

Retraction

Retracted: Modeling and Simulation of Real Estate Valuation Based on Genetic Neural Network Model

Journal of Function Spaces

Received 15 August 2023; Accepted 15 August 2023; Published 16 August 2023

Copyright © 2023 Journal of Function Spaces. 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.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets 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 own 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. Shen, "Modeling and Simulation of Real Estate Valuation Based on Genetic Neural Network Model," *Journal of Function Spaces*, vol. 2022, Article ID 1078057, 9 pages, 2022.

Research Article

Modeling and Simulation of Real Estate Valuation Based on Genetic Neural Network Model

Shuli Shen 

School of Public Administration, Zhejiang Gongshang University, Hangzhou 310018, China

Correspondence should be addressed to Shuli Shen; shenshuli@mail.zjgsu.edu.cn

Received 7 June 2022; Revised 30 July 2022; Accepted 2 August 2022; Published 17 August 2022

Academic Editor: Miaochoao Chen

Copyright © 2022 Shuli Shen. 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.

China's real estate market is developing rapidly, but the house price is abnormal. The nonlinear relationship between housing characteristics and real estate value is difficult to calculate, resulting in the difficulty of house price prediction. Based on this, the relationship model between characteristic variables and house prices is constructed by using the machine learning method. At the same time, genetic algorithm is used to screen the specific values. The experimental results show that the optimized model converges in 56 iterations; in the application test, the research model found that in 90% of the test samples, the error between the predicted value and the actual measured value shall not exceed 10%. Experiments show that the genetic algorithm is effective in optimizing the BP house price valuation model and improves the calculation efficiency and valuation accuracy of the valuation model.

1. Introduction

Real estate economic activities consist of real estate production, real estate circulation, and real estate consumption. Among them, real estate production is the direct production process of real estate, and real estate circulation is the reproduction and realization process of real estate. Real estate consumption is the realization of the purpose of real estate production and the continuation of some direct production processes in the consumption link. The circulation of real estate production, circulation, and consumption needs corresponding economic operation mechanism and economic system conditions. The real estate economy can be investigated from three aspects: microeconomics, meso-economy, and macro-economy. All economic behaviors taking real estate enterprises as economic units and the operation of real estate development and management projects belong to the micro-level. The real estate sector economy and the real estate regional economy basically belong to the meso-economy. As a kind of industrial economy, the real estate sector economy has many characteristics and special laws beyond its commonness. Real estate regional economy mainly refers to the real estate economy of each province,

city, and district, which belongs to a department of regional industrial structure. Considering the whole national economy as an object belongs to the macro level. The real estate economy is not only the object of national macro-control but also an important part of the macro-economy. Real estate not only includes the material form of buildings and land but also has legal and economic characteristics due to the real rights generated by entities and economic activities generated by transactions. Real estate value is generally the sum of building value and land value, but due to its legal and economic characteristics, region, population, consumption psychology, policy, and even buyer gender have also become important factors affecting the change of real estate value [1]. At the same time, real estate transactions often drive the development of other investment and trading markets, increasing the risk of market professional portfolio, especially when purchasing investment real estate and using extensive leverage. Therefore, maintaining the healthy and stable market is an important measure for China's economy and people's livelihood [2]. The value of real estate is ultimately reflected in the price paid by the real estate demander to the supplier. Therefore, building a valuation model based on predicting the transaction price of real estate is a common evaluation method in the trading market.

Traditional asset appraisal has objective disadvantages in terms of data volume and appraisal funds. Asset appraisers have limited ways to obtain information and data and can only consult and learn through field investigations or cases filed by their own companies. Moreover, the traditional asset evaluation methods inevitably have a lot of subjectivity, which has a certain impact on the objectivity and fairness of the evaluation results. How to eliminate the subjectivity in the evaluation process to the greatest extent and establish an objective evaluation system to the greatest extent is also a major requirement of the economic and social development in the new era. The research method of this paper is to use a specific algorithm to confirm the predicted value. Finally, the algorithm value after the selection is obtained, design the fitness function to specify the iteration scale of the algorithm, and finally classify and input various indicators affecting house prices. Through the above optimization, the real estate valuation model is constructed.

The innovation of the research is that it not only considers the use of genetic algorithm to improve the fitness function design of valuation model and the parameters of neural network threshold but also inputs the characteristic variables affecting real estate value. Before the experiment, the research classifies various characteristic factors into location indicators and building indicators and quantifies and normalizes these variables, so as to avoid the negative impact on the model caused by the non-uniform unit and large difference of numerical magnitude between variables.

The second part briefly describes the theoretical achievements of domestic and foreign researchers on real estate market valuation and investment in recent years, as well as the application experience of genetic algorithm and neural network in different industries. In the third part, based on the characteristic variables of real estate transactions, a neural network model for valuation is proposed, and the optimization process of genetic algorithm for neural network is studied. The last part is the training and testing experiment of the model. The effectiveness of the experimental results is analyzed by using algorithm tools. In terms of measurement accuracy of experimental results, different models are judged differently, and different results are obtained.

2. Related Works

In recent years, although the real estate market is in the development trend of continuous expansion, the real estate price has been in an unstable state due to the influence of real estate speculation, investment, and other factors. Therefore, more and more people begin to pay attention to the real estate market, and professional scholars have put forward different views on the prediction of real estate and house prices, and studied buyers' expectations and investment potential for sustainable residential real estate by capturing four variables: economy, society, environment, and system. The results show that countries adhere to the sustainable development goal (SDG) as the economic strategy. Due to the changing market environment, this paper makes a highly sensitive detection of the powerful environment [3]. Dobrowskien et al. constructed an evaluation model of real estate

sustainability index (resi) at the technical level based on the multi-criteria decision-making method. Experiments show that the model can promote the consideration of sustainable investment in new technologies [4]. Hodoshima et al. analyzed the real estate investment trust market in Tokyo based on four performance indicators: internal risk aversion rate (Irta), sharp ratio, Sotino et al. found that Irta is more sensitive to potential risks and more relevant to risk averse investors [5]. It is found that the total income rate of return initially fluctuated around 5%, but it showed a downward (upward) trend. The results show that the annualized actual total return after deducting costs ranges from about 2.3% of residential real estate to 4.5% of agricultural real estate [6]. D'acci quantifies regional characteristics based on market comparison method. The relationship between factors in different positions was studied. After calculation, the results show that there is a great relationship between the increase of land price and distance [7]. Ulbl et al. constructed the batch valuation model of apartment market based on the generalized additive model method. Experiments show that the model can accurately analyze the impact of apartment market heterogeneity on apartment transaction price [8]. Wang et al. introduced the application trend and classification of batch evaluation according to the group evaluation models and methods from 2000 to 2018 and emphasized the 3I trend. Many experiments in the past have shown that different spatial expressions are very different based on different Internet environments [9]. He et al. designed and described a method of generating finite difference (rbf-fd) based on multi quadratic radial basis function of local RBF format. The results show that rbf-fd method is effective and stable in terms of pre-real estate valuation accuracy [10]. Sabina et al. applied fuzzy logic theory to estimate the land value of agricultural land market. The results show that the real estate is in the process of initial development. The research method of this paper has certain utility. This increases the opacity of the market [11].

Chen et al. prove the accurate prediction effect and stable operation performance of the model [12]. Wang et al. measured the relationship between pupil response and objective pain based on the artificial neural network machine learning method optimized by genetic algorithm. The results show that pupil response and machine learning algorithm are a promising objective pain level evaluation method, which can improve the patient's experience of measuring pain in long-range medical care [13]. Xie et al. believe that genetic algorithm does not guarantee that you can obtain the optimal solution of the problem, but the biggest advantage of using genetic algorithm is that you do not have to understand and worry about how to "find" the optimal solution. As long as you simply "deny" some individuals who perform poorly. This is the essence of genetic algorithm [14]. Jiang et al. taking the gearbox of a tractor as the engineering background, the BP neural network based on genetic algorithm is used to diagnose the gearbox fault. Statistics show that about 60% of gearbox faults are caused by gearbox faults, so only gear faults are studied here. Here, several characteristic quantities in the frequency domain are selected. Gear faults in the frequency domain are more obvious in the edge band

at the meshing frequency [15]. Combined with the analysis of the corresponding training results, using genetic algorithm to optimize the connection weight of BP neural network is more effective than the traditional algorithm, but it also needs to try the network structure. In other words, it is to select the appropriate number of hidden layer neurons [16]. Moreover, using genetic algorithm to optimize the weight and structure of neural network at the same time is more intelligent. It is possible to find the appropriate initial weight and excellent network structure. However, if the data is cumbersome, the search speed will slow down [17].

To sum up, relevant scholars have a variety of research methods on real estate valuation and housing market investment. Based on the different development of real estate market in different regions, the valuation methods adopted also change accordingly. In some foreign regions, the real estate transaction information is not transparent on the Internet or the network transaction is not developed. Therefore, the methods of some foreign scholars are still in the market comparison method of collecting offline transaction information, while China's transaction market has tended to be offline and online parallel. Therefore, the evaluation method also needs to be improved. The model of genetic algorithm optimized neural network has been shown in medicine, geological survey, engineering, and so on. In view of this, this paper studies the nonlinear machine learning method based on BP neural network optimized by genetic algorithm and discusses the impact of location variables and construction variables on house prices, hoping to provide a stable valuation idea for China's real estate trading market.

3. BP Neural Network House Price Valuation Model Based on Genetic Algorithm

3.1. Establishment of BP Neural Network Algorithm Model. There are three traditional real estate valuation methods: Among them, market comparison method is the most widely used because of the frequent transactions. The marketing mode of the real estate industry has also changed from offline dominance to online and offline combination. The traditional market comparison method also has too large deviation in valuation results due to personnel interference and scattered transaction information. Therefore, genetic algorithm (GA) is introduced to optimize the back propagation (BP) neural network threshold. The nonlinear machine learning method of BP neural network is used to replace the traditional calculation to establish the general model of real estate value evaluation [18]. The basic topology of BP neural network is shown in the figure below.

As shown in Figure 1, the weight optimization among the three depends on machine learning adjustment and is not affected by personal experience [19]. At the same time, the nonlinear mapping functions so BP neural network is used to construct house price evaluation model. Firstly, describe the forward propagation model in the network structure. Set the hidden layer nodes as n , and the calculation formula of hidden layer output is shown in formula (1).

$$z_b = f_1 \left(\sum_{a=0}^m v_{ba} \cdot x_a \right). \quad (1)$$

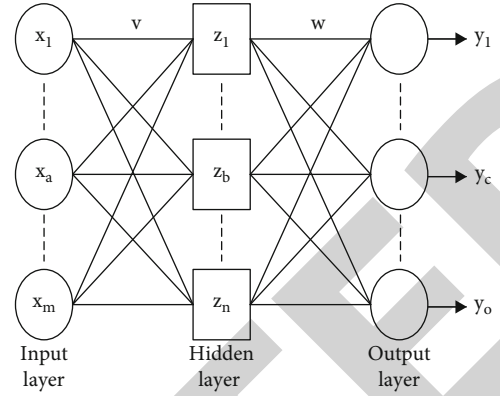


FIGURE 1: Structure and parameter of BP neural network.

In formula (1), x_a represents the a input layer parameter, z_b is its corresponding hidden layer parameter, f_1 is the hidden layer transfer function, indicating the transfer after mapping the input layer variables of weighted x , and v_{ab} indicates the weight between input layer x_a and hidden layer z_b . Similarly, the output formula of the output layer node is shown in equation (2).

$$y_c = f_2 \left(\sum_{b=0}^n w_{cb} \cdot z_b \right). \quad (2)$$

In equation (2), y_c is the calculated output layer signal, $c = 1, 2, 3 \dots o$ and f_2 are the output layer transfer function, and w_{cb} is the weight between the hidden layer z_b which is defined by mean square error, which is defined as the following equation (3).

$$E_g = \frac{1}{2} \sum_{c=1}^o (t_c^g - y_c^g)^2. \quad (3)$$

In equation (3), E_g indicates the error between the trained output of the g learning sample and the expected output t . According to the error function, the global error formula is equation (4).

$$E = \sum_{g=1}^G E_g = \frac{1}{2} \sum_{g=1}^G \sum_{c=1}^o (t_c^g - y_c^g)^2. \quad (4)$$

In equation (4), E represents the global error of all G learning samples, and the weight is adjusted according to the obtained error. The model of weight adjustment of hidden layer is described in equation (5) below.

$$\Delta v = \sum_{g=1}^G \sum_{c=1}^o \eta (t_c^g - y_c^g) f_2'(S_c) w_{bc} f_1'(S_b) x_a. \quad (5)$$

In equation (5), Δv represents the adjusted value of the hidden layer weight, η represents the learning efficiency constant, f_1', f_2' represents the initial transfer, S_c represents the net input of the c node of the output layer, and its calculation

formula is $S_c = \sum_{b=0}^n w_{cb} \cdot z_b$; S_b represents the net input of the b , and its calculation formula is $S_b = \sum_{a=0}^m v_{ba} \cdot x_a$. Finally, the neuron weight adjustment formula of the output layer is shown in equation (6).

$$\Delta w = \sum_{g=1}^G \sum_{c=1}^o \eta (t_c^g - y_c^g) f_2'(S_c) z_b. \quad (6)$$

3.2. Construction of GA-BP Real Estate Valuation Model. This model's advantage is the treatment of the nonlinear relationship between influencing factors and results, but its disadvantage is that it is easy to fall into local minimization and the calculation convergence speed is slow. The former problem comes from the randomized initial weight of machine learning, and the second problem is the sawtooth phenomenon common to gradient descent method [20]. The genetic algorithm and genetic algorithm are introduced to optimize the structure of the two BP networks, as shown in Figure 2.

As shown in Figure 2, encode all individuals of the population in the initial solution space, set the definition domain of variable x as $[j, k]$, and the accuracy requirement is 10^{-5} , then the definition domain $[j, k]$ can be divided into $(k - j) \times 10^5$ equal length intervals, all intervals are represented by binary strings, and the length of the string is set as l , then each point in the definition domain can be represented by $(k_{l-1}k_{l-2} \cdots k_0)$. After the conversion from the above real variable space to binary bit string is completed, convert binary string into decimal system, and its formula is shown in formula (7).

$$\begin{cases} x' = \sum_{i=0}^{l-1} k_i \times 2^i \\ x = j + x' \frac{(k-j)}{(2^l - 1)} \end{cases} \quad (7)$$

The first line of equation (7) is the decimal formula for binary string conversion, and the second line is the calculation formula for the value of the variable. Based on binary coding and supplemented by gray coding, the mathematical expression of its transformation is shown in equation (8).

$$j_i \begin{cases} k_i, & i = 1 \\ k_{i-1} \oplus k_i, & i > 1 \end{cases} \quad (8)$$

In equation (8), j_i represents gray string corresponding to binary string k_i , and \oplus symbol represents modulo two addition operation. Because the individual's adaptability to the environment determines whether the individual can be retained in the new group, the fitness is the basis for forming optimization.

$$Fun = \sum_{i=1}^N a_i \cdot E_g + a \cdot E. \quad (9)$$

In equation (9), Fun is the fitness function constructed by the weight coefficient method, a is the proportion of all sample errors in the fitness function, and a_i is the proportion of each

sample in the fitness function. The third step of genetic algorithm optimizing neural network depends on the operation of selection, and the individuals with the best performance form a new population. The selection operation is repeated until the size of the population reaches the specified size. The mathematical model is shown in equation (10).

$$p = \frac{f_i}{\sum_{i=1}^N f_i}. \quad (10)$$

The function of crossover behavior is to maintain the nature of excellent individuals and obtain new excellent individuals through hybridization. The offspring production of single point crossover can be expressed as formula (11).

$$\begin{cases} x' = [x_1, x_2, \dots, x_k, y_{k+1}, y_{k+2}, \dots, y_n] \\ y' = [y_1, y_2, \dots, y_k, x_{k+1}, x_{k+2}, \dots, x_n] \end{cases} \quad (11)$$

In equation (11), the gray strings of the two parents are $x = [x_1, x_2, \dots, x_n]$ and $y = [y_1, y_2, \dots, y_n]$, respectively, the intersection point of the two is k , and there is only one chromosome breakpoint for single point crossing. When the string length of the parent is n , there are $n - 1$ kinds of crossing results. In the two-point crossover, if the parent string length is n , there are $((n - 1) \times (n - 2)) / 2$ kinds of crossover results, and the mathematical description of the two-point crossover is equation (12).

$$\begin{cases} x' = [x_1, x_2, \dots, x_i, y_{j+1}, y_{j+2}, \dots, y_j, x_{j+1}, x_{j+2}, \dots, x_n] \\ y' = [y_1, y_2, \dots, y_i, x_{j+1}, x_{j+2}, \dots, x_j, y_{j+1}, y_{j+2}, \dots, y_n] \end{cases} \quad (12)$$

In equation (12), the gray string of the parent is $x = [x_1, x_2, \dots, x_n]$ and $y = [y_1, y_2, \dots, y_n]$, and the two intersections are set to i, j and $1 < i < j$ at the same time. The introduction of cross behavior can not only retain the advantages of developing individual groups but also reduce the number of iterations of the model and accelerate the convergence. Finally, the mutation behavior is introduced to expand the diversity of the population. Its specific operation is to randomly select individuals, give the mutation probability, and change their chromosome coding. At the same time, the structure of BP neural network is optimized based on empirical formula (13).

$$n = \sqrt{m + o} + c. \quad (13)$$

Finally, the characteristic variables required for the construction of neural network model include location factors and architectural characteristic factors. After quantification, the difference of characteristic variables will affect the training results of the model, so the maximum and minimum method of normalization function is used to process the characteristic variables, so that the distribution trend of variable data is transformed into a conventional function, and its formula is shown in formula (14).

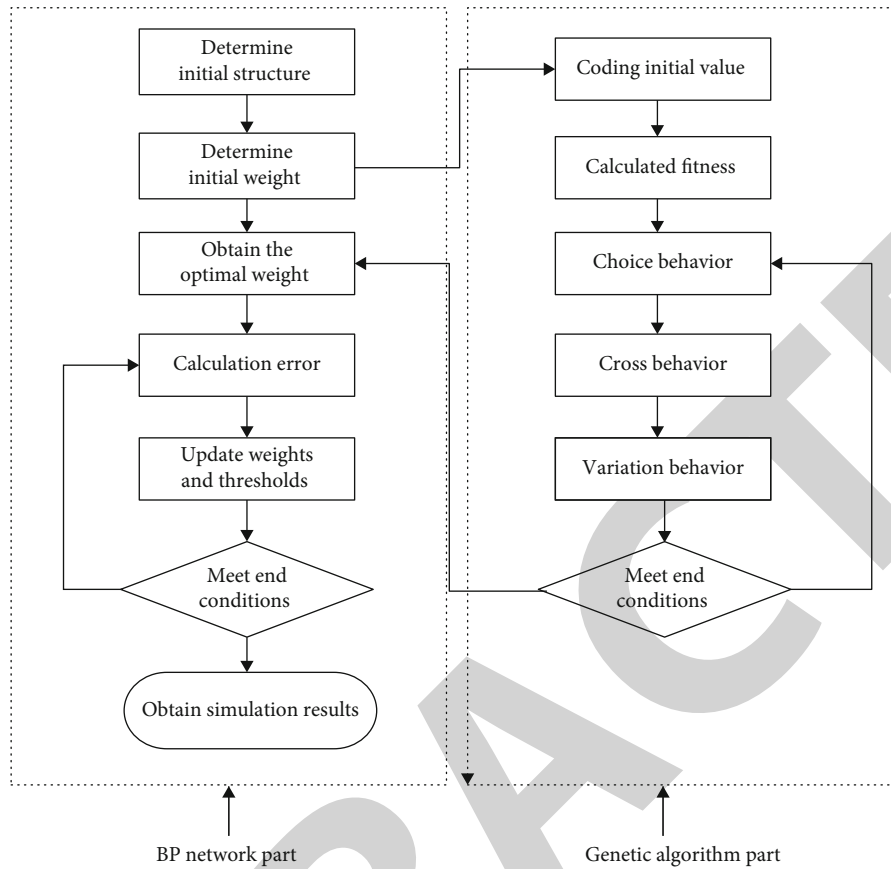


FIGURE 2: Optimization of weight flow of BP network by genetic algorithm.

TABLE 1: Characteristic variables affecting house price and variables' input processing.

Index	Independent variable	Input process	Quantification or unit
Azimuth index	Position	Distance from the city center	Km
	Traffic	Convenience grading evaluation ([1-5])	/
	Greening rate	No process	
	Plot ratio	No process	
	Number of households	No process	(Num)
	Property fee	No process	Yuan/m ² month
	Supporting facilities	Convenience grading evaluation ([1-5])	/
Building index	House type	Type number ([1-5])	/
	Measure of area	No process	m ²
	Orientation	Number by direction ([1, 2])	/
	Residential type	Type number ([1, 2])	/
	Renovation	Type number ([1-3])	/
	Building scale	Number by size ([1-5])	/

$$X_{\text{norm}} = 1 - \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}. \quad (14)$$

In equation (14), X is the characteristic variable data, and $X_{\text{max}}, X_{\text{min}}$ is the maximum and minimum value of the data. After the model is constructed, the mean square

error is used to evaluate the performance of the model, and its formula is shown in formula (15).

$$\begin{cases} MSE_t = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} \\ MSE_p = \frac{\sum_{j=1}^m (y_j - \hat{y}_j)^2}{m} \end{cases} \quad (15)$$

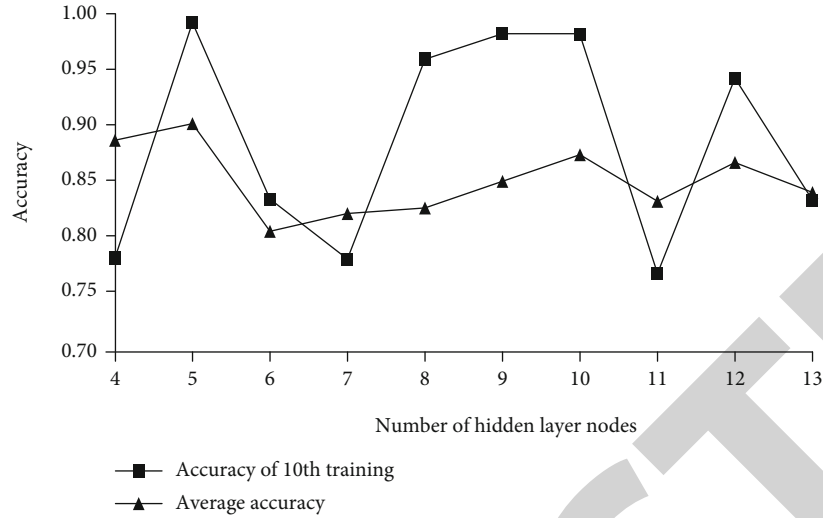


FIGURE 3: Prediction accuracy of nodes in different hidden layers.

TABLE 2: Accuracy of different fitness functions.

Fitness function	1	2	3	4	5	6	7	8	9	10	Average
Genetic algorithm optimization	0.985	0.996	0.975	0.790	0.756	0.706	0.952	0.963	0.728	0.986	0.884
Maximum objective function	0.881	0.742	0.889	0.767	0.985	0.723	0.910	0.870	0.899	0.778	0.844

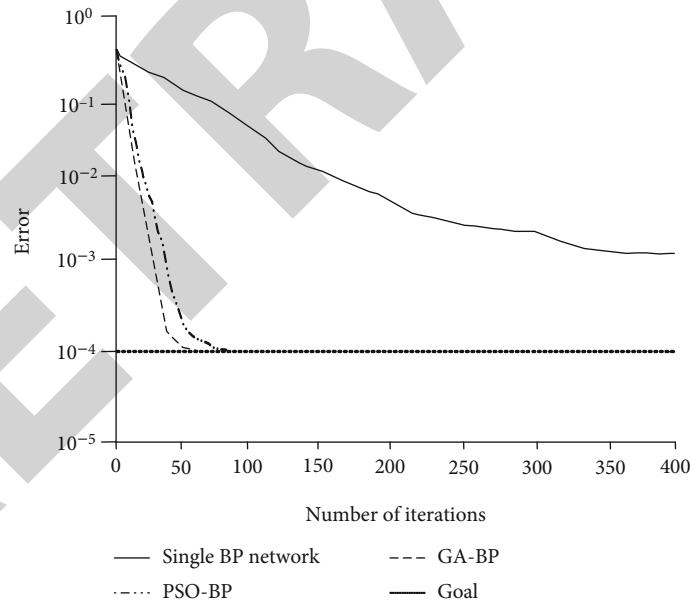


FIGURE 4: Comparison diagram of iteration times of algorithm model.

In equation (15), MSE_t and MSE_p represent the training mean square error and test mean square error of the model, respectively, m, n represents the number of samples tested and the number of samples trained, y_i, \hat{y}_i represents the real value and predicted value of the dependent variable data in the training set, and y_j, \hat{y}_j represents the real value and predicted value of the dependent variable data in the test set.

4. Training and Application Experiment of GA-BP Real Estate Valuation Model

4.1. GA-BP Model Training and Algorithm Optimization Verification. GA-BP algorithm is an iterative algorithm for symmetric diagonally dominant linear equations. It is a probabilistic reasoning algorithm based on recursive update, which has low computational complexity and high parallelism.

Because of these two properties, GA-BP algorithm is very suitable for dealing with large-scale sparse linear equations. GA-BP algorithm is different from classical iterative algorithm and Krylov subspace algorithm. GA-BP algorithm has good convergence for solving symmetric diagonally dominant linear equations. Based on the current market information, this paper makes assumptions on the relevant model market. The study randomly selected some samples and sets up many new samples, of which the number of test samples is 80. At the same time, 13 characteristic variables that have the greatest impact on house prices are extracted. The details of characteristic variables are shown in Table 1.

As shown in Table 1, the house type in the characteristic variable will be numbered from 1 to 5 in turn from one room and one living room, one room and two living rooms, and two living rooms, three rooms and two living rooms, four rooms and two living rooms; The residential types are divided into large flat (1) and ordinary residential (2). The decoration variables number the blank, simple decoration and hard decoration as 1-3 in turn. According to the above 13 characteristic variables, different numbers of hidden layer nodes are brought into the model. After 10 repeated training, the different accuracy of the model is shown in Figure 3.

The accuracy in Figure 3 is the proportion of the number of samples within 10% of the prediction error in the total number of training samples. It can be seen from the figure that after ten training for the same sample, when the number of hidden layer nodes is 5, the accuracy of the model is the highest, and the average accuracy and 10 repeated training accuracy are 0.901 and 0.992, respectively. At the same time, the effects of the maximum objective fitness function and the optimized fun function on the training model are compared, and the results are shown in Table 2.

It can be seen from Table 2 that after taking the traditional maximum objective function as the fitness function of the real estate valuation model, the average accuracy of 10 repeated training is 0.844, while the average accuracy of the optimized fitness function is 0.884. The experiment shows that after the optimization of the fitness function, through the above model, we can combine all the projects of real estate enterprises as one project. The land acquired in that year was also merged into a project. In this way, we can simulate the appropriate sales progress and construct a similar cash flow to estimate their net asset value (NAV) after the year. Then, get the net present value or valuation through an appropriate discount rate.

4.2. Research on Application Performance of GA-BP Real Estate Valuation Model. In order to calculate the performance of the real estate model, 100 samples were selected as the test target. The BP neural network optimized by GA-BP neural network model and particle swarm optimization algorithm is trained and tested. The maximum number of iterations is 400, the learning rate is 0.1, the hidden layer transfer function is logsig function, and the target accuracy is 0.0001. The error accuracy comparison of the three algorithm models is shown in Figure 4.

As shown in Figure 4, when the training samples are the same, the first to achieve the target error is the neural net-

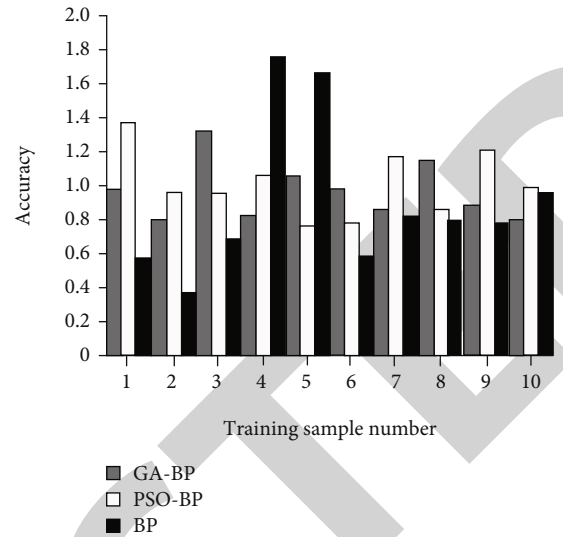


FIGURE 5: Training output error of different models.

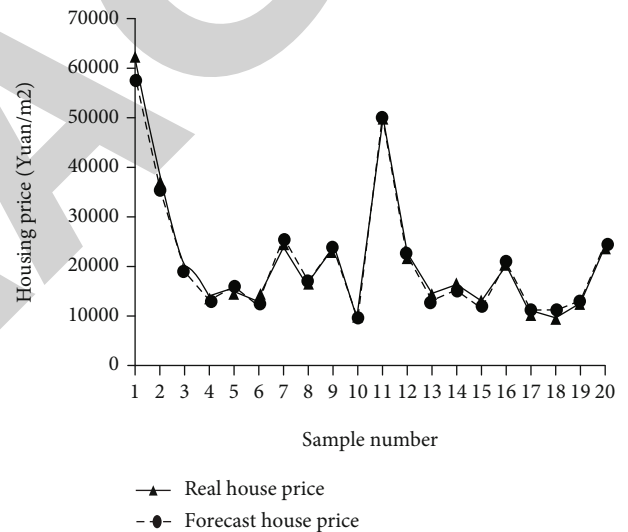


FIGURE 6: The error between the predicted house price and the actual value of the model.

work model after genetic optimization. The GA-BP curve tends to be stable when it is close to 50 iterations, and the convergence is completed only in 56 iterations; at the 84th iteration, it reaches the target value of 0.0001; finally, when the traditional BP neural network completes the maximum number of iterations of 400 times, it still fails to achieve convergence and falls into the result of local extremum. Experiments show that after the optimization, the anti-interference of BP neural network is improved, separated from the local extreme value, and the estimation efficiency of the model is also improved. However, the convergence times of GA-BP model are 28 times less than that of PSO-BP model, which proves that the calculation efficiency of Y genetic algorithm is higher. In the iterative training process of the three algorithms, 10 groups of data are randomly selected for comparison, as shown in Figure 5.

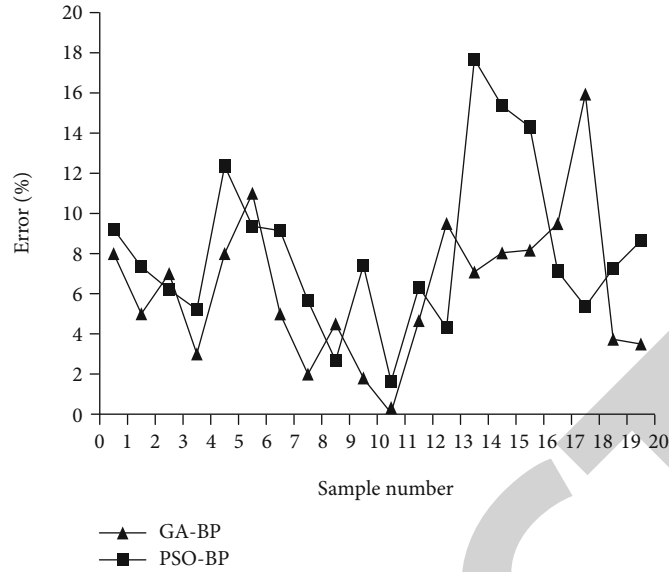


FIGURE 7: House price prediction error of PSO-BP and GA-BP.

The accuracy calculation formula in Figure 5 is the predicted house price/actual house price output by the model. The maximum ratio is 1.75 and the minimum ratio is 0.3; the training result of PSO-BP model is small, with the maximum ratio of 1.36 and the minimum ratio of 0.76. The predicted house price is closest to the actual house price, with the maximum ratio of 1.32 and the minimum ratio of 0.78. Experiments show that the accuracy of genetic algorithm is higher and higher than that of traditional BP model. After completing the training of the model and extracting the optimal initial value and threshold, number 20 test samples bring them into the trained GA-BP neural network system for simulation experiments to compare the actual house price with the error of the estimated house price of the model, as shown in Figure 6.

In Figure 6, the broken line graph of the actual house price is basically consistent with the broken line graph of the predicted house price. Among the 20 test samples, the sample No. 18 has the largest estimation error. The actual house price is 9517 yuan/m² and the estimated house price is 11034 yuan/m². Calculated according to the formula (actual house price - estimated house price)/actual house price, the prediction error is 16%. The actual house price of sample No. 6 is 13988 yuan/m², the model predicts that the house price is 12449 yuan/m², and the prediction error is 11%. The prediction error of other samples is less than 10%. The test sample No. 11 has the smallest error. The actual house price of the sample is 49969 yuan/m², and the model predicts that the house price is 50012 yuan/m², with an error of only 0.08%. If the prediction error of 10% is qualified, the prediction accuracy of the model is 90%. At the same time, the same 20 groups of test sample data are brought into the PSO-BP model, and the prediction errors of the two models are compared. The specific results are shown in Figure 7.

As can be seen from Figure 7, the overall error of particle swarm optimization model is greater than that of genetic

algorithm optimization model. Most of the prediction errors of GA-BP are in the range of 2%-8%, and the overall error of PSO-BP model is in the range of 6%-15%. The average error of GA-BP model is 6.27%. Compared with 8.12% of the average error of PSO-BP model, the relative error optimized by genetic algorithm is reduced by 1.85%. If the prediction error is less than 10% as the qualified accuracy, the prediction accuracy of GA-BP model is 90%, while the prediction error of PSO-BP model is more than 10% in four places, so the prediction accuracy is only 80%. From the experimental data, it can be concluded that the BP neural network model optimized by genetic algorithm is more neutral in real estate valuation. The advantages of the algorithm are not only reflected in higher prediction accuracy and stable prediction qualified rate but also reflected in less iterative training times of the model. Through MATLAB programming, the accuracy of GA-BP neural network, PSO-BP neural network, and traditional BP neural network is compared. In addition, through calculation, it is found that although the prediction accuracy of PSO-BP neural network has been improved, the mean square deviation of multiple runs is lower than that of GA-BP neural network, which can be considered that the model optimization has achieved good results.

5. Conclusion

This paper uses machine learning method to establish the relationship model between characteristic variables and house prices. From the perspective of hidden layer node optimization and fitness function optimization, the calculation results show that when the number of hidden layer nodes is 5, the prediction accuracy of the model is the highest. The model training experiment and model performance test simulation experiment are carried out successively by using MATLAB tools. The former is 4 percentage points higher than the latter, which proves the accuracy and performance of the real estate appraisal model. In the part of

model performance simulation experiment, the iteration times of the three algorithm models are compared. At the same time, the performance of fitness function and maximum objective function optimized by genetic algorithm is compared. It is found that GA-BP converges in 56 iterations, PSO-BP model needs 84 iterations, while traditional BP fails to achieve the target error in 400 iterations. The results show that the qualified rate of GA-BP model is 5% higher than that of PSO-BP model. Finally, the proportion of the predicted house price and the actual house price of the experimental calculation model shows that the proportion of GA-BP model is closer to 1 than that of traditional BP and PSO-BP models. However, this experiment has certain limitations. The number of experimental samples is only 100, which cannot cover all housing characteristic variables. The performance test shows that the prediction accuracy and calculation efficiency of GA-BP model in real estate appraisal are better than PSO-BP model.

Data Availability

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

Conflicts of Interest

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

References

- [1] S. Andersen, J. Marx, K. M. Nielsen, and L. Vesterlund, "Gender differences in negotiation: evidence from real estate transactions," *The Economic Journal*, vol. 131, no. 638, pp. 2304–2332, 2021.
- [2] D. C. Ling, Y. Lu, and S. Ray, "How do personal real estate transactions affect productivity and risk taking? Evidence from professional asset managers," *Real Estate Economics*, vol. 49, no. S1, pp. 7–40, 2021.
- [3] M. Mironiuc, E. Ionacu, M. C. Huian, and A. Țaran, "Reflecting the sustainability dimensions on the residential real estate prices," *Sustainability*, vol. 13, no. 5, article 2963, 2021.
- [4] N. Dobrovolskien, A. Pozniak, and M. Tvaronavien, "Assessment of the sustainability of a real estate project using multi-criteria decision making," *Sustainability*, vol. 13, no. 8, article 4352, 2021.
- [5] J. Hodoshima, "Evaluation of performance of stock and real estate investment trust markets in Japan," *Empirical Economics*, vol. 61, no. 1, pp. 101–120, 2021.
- [6] C. David, S. Christophe, and A. S. Ev, "The rate of return on real estate: long-run micro-level evidence," *The Review of Financial Studies*, vol. 34, no. 8, pp. 3572–3607, 2021.
- [7] L. D'Acci, "Quality of urban area, distance from city centre, and housing value. Case study on real estate values in Turin," *Cities*, vol. 91, pp. 71–92, 2019.
- [8] M. Ulbl, M. Verbič, A. Lisec, and M. Pahor, "Proposal of real estate mass valuation in Slovenia based on generalised additive modelling approach," *Geodetski Vestnik*, vol. 65, no. 1, pp. 46–81, 2021.
- [9] D. Wang and V. J. Li, "Mass appraisal models of real estate in the 21st century: a systematic literature review," *Sustainability*, vol. 11, no. 24, pp. 7006–7014, 2019.
- [10] X. He and P. Gong, "A radial basis function-generated finite difference method to evaluate real estate index options," *Computational Economics*, vol. 55, no. 3, pp. 999–1019, 2020.
- [11] S. Zrobek, O. Kovalyshyn, M. Renigier-Bilozor, S. Kovalyshyn, and O. Kovalyshyn, "Fuzzy logic method of valuation supporting sustainable development of the agricultural land market," *Sustainable Development*, vol. 28, no. 5, pp. 1094–1105, 2020.
- [12] L. Chen, W. Lin, P. Chen, S. Jiang, L. Liu, and H. Hu, "Porosity prediction from well logs using back propagation neural network optimized by genetic algorithm in one heterogeneous oil reservoirs of Ordos Basin, China," *Journal of Earth Science*, vol. 32, no. 4, pp. 828–838, 2021.
- [13] L. Wang, Y. Guo, B. Dalip, Y. Xiao, R. D. Urman, and Y. Lin, "An experimental study of objective pain measurement using pupillary response based on genetic algorithm and artificial neural network," *Applied Intelligence*, vol. 52, no. 2, pp. 1145–1156, 2022.
- [14] Y. Xie, Y. Ishida, J. Hu, and A. Mochida, "A backpropagation neural network improved by a genetic algorithm for predicting the mean radiant temperature around buildings within the long-term period of the near future," *Building Simulation*, vol. 15, no. 3, pp. 473–492, 2022.
- [15] C. S. Jiang and G. Q. Liang, "Modeling shear strength of medium- to ultra-high-strength concrete beams with stirrups using SVR and genetic algorithm," *Soft Computing*, vol. 25, no. 16, pp. 10661–10675, 2021.
- [16] Y. Cao, X. Fan, Y. Guo, S. Li, and H. Huang, "Multi-objective optimization of injection-molded plastic parts using entropy weight, random forest, and genetic algorithm methods," *Journal of Polymer Engineering*, vol. 40, no. 4, pp. 360–371, 2020.
- [17] W. Zhang, F. Wang, and N. Li, "Prediction model of carbon-containing pellet reduction metallization ratio using neural network and genetic algorithm," *ISIJ International*, vol. 61, no. 6, pp. 1915–1926, 2021.
- [18] Y. Ning, S. Niu, Y. Wang, J. Zhao, and C. Lu, "Sono-modified halloysite nanotube with NaAlO₂ as novel heterogeneous catalyst for biodiesel production: optimization via GA_BP neural network," *Renewable Energy*, vol. 175, pp. 391–404, 2021.
- [19] Y. Wang and P. Fu, "Integration performance statistics of green suppliers based on fuzzy mathematics and BP neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 2, pp. 2083–2094, 2021.
- [20] W. Zhu, H. Wang, and X. Zhang, "Synergy evaluation model of container multimodal transport based on BP neural network," *Neural Computing and Applications*, vol. 33, no. 9, pp. 4087–4095, 2021.