security in cloud via stream cipher and clustering model," *Wireless Personal Communications*, vol. 123, no. 4, pp. 3483–3496, 2022.

[2] G. Kapil, A. Agrawal, and R. A. Khan, "Big data security challenges: Hadoop perspective[J]," *International Journal of Pure and Applied Mathematics*, vol. 120, no. 6, pp. 11767–11784, 2020.

[3] S. Zhu, "Financial classification of listed companies in China based on BP NN method[J]," *Journal of Financial Risk Management*, vol. 05, no. 3, pp. 171–177, 2016.

[4] A. Pandey and A. Mishra, "Application of artificial neural networks in yield prediction of potato crop," *Russian Agricultural Sciences*, vol. 43, no. 3, pp. 266–272, 2017.

[5] N. Maknickiene, I. Lapinskaite, and A. Maknickas, "Application of ensemble of recurrent neural networks for forecasting of stock market sentiments," *Equilibrium*, vol. 13, no. 1, pp. 7–27, 2018.

[6] M. Memarianfard and A. M. Hatami, "Artificial NN forecast application for fine particulate matter concentration using meteorological data[J]," *Global Journal of Environmental Science & Management*, vol. 3, no. 3, pp. 333–340, 2017.

[7] J. Abbot and J. Marohasy, "Forecasting extreme monthly rainfall events in regions of Queensland, Australia using artificial NNs[J]," *International Journal of Sustainable Development and Planning*, vol. 12, no. 7, pp. 1117–1131, 2017.

[8] E. Magaña-Villegas, J. M. Carrera-Velueta, and S. Ramos-Herrera, "Clustering approach applied on an artificial NN model to predict PM10 in mega cities of Mexico[J]," *International Journal of Sustainable Development and Planning*, vol. 11, no. 4, pp. 566–577, 2016.

[9] A. N. Deshmukh and M. C. Deo, P. K. Bhaskaran, P. K. Bhaskaran, T. M. B. Nair, and K. G. Sandhya, Neural-network-based data assimilation to improve numerical ocean wave forecast," *Ieee Journal of Oceanic Engineering*, vol. 41, no. 4, pp. 944–953, 2016.

[10] L. I. Weihong, W. Lei, and C. Yebin, "Property crime forecast based on improved GA-BP NN model[J]," *Geomatics & Information Science of Wuhan university*, vol. 42, no. 8, pp. 1110–1116, 2017.

[11] S. Nayak, B. B Misra, and H. Behera, "An adaptive second order neural network with genetic-algorithm-based training (ASONN-GA) to forecast the closing prices of the stock market," *International Journal of Applied Metaheuristic Computing*, vol. 7, no. 2, pp. 39–57, 2016.

[12] W. Wu, S. An, and J. Guo, "[Application of nonlinear autoregressive NN in predicting incidence tendency of hemorrhagic fever with renal syndrome][J]," *Zhonghua Liuxingbingxue Zazhi*, vol. 36, no. 12, pp. 1394–1396, 2016.

[13] L. F. Gómez-Ossa and V. Botero-Fernández, "Application of artificial NNs in modeling deforestation associated with new road infrastructure projects[J]," *Dyna*, vol. 84, no. 201, p. 68, 2017.

[14] A. Mehrizad and P. Gharbani, "Application of central composite design and artificial neural network in modeling of reactive blue 21 dye removal by photo-ozonation process," *Water Science and Technology*, vol. 74, no. 1, pp. 184–193, 2016.

[15] Y. Zhang, "Application of fuzzy NN in music recognition[J]," *Revista de la Facultad de Ingenieria*, vol. 32, no. 6, pp. 171–180, 2017.

[16] H. Yu, G. Rossi, A. Braglia, and G. Perrone, "Application of Gaussian beam ray-equivalent model and back-propagation artificial neural network in laser diode fast axis collimator assembly," *Applied Optics*, vol. 55, no. 23, pp. 6530–6537, 2016.

[17] L. Huang and C. S Du, "Application of recurrent NNs in text classification[J]," *Journal of Beijing university of Chemical Technology(natural Science Edition)*, vol. 44, no. 1, pp. 98–104, 2017.

[18] C. Bergmeir and M. José, "Benítez. RSNNS: NNs in R using the stuttgart NN simulator (SNNS)[J]," *Carpathian Journal of Electronic & Computer Engineering*, vol. 46, no. 2, pp. 3–6, 2017.

[19] A. Asiamah, N Nimako, Q. Quaye, and S. Buame, "Implicit and explicit loyalty: the role of satisfaction, trust and brand image in mobile telecommunication industry," *International Journal of Business and Emerging Markets*, vol. 8, no. 1, p. 94, 2016.

[20] H. Oh and S. Lim, "Light-weight routing protocol in IoT-based inter-device telecommunication wireless environment," *International Journal of Electrical and Computer Engineering*, vol. 6, no. 5, pp. 2352–2361, 2016.