

Research Article

Design of Macroeconomic Growth Prediction Algorithm Based on Data Mining

Hongxiang Sun,¹ Zhongkai Yao,¹ and Qingchun Miao ²

¹*Xi'an University of Science and Technology, Xian 710000, China*

²*Shihezi University, Shihezi 832003, China*

Correspondence should be addressed to Qingchun Miao; 18217005026@stu.xust.edu.cn

Received 27 July 2021; Accepted 26 August 2021; Published 2 September 2021

Academic Editor: Fazlullah Khan

Copyright © 2021 Hongxiang Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of information technology and globalization of economy, financial data are being generated and collected at an unprecedented rate. Consequently, there has been a dire need of automated methods for effective and proficient utilization of a substantial amount of financial data to help in investment planning and decision-making. Data mining methods have been employed to discover hidden patterns and estimate future tendencies in financial markets. In this article, an improved macroeconomic growth prediction algorithm based on data mining and fuzzy correlation analysis is presented. This study analyzes the sequence of economic characteristics, reorganizes the spatial structure of economic characteristics, and integrates the statistical information of economic data. Using the optimized Apriori algorithm, the association rules between macroeconomic data are generated. Distinct features are extracted according to association rules using the joint distribution characteristic quantity of macroeconomic time series. Moreover, the Doppler parameter of macroeconomic time series growth prediction is calculated, and the residual analysis method of the regression model is used to predict the growth of macroeconomic data. Experimental results show that the proposed algorithm has better adaptability, less computation time, and higher prediction accuracy of economic data mining.

1. Introduction

With the continuous improvement of China's market economic system, the prediction of social and economic development has become one of the hot issues. To develop effective implementation policies and plans, relevant research institutions must make accurate forecasts of future macroeconomic trends, based on which to decide whether the size of the economy needs to be stimulated or suppressed [1]. But future macroeconomic trends, influenced by many factors within and around them, are nonlinear and complex systems that change in real time. To ensure scientific and precise financial management and optimize the allocation of social resources, it is necessary to make effective and accurate predictions of financial revenue [2].

Many financiers use various sources of information to forecast target values and make strategies to achieve

superiority in competitions. Earlier studies were based on fundamental and technical analysis, such as assessing the value of a company using its performance or estimating future target values based on historical data. Recently, owing to the advancement in computing, different data mining models are proposed. Since the financial markets produce big data in real time, an advantage in competition is the capability to use different modes of data in real time and depict the financial market in the form of data-driven guidelines that are deliverable to business decision-makers [3]. Therefore, it is essential to develop a standard process capable of deriving data-driven insights that are directly related to performance in the financial markets.

Data mining techniques can uncover hidden patterns and predict future trends and behaviors in financial markets. They create opportunities for companies to make proactive and knowledge-driven decisions to gain a competitive

advantage. Data mining has been applied to many financial applications, including the development of trading models, investment selection, loan assessment, and portfolio optimization. The competitive advantages achieved by data mining include better revenue, reduced cost, and improved marketplace awareness.

Recently, several studies conducted on data mining frameworks have been proposed for economic growth prediction. Jiao et al. [4] proposed a macroeconomic forecasting algorithm based on an optimized wavelet neural network. Different from the existing backpropagation (BP) neural network prediction model, the wavelet neural network model was used for macroeconomic prediction. In addition, the wolf colony algorithm was applied to optimize the weights of the wavelet neural network model. The normalized economic data was used to examine and train the proposed model. Changsheng et al. [5] combined the improved whale optimization algorithm (IWOA) with adaptive noise-based complete set empirical mode decomposition (CEEMD) and Elman neural network prediction models to depict the stock fluctuation trend. Initially, the Internet public opinion of 180 shares in SSE was mined and quantified by text mining techniques, and the Boruta algorithm was used to screen important attributes and reduce the complexity of the attribute set. Next, the CEEMD algorithm was employed to add a certain amount of white noise with specific variance to the attribute set to realize the decomposition and noise reduction of the attribute sequence. At the same time, the adaptive weight was used to improve the whale optimization algorithm (WOA) to enhance its global search and local mining ability. Finally, WOA is used to optimize the initial weights and thresholds of the Elman neural network. The author in [6] estimated the stock market crisis using logistic regression (LR) and extreme gradient boosting. In some studies, linear models in time series prediction, such as autoregressive integrated moving average (ARIMA), and the autoregressive moving average (ARMA) are combined with data mining methods to predict the stock market. Lin and Pai [7] combined support vector machine (SVM) with ARIMA to forecast the stock price. The combined model enhanced the prediction accuracy by employing the distinct features of linear and nonlinear structures. Likewise, Pan et al. [8] predicted three well-known time series datasets using ARIMA and an artificial neural network (ANN) hybrid model and extracted prominent patterns from time series data. Mostafa [9] combined generalized regression neural network (GRNN) and multilayer perceptron (MLP) that avoid predicting many hidden units to forecast price and made perfect forecasts and better understood the inner dynamics of the stock market in Kuwait. A heterogeneous model based on BP neural network and ARIMA was applied to the Growth Enterprise Market of the Chinese stock market [10]. They verified the strength and generalization of their model. Similarly, Preis and colleagues [11] recognized online precursors of stock prices from search volume data collected by Google Trends to categorize optimal trading strategies. Likewise, Moat et al. [12] studied the relationship 120 between Wikipedia's views change and total stock market

movements. However, the traditional prediction algorithm ignores the acquisition of financial redundancy and equity capital time series and has the problems of multicollinearity and error series correlation, which leads to the prediction accuracy being not ideal. To solve this problem, this paper designed a new macroeconomic growth prediction algorithm based on data mining. We analyzed the sequence of economic characteristics, reorganized the spatial structure of economic characteristics, and integrated the statistical information of economic data. Through the optimization of the Apriori algorithm, the association rules between macroeconomic data are also generated. The algorithm has high accuracy and can provide a reliable basis for macroeconomic change trends.

We organized the remaining paper as follows. In Section 2, we present the proposed economic growth prediction algorithm. The results and corresponding discussion are presented in Section 3. Finally, Section 4 summarizes the proposed work.

2. Designing Prediction Modeling Algorithm

In this section, we propose an improved data mining algorithm for macroeconomic growth prediction. The proposed algorithm is described in the following section.

2.1. Statistical Information Analysis and Fusion of Economic Data. To realize the design of the macroeconomic data growth prediction algorithm, combined with the fuzzy correlation analysis method [13], we analyzed the economic characteristic series and presented the constraint variable factor of macroeconomic time series. Using the analysis methods of explanatory variables and statistical variables, the sample set of fuzzy characteristic sampling information of the characteristic point of macroeconomic time series distribution at time t was formulated and is expressed as $(y_{1,j}, y_{2,j}, \dots, y_{t,j})$, where t is the steady-state parameter of the characteristic distribution of macroeconomic data. Under the control of the stable growth trend model, the statistical characteristic quantity of macroeconomic data feature parameter fusion $y_{t,j}$ was combined with the template feature matching method, and the weighted coefficient for the macroeconomic data growth prediction was computed. Based on the analysis of supplier concentration, the fuzzy similarity feature set of macroeconomic data growth was calculated. The standard quantitative distribution coefficient of macroeconomic data is calculated using the following equation:

$$X = \frac{f(x)}{t}, \quad (1)$$

where $f(x)$ is the fuzzy characteristic clustering function of macroeconomic time series at time t . Applying semantic detection, the joint characteristic distribution value of the macroeconomic time series is obtained as follows:

$$P = \frac{X}{[d_i + d_s]}, \quad (2)$$

where d_i and d_s are the correlation coefficients of the negative impact of concentration. Using the fuzzy clustering method, the growth of macroeconomic data is predicted as

$$K(x) = \sum_{t=1} P(t) + f(x). \quad (3)$$

Based on the theory of information asymmetry and earnings growth prediction method, the joint detection statistical characteristic quantity of macroeconomic time series is obtained as follows:

$$R = \frac{c \cdot (K(x) + a(t))}{\phi(t)} + e(a), \quad (4)$$

where $\phi(t)$ is the time interval distribution difference function, $e(a)$ is the decentralized control parameter of economic data, and c is the weighting coefficient of macroeconomic data growth prediction. Based on the joint autocorrelation feature analysis method, the mean membership degree of each class of macroeconomic time series samples is

$$E_j = \sum \frac{F_{ij}}{K_j}, \quad (i \in K_j, j = 1, 2, \dots, N), \quad (5)$$

where K_j is the total number of macroeconomic time series samples of class j , the joint feature distribution function is extracted, and the statistical fusion center block function is S . The first cluster center z_1 is the distribution value s_i of joint statistical parameters of macroeconomic time series to realize the statistical information analysis of macroeconomic data.

Combined with the analysis method of feature detection of accumulated financial redundant resources [14], the equity capital feature matching of macroeconomic data is realized. Applying the separation degree of macroeconomic data μ , the fuzzy autocorrelation feature matching set of macroeconomic time series $s(c)$ and the joint feature distribution set of macroeconomic time series are computed as follows:

$$I(a_t) = \frac{s(c) + N_s + X}{N} + u, \quad (6)$$

where N is the reliability characteristic quantity of macroeconomic time series and N_s is the joint distribution parameter of macroeconomic time series in S .

Using the method of multiple growth prediction, the macroeconomic dataset X can be represented as

$$X = \{x_1, x_2, \dots, x_n\}, \quad (7)$$

where n is the number of datasets and each element X is represented as the p -dimensional vector of economic fluctuation risk in the same year. X is comprised of n categories, and the economic fluctuation parameter of enterprises in the first category is $v_i = \{v_{i1}, v_{i2}, \dots, v_{ip}\}$. Using joint distributed detection, the similarity characteristic variable of macroeconomic time series at the i^{th} moment is computed as $\{x_{i1}, x_{i2}, \dots, x_{im}\}$, and the corresponding correlation distribution type of macroeconomic time series is the value 1 or -1 of y_i , where $y_i = 1$ represents normal and $y_i = -1$ represents abnormal growth.

$$A = y_i \cdot x_i + \{v_{i1}, v_{i2}, \dots, v_{im}\}. \quad (8)$$

Based on the results of information fusion, the big data distribution characteristics of macroeconomic time series

are extracted to improve the accuracy of macroeconomic growth prediction.

2.2. Association Rules Mining Based on Improved Apriori Algorithm. The traditional Apriori algorithm [15] not only needs to scan the macroeconomic database repeatedly but also produces a large number of candidate itemsets. Combining the Map-Reduce parallel technique and matrix multiplication, the parallel processing ability of the Apriori algorithm was strengthened, and the algorithm performance was improved. The Map-Reduce is a computation technique used for big data computation. The steps are as follows.

Parallelization. The macroeconomic database was divided into several blocks and was represented in the form of a matrix, and finally the matrixed data was allocated to the corresponding itemset for map and reduced calculation.

Matrix. The transaction database was scanned, and transformed information was stored in a matrix. Each row in the matrix represented an object, and each column is an item in the transaction. The numbers "0" and "1" represent whether the item exists in the current transaction. If it exists, it is represented as "1", and if it does not exist, then it is "0".

Calculate Support. The items were sorted according to the itemset. For two frequent itemsets A and B , if their first k items were the same, then they can connect to generate a candidate k itemset C . If the first k items of C come from their first k items, then two candidate sets a' and b' will be generated. The complexity of the matrix was obtained by logic calculation. This method reduced the number of scans on the database. After modification, the database was only required to be scanned twice, and the frequent itemsets and association rules were obtained quickly.

The improved Apriori algorithm mining data can make the mining rules more accurate [16]. This method not only meets the needs of users but also reduces the scale of the frequent itemset, saves data mining time, and improves efficiency. Let the mining target item be NT , and the mining process is as follows.

- (i) Get a frequent itemset L_1 . If L_1 contains two items NT , continue; otherwise, end.
- (ii) Calculate the support degree of NT the candidate itemset; stop mining if it is less than the minimum support degree; otherwise, continue mining.
- (iii) Obtain the candidate item h by connecting the target item and the itemset of a frequent itemset L_1 that does not belong to N and T . At this time, all itemsets in A_3 are $NT * -$ type to obtain support.
- (iv) Select the frequent itemset L_3 which is greater than the minimum support; then, each itemset of L_3 contains the target item NT . Then, the candidate itemset connecting the frequent itemset L_3 will

contain 0 items, and finally, the mining rules obtained will also contain NT items.

Using these steps to mine data through the modified Apriori algorithm improves the accuracy of mining to a greater extent. The specific process of mining association rules using the improved Apriori algorithm is as follows.

- (i) Scan the database to get frequent itemsets L_1 for parallel and matrix processing. Use this matrix to distinguish whether the macroeconomic data is in the current target item to analyze whether the data need to be deleted or divided into the corresponding project set.
- (ii) Connect the frequent itemset L_1 with all the data that do not belong to the target item in the target itemset, and each itemset in the obtained itemset $L(t+1)$ contains the target item, where t is the number of target items, to achieve the purpose of the algorithm.
- (iii) Use the improved Apriori algorithm to mine frequent itemsets, and use the logical operation of itemsets to obtain the local support of target items.
- (iv) Use Reduce operation to obtain the local support of each itemset and candidate set after integration, and finally obtain the global frequent itemset.

The improved Apriori algorithm only needs to scan the macroeconomic database twice. The first scan generates frequent itemsets, and the second scan transforms other data into a matrix for a logical operation to generate association rules. Compared with the traditional Apriori algorithm, it reduces the number of database scanning processes, prunes the candidate set by Map-Reduce, and connects the target item L_1 and the item A_3 other than the target item, so that each item of the frequent itemset contains the target item, which makes the improved Apriori algorithm more decisive. Moreover, it reduces the candidate set and accelerates the speed of data mining.

2.3. Macroeconomic Feature Extraction. Before using a large amount of data in prediction models, we use mining association rules and apply the kernel principal component analysis (KPCA) method to extract the features of economic data [17]. KPCA is an extension of PCA using kernels. Using a kernel, the originally linear operations of PCA are performed in a reproducing kernel Hilbert space. Suppose x_1, x_2, \dots, x_a is training data and $\{x_1\}$ is input space. Let the corresponding mapping be Φ ; we combine the mapping Φ with kernel function to realize the high-dimensional spatial mapping F of features, so that the macroeconomic data can meet the centralization conditions needed in the feature space.

Suppose there are M macroeconomic data samples with N features. Using KPCA, we extract the features of M macroeconomic data. The steps are as follows:

The N feature obtained from macroeconomic data with M the sample is represented in the matrix and the $(M \times N)$ -dimensional data matrix 'A' is obtained as

$$A = \begin{Bmatrix} j_{11} & \cdots & j_{1n} \\ \vdots & \ddots & \vdots \\ j_{m1} & \cdots & j_{mn} \end{Bmatrix}, \quad (9)$$

j represents macroeconomic data.

The matrix K is calculated using the following equation:

$$K = (\Phi(j)^2), \quad (10)$$

Φ is the mapping between economic data.

The function parameter M and Gaussian radial basis function K' in the matrix are obtained by modifying the kernel matrix as follows:

$$K' \rightarrow M - \frac{1}{m} \left(\sum_{m=1}^M K \right). \quad (11)$$

Using the modified matrix, the eigenvector v_0, \dots, v_M KL and the eigenvalue $\lambda_0, \dots, \lambda_M$ of its matrix are tested and calculated. The eigenvalue $\lambda_0, \dots, \lambda_M$ is processed in descending order, and the corresponding eigenvector is modified. Moreover, Schmidt orthogonalization is used to process eigenvectors.

The cumulative contribution rate B_1, \dots, B_M and the specified information extraction rate P are calculated according to the eigenvector. If $B_1 \geq P$ and $B_{1-t} < P$, the previous feature vector β_1, \dots, β_m is extracted.

The projection on the eigenvector is obtained by calculating the corrected kernel matrix.

Based on the principle of function replacement, KPCA is used to reduce the dimension of economic data to obtain the feature vector $\beta = (\beta_1, \dots, \beta_l)$ of economic data. The data mapping result is described using the following equation:

$$(\nu^k \cdot \Phi(x)) = \sum_{i=1}^M (\alpha_i) \cdot K, \quad (12)$$

where ν^k is the spatial projection of the eigenvector and x is the input space.

Based on the above, through the warrant analysis method of equity capital time series detection, it is assumed that the initial frequency mean of the associated data in the macroeconomic time series is $\hat{\mu}$, and the standard deviation is $\hat{\sigma}$. The training samples update the iterative abnormal state space using the following equation:

$$\Delta q = \hat{\mu} \cdot \cos A + \hat{\sigma}. \quad (13)$$

According to the optimal grouping detection results of macroeconomic time series, the center point of the cluster is updated, and the correlation characteristic component of the associated data in the macroeconomic time series is obtained. Through mutual information parameter identification of macroeconomic time series [18], the joint characteristic parameter distribution is obtained as follows:

$$B(s) = p_v \frac{w(e_i)}{\Delta q}, \quad (14)$$

where $w(e_i)$ is the characteristic parameter associated with data sample and p_v is the mutual information identification parameter of economic time series [19]. Using the above steps, the information characteristic model of the macroeconomic time series is constructed. Combined with quantitative regression analysis and subsection test method, the joint probability density distribution of macroeconomic time series is obtained as follows:

$$\Gamma(a) = \frac{\sum_{s=1} B(s) + w_n(t)}{P_v}, \quad (15)$$

where $w_n(t)$, ($n = 1, 2, \dots, N$) is a complex Gaussian process with zero mean and one variance. The distribution of associated data in macroeconomic time series is computed as

$$E(s) = \sqrt{S_d(f)} + f(x) + \Gamma(a), \quad (16)$$

where $S_d(f)$ is the Doppler parameter of macroeconomic time series distribution and $f(x)$ is the joint characteristic distribution parameter, which is obtained by calculating the information focus parameter of associated data flow in macroeconomic time series. By associating the training set with the big data of macroeconomic time series, the gain correlation characteristic quantity of the associated data in macroeconomic time series is calculated using the following equation:

$$G_{\text{features}} = \frac{[f(x) \cdot c_i]}{E(s)}, \quad (17)$$

where c_i is the big data information control function of economic time series.

2.4. Macroeconomic Growth Prediction Algorithm. With the recent development of data mining techniques, rather complex models are representing good performance. We obtained the joint distribution characteristic quantity $w_N(t)$ of the macroeconomic time series, the gain correlation characteristic quantity, and the joint parameter fusion of the associated data [20]. The joint autocorrelation characteristic analysis is performed to obtain the Doppler parameter σ_n of the macroeconomic time series. We jointly estimated the macroeconomic time series and obtained the maximum likelihood estimates \hat{s} and $\hat{\tau}$ using the residual analysis method of the regression model $s_0 \tau_0$. The logarithm of the natural distribution of the macroeconomic time series is given as

$$L = \frac{w_N(t)}{G + \sigma_n} + \{\hat{s} + \hat{\tau}\}. \quad (18)$$

We set the measurement index parameter B_L of environmental uncertainty and obtained the category number of macroeconomic time series samples (e.g., the i ($i \in p$)) and the distribution coefficient of economic time series samples [20], using the method of autocorrelation feature distributed

estimation, to calculate the difference between the maximum membership degrees of each type of economic time series samples and get the detection statistical distribution coefficient value of macroeconomic time series as B_L . Using the joint parameter analysis method, the macroeconomic time series prediction results are computed as follows:

$$Y = L + \frac{B_L}{n + x_n}, \quad (19)$$

where n is the number of training samples of economic time series big data distribution and x_n is the standard deviation of economic time series distribution.

3. Experimental Design and Result Analysis

To verify the application performance of the proposed algorithm in the prediction of macroeconomic growth, the simulation test experiment was designed using the MATLAB R2021a platform. Combined with SPSS 17.0 statistical analysis software, the data sample detection model of macroeconomic growth prediction was developed. We used 1024 data samples of the economic data series with 214 training data samples and 120 data samples as the statistical characteristic distribution set. The autocorrelation distribution coefficient was 0.212. We employed the prediction algorithm based on the optimized wavelet neural network and improved the whale optimization algorithm. According to the above parameters, the experimental process was designed as follows.

3.1. Efficiency Test of Macroeconomic Growth Prediction Algorithm. We constructed the descriptive statistical analysis model of macroeconomic growth prediction for the years 2008–2020. The experimental results are shown in Table 1.

According to the descriptive statistical analysis of the macroeconomic time series given in Table 1, the input dispersion diagram of macroeconomic data distribution is obtained, as shown in Figure 1.

According to the distribution of macroeconomic input, the growth of the macroeconomy is predicted, and the predicted output is shown in Figure 2.

Figure 2 shows the year-wise return on net assets. It is evident that the proposed algorithm can effectively predict the macroeconomic growth of economic time series.

3.2. Computation Time of Macroeconomic Data Mining Algorithm. We compared the computation time of the proposed prediction model with the computation time of different algorithms to mine the economic data. The experimental results of the computation time for different data mining algorithms are shown in Figure 3.

We compared the computation time of the proposed data mining algorithm with the algorithm of the optimized wavelet neural network given in [4] and the method based on the improved whale optimization algorithm given in [5]. Figure 3 shows that the proposed algorithm is obviously more efficient in terms of data mining time than the other

TABLE 1: Results of the descriptive statistical analysis of the macroeconomy.

Year	Economic input cost/million yuan	Return on net assets	Economic growth level
2008	65.3	0.6	0.6
2009	75.3	0.8	0.5
2010	45.1	0.7	0.6
2011	44.2	0.5	0.6
2012	67.4	0.7	0.5
2013	65.3	0.7	0.7
2014	33.3	0.9	0.6
2015	65.4	0.6	0.6
2016	75.1	0.9	0.9
2017	95.1	0.5	0.4
2018	45.2	0.3	0.3
2019	32.3	0.7	0.5
2020	31.1	0.5	0.3

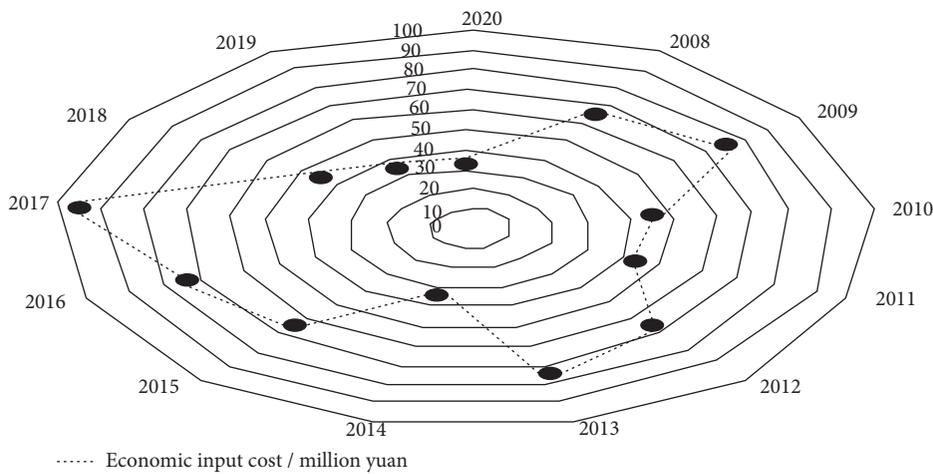


FIGURE 1: Input distribution of macroeconomic data.

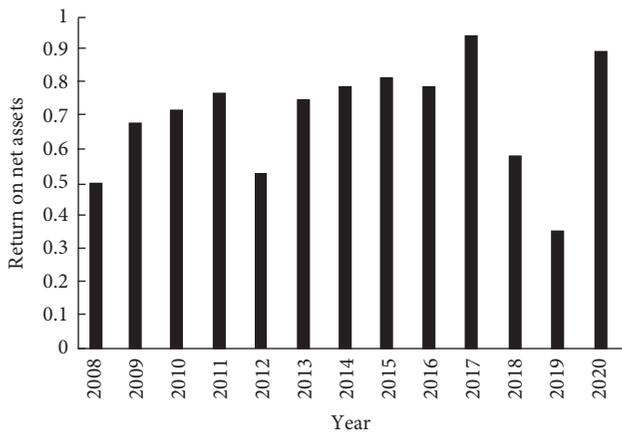


FIGURE 2: Growth forecast results of economic time series.

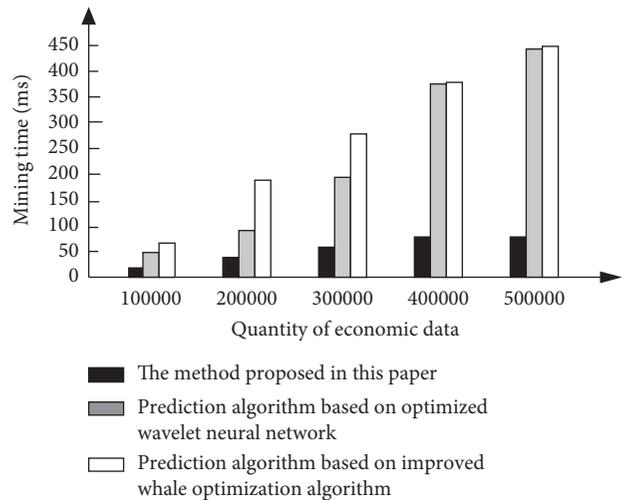


FIGURE 3: Mining time test results.

two algorithms. This is because the proposed algorithm improves the traditional Apriori algorithm by matrix multiplication so that it only needs to scan the macroeconomic database twice. We can then get the association rules between the data and realize the mining of macroeconomic data based on association rules, which leads to time reduction in mining data and improves the mining efficiency.

3.3. Comparative Test of Macroeconomic Data Mining Rate. Data mining extracts useful data from a large set of raw data. Figure 4 shows the comparative results for the data mining rate of the proposed algorithm with optimized wavelet neural network and improved whale optimization algorithms.

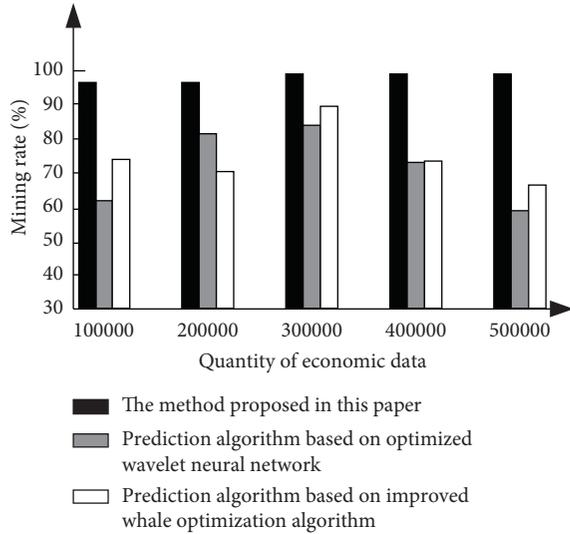


FIGURE 4: Mining rate test results.

The mining rate of the proposed algorithm is higher than that of the optimized wavelet neural network and improved whale optimization algorithm. In addition, the mining rate of the proposed algorithm decreases more smoothly with the increase of detection times. The main reason is that the proposed algorithm uses the Apriori algorithm to quickly mine association rules in a short time and uses the Map-Reduce method to prune the candidate set in the mining process. With the help of pruning, the precision of association rules is improved, and the high-precision association rules are applied to extract the distinct features of macroeconomic data and consequently improve the mining rate of the proposed algorithm.

3.4. Comparative Analysis of False-Positive Rate of Different Economic Data Mining Algorithms. Each data mining technique has its intrinsic characteristics and underlying assumptions that make it a better choice for macroeconomic growth prediction. We also computed the false-positive rate of the proposed algorithm, optimized wavelet neural network, and improved whale optimization algorithms. The results are shown in Figure 5.

The false-positive rate of the proposed algorithm is relatively higher than that of the optimized wavelet neural network and improved whale optimization algorithms. However, the proposed algorithm is more stable than the other two methods. In the proposed method, the improved Apriori algorithm is applied in the process of feature extraction, and the kernel matrix is modified to obtain a new feature vector, which makes the feature extraction more targeted, which reduces the false-positive rate in the process of macroeconomic data mining.

3.5. Comparison of Economic Growth Prediction Accuracy of Different Algorithms. We investigated the accuracy of different algorithms used for economic growth prediction. Figure 6 shows the prediction accuracy achieved by different

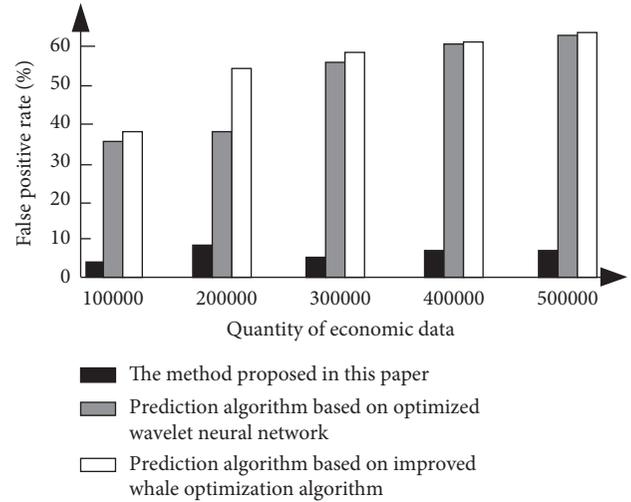


FIGURE 5: Mining false alarm rate test results.

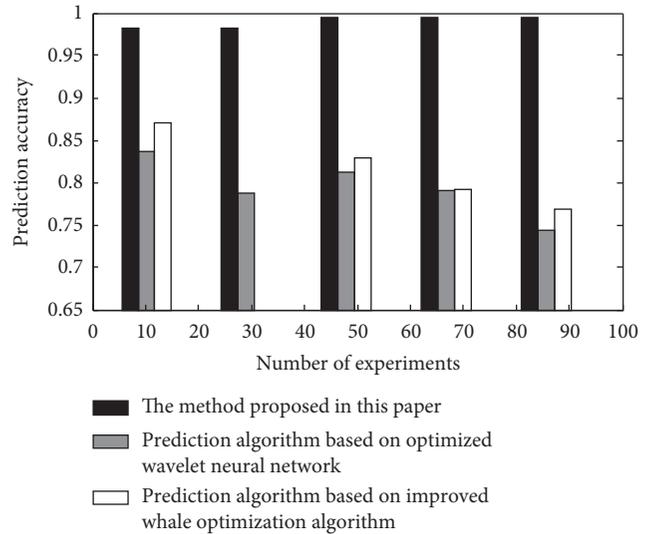


FIGURE 6: Comparison of economic time series growth prediction accuracy.

macroeconomic growth prediction algorithms. It can be seen that the economic time series growth prediction accuracy of the proposed algorithm is higher than the other two algorithms.

The proposed algorithm constructs the financial redundancy and equity capital time series detection model of macroeconomic data growth prediction. Through the warrant analysis of equity capital time series detection, the accuracy of the growth forecast is effectively improved, and the application effect of the algorithm is optimized.

4. Conclusion

To solve the problem of large deviation of traditional economic growth prediction algorithms, a new macroeconomic growth prediction algorithm based on data mining is developed. This study analyzes the sequence of economic

characteristics, reorganizes the spatial structure of economic characteristics, and integrates the statistical information of economic data. Using the optimized Apriori algorithm, the association rules between macroeconomic data are generated. The residual analysis method of the regression model is used to predict the growth of macroeconomic data. To verify the effectiveness of the algorithm, a simulation experiment is designed. The experimental results show that the proposed algorithm has ideal application adaptability, and the economic growth prediction results are in line with the actual situation, which shows that the algorithm has high accuracy and can provide a reliable basis for macroeconomic change trends.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] W. Wu, W. Wang, and Z. Zhang, "Using China's provincial panel data exploring the interaction between Socio-economic and Eco-environment system," *Ecological Complexity*, vol. 44, no. 2, pp. 100–108, 2020.
- [2] L. Mauler, F. Duffner, and J. Leker, "Economies of scale in battery cell manufacturing: the impact of material and process innovations," *Applied Energy*, vol. 286, no. 15, pp. 116–121, 2021.
- [3] J. S. H. van Leeuwen, B. Mathijssen, and B. Zwart, "Economies-of-Scale in many-server queueing systems: tutorial and partial review of the QED halfin--whitt heavy-traffic regime," *SIAM Review*, vol. 61, no. 3, pp. 403–440, 2019.
- [4] F. Jiao and L. Hongpu, "Research on macroeconomic forecasting technology based on optimized wavelet neural network," *Modern Electronics Technique*, vol. 42, no. 7, pp. 189–191, 2019.
- [5] Z. Changsheng, K. Langhe, and F. Wenfang, "Stock closing price prediction algorithm using adaptive whale optimization algorithm and Elman neural network," *Journal of Computer Applications*, vol. 40, no. 5, pp. 1501–1509, 2020.
- [6] S. P. Chatzis, V. Siakoulis, A. Petropoulos, E. Stavroulakis, and N. Vlachogiannakis, "Forecasting stock market crisis events using deep and statistical machine learning techniques," *Expert Systems with Applications*, vol. 112, pp. 353–371, 2018.
- [7] P.-F. Pai and C.-S. Lin, "A hybrid ARIMA and support vector machines model in stock price forecasting," *Omega*, vol. 33, no. 6, pp. 497–505, 2005.
- [8] Y. Pan, Z. Xiao, X. Wang, and D. Yang, "A multiple support vector machine approach to stock index forecasting with mixed frequency sampling," *Knowledge-Based Systems*, vol. 122, pp. 90–102, 2017.
- [9] M. M. Mostafa, "Forecasting stock exchange movements using neural networks: empirical evidence from Kuwait," *Expert Systems with Applications*, vol. 37, no. 9, pp. 6302–6309, 2010.
- [10] L. Xiong and Y. Lu, "Hybrid ARIMA-BPNN model for time series prediction of the Chinese stock market," in *Proceedings of the 2017 3rd International conference on information management (ICIM)*, pp. 93–97, IEEE, Chengdu, China, April 2017.
- [11] T. Preis, H. S. Moat, and H. E. Stanley, "Quantifying trading behavior in financial markets using google trends," *Scientific Reports*, vol. 3, no. 1, Article ID 1684, 2013.
- [12] P. Ejegwa, S. Wen, Y. Feng, W. Zhang, and N. Tang, "Novel Pythagorean fuzzy correlation measures via Pythagorean fuzzy deviation, variance and covariance with applications to pattern recognition and career placement," *IEEE Transactions on Fuzzy Systems*, vol. 10, p. 1, 2021.
- [13] Y. Zhou, H. Jin, and B. Liu B, "An improved feature parameter extraction algorithm of composite detection method based on the fusion theory," *Journal of Sensors*, vol. 2021, no. 18, 10 pages, Article ID 8898991, 2021.
- [14] M. Y. Idna Idris, N. B. Abd Warif, H. Arof, N. Mohamed Noor, A. W. Abdul Wahab, and Z. Razak, "Accelerating fpga-surf feature detection module by memory access reduction," *Malaysian Journal of Computer Science*, vol. 32, no. 1, pp. 47–61, 2019.
- [15] X. Yu, D. Zhan, L. Liu, H. Lv, L. Xu, and J. Du, "A privacy-preserving cross-domain healthcare wearables recommendation algorithm based on domain-dependent and domain-independent feature fusion," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, p. 1, 2021.
- [16] M. Yu, T. Quan, Q. Peng, X. Yu, and L. Liu, "A model-based collaborate filtering algorithm based on stacked AutoEncoder," *Neural Computing & Applications*, vol. 204, 2021.
- [17] B. Bouaita, A. Moussaoui, and N. E. I. Bachari, "Rainfall estimation from MSG images using fuzzy association rules," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 1, pp. 1357–1369, 2019.
- [18] Y. Qiuxiang, W. Guannan, and W. Ting, "Simulation of multi-source log security data mining based on time series," *Computer Simulation*, vol. 36, no. 02, pp. 307–311, 2019.
- [19] R. Shirani 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.
- [20] A. Al-Qaisi, M. S. Altarawneh, A. Elsaid, and Z. Alqadi, "A hybrid method of face feature extraction, classification based on MLBP and layered-recurrent network," *Traitement du Signal*, vol. 37, no. 4, pp. 555–561, 2020.