Enhanced Bank Marketing System Using Optimized GSA-BP Neural Network for Improved Customer Targeting

Ziyi Jiao
Henan Polytechnic Institute Nanyang, Henan, China
E-mail: m15237711330@163.com

Keywords: gravity search algorithm, BP neural network, bank, marketing system, integrated learning algorithm

Received: October 20, 2024

*To more accurately predict the needs of bank customers and improve the marketing success rate of bank wealth management products. Study designed a bank marketing system that can mine potential customers by analyzing customer data. The discovery of potential customers is mainly achieved by constructing the BA GSA-BP model, which optimizes the structure of the backpropagation neural network through gravity algorithm and integrates the optimized neural network using Bagging set learning algorithm. The study conducted simulation tests on the performance of the model. In the Bank Marketing dataset, after 225 iterations, the MES of GSA-BP remained stable at 0.098; In the Credit Card Default dataset, after 164 iterations, the MES of GSA-BP remained stable at 0.093, In terms of accuracy and recall, GSA-BP also performs well. For example, in the PR curve of the Bank Marketing dataset, the recall rate is about 0.896 when the accuracy is 0.8, and about 0.800 when the accuracy is 0.9. After using the Bagging ensemble algorithm, the performance of GSA-BP was further improved, with the largest area under the ROC curve, indicating strong generalization ability and low risk of overfitting. When the number of users is 200, the response time of the research system is 0.24 seconds, and the prediction success rate is about 60%. The above data illustrates that the method used in the study improves the over-fitting problem of traditional BP neural network and enhances the generalization ability of the algorithm. The constructed bank marketing system has the characteristics of short response time and high prediction success rate. The results of this study can accurately predict whether customers will order bank promotional products, conduct targeted marketing for customers, reduce marketing risks, and improve the success rat*e of *product marketing and provide reliable technical support for predictions in other industries.*

Povzetek: Predlagan je izboljšan bančni marketinški sistem z optimizirano nevronsko mrežo GSA-BP, ki učinkovito napoveduje potencialne stranke, skrajša odzivni čas in izboljša uspešnost ciljnega trženja produktov.

1 Introduction

In the context of the Internet era, the main challenge facing the marketing of commercial activities of banks is how to accurately locate customer needs, improve the success rate of marketing activities, and thus enhance the competitiveness of commercial banks [1-2]. Therefore, how to effectively identify and attract potential customers and design efficient marketing strategies to improve marketing resources has become the primary issue that banks need to address. Traditional marketing methods rely on manual planning and market research, but have limitations in efficiency and accuracy [3]. Artificial planning and market research rely on manual collection and analysis of data, with slow processing speeds that make it difficult to respond to rapidly changing market demands. According to industry reports, the conversion rate of this method is generally below 10%, and the customer churn rate is relatively high. The strategy of telephone direct sales requires a large amount of human resources and is costly. And frequent sales calls can easily trigger customer resentment and resistance, leading to a decrease in customer satisfaction. According to customer surveys, over 60% of customers express dissatisfaction with telephone direct sales, and the marketing success rate of this method is usually less than 20%. With the development of modern heuristic algorithms, how to apply these algorithms to improve the effectiveness of bank marketing systems has become a hot research topic. The Gravitational Search Method (GSA), as an optimization algorithm, is widely used to solve multi-dimensional data optimization problems [4]. At the same time, the backpropagation (BP) neural network serves as a powerful predictive model that can learn complex relationships in data. However, BP neural network has the problems of easily falling into local optimal solutions and slow training speed, which limits its effectiveness in practical applications. By simulating the gravitational interaction between objects, GSA can perform global search in the entire search space, effectively avoiding the problem of BP neural network getting stuck in local optimal solutions. After finding the potential global optimal solution, GSA can further improve the accuracy of the model by adjusting the search strategy and conducting fine searches within the local range. The optimization capability of GSA can compensate for the shortcomings of

BP neural network in training speed and accuracy, while the powerful predictive ability of BP neural network can fully utilize the optimized model parameters of GSA to achieve more efficient potential customer identification and personalized marketing. Compared to genetic algorithms, grey wolf algorithms, etc., the universal gravitation algorithm is logically easier to understand and interpret, and has fewer parameter adjustments, resulting in faster convergence speed. Therefore, the research uses the global search ability and better local search ability of the gravity search algorithm to optimize the training process of BP neural network, thereby optimizing its network structure. Therefore, the gravity search method is combined with the BP neural network in order to improve the effect of potential customer identification and personalized marketing in the bank marketing system. The first section of the research mainly talks about the research of sales promotion system by many experts and scholars and the application of neural network and gravity search algorithm. The second section mainly describes the construction process of the bank marketing system prediction model based on the BA-GSA-BP algorithm. The third section mainly focuses on verifying the effectiveness of the algorithm used in the study and the system constructed. Section 4 mainly discusses the results of the proposed model and the SOTA method reviewed in related work. Section 5 mainly analyzes the simulation results and points out the shortcomings of the research.

2 Related works

In the Internet era, more insightful marketing methods are needed to accurately locate customer needs, improve the success rate of product marketing of commercial activities of banks, and enhance the competitiveness of commercial banks. Therefore, the accuracy of bank promotion systems in predicting intended customers is crucial. To solve this problem, many experts and scholars have conducted relevant research on promotion systems in a general sense. Arun artificially analyzed bank data information and proposed a classifier model using the data mining tool WEKA. The model improves the classification model through data mining analysis and parameter changes. The results show that the model has better classification results when analyzing binary data sets [5]. Hosseini proposed the BN decision support system in order to predict the success rate of bank long-term deposit telemarketing. The system network is treated as an output variable, and the Nave Bayes method is used to determine the causal relationship between customer attributes and outcomes. The results show that the adopted system can predict subscribers with specific characteristics [6]. Gao proposed a time series based on KNN and SVR to predict the cyclical trend of consumer behavior in order to predict the time when users will re-purchase the product next time. Estimate the user's current life stage based on the user's recent behavior and recommend new products suitable for the user's estimated life stage. Experimental results show that this method significantly improves the effectiveness of recommendations [7]. In order to plan the customer purchase journey and develop decision support capabilities, Ma and Sun have developed a unified conceptual framework and multi-faceted research methods. These methods extract insights from large-scale unstructured, tracking and network data, using them in a transparent manner for descriptive, causal and prescriptive analysis. The results show that this method can plan the customer's buying journey [8]. Liu built a precision marketing model based on neural networks to address multi-faceted issues such as environmental pollution of marketing data and unscientific algorithms, and focused on data mining to study the two most important factors affecting marketing revenue. Research results show that precision marketing methods based on big data information platforms need to be more detailed and comprehensive [9]. From the above information, it can be found that machine learning models have high reliability, security, accuracy, and efficiency in processing data information, and have good performance in marketing issues. Therefore, machine learning technology can be applied to banking systems to obtain effective information from a large amount of raw data, extract knowledge from the data, and use this knowledge to design strategic plans, thereby improving the market competitiveness of banks.

With the development of modern heuristic algorithms, how to use these algorithms to improve the effectiveness of prediction models has become a hot research topic. Geng et al. proposed a SOC estimation method based on artificial neural networks and k-fold cross-validation in order to solve the main limitation of neural networks, which requires a large amount of data. The results show that compared with traditional BP neural network prediction, this method has a prediction accuracy of 91% [10]. In order to develop a wind speed prediction model, Mohammed et al. proposed a method that integrates deep learning and neural networks. Mapping wind speeds and predicting future propagation using recurrent neural networks. The results show that this method can accurately predict wind speed changes in Erbil City [11]. Ramos, F. R et al. proposed using a neural network with long-term memory to model the dynamic changes in sequence data when constructing a predictive time series model. The results show that long-term memory can significantly improve prediction performance in a time series [12]. In order to select the minimum number of features without affecting the classification accuracy, Yadav et al. proposed a method of merging fuzzy logic and gravity search algorithms. Use this method to explore feature selection in automatic disease prediction. The results show that this method not only reduces the feature set by 67.66% on average, but also improves the accuracy by 12% on average [13]. Tochaiwat et al. developed a sales rate prediction model using artificial neural network technology to study the monthly sales unit quantity of independent residential projects. The results showed that the root mean square error (RMSE) of the model was ± 6.296, and the slope and R2 values of the linear regression analysis between the prediction rate and the actual rate were 0.620 and 0.571, respectively. This confirms the superiority of artificial neural network technology in dealing with limited data problems [14]. Compare the tree model with the research model, and the results are shown in Table 1.

Table 1: Model performance comparison

Researcher	Algorithm model	Test dataset	Accuracy	Convergence iteration times	Predict success rate
Geng et al. [10]	SOC estimation method based on artificial neural network and k-fold cross validation	NASA Battery Dataset	91%	324	92%
Mohammed et al. [11]	A model based on recursive neural network	Wind speed test dataset	90%	354	93%
Ramos, F.R et al. [12]	Prediction model based on long-term memory neural network	Climate Shared Dataset	92%	356	90%
Yadav et al. [13]	Integrating Fuzzy Logic and Gravity Search Algorithm Model	Disease dataset	91%	265	92%
Tochaiwat et al. [14]	Prediction Model Based on Artificial Neural Network Technology	Residential project sales dataset	93%	312	94%
This study	A model combining gravity search method and neural network	Bank Marketing Dataset	98%	225	98%

From the research of scholars both inside and outside China in Table 1, it can be seen that, neural network algorithms and gravity algorithms are currently widely used in many fields and it has achieved good results, but if it is directly applied to existing bank prediction systems, the prediction accuracy and model iteration times both need to be improved. This is because although Geng et al.'s model uses artificial neural networks and k-fold cross validation to improve the accuracy of SOC estimation, it is limited by the structure and parameter settings of the neural network, making it difficult to achieve higher accuracy when processing complex bank datasets. Mohammed et al.'s recursive neural network model performs well in processing time series data, but the recursive structure makes the model more prone to gradient vanishing or exploding during training, which affects the convergence speed and prediction performance of the model. Although Ramos et al.'s long-term memory neural network model can capture the long-term dependencies of time series, its high complexity can make it difficult to converge quickly in situations where computing resources are limited. The model developed by Yadav et al., which combines fuzzy logic and gravity search algorithm, has advantages in dealing with uncertainty problems. However, the rule making of fuzzy logic and parameter adjustment of gravity search algorithm may be difficult and prone to getting stuck in local optimal solutions. Tochaiwat et al.'s model achieved high accuracy in predicting sales of residential projects, but may have relied too heavily on the features of specific datasets, leading to a decrease in performance when generalizing to other datasets. In view of this, the study will combine the predictive needs of the banking system with the advantages of machine algorithms proposes a bank marketing system prediction model based on the gravity search method combined with neural networks. It is expected that it can mine the purchase intention of potential users and provide more accurate prediction results and decision support for bank marketing decisions.

3 Construction of bank marketing system prediction model based on BA-GSA-BP algorithm

This section mainly designs a bank marketing system for marketing of bank-related products. It mainly consists of four major modules: system management, marketing management, project management, and personal leave management. And respectively elaborate on the functions and implementation of these four models. This marketing system needs to use algorithms to predict potential intended customers. In view of the shortcomings of the traditional back propagation (BP) neural network, it is proposed to use the Gravity Algorithm (GSA) and Bagging

(BA) integrated algorithms to improve it, and elaborates on the detailed improvement method.

3.1 Construction of bank marketing system

The banking industry is a highly competitive industry, with various banks competing for limited customer resources. In order to maintain a competitive advantage in the market, banks need to attract and retain customers through effective marketing strategies and provide personalized products and services [15]. Building a bank marketing system can use advanced algorithm technology and data analysis capabilities to achieve better customer management and marketing effects. The overall architecture of the bank marketing system studied and designed is shown in Figure 1.

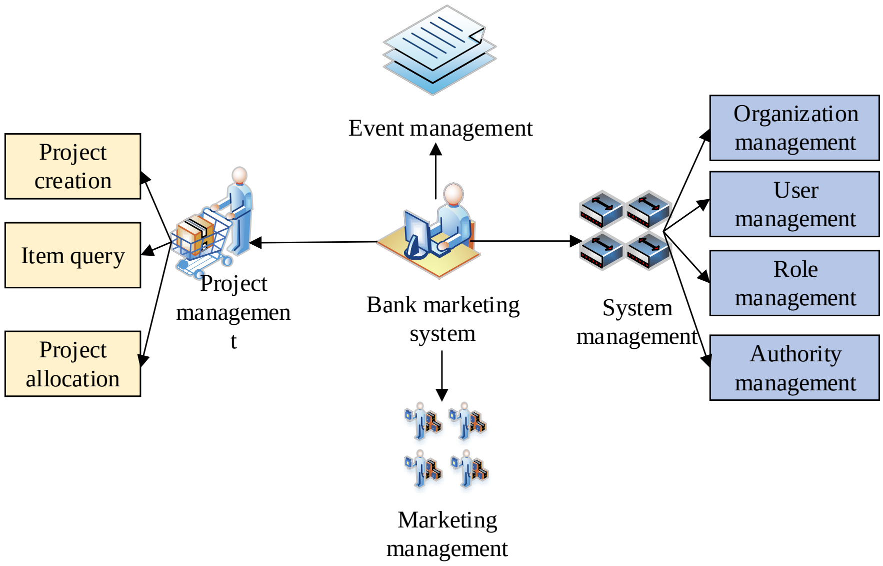


Figure 1: Overall architecture diagram of bank marketing system

As shown in Figure 1, the bank's marketing system mainly consists of four major modules: system management, marketing management, project management, and personal leave management. The system management module is composed of four parts: organization management, user management, role management, and authority management. The implementation of the functions of each module of the system is based on the Springboot framework. The Springboot framework organizes the collaborative work of each module in the system by separating interface display, data and business logic. It has the advantages of low coupling, high reusability, fast construction, low life cycle, and easy maintenance. Institutional management allows bank administrators to create, query, delete and modify institutional information. The system institution management process is shown in Figure 2. The administrator interacts with the service class layer through the operation interface, passes instructions to the business class layer to perform specific business operations, and returns the execution results to the user interface for display. And the administrator can create a new institution, enter the basic information of the institution and add it to the database; can understand the details of the banking institution by accurately querying or browsing all institution information; can delete institution information that is no longer used; and can also modify parts of the institution or All information will be modified to maintain the accuracy and timeliness of the information.

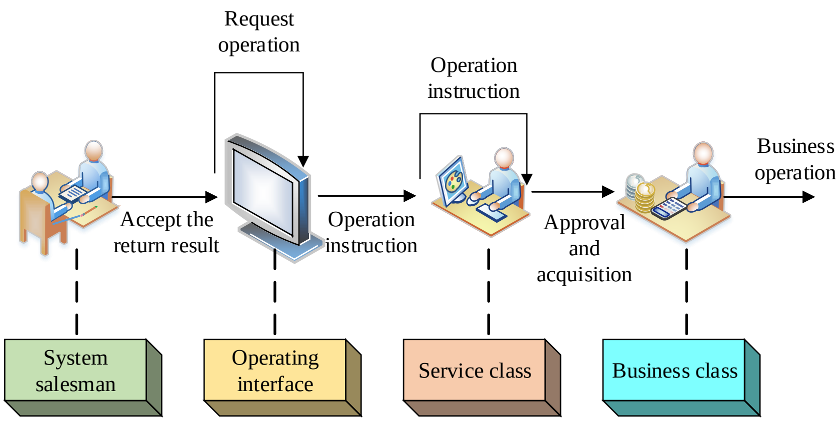


Figure 2: Organization management flow chart

A functional module used by system administrators to manage and maintain basic information about system users. It provides nine operations such as user query, user information export, user deletion, user deactivation, user activation, new user, modification of user information, reset password and user role configuration. Through these operations, system administrators can easily manage and adjust the status, permissions, and information of system users to ensure the information security and compliance of system users. Role management can achieve management and control of users with different permissions by assigning different "roles" to customers, including querying, deleting, modifying role information, etc. Permission management is mainly implemented using the Role-Based Access Control (RBAC) model [16]. The model divides users into three roles, namely account manager, account executive and system administrator. Different customer roles enjoy Different operating permissions. The project management module is mainly responsible for marketing project information in the entire bank marketing system, and mainly includes three functions: project creation, project query, and project allocation. The basic process of realizing marketing project management is shown in Figure 3.

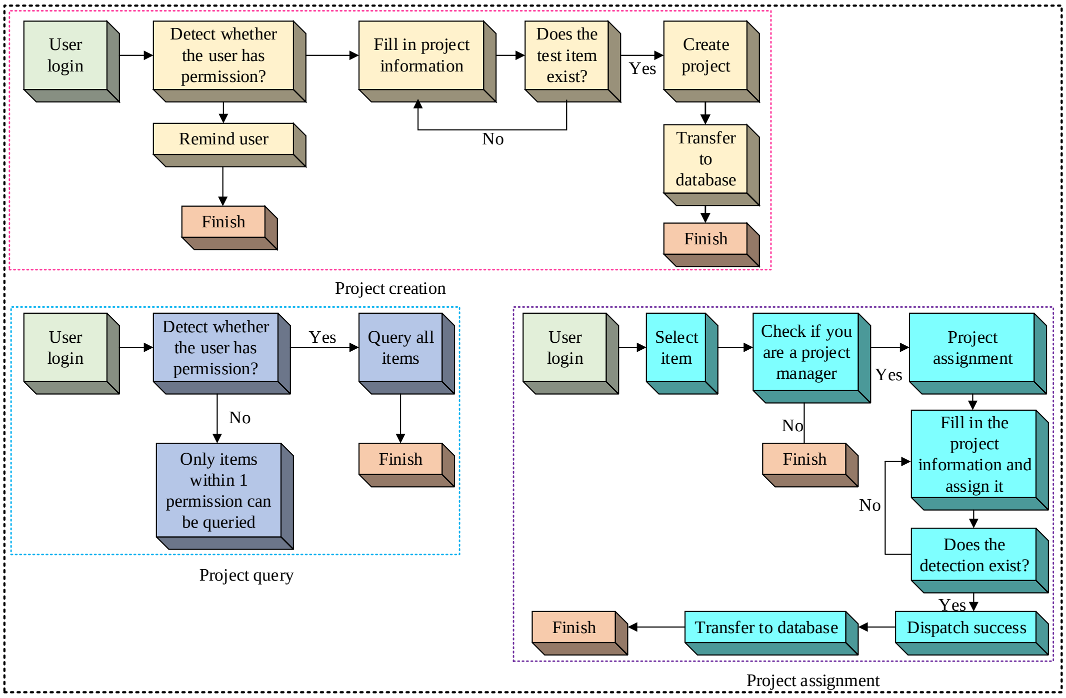


Figure 3: Basic process of marketing project management

The marketing management module is mainly responsible for helping bank business personnel complete more marketing activities, including three functions: task list query, marketing result prediction and customer marketing follow-up. After logging into the system, bank clerks can query the current task list and select a sub-project to predict whether customers are willing to order products. The system will display customer information for which the prediction result is "yes", and employees will conduct marketing promotions and update customer records. In addition, employees can also select sub-projects for customer marketing follow-up, query customer information and follow-up work, and update records. The overall process is concise and clear, allowing employees to easily complete marketing management tasks. In this process, the most important thing is marketing result prediction, and its function sequence diagram is shown in Figure 4.

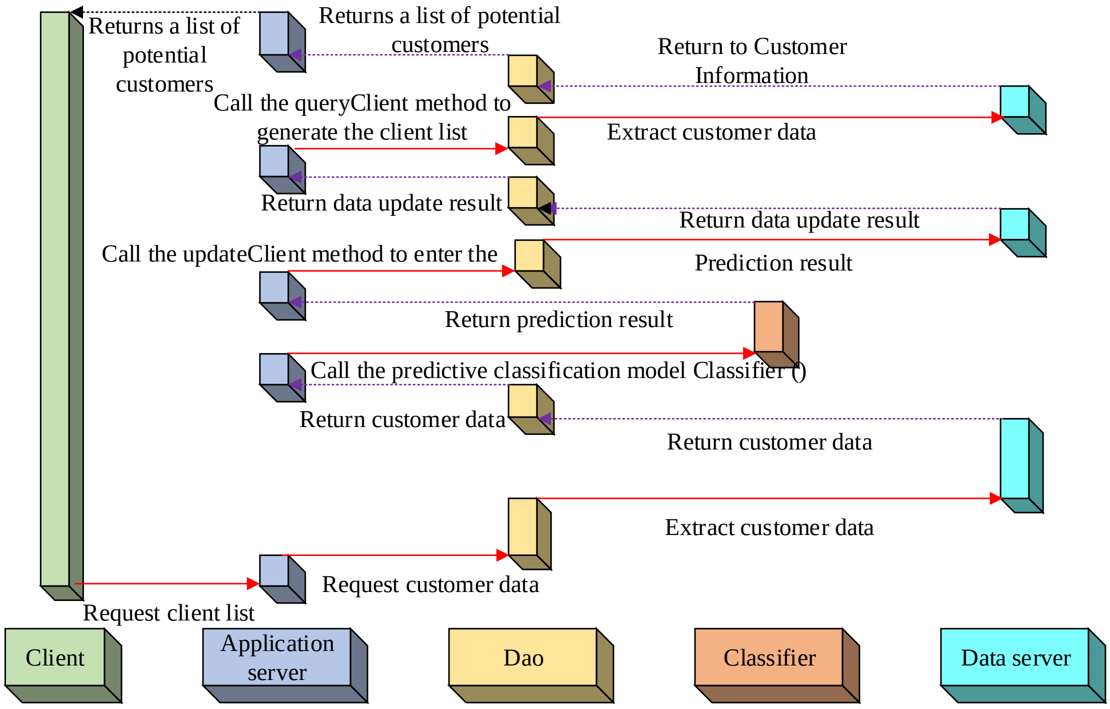


Figure 4: Marketing result prediction function sequence diagram

3.2 Construction of prediction model integrating BA, GSA and BP neural network

For the bank marketing system constructed by the research institute, the BP neural network algorithm can be used to select the main marketing objects. However, the traditional BP neural network model has many problems such as slow convergence speed and uncertain network structure [17]. Therefore, the study used GSA to optimize the BP neural network structure, and selected the Bagging integration algorithm to integrate the optimized BP neural network [18]. The main purpose of GSA to optimize the BP neural network is to use GSA's powerful global search capability and convergence speed to calculate the optimal structure of the hidden layer of the BP neural network. The specific process of GSA optimizing BP neural network is shown in Figure 5.

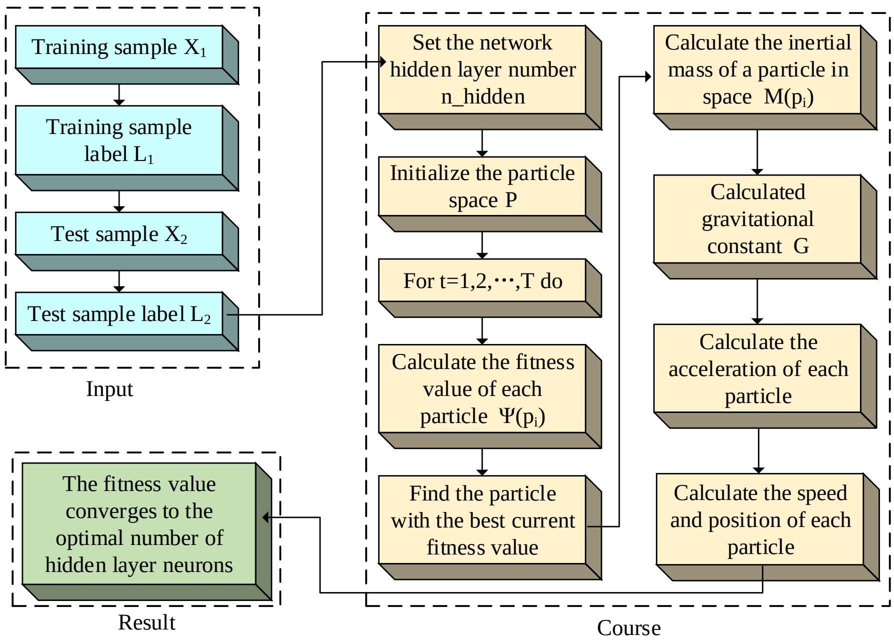


Figure 5: GSA-BP algorithm flow chart

In Figure 5, the particle space is initialized, see Formula (1).

$$\begin{cases} P=\{p_1,p_2,...p_N\} \\ p_i=\left(p_i^{1},p_i^{2},...p_i^{m}\right), i=1,2,...,N \end{cases} \quad (1)$$

In Formula (1), P is the initial particle space; i is a particle; P_i is the initial i position of the particle in this space, and the initial position of the particle is one of the options for the neural hidden layer of the network. Then calculate the fitness value of the particles, see Formula (2).

$$\psi(p_i)=\frac{1}{n}\sum_{j=1}^{n}\left(\hat{y}_j-y_j\right)^2 \quad (2)$$

In Formula (2), $\psi(p_i)$ is p_i the fitness value of the particle, which is also the mean square error of the neural network; j is a particle; $\hat{y}_j$ is the prediction result of the neural network; y_j is the actual label; n is the number of samples. Then find the optimal fitness of the particles, see Formula (3).

$$\begin{cases} best\_P=\max\left(\psi(p_i)\right), i=1,2,...,N \\ worst\_P=\min\left(\psi(p_i)\right), i=1,2,...,N \end{cases} \quad (3)$$

In Formula (3), $best\_P$, it is the particle with the best fitness in the current iteration; $worst\_P$ it is the particle with the worst fitness in the current iteration. Calculate the inertial mass of the particle according to Formula (3), see Formula (4).

$$\begin{cases} m(p_i)=\dfrac{\psi(p_i)-worst\_P}{best\_P-worst\_P} \\ M(p_i)=\dfrac{m(p_i)}{\sum_{i=1}^{N}m(p_i)} \end{cases} \quad (4)$$

In Formula (4), $M(p_i)$ is P the inertial mass of all particles in space. Then calculate the gravitational constant G. The calculation method is shown in Formula (5)

$$G=G_0\left(-\alpha\frac{t}{T}\right) \quad (5)$$

In Formula (5), G_0 is the gravitational constant value in the traditional sense, α which is a parameter, t is the current number of iterations, T is the maximum number of iterations. Then, based on the traditional law of universal gravitation, When iterating t times, The force exerted by particle i on particle j in the n dimension, as shown in Formula (6).

$$\begin{cases} F_{ij}^{n}=G(t)\dfrac{M_a(t)M_p(p_i)}{R_{ij}(t)-\varepsilon}\left(p_j^n(t)-p_i^n(t)\right) \\ R_{ij}(t)=\left\|p_i(t),p_j(t)\right\|_2 \end{cases} \quad (6)$$

In Formula (6), F_{ij}^{n} is the acting force , M_a is p_j the active gravitational mass of the particle; M_p is p_i the passive gravitational mass of the particle; ε is a constant; $R_{ij}(t)$ is t the distance between p_i and p_j at the first iteration. p_i Throughout the space, there is p_i more than one effect p_j. Therefore, assuming that p_i the particles that act on the particle are a set kB, t the resultant force on $F_i^n(t)$ the particle in the first iteration is i shown kB in Formula (7).

$$F_i^n(t)=\sum_{j\in kB, j\neq i} rand_j F_{ij}^{d}(t) \quad (7)$$

In Formula (7), $rand_j$ it is the random variable in [0,1]. The particle acceleration can be obtained from the resultant force $a_i^n(t)$, see Formula (8).

$$a_t^n(t)=\frac{F_i^n(t)}{M(p_i)} \quad (8)$$

Finally, the velocity of the particle is updated according to the acceleration in $v_i^n(t+1)$ Formula (6) , see Formula (9).

$$v_i^n(t+1)=rand_i\times v_i^n(t)+a_i^n(t) \quad (9)$$

The position of the particle can be calculated according to the particle speed in $x_i^n(t+1)$ Formula (7) , see Formula (10)

$$x_i^n(t+1)=x_i^n(t)+v_i^n(t+1) \quad (10)$$

Repeat the steps from Formula (2) to Formula (10) until the algorithm reaches convergence. In the GSA-BP model, fitness values are used to evaluate the strengths and weaknesses of individuals (i.e. the parameter configuration of the BP neural network). As the number of iterations increases, the GSA algorithm continuously searches and optimizes to gradually approach the optimal parameter configuration of the BP neural network, resulting in a gradual decrease in fitness values. When the fitness value reaches the preset error threshold or the number of iterations reaches the upper limit, the algorithm stops iterating and outputs the optimal solution, that is, calculates the optimal structure of the hidden layer of the BP neural network.

The bagging algorithm obtains multiple different sampling sets through sampling with replacement, and then trains a base learner for each sampling set, and improves the generalization ability by integrating the output of these learners. For unstable learning models, such as BP neural networks, the Bagging algorithm can reduce the risk of overfitting. Therefore, the bagging algorithm is used to integrate the GSA-optimized BP neural network to further improve the model performance. The integration process is detailed in Figure 6.

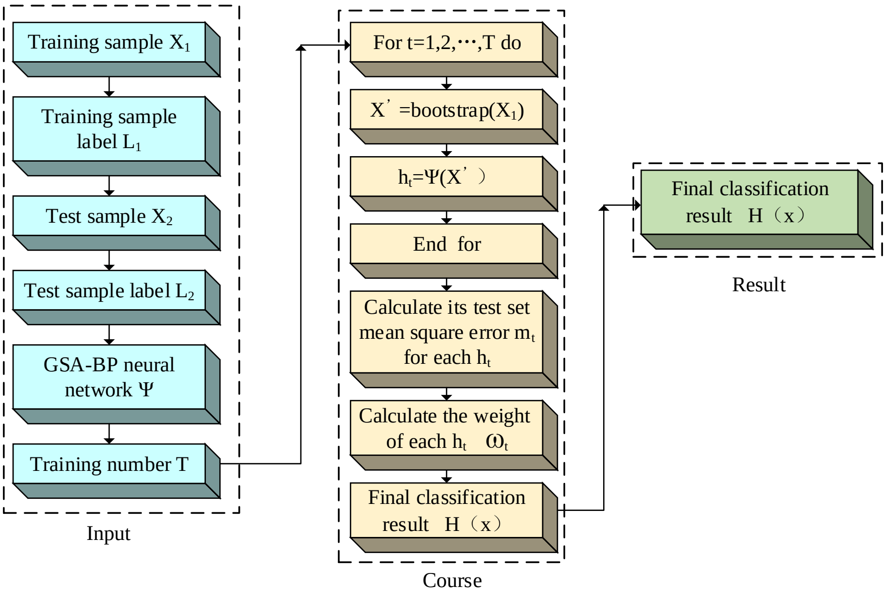


Figure 6: BA -GSA-BP algorithm flow chart

In the algorithm flow of Figure 6, first set h_t the number of base classifiers to T ; then for each base classifier h_t ; sample from the original training set with replacement, and extract the same number of samples as the original training set samples to form A new training set subset is used; then the training set subset is used to train the base classifier h_t ; and finally the base classifier is combined using a combination strategy of the weighted voting method. T The weight of each classifier is shown in h_t Formula (11).

$$\begin{cases} m_t = \dfrac{n}{\sum_n^{i=1}\left(h_t\left(X_2^i\right) - L_2^i\right)}, t = 1,2,3,\ldots,n \\ \omega_t = \dfrac{m_t}{\sum_T^{t=1} m_t} \end{cases} \quad (11)$$

In Formula (11), m_t it is the mean square error; ω_t it is the weight. The classification result of weighted voting is shown in Equation (12).

$$H(x) = c_{\arg\max} \sum_{t=1}^{T} \omega_t h_i^j(x) \quad (12)$$

In Formula (12) $h_i^j(x)$ is the output $h_i(x)$ on the category label c_j . The study used mean square error (MSE) and ROC curve to evaluate the constructed prediction model . The calculation method of mean square error MSE is shown in Formula (13).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2 \quad (13)$$

The abscissa in the ROC curve is FPR, and the ordinate is TPR. The calculation method is shown in Formula (14).

$$\begin{cases} FPR = \dfrac{FP}{TN+FP} \\ TPR = \dfrac{TP}{TP+FN} \end{cases} \quad (14)$$

4 Performance analysis and simulation application of bank marketing system prediction model based on BA- GSA-BP algorithm

This section mainly focuses on verifying the validity of the method used in the study. First, the selection of the simulation experiment environment and data set is explained, and then comparative experiments are conducted on the fitness value of GSA, the PR curve of GSA-BP, the ROC curve of BA-GSA-BP, the response time of the marketing system, and the prediction success rate. The data generated were analyzed and the superiority of the method used in the study was demonstrated.

4.1 Performance analysis of bank marketing system prediction model

In order to verify the effectiveness of the bank marketing system prediction model based on the GSA-BP algorithm, the study first tested the performance of the constructed system. The basic hardware environment settings of the experiment are shown in Table 2.

Table 2: Software and hardware configuration of experimental environment and model parameter settings

Project	Configuration item	Parameter
Software and hardware configuration of experimental environment	operating system	Windows 10
	CUP	i9-11900K
	Memory	GDDR6X
	Disk	Samsung 980 PRO
	network adapter	Intel Ethernet Converged Network Adapter X710
	underlying database	MySQL5.7
Model parameter settings	Population Size	30
	Iteration Number	500
	Gravitationa Constant	1.1
	Learning Rate	0.001

This study first selected the improved genetic algorithm (IGA) from reference [19] and the improved simulated annealing algorithm (ISA) from reference [20]. with the same experimental experience to compare the performance with the Gravity Search Algorithm (GSA) used in the study. In order to ensure the fairness and rationality of the experiment, the number of iterations of each algorithm was set to 500. Then the Bank Marketing data set and the Credit Card Default data set were selected to test the performance indicators of the model. The Bank Marketing Data Set includes a total of 324 pieces of customer basic information (customer ID, age, occupation, marital status, etc.), 296 pieces of deposit and loan information (housing loans, personal loans, etc.), 568 pieces of bank marketing information for customers (contact frequency, previous activity results, call duration, etc.), and 154 pieces of customer subscription information. The Credit Card Default Data Set includes a total of 1324 pieces of customer basic information, credit data, and default information. For the information of two datasets, first check and process the missing values, outliers, and duplicate values in the data, and then use the mean normalization method to standardize the data. Afterwards, the model will be trained and tested in a ratio of 4:1 between the training set and the testing set. The fitness changes of the algorithm are shown in Figure 7.

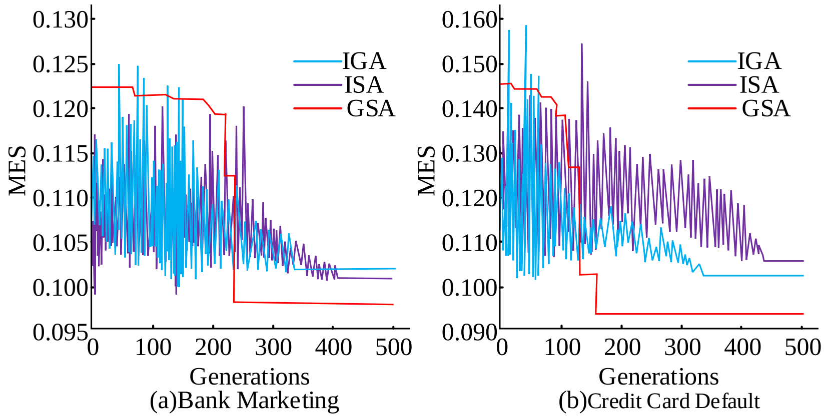


Figure 7: Comparison of fitness values of different data sets

As shown in Figure 7, Figure 7(a) shows the fitness value convergence process of the three algorithms in the Bank Marketing data set. It can be seen that the convergence curves of the two algorithms IGA and ISA fluctuate greatly at the beginning, and there is great instability. Then as the number of iterations increases, the curve fluctuations become smaller and smaller. In the end, IGA iterated 336 times, and its MES value became stable, about 0.103; ISA iterated 406 times, and its MES value became stable, about 0.101. In contrast, the convergence curve of GSA is smoother and declines in a ladder-like manner. At the final iteration of 225 times, the MES value stabilizes at about 0.098. Figure 7(b) shows the fitness value convergence process of the three algorithms in the Credit Card Default data set. It can be seen that the curve change trend is similar to Figure 7(a). The curves of IGA and ISA fluctuate greatly and are unstable, with 341 and 443 iterations respectively. The MES values tend to be stable, with stable values of approximately 0.101 and 0.107 respectively. The GSA convergence curve still declines in a relatively stable step-like manner, and its MES value tends to be stable at about 0.093 when iterating 164 times. It shows that compared with the other two algorithms, GSA has faster convergence speed and better fitness value. This is because GSA searches the solution space of the problem by simulating the gravitational interaction between celestial bodies. This mechanism enables GSA to explore the whole search space more effectively. Once GSA is close to the global optimal solution, its gravitational action mechanism will become more refined, which is helpful to fine tune in the local region, so as to find a more accurate solution. Compared with GA, GSA avoids the trap of local optimal solution caused by crossover and mutation operation, because gravity can naturally guide the solution to move to a better region. Compared with SA, the search process of GSA is more systematic and structured, rather than relying solely on the acceptance probability and temperature parameters to control the search direction.Then use the Bank Marketing data set as the main data set to conduct extended performance testing. For the precision rate and recall rate of the three algorithms GAS-BP, GAS-IGA, and GAS-ISA, the PR curves are shown in Figure 8.

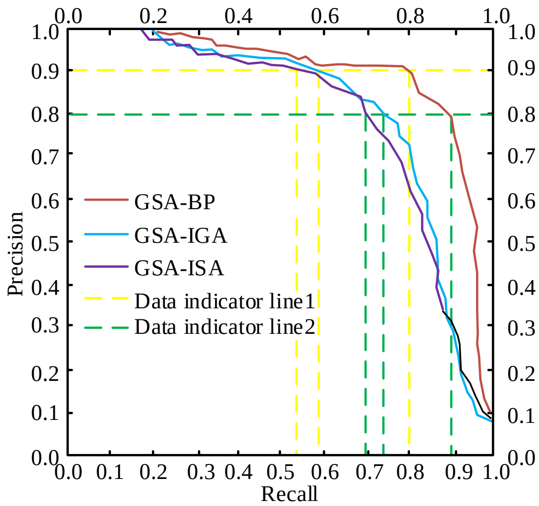
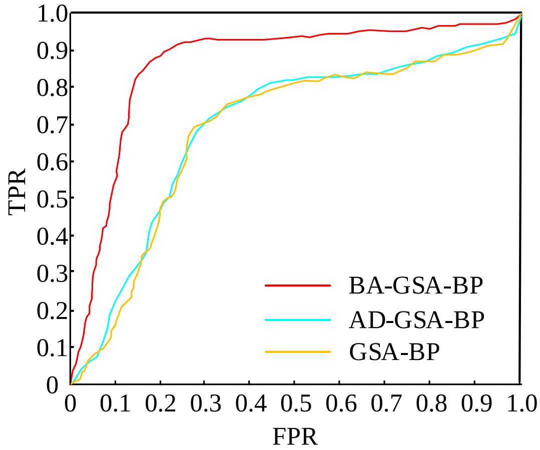
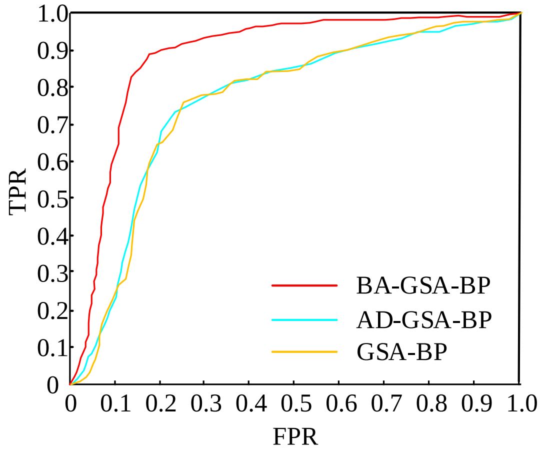


Figure 8: PR curves of three algorithms

As shown in Figure 8, it can be seen that for GSA-BP, when the precision rate is 0.8, the corresponding recall rate is approximately 0.896. This means that under a given threshold, the GSA-BP algorithm can find most relevant samples with a high accuracy. When the precision rate is 0.9, the corresponding recall rate is about 0.800. This shows that the GAS-BP algorithm may miss some relevant samples under higher precision, but still has a high recall rate. For GSA-IGA, when the precision rate is 0.8, the corresponding recall rate is approximately 0.731. Compared with the GSA-BP algorithm, the GSA-IGA algorithm can find fewer relevant samples with the same accuracy. When the precision rate is 0.9, the corresponding recall rate is approximately 0.591. This further shows that the GSA-GA algorithm misses more relevant samples with higher accuracy. For GSA-ISA, when the precision rate is 0.8, the corresponding recall rate is approximately 0.695. Compared with the GSA-BP algorithm and the GSA-IGA algorithm, the GSA-ISA algorithm can find fewer relevant samples. When the precision rate is 0.9, the corresponding recall rate is approximately 0.534. Although the GSA-ISA algorithm has a lower recall rate at higher precision, it still performs better than the GSA-IGA algorithm. After determining the optimal hidden layer structure of the BP neural network, use the Bagging integration algorithm to integrate the GSA-BP network, and use test samples in the Bank Marketing data set and Credit Card Default data set to verify the training results. In order to further illustrate the superiority of the BA-GSA-BP algorithm, the study also selected the Abdoost ensemble algorithm as a comparison. The ROC curve results of BA-GSA-BP, AD-GSA-BP, and GSA-BP are shown in Figure 9.



(a)Bank Marketing



(b)Credit Card Default

Figure 9: Comparison of ROC curve results of three algorithms

As shown in Figure 9, Figure 9(a) and (b) are the ROC curves tested in the data sets Bank Marketing and Credit Card Default respectively. It can be seen that in different data sets, the curve trends of the three algorithms are roughly the same. It shows that their classifier performance is similar in different data sets. In addition, it can be seen that through the BA-GSA-BP algorithm based on Bagging integration, the area under the ROC curve reaches the maximum, indicating that this algorithm is very effective in reducing the over-fitting risk of the GSA-BP algorithm and has stronger generalization ability. This shows that the BA-GSA-BP algorithm can better cope with the challenges brought by different data sets, avoid over-fitting problems, and obtain a more reliable classification model. However, the AD-GSA-BP algorithm based on Adaboost integration does not significantly improve the generalization ability of the GSA-BP algorithm. This means that the use of Adaboost integration does not effectively reduce the overfitting risk of the GSA-BP algorithm, nor does it bring significant performance improvements. To verify the importance of each index performance of the constructed model compared with other models. Statistical software SPSS was used to make statistics on the above data and conduct independent sample t-test, as shown in Table 3.

Table 3: T-test of model performance comparison

Algorithm comparison	Sample mean (MES value)	Average number of iterations	t-values	P-value	Confidence interval
GSA vs IGA (Bank Marketing)	GSA: 0.098, IGA: 0.103	GSA: 225, IGA: 336	0.086	P<0.05	98%
GSA vs ISA (Bank Marketing)	GSA: 0.098, ISA: 0.101	GSA: 225, ISA: 406	0.074	P<0.05	99%

GSA vs IGA (Credit Card Default)	GSA: 0.093, IGA: 0.101	GSA: 164, IGA: 341	0.085	P<0.05	99%
GSA vs ISA (Credit Card Default)	GSA: 0.093, ISA: 0.107	GSA: 164, ISA: 443	0.084	P<0.05	98%
Algorithm comparison	Sample mean (ROC value)	t-values		p-value	Confidence interval
BA-GSA-BP vs AD-GSA-BP (Bank Marketing)	BA-GSA-BP:0.904, AD-GSA-BP:0.714	0.075		P<0.05	98%
BA-GSA-BP vs GSA-BP (Bank Marketing)	BA-GSA-BP:0.904, AD-GSA-BP:0.701	0.063		P<0.05	99%
BA-GSA-BP vs AD-GSA-BP (Credit Card Default)	BA-GSA-BP:0.894, AD-GSA-BP:0.764	0.075		P<0.05	99%
BA-GSA-BP vs GSA-BP (Credit Card Default)	BA-GSA-BP:0.894, AD-GSA-BP:0.741	0.081		P<0.05	98%

It can be seen from table 3 that the p-value of performance comparison between models is less than 0.05, indicating that these performances have significant differences. At the same time, the confidence interval is more than 95%, indicating that the estimation of model performance difference is relatively accurate.

4.2 Simulation application of bank marketing system prediction model

In order to verify the practical application effect of the constructed system, the study selected data from two financial bank databases A and B for simulation experiments. This study mainly focused on the system function integration test of the model built by the institute and the response time and prediction success of different clients. The efficiency was verified, and the traditional prediction model, AD-GSA-BP, and GSA-BP models under the same experimental conditions were simultaneously tested as references. First, conduct integration testing on the four major modules of the bank marketing system: system management, marketing management, project management, and personal leave management (integration testing is to test whether the system can work normally after several parts of the built system are combined together). After testing, the expected results of each part of the system constructed by the institute are consistent with the actual results. It was proved that the research model can operate normally. Then, in order to meet the work needs of various banking personnel, the response time of the model was tested. The specific results are shown in Figure 10.

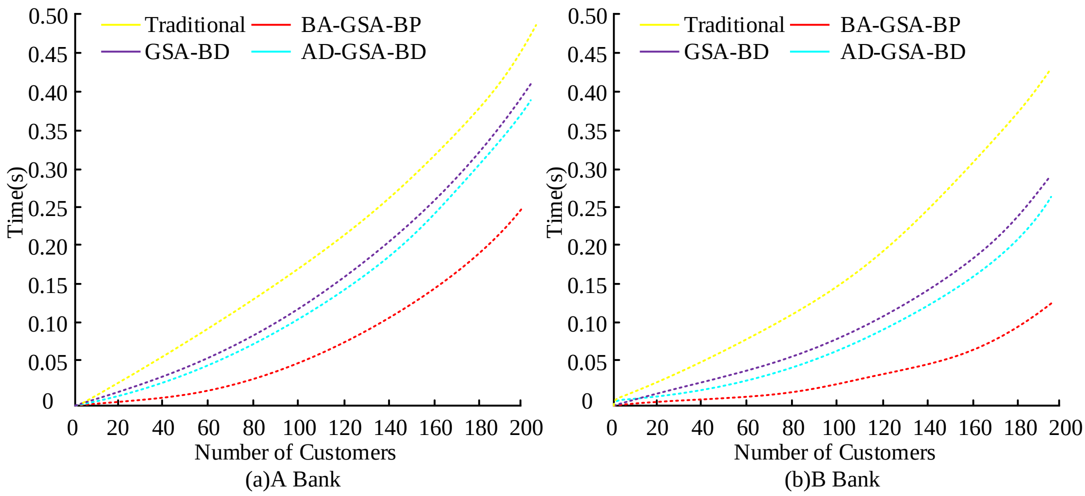


Figure 10: Comparison of response times of four models

It can be seen from Figure 10(a) that in the Bank Marketing data set, as the number of clients increases, the response time of each system begins to increase. When the number of clients is 200, the response time of the BA-GSA-BD model is the smallest, about 0.24s, less than 0.3s, which meets the work needs of banking personnel. The response times of the traditional prediction model, AD-GSA-BP, and GSA-BP are 0.48s, 0.39s, and 0.41s respectively. It can be seen from Figure 10(b) that in the Morningstar data set, the response time of each model has decreased, but the AD-GSA-BD model still has the smallest response time, which is about 0.12. In comparison, the response times of other models are longer. The response times of the traditional prediction model, AD-GSA-BP, and GSA-BP are 0.26s, 0.29s, and 0.43s respectively. It shows that the response time of the prediction model is faster when processing simple data sets. This is because ba-gsa-bp model optimizes the structure of BP neural network through GSA, and also combines bagging ensemble learning algorithm. This optimization makes the model more efficient in feature selection, parameter adjustment and model combination, thus reducing the response time. However, ad-gsa-bp and gsa-bp models lack the optimization strategy in ba-gsa-bp model, resulting in more time to complete the calculation when processing the same data set. Finally, the prediction success rates of various models were tested 10 times using the data of banks A and B. The specific results are shown in Figure 11.

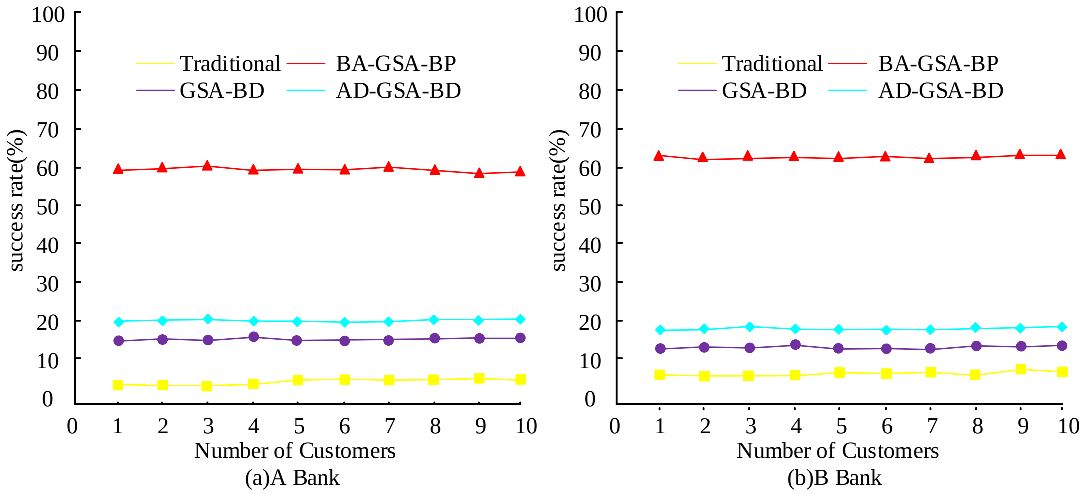


Figure 11: Comparison of prediction success rates of four models

As shown in Figure 11(a), under the data test of Financial Bank A, the BA-GSA-BP prediction model conducts marketing promotions for all customers, and the marketing success rate is as high as about 60%. The marketing success rates for all customers of the traditional prediction model, AD-GSA-BP, and GSA-BP are approximately 5%, 20%, and 15% respectively. Looking at Figure 11(b) again, we can see that the marketing success rates of the four forecast models BA-GSA-BP, traditional forecast model, AD-GSA-BP, and GSA-BP for all customers are approximately 62%, 8%, and 18% respectively. and 12%. Based on the above data, the performance comparison of different models in the actual banking system is summarized in Table 4.

Table 4: Performance comparison of different models in actual banking system

Data application scenarios	Model	Response time (seconds)	Prediction success rate (%)	Mean squared error
Bank A	BA-GSA-BP	0.24s	60	0.001
	GSA-BP	0.41s	15	0.012
	AD-GSA-BP	0.39s	20	0.024
	Traditional BP	0.48s	5	0.052
Bank B	BA-GSA-BP	0.12s	62	0.002
	GSA-BP	0.29s	12	0.011
	AD-GSA-BP	0.26s	18	0.024
	Traditional BP	0.43s	8	0.045

Table 4 shows that the prediction model constructed by the research can screen out customers who are more interested in ordering bank marketing products, thus improving the marketing efficiency and success rate of the entire bank. However, the interest and demand of the same customer for bank marketing products are constantly changing. With the change of market trend, personal financial status or life events (such as house purchase, marriage, birth, etc.), customers may change from interested in one product to interested in another product. The ba-gsa-bp model can adapt to the changes of customer behavior by regularly updating the training data and retraining the model.

5 Discussion

The models proposed by other individuals reviewed in the relevant work have achieved certain results on their respective test datasets, but still have limitations. Geng et al.'s artificial neural network model achieved an accuracy of 91% on the NASA battery dataset, but had a convergence iteration of 324 times, indicating that its training efficiency needs to be improved [10]; Mohammed et al.'s recurrent neural network has slightly lower accuracy (90%) on wind speed test datasets, and has more convergence iterations (354 times), due to the high complexity of the model leading to increased training costs [11]; Ramos et al.'s long-term memory neural network had a low prediction success rate (90%) on climate shared datasets, and also had a high number of convergence iterations (356 times), reflecting the model's insufficient generalization ability [12]; The model developed by Yadav et al., which combines fuzzy logic and gravity search, has decent accuracy on disease datasets (91%), but the reduction in convergence iterations did not significantly improve the prediction success rate (92%), indicating that the fusion effect of the algorithm is limited [13]; Tochaiwat et al.'s artificial neural network technology has high accuracy (93%) on the sales dataset of residential projects, but there is still room for further improvement in convergence speed and prediction success rate. In contrast, the gravity based search method (GSA) combined with neural networks proposed in the study showed significant superiority on bank marketing datasets. The model not only achieved an accuracy of up to 98% and a prediction success rate, but also significantly reduced the number of convergence iterations to 225, mainly due to the new combination of GSA and neural network and the strategy of binding ensemble [14]. Compared with Geng et al.'s method based on artificial neural networks, the study optimized the weights of the neural network by introducing GSA, avoiding the problem of traditional BP (backpropagation) algorithm easily falling into local optima, thereby improving the accuracy and training efficiency of the model. At the same time, the binding integration strategy further enhances the model's generalization ability and robustness, making it more advantageous when facing complex and ever-changing bank marketing data. These enhancements are particularly important in the field of banking marketing systems, as accurate customer forecasting and rapid response to market demand are key to improving marketing effectiveness and customer satisfaction. In the context of a large amount of data and changing customer behavior, the proposed method can better capture market trends, achieve precise marketing, and thus gain an advantage in the fiercely competitive market environment.

6 Conclusion

Accurately locating customer needs and achieving targeted marketing to customers are important aspects of bank development. The study first established a suitable bank marketing system prediction model based on its own needs. Then the BP neural network is introduced into the model framework as the main learning and fitting tool for the prediction model. Then the global search capability of the GSA algorithm is used to optimize the BP neural network structure, and finally the Bagging (BA) ensemble learning algorithm is used to integrate the optimized neural network. The results show that in the Bank Marketing data set, when GSA, SA, and GA are iterated 225 times, 406 times, and 336 times respectively, the MES values tend to be stable, and the MES values are approximately 0.098, 0.101, and 0.103 respectively. The method used in the study has a faster convergence speed, better fitness value, and the search strategy and parameter settings are more suitable for optimizing the BP neural network structure, which can converge to an optimal solution faster. When the accuracy rates are 0.9 and 0.8 respectively, the corresponding summoning rates of GAS-BP are 0.800 and 0.731 respectively. The other two algorithms correspond to the highest summoning rates of 0.591 and 0.695 respectively. Compared to the other two algorithms, the research method has higher accuracy and reliability in processing data, is more capable of identifying samples related to the problem, and provides better classification performance. This has advantages in handling banking system information. In addition, the BA-GAS-BP model has the largest ROC curve area, indicating that the method has better generalization performance on the same dataset, can better adapt to new data, and reduce overfitting risk of the model. The prediction model constructed using this method can adapt and process the vast dataset of the banking system at an extremely fast speed. When the number of customers is 200, the response time of the model built by the institute is 0.24s, and the prediction success rate is about 60%. The other three models require 0.39a at the fastest, and the highest prediction success rate is 12%. It shows that the system built by the institute is efficient and accurate in predicting potential bank users. However, the dataset used in this experiment is all historical data and cannot be changed according to market conditions and specific user situations. Subsequent research will further explore the adaptive updating of prediction models and improve the automatic updating function of prediction results. Subsequent research will further improve the adaptive ability of the prediction model, especially by integrating real-time data streams, dynamically adjusting model parameters and introducing advanced machine learning algorithms, so that the model can accurately adapt to the characteristics of different customer demographic data and the rapidly changing market environment. This will include developing a mechanism to automatically update the training data set on a regular basis or based on trigger conditions, and retraining or fine-tuning the model to ensure that it is always in sync with market trends and customer needs.

Ethics approval and consent to participate

Not Applicable.

Consent for publication

All authors read and approved the final manuscript.

Availability of data and material

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Competing interests

The authors have declared that no competing interests exist.

Funding statement

This paper is supported by the Philosophy and Social Science Planning Project of Henan Province (2022BJJ023), Soft Science Research Planning Project of Henan Province (232400411048), Key Research Project of Colleges and Universities of Henan Province (23A790023) and Research Start-up Foundation of Henan Finance University (2021BS008).

Authors' contributions

JZ wrote the whole paper.

Acknowledgements

Not Applicable.

References

[1] Wang T, Zhao S, Zhu G, Zheng H. A machine learning-based early warning system for systemic banking crises. Applied economics, 2021, 53(26): 2974-2992. doi.org/10.1080/00036846.2020.1870657

[2] Anand M, Velu A, Whig P. Prediction of loan behaviour with machine learning models for secure banking. Journal of Computer Science and Engineering (JCSE), 2022, 3(1): 1-13. doi.10.36596/jcse.v3i1.237

[3] Paiva B M R, Ferreira F A F, Carayannis E G, Zopounidis C, Ferreira J J, Pereira L F, Dias P J. Strategizing sustainability in the banking industry using fuzzy cognitive maps and system dynamics. International Journal of Sustainable Development & World Ecology, 2021, 28(2): 93-108. doi.org/10.1080/13504509.2020.1782284

[4] Tan X, Wang Q, Feng J, Huang, Y, Huang, Z. Fast modeling of gravity gradients from topographic surface data using GPU parallel algorithm. Geodesy and Geodynamics, 2021, 12(4): 288-290. doi.10.1016/j.geog.2021.04.003.

[5] Arun R. A Novel Classifier Model Efficient Prediction of Success Rate in Bank Marketing. International Journal of Advanced Science and Technology, 2020, 29(9): 6955-6961. doi. 10.3233/JIFS-189491

[6] Hosseini S. A decision support system based on machined learned Bayesian network for predicting successful direct sales marketing. 2021, 8(2): 295-315. doi.10.1080/23270012.2021.1897956

[7] Gao H, Kuang L, Yin Y, Guo B, Dou K. Mining consuming Behaviors with Temporal Evolution for Personalized Recommendation in Mobile Marketing Apps. Mobile Networks and Applications, 2020, 25(4): 1233-1248. doi.10.1007/s11036-020-01535-1.

[8] Ma L, Sun B. Machine learning and AI in marketing--Connecting computing power to human insights. International Journal of Research in Marketing, 2020, 37(3): 481-50. doi. 10.1016/j.ijresmar.2020.04.005

[9] Liu H. Big data precision marketing and consumer behavior analysis based on fuzzy clustering and PCA model. Journal of Intelligent and Fuzzy Systems, 2020, 40(4): 6529-6539. doi. 10.3233/JIFS-189491

[10] Geng P, Wang J, Xu X, Zhang Y, Qiu S. SOC Prediction of power lithium battery using BP neural network theory based on keras. International Core Journal of Engineering, 2020, 6(1) :171-181. doi. 10.6919/ICJE.202001_6(1).0025

[11] Mohammed AS, Mohammed AS, Kareem SW. Deep Learning and Neural Network-Based Wind Speed Prediction Model. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2022, 30(3): 403-425. doi.org/10.1142/S021848852240013X

[12] Ramos F R, Pereira M T, Oliveira M, Rubio L. The memory concept behind deep neural network models: An application in time series forecasting in the e-Commerce sector. Decision Making: Applications in Management and Engineering, 2023 6(2), 668-690. doi.org/10.31181/dmame622023695

[13] Yadav U, Tyagi T, Nagpal S. Fuzzy Optimized Gravitational Search Algorithm for Disease Prediction. International Journal of Swarm Intelligence Research (IJSIR), 2020, 11(3): 89-103. doi. 10.4018/IJSIR.2020070106

[14] Tochaiwat K, Pultawee P, Rinchumphu D. Sales rate forecasting of single-detached houses using artificial neural network technique. Decision Making: Applications in Management and Engineering, 2023, 6(2), 772-786. doi.org/10.31181/dmame622023707

[15] Shakya S, Smys S. Big data analytics for improved risk management and customer segregation in banking applications. Journal of IoT in Social, Mobile, Analytics, and Cloud, 2021, 3(3): 235-249. doi: 10.36548/jismac.2021.3.005

[16] Gams M , Kolenik T .Relations between Electronics, Artificial Intelligence and Information Society through Information Society Rules. Electronics, 2021, 10(4):514. DOI:10.3390/ELECTRONICS10040514.

[17] Deghbouch H , Debbat F .Hybrid Algorithm for Node Deployment with the Guarantee of Connectivity in Wireless Sensor Networks.Informatica (Slovenia), 2022, 46. doi:10.31449/inf.v46i8.3370.

[18] Dzemyda G , Sabaliauskas M , Medvedev V .Geometric MDS Performance for Large Data Dimensionality Reduction and Visualization.Informatica, 2022, 33:299-320. doi:10.15388/22-infor491.

[19] Saffaran A, Moghaddam M A, Kolahan F. Optimization of backpropagation neural network-based models in EDM process using particle swarm optimization and simulated annealing algorithms. Journal of the Brazilian Society of Mechanical Sciences and Engineering, 2020, 42(1):1-14. doi:10.1016/j.ijpvp.2021.104531.

[20] Safaei S, Ghasemi P, Goodarzian F, Momenitabar, M. Designing a new multi-echelon multi-period closed-loop supply chain network by forecasting demand using time series model: a genetic algorithm. Environmental Science and Pollution Research, 2022, 29(53): 79754-79768. doi:10.1007/s00170-013-4953-6