Available online at www.HighTechJournal.org

HighTech and Innovation Journal

ISSN: 2723-9535

Vol. 4, No. 2, June, 2023

The Prediction of Douyin Live Sales based on Neural Network Algorithms

Jinman Sun [1], Xinling De [1*], Hui Zheng [1]

[1] *School of Economics and Management, Cangzhou Jiaotong College, Huanghua, Hebei 061199, China.*

Received 26 January 2023; Revised 19 May 2023; Accepted 23 May 2023; Published 01 June 2023

Abstract

This paper aims to optimize neural network algorithms in order to improve their predictive performance and meet the demand for Douyin live sales forecasting. The Douyin live data of Florasis between January 2022 and June 2022 were collected from Huitun data and preprocessed. Using the Pearson correlation coefficient, six factors that are highly correlated with live sales were selected for subsequent prediction. This paper briefly introduces the back-propagation neural network (BPNN) algorithm and analyzes its parameter optimization methods, including particle swarm optimization (PSO), the artificial bee colony (ABC) algorithm, and the beetle antler search (BAS) algorithm. Then, an improved beetle antennae search (IBAS) algorithm was developed by introducing inertia weight and used to construct an IBAS-BPNN model for predicting sales volume in Douyin live streaming. The results showed that compared with the BPNN, PSO-BPNN, and ABC-BPNN algorithms, the IBAS-BPNN algorithm had better prediction performance, with a root-mean-square error of 335.6694, a mean absolute percentage error of 0.0532%, an equilibrium coefficient of 0.9889, and a shorter training time of 90.07 s. The experimental results demonstrate the reliability of the IBAS-BPNN algorithm for predicting sales volume in Douyin live streaming, providing new insights into parameter optimization of BPNN and offering references for further research on BPNN parameter optimization. It also provides an effective method with both timeliness and high accuracy for predicting sales volume in Douyin live streaming in practical applications.

Keywords: Neural Network; Douyin; E-Commerce Live; Sales Prediction; Beetle Antennae Search Algorithm.

1. Introduction

Influenced by the Internet, mobile technology, and other factors, the e-commerce industry has entered a brand-new phase of growth. E-commerce live streaming, as an interactive and flexible sales method [1], is increasingly preferred by users and can generate significant revenue for e-commerce companies [2]. Douyin, as one of the most popular short video platforms [3], has demonstrated enormous potential and value as an e-commerce live streaming platform. Sales forecasting plays an important role in inventory control and production plan adjustments. With advancements in machine learning and other technologies, sales forecasting has become increasingly intelligent. Sharma et al. [4] used machine learning techniques to forecast the auto sales of an Indian automobile company during the COVID-19 pandemic. The experiment showed that the model could effectively predict changes in sales, providing guidance for the company's financial preparation.

Boone et al. [5] forecasted sales for retailers and combined user-generated data from Google Trends search queries with operational data. Through experiments, they found that the addition of Google Trends data could reduce the mean

* Corresponding author: 1xlde@czjtu.edu.cn

http://dx.doi.org/10.28991/HIJ-2023-04-02-09

absolute percentage error (MAPE) by 2.2% to 7.7%. To predict housing sales in São Paulo, Moro et al. [6] modeled the index softening, Box-Jenkins, and artificial neural network models and compared their effects using different combinations. Ma et al. [7] applied the genetic algorithm-back-propagation neural network (GA-BPNN) algorithm to the sales forecast of electric vehicles. They found out through experiments that A-class electric vehicles will become the main potential model in 2020.

Wang et al. [8] proposed a sales forecasting model, M-GNA-XGBOOST, based on time series prediction that can provide short-term predictions for the sales of each product in an online store. Through experiments, it was found that the root-mean-square error (RMSE) and mean absolute error (MAE) of this model were approximately 11.9 and 8.23, respectively. Raiyani et al. [9] utilized sales data from ten stores and 100 different products over a period of five years to compare the effectiveness of several machine learning models in predicting future sales. Through experimentation, they found that the hybrid model yielded superior results compared to individual models. Takahashi & Goto [10] designed a double exponential smoothing method for predicting the inventory/sales of products. They found that this method had high accuracy in forecasting sales for each type of medication. Zhan et al. [11] developed a Bayesian linear regression method to predict product sales and found that this method performed well in terms of MAE and RMSE. Liu et al. [12] designed a support vector regression model to predict drug sales and improve inventory management, achieving an accuracy rate of 91% for this method.

Current research primarily focuses on sales prediction for physical products, with limited studies on e-commerce sales prediction and no mention of live-streaming sales prediction in e-commerce. However, given the rapid development of live-streaming in e-commerce, it has become a significant form of online sales, making research on live-streaming sales prediction in e-commerce crucial. For Douyin e-commerce live streaming, sales forecasting can help e-commerce companies better understand the reasons for changes in sales and adjust their live streaming methods in a timely manner. Therefore, this paper designed a neural network-based sales forecasting algorithm for Douyin e-commerce live streaming and experimentally proved its effectiveness, providing a new method of sales forecasting for e-commerce live streaming in practice.

2. Douyin E-Commerce Live Analysis

The sales generated by live streaming are closely related to inventory management and production arrangements for e-commerce companies. If sales are good, inventory should be greatly reduced, and production arrangements should be increased accordingly. Conversely, if the products are not selling well, production plans should be promptly adjusted to avoid greater losses. Accurate sales forecasting can help e-commerce companies better manage product production and marketing, thus providing better services for e-commerce live streaming.

Taking Douyin as an example, e-commerce live streaming generates massive data, such as the number of commodities, follower growth, likes, and comments. This article obtained relevant data on Douyin e-commerce live streaming from Huitun Data (*https://www.huitun.com/*) and examined the sales forecast of Florasis, a well-known beauty brand. Founded in 2017, Florasis specializes in cosmetics, skin care, and other products and is highly active on the Douyin platform. Through collaborations with professionals and internet celebrities, the brand has successfully established effective communication and interaction with users. With robust capabilities in product development and research, it has achieved excellent results in Douyin live streaming.

The Douyin live streaming data for Florasis from January to June 2022 was obtained from Huitun Data, with duplicates and anomalies removed and missing values filled in with the mean. To facilitate the subsequent sales forecast, all data were normalized, and the processed live streaming data is demonstrated in Table 1.

Table 1. Normalized live data

Time	January 1, 2022	January 2, 2022	January 3, 2022	......	June 30, 2022
Duration of live broadcast	0.7468	0.6485	0.6845	......	0.7168
Times of viewing	0.5126	0.6715	0.3158	......	0.6874
Peak number of people	0.8525	0.8568	0.3696	......	0.7756
Per capita stay	0.7878	0.9025	0.4285	......	0.2168
Number of products	0.7965	0.7548	0.1632	......	0.3256
Number of likes	0.3658	0.7744	0.9258	......	0.5148
Number of new fans	0.6952	0.5968	0.2148	......	0.4851
Interaction rate	0.6241	0.5326	0.3236	......	0.3658
Conversion rate	0.3365	0.4958	0.7516	......	0.1529
Live sales	0.7853	0.8222	0.8584	......	0.6258

The correlation between factors such as the duration of live broadcast and the times of viewing in Table 1 and live sales was further analyzed using the Pearson correlation coefficient [13]. The corresponding Equation is:

$$\rho_{X,\ Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} \tag{1}$$

where $\mathrm{cov}(\mathrm{X}, \mathrm{Y})$ refers to the covariance of X and Y and σ is the variance. The results obtained after calculation are shown in Table 2.

Table 2. Results of correlation coefficient calculation

Factor	Correlation coefficient
Duration of live broadcast	**0.0263**
Times of viewing	0.6785
Peak number of people	**0.0745**
Per capita stay	0.6447
Number of products	**0.0321**
Number of likes	0.7158
Number of new fans	0.6845
Interaction rate	0.7715
Conversion rate	0.6162

For factors with correlation coefficients less than 0.1, they were considered to have a weak correlation with live sales and were eliminated, and the remaining six factors were used for subsequent sales forecasting. Therefore, when using neural network algorithms for sales forecasting, the input data should include six variables from Table 2: number of views, average duration per person, number of likes, increase in followers, interaction rate, and conversion rate. The output was the prediction result for live sales.

3. Sales Forecasting based on Neural Network Algorithm

3.1. Back-propagation Neural Network Algorithm

In data prediction, the commonly used methods include linear regression [14], decision trees [15], etc., and the back-propagation neural network (BPNN) algorithm is also one of them [16]. The BPNN algorithm has good performance in processing nonlinear data [17] and is flexible and adaptable, making it widely used in finance and health care for data prediction [18]. Therefore, this paper used the BPNN algorithm to implement the prediction of Douyin e-commerce live sales. However, during training, the BPNN algorithm is prone to getting stuck in local optima [19], mainly due to the influence of weights and thresholds. In order to improve this flaw and enhance the effectiveness of the BPNN algorithm in live sales forecasting, this paper analyzed several methods for parameter optimization and validated them using the data obtained in the previous section.

For a simple three-layer BPNN, assume that the i-th input is x_i, then the input of its implicit layer can be written as: $H_j = f\left(\sum_{j=1}^{i} w_{ij} x_i + a_j\right)$, where f is the activation function, w_{ij} is the weight between the input layer and the implied layer, and a_j is the threshold value.

The output of the output layer can be written as: $O_k = \sum_{i=1}^{i} H_i w_{ik} + a_k$, where w_{ik} is the weight between the implied layer and the output layer and a_k is the threshold value.

If the desired output is Y_k, then the calculation formula of error is: $E = \frac{1}{2}\sum_{k=1}^{m}(Y_k - O_k)^2$. Denote $Y_k - O_k = e_k$, then the equation can be rewritten as: $E = \frac{1}{2}\sum_{k=1}^{m} e_k^2$. The BPNN algorithm updates the weights and thresholds by back-propagating the errors. The weights are updated by the following equation:

$$\begin{cases} w_{ij} = w_{ij} + \eta H_j\left(1 - H_j\right) x_i \sum_{k=1}^{m} w_{jk} e_k \\ w_{jk} = w_{jk} + \eta H_j e_k \end{cases} \tag{2}$$

The threshold value is updated by the following equation:

$$\begin{cases} a_j = a_j + \eta H_j\left(1 - H_j\right) x_i \sum_{k=1}^{m} w_{jk} e_k \\ a_k = a_k + \eta a_k \end{cases} \tag{3}$$

3.2. Optimization Methods of BPNN Parameters

Several algorithms were used to optimize the parameters of the BPNN algorithm, as follows.

1) *Particle Swarm Optimization Algorithm*

The particle swarm optimization (PSO) algorithm was developed by studying birds' foraging habits [20]. Each bird is considered as a particle and the optimal solution is found by updating the position and velocity of the particle:

$$V_i^{k+1} = wV_i^k + c_1 r_1 \left(P_{ibest}^k - X_i^k\right) + c_2 r_2 \left(P_{gbest}^k - X_i^k\right) \tag{4}$$

$$X_i^{k+1} = X_i^k + V_i^{k+1} \tag{5}$$

where V_i^k stands for particle velocity at the k-th iteration, V_i^{k+1} stands for particle velocity at the $k+1$-th iteration, c_1 and c_2 are acceleration factors, r_1 and c_2 are random numbers in [0,1], P_{ibest}^k represents the individual extreme, and P_{gbest}^k represents the population extreme.

The parameters of the BPNN algorithm were improved using the PSO algorithm to obtain the PSO-BPNN algorithm. The process is as follows. After determining the structure of the BPNN algorithm, the parameters of the PSO algorithm were initialized, and continuously updating particles helped obtain optimal weights and thresholds for the BPNN algorithm. These weights and thresholds were input into the BPNN algorithm to complete its training.

2) *Artificial Bee Colony Algorithm*

The artificial bee colony (ABC) algorithm is an imitation of the honey harvesting process of a bee colony [21]. It is assumed that the nectar source location is: $X_i = (X_{i1}, X_{i2}, \cdots, X_{iD})$, then the initial solution is obtained: $X_i^j = X_i^{min} + r_i^j \times \left(X_i^{max} - X_i^{min}\right)$, where X_i^{min} and X_i^{max} refer to the upper and lower limits of the the j-th dimensional solution space and r_i^j refers to a random number in [0,1].

The fitness of the i-th solution can be written as: $F_i = \begin{cases} \frac{1+}{1+f_i}, f_i \geq 0 \\ 1+|f_i|, f_i < 0 \end{cases}$, where f_i refers to the objective value of the problem.

In the starting phase, scout bees search for nectar sources according to the following formula: $V_i^j = X_i^j + r_i^j \left(X_m^j - X_k^j\right)$, where r_i^j represents a random number in [0,1]. Subsequently, the follower bees choose whether to follow or not according to the nectar source message shared by the scout bees, and the probability is: $p_i = \frac{f_i t_i}{\sum_{i=1}^{NP} f_i t_i}$, where $f_i t_i$ is the fitness value of possible solution X_i.

The scout and follower bees continuously search for nectar sources to find the optimal solution, which is the best nectar source. The ABC algorithm was applied to optimize the parameters of the BPNN algorithm, resulting in the ABC-BPNN algorithm.

3) *Beetle Antennae Search Algorithm*

The beetle antennae search (BAS) algorithm is based on the process by which the beetle relies on its tentacles to search for food [22]. In the BAS algorithm, a random vector is used to represent the initial orientation of beetle's antennae: $\vec{b} = \frac{rands(k,1)}{\|rands(k,1)\|}$, where $rands$ is the random function and k is the spatial dimension. The spatial coordinates of the left and right antenna of the beetle can be written as: $\begin{cases} x_{rt} = x^t + d_0 \times \frac{\vec{b}}{2} \\ x_{lt} = x^t - d_0 \times \frac{\vec{b}}{2} \end{cases}$, where t refers to the number of iterations, x_{rt} and x_{lt} are the coordinates of the left and right beetle antenna, x^t is the central coordinate of the left and right beetle antenna obtained after t times of iterations (the distance between them is d_0).

The beetle antenna find food according to the odor, and the intensity of the odor is recorded as the fitness value. According to the intensity of the odor, the process of updating the position of the beetle can be written as: $x^{t+1} = x^t - \delta^t * \vec{b} * sgn[f(x_{rt}) - f(x_{lt})]$, where δ^t is the step size factor and is updated in a decreasing manner: $\delta^{t+1} = \delta^t * eta$, where eta is the decay factor, $eta \in [0,1]$.

Although the BAS algorithm has the advantage of fast convergence, it also has the disadvantage of easily falling into local optimum. To address this point, this paper introduced inertia weight ω to develop an improved BAS (IBAS) algorithm. The beetle position update process is improved as:

$$x^{t+1} = \omega x^t - \delta^t * \vec{b} * sgn[f(x_{rt}) - f(x_{lt})] \tag{6}$$

$$\omega = \omega_{min} + \frac{(\omega_{max} - \omega_{min})(t_{max} - t)}{t_{max}} \tag{7}$$

where t_{max} is the maximum value of the number of iterations and ω_{max} and ω_{min} are the maximum and minimum values of ω.

The IBAS algorithm was used to optimize the parameters of the BPNN algorithm to obtain the IBAS-BPNN algorithm. The flowchart of the designed Douyin live sales prediction method is shown in Figure 1.

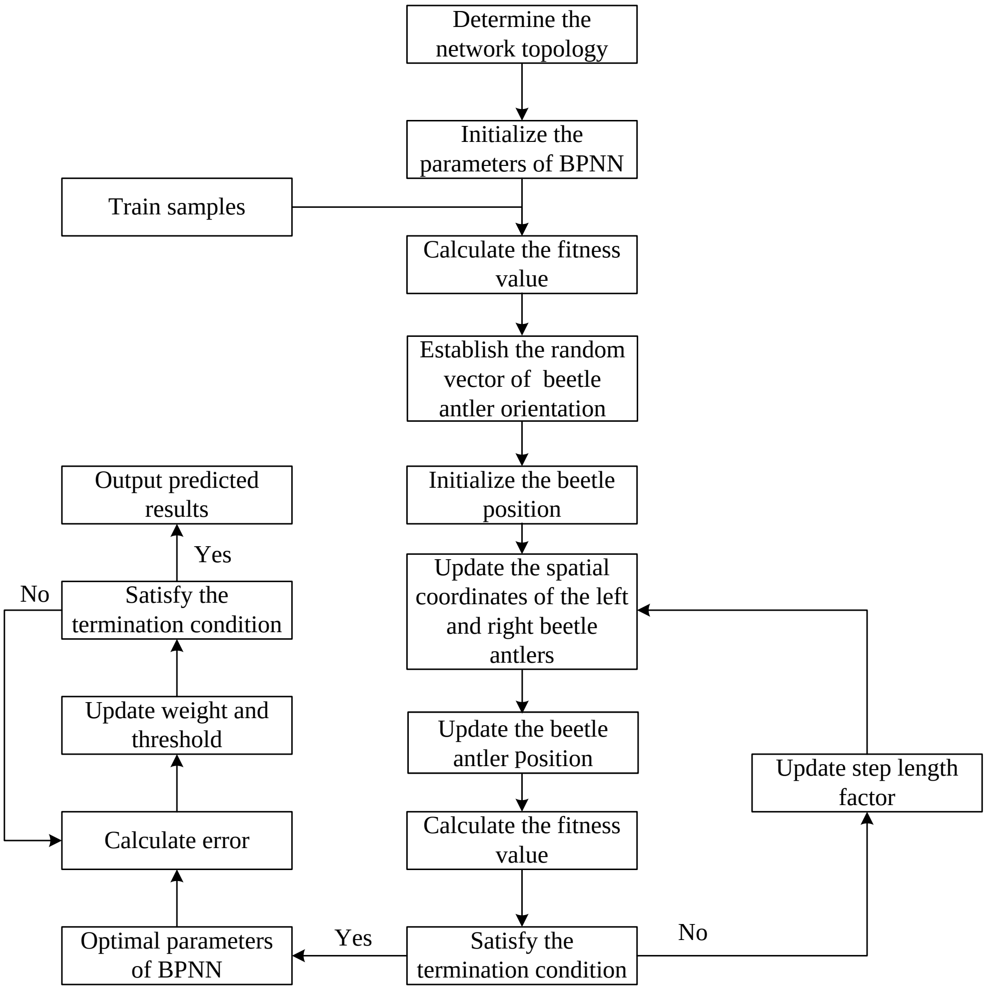

Figure 1. The IBAS-BPNN-based Douyin live sales prediction method

As shown in Figure 1, after determining the structure of the BPNN algorithm, the optimal parameters of the BPNN algorithm were obtained through the IBAS algorithm and input into the BPNN algorithm. The prediction of Douyin live sales volume prediction was realized by constantly updating the error.

4. Results and Analysis

4.1. Experimental setup

For the experimental data in Table 1, 80% of them were used for training and 20% for testing. It was seen from Table 2 that the input of the BPNN algorithm consisted of the remaining six indicators after removing three indicators with low correlation coefficients. Therefore, the number of nodes in the input layer of the BPNN algorithm was 6, and the number of nodes in the output layer was 1, i.e., live sales. The number of nodes for the implied layer was determined as 5 using a trial-and-error method, resulting in a BPNN algorithm structure of 6-5-1. The Tansig function was used in the implied layer, while the RULE function was used in the output layer. The learning rate was 0.005, and the maximum number of iterations was limited to 1,000.

After repeated training of the PSO-BPNN algorithm, the population size with the smallest MSE value in the training set was used as the population size for testing. The MSE of the corresponding population size is shown in Table 3.

Table 3. The training error of the PSO-BPNN algorithm under different population sizes

Population size	MSE
10	0.0083
20	0.0076
30	0.0072
40	0.0075
50	0.0077
60	0.0081
70	0.0084
80	0.0079
90	0.0083
100	0.0082

According to Table 3, the population size of the PSO-BPNN algorithm was 30, acceleration factors c_1 and c_2 were both 1.5, and the maximum iteration count was set at 3,000.

Considering the complexity of the ABC-BPNN algorithm, the population size was determined within the range of [10,50]. Through repeated training, the training errors are shown in Table 4.

Table 4. The training error of the ABC-BPNN algorithm under different population sizes

Population size	MES
10	0.0075
20	0.0061
30	0.0068
40	0.0069
50	0.0071

According to Table 4, the population size of the ABC-BPNN algorithm was determined as 20, and the maximum iteration count was set at 2,000.

For the BAS-BPNN and IBAS-BPNN algorithms, to ensure the search ability of the beetle, the initial step length should be as large as possible. Therefore, the initial step length of the beetle was set at 1, eta was set at 0.95, ω_{max} was set as 0.9, and ω_{min} was set at 0.4.

4.2. Evaluation Indicators

For m samples, let the live sales predicted by the algorithm be $\hat{y}$ and the actual live sales be y. The evaluation indicators used in this paper are as follows.

1) Root mean square error (RMSE): it evaluated the accuracy of the algorithm's prediction of live sales, and the corresponding equation is:

$$RMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - \hat{y}_i)^2} \tag{8}$$

2) MAPE: it was used to describe the accuracy of the algorithm's prediction of live sales, and the corresponding equation is:

$$MAPE = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{9}$$

3) Equilibrium coefficient (EC): it was used to describe the trend fit of the predicted value curve to the actual value curve, and the corresponding formula is:

$$EC = 1 - \frac{\sum_{i=1}^{m}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{m}(y_i + \hat{y}_i)^2} \tag{10}$$

4.3. Analysis of Results

Firstly, to determine the performance of several parameter optimization methods, tests were conducted on five single-peak functions and five multi-peak functions on the CEC2013 test set [23] to understand the performance of PSO, ABC, BAS, and IBAS for optimization search. The dimension of the search space was set to 30, and the search range was [-100,100]. The results are presented in Table 5.

Table 5. Results of the average fitness value of the algorithms for the test set (bold refers to the optimal result)

	f_1	f_2	f_3	f_4	f_5
PSO	1.45E+00	5.77E+06	2.44E+07	7.88E+02	3.88E+00
ABC	8.37E-15	**9.48E+05**	**1.81E+06**	1.82E+02	5.95E+01
BAS	1.27E-03	4.26E+06	1.56E+07	1.87E+02	8.55E-01
IBAS	**0.00E+00**	1.87E+06	1.52E+07	**4.54E+01**	**0.00E+00**
	f_6	f_7	f_8	f_9	f_{10}
PSO	9.98E+01	7.46E+01	2.35E+01	4.87E+01	1.07E-01
ABC	4.84E-04	2.85E+01	2.13E+01	2.36E+01	2.16E+01
BAS	3.38E+01	4.55E+01	**2.09E+01**	1.45E+01	9.18E-02
IBAS	**1.21E+01**	**3.84E+00**	**2.09E+01**	**9.81E+00**	**6.85E-02**

From Table 5, it can be observed that IBAS achieved the best average results among the functions, except for f_2 and f_3. This indicated that the IBAS algorithm performed well on both single-peak and multi-peak functions, demonstrating excellent capabilities in solving optimization problems and outperforming other parameter optimization methods.

The effectiveness of the BPNN, PSO-BPNN, ABC-BPNN, BAS-BPNN and IBAS-BPNN algorithms in live sales prediction was compared. Firstly, the comparison of the RMSE is displayed in Figure 2.

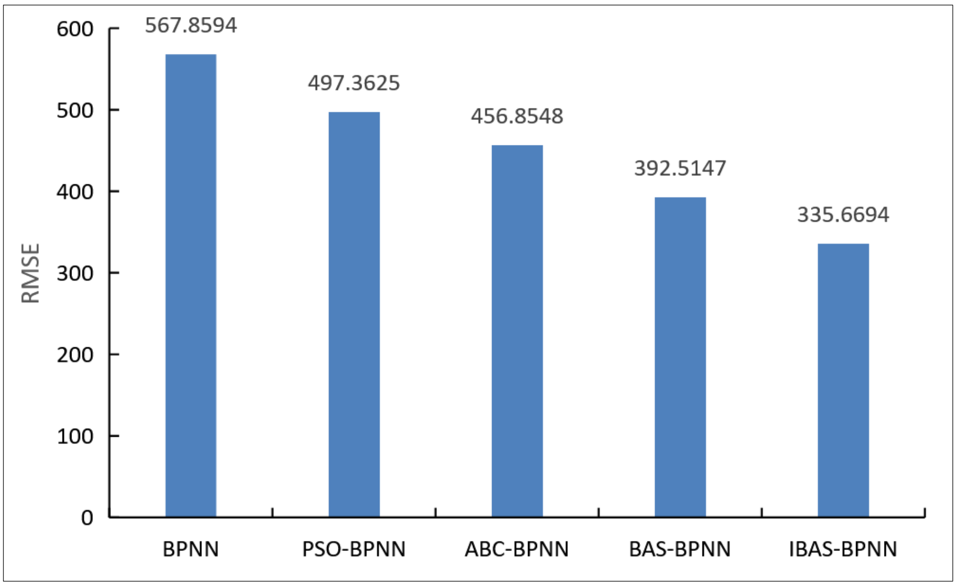

Figure 2. Comparison of the RMSE

From Figure 2, it was seen that in Douyin live streaming sales forecasting, when the BPNN algorithm was used, the RMSE reached its highest value at 567.8594. However, after optimizing the parameters, the RMSE decreased significantly. The RMSE of the PSO-BPNN algorithm was reduced by 12.41% compared with the traditional BPNN algorithm, while the ABC-BPNN algorithm showed a reduction of 19.55%, and the BAS-BPNN algorithm demonstrated a reduction of 30.88%. These results highlight the effectiveness of parameter optimization in improving the accuracy of the BPNN algorithm. The BAS algorithm was proven to have the best performance among these parameter optimization methods. Furthermore, the RMSE of the IBAS-BPNN algorithm was 335.6694, which decreased by 40.89% compared to the BPNN algorithm and by 14.48% compared to the BAS-BPNN algorithm, demonstrating the reliability of improvement achieved by using the BAS algorithm.

The comparison of the MAPE among the five algorithms is shown in Figure 3.

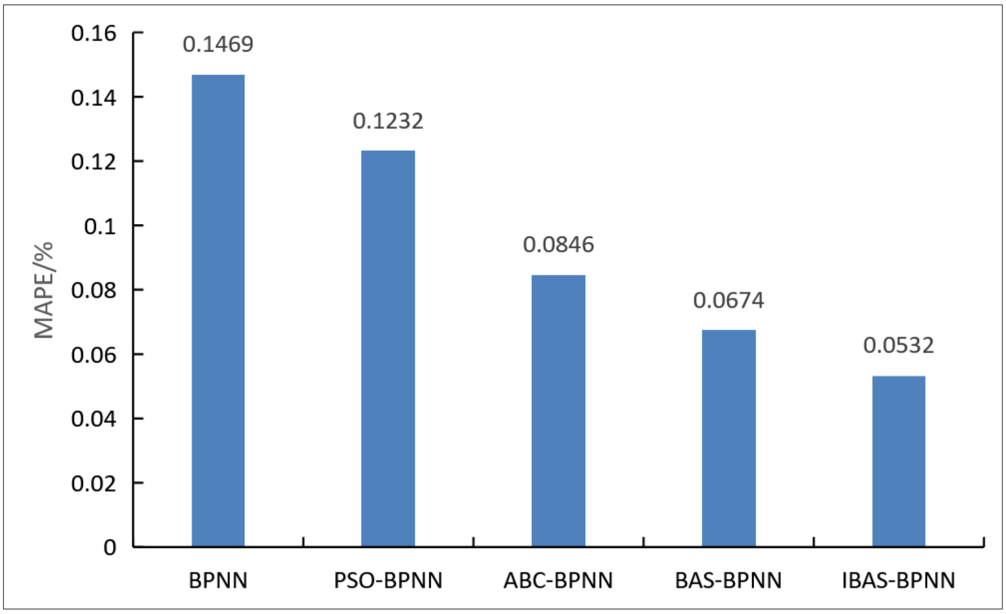

Figure 3. Comparison of the MAPE

From Figure 3, it was seen that these algorithms showed similar results for MAPE, and the MAPE value of the BPNN algorithm reached its highest value at 0.1469%, indicating its poor performance in forecasting Douyin live streaming sales. Among the comparison of several optimization methods, the MAPE value of the BAS-BPNN algorithm was 0.0674%, which was much lower than that of the PSO-BPNN and ABC-BPNN algorithms, resulting in a reduction of 0.0795% compared to the BPNN algorithm. Finally, the comparison between the BAS-BPNN and IBAS-BPNN algorithms showed that the latter had an MAPE value of 0.0532%, which was 0.0142% lower than the former, demonstrating the effectiveness of the IBAS-BPNN algorithm in forecasting Douyin live streaming sales.

The comparison of the EC is illustrated in Figure 4.

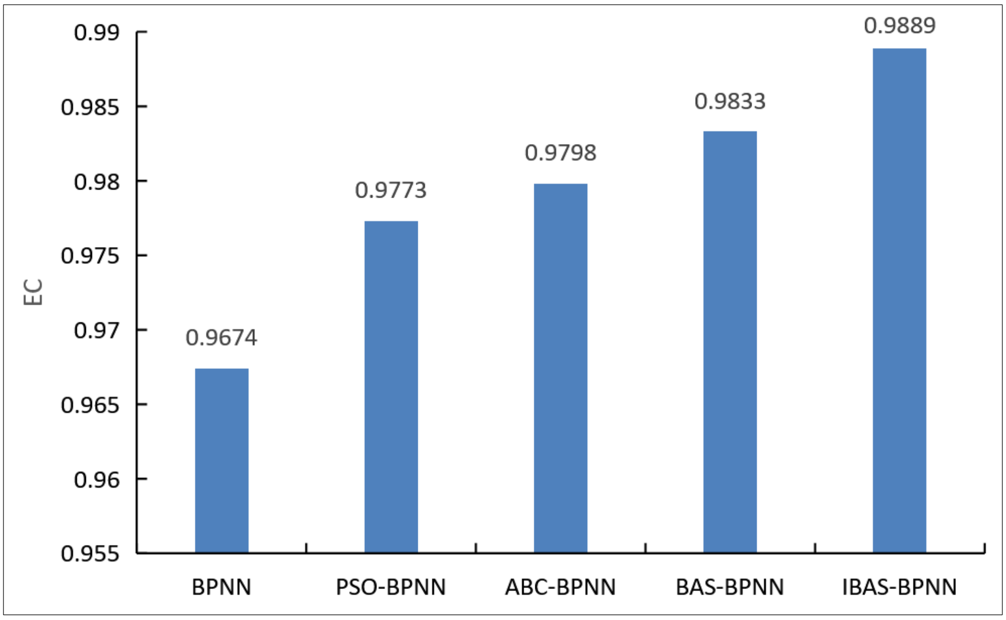

Figure 4. Comparison of the EC

The closer the EC value was to 1, the closer the predicted result of the algorithm was to the actual value. According to Figure 4, the EC value of the BPNN algorithm was 0.9674, which was the lowest, while the EC value of the IBAS-BPNN algorithm was 0.9889, which was the highest and represented a 2.22% improvement over to the BPNN algorithm and a 0.57% improvement over the BAS-BPNN algorithm. This demonstrated that the IBAS-BPNN algorithm had superior predictive performance in Douyin live streaming forecasting.

Finally, the training time of these algorithms was compared, and the results are presented in Figure 5.

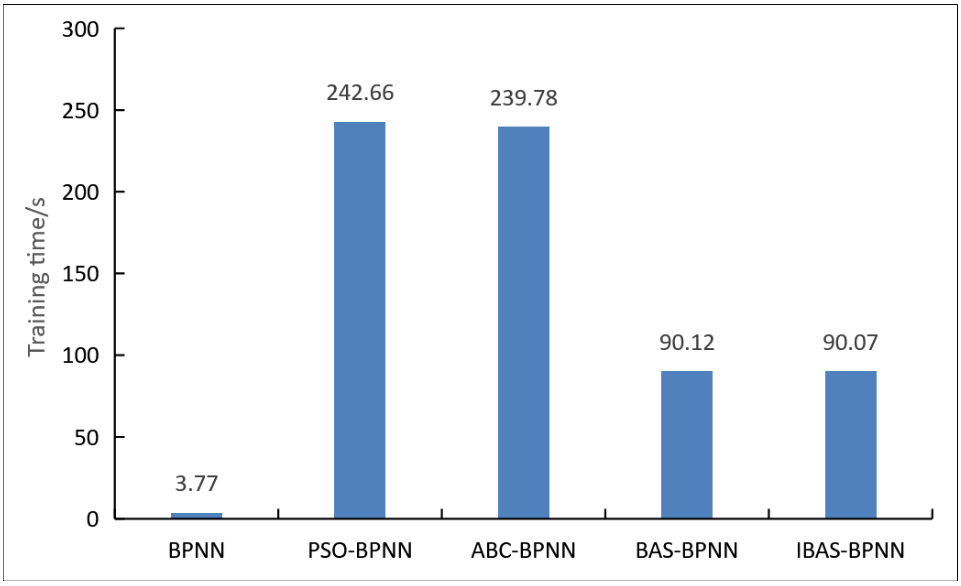

Figure 5. Comparison of the training time

The BPNN algorithm had the shortest training time, only 3.77 s, as observed from Figure 5. After parameter optimization, the training time of the BPNN algorithm increased significantly. However, it was found that the training time of the BAS-BPNN algorithm was much lower than that of the PSO-BPNN and ABC-BPNN algorithms. In addition, no significant difference was seen in the training time between IBAS-BPNN and BAS-BPNN algorithms, indicating that the improvement of the BAS algorithm did not increase the computational burden and demonstrating the reliability of this method.

The IBAS-BPNN algorithm was compared with other data prediction methods: Bayesian neural network [24], probability graph convolution model (PGCM) [25], and multi-dimensional recurrent neural network (MDRNN) [26].

The prediction performance of different algorithms for the same dataset is displayed in Table 6.

Table 6. Comparison results of the IBAS-BPNN algorithm and other data prediction methods

	RMSE	MAPE	EC
BNN	401.2362	0.0677	0.9837
PGCM	375.1254	0.0578	0.9852
MDRNN	345.1285	0.0551	0.9884
IBAS-BPNN	335.6694	0.0532	0.9889

From Table 6, it can be observed that compared to other data prediction methods, the IBAS-BPNN algorithm demonstrated superior performance in terms of RMSE, MAPE, and EC. Taking the comparison with the MDRNN model as an example, the IBAS-BPNN algorithm exhibited a reduction of 2.74% in RMSE, a decrease of 3.45% in MAPE, and an improvement of 0.0005 in EC. This further confirmed the advantage of the designed IBAS-BPNN model in data prediction and its ability to achieve high accuracy in live sales forecasting.

5. Conclusion

This article focused on predicting sales for Douyin live streaming. Using Florasis's Douyin live streaming data as an example, an experimental analysis was conducted using the BPNN method. An improved IBAS-BPNN method was designed to optimize the parameters. A comparison was made between the IBAS-BPNN method and traditional methods such as BPNN, PSO-BPNN, ABC-BPNN, and BAS-BPNN. The results showed that the IBAS-BPNN algorithm outperformed traditional methods such as BPNN, PSO-BPNN, ABC-BPNN, and BAS-BPNN in all aspects. Specifically, the IBAS-BPNN algorithm achieved an RMSE of 335.6694, an MAPE of 0 .0532, and an EC of 0 .9889, which were significantly better than those obtained by the other algorithms. This indicated that the sales forecast obtained by the IBAS-BPNN algorithm was very close to the actual values, i.e., it provided accurate predictions for Douyin live streaming sales. Additionally, this method had a short training time of 90.07 s, demonstrating its ability to offer real-time results efficiently and quickly in practical applications. The performance comparison with other data prediction methods further demonstrated the superiority of the IBAS-BPNN algorithm. This study proved the effectiveness of improving BAS and provides a new method for parameter optimization of the BPNN algorithm, which is conducive to conducting more comprehensive and in-depth research on parameter optimization for the BPNN algorithm. The feasibility of improving the performance of the BPNN algorithm by further optimizing the optimization algorithm was discussed in this article. Additionally, a reliable method for predicting sales on the Douyin live streaming platform was provided. In future research, further analysis and screening can be conducted on Douyin live streaming data to compare the performance of additional prediction methods in live sales forecasting. This will deepen the study in this field and provide a theoretical basis for better promoting e-commerce livestreaming.

6. Declarations

6.1. Author Contributions

Conceptualization, J.S. and X.D.; methodology, J.S.; software, J.S. and X.D.; validation, J.S., X.D., and H.Z.; formal analysis, J.S.; investigation, J.S. and X.D.; resources, J.S. and X.D.; data curation, J.S.; writing—original draft preparation, J.S.; writing—review and editing, J.S.; visualization, J.S. and X.D.; supervision, J.S.; project administration, J.S. and H.Z.; funding acquisition, X.D. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

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

7. References

[1] Peng, L., Lu, G., Pang, K., & Yao, Q. (2021). Optimal farmer's income from farm products sales on live streaming with random rewards: Case from China's rural revitalization strategy. Computers and Electronics in Agriculture, 189, 106403. doi:10.1016/j.compag.2021.106403.

[2] Liu, K., Liu, B., Xu, H., He, Y., & Cao, Y. (2020). Research on e-commerce live broadcasts helping poverty alleviation under the influence of the COVID-19 : Take Xinhua County, Hunan Province as an example. 2020 2nd International Conference on Economic Management and Model Engineering (ICEMME). doi:10.1109/icemme51517.2020.00176.

[3] Cheng, H., Huang, Y. T., & Huang, J. (2022). The Application of DEMATEL-ANP in Livestream E-Commerce Based on the Research of Consumers' Shopping Motivation. Scientific Programming, 2022, 1–15. doi:10.1155/2022/4487621.

[4] Sharma, P., Khater, S., & Vashisht, V. (2021). Sales Forecast of Manufacturing Companies using Machine Learning navigating the Pandemic like COVID-19. 2021 2nd International Conference on Computation, Automation and Knowledge Management (ICCAKM). doi:10.1109/iccakm50778.2021.9357751.

[5] Boone, T., Ganeshan, R., Hicks, R. L., & Sanders, N. R. (2018). Can Google Trends Improve Your Sales Forecast? Production and Operations Management, 27(10), 1770–1774. doi:10.1111/poms.12839.

[6] Moro, M. F., Weise, A. D., & Bornia, A. C. (2020). Model hybrid for sales forecast for the housing market of São Paulo. Real Estate Management and Valuation, 28(3), 45–64. doi:10.1515/remav-2020-0023.

[7] Ma, J., Yu, J., & Gong, Z. (2019). EV Regional Market Sales Forecast Based On GABP Neural Network. IOP Conference Series: Earth and Environmental Science, 252, 032098. doi:10.1088/1755-1315/252/3/032098.

[8] Wang, S., & Yang, Y. (2021). M-GAN-XGBOOST model for sales prediction and precision marketing strategy making of each product in online stores. Data Technologies and Applications, 55(5), 749–770. doi:10.1108/DTA-11-2020-0286.

[9] Raiyani, A., Lathigara, A., & Mehta, H. (2021). Usage of time series forecasting model in Supply chain sales prediction. IOP Conference Series: Materials Science and Engineering, 1042(1), 012022. doi:10.1088/1757-899x/1042/1/012022.

[10] Takahashi, K., & Goto, Y. (2022). Embedding-Based Potential Sales Forecasting of Bread Product. Journal of Advanced Computational Intelligence and Intelligent Informatics, 26(2), 236–246. doi:10.20965/jaciii.2022.p0236.

[11] Zhan, M., Gao, H., Liu, H., Peng, Y., Lu, D., & Zhu, H. (2021). Identifying market structure to monitor product competition using a consumer-behavior-based intelligence model. Asia Pacific Journal of Marketing and Logistics, 33(1), 99–123. doi:10.1108/APJML-08-2019-0497.

[12] Liu, Y., Yang, X., Zhu, C., & Meng, J. (2021). Drugs Sale Forecasting Based on SVR Integrated Promotion Factors. Journal of Physics: Conference Series, 1910(1), 012056. doi:10.1088/1742-6596/1910/1/012056.

[13] Prakash, K. V., & Yogeswari, K. (2022). A Structural Equation Modelling (SEM) approach on Construction Noise Related Perception and Behavior among the Workforces. Civil Engineering Journal, 8(12), 2721-2733. doi:10.28991/CEJ-2022-08-12-04.

[14] Kumar, P., & Singh, A. K. (2022). A comparison between MLR, MARS, SVR and RF techniques: hydrological time-series modeling. Journal of Human, Earth, and Future, 3(1), 90-98. doi:10.28991/HEF-2022-03-01-07.

[15] Mendonça, Y. V., Naranjo, P. G. V., & Pinto, D. C. (2023). The Role of Technology in the Learning Process: A Decision Tree-Based Model Using Machine Learning. Emerging Science Journal, 6(Special Issue), 280-295. doi:10.28991/ESJ-2022-SIED-020.

[16] Rusen, V., Krukonis, A., & Plonis, D. (2021). Prediction of Parameters of Semiconductor Band-pass Filters using Artificial Neural Network. 2020 IEEE 8th Workshop on Advances in Information, Electronic and Electrical Engineering (AIEEE). doi:10.1109/aieee51419.2021.9435748.

[17] Asaad, R. R., & Ali, R. I. (2019). Back Propagation Neural Network (BPNN) and Sigmoid Activation Function in Multi-Layer Networks. Academic Journal of Nawroz University, 8(4), 216–221. doi:10.25007/ajnu.v8n4a464.

[18] Wibawa, A. P., Izdihar, Z. N., Utama, A. B. P., Hernandez, L., & Haviluddin. (2021). Min-Max Backpropagation Neural Network to Forecast e-Journal Visitors. 2021 International Conference on Artificial Intelligence in Information and Communication (ICAIIC). doi:10.1109/icaiic51459.2021.9415197.

[19] Jin, D., Liu, Y., Ma, X., & Song, Q. (2021). Long time prediction of human lower limb movement based on IPSO-BPNN. Journal of Physics: Conference Series, 1865(4), 042099. doi:10.1088/1742-6596/1865/4/042099.

[20] Rezaei, S., & Behnamian, J. (2022). Benders decomposition-based particle swarm optimization for competitive supply networks with a sustainable multi-agent platform and virtual alliances. Applied Soft Computing, 114, 1–32. doi:10.1016/j.asoc.2021.107985.

[21] Singh, S., Singh, N. J., & Gupta, A. (2022). System sizing of hybrid solar-fuel cell battery energy system using artificial bee colony algorithm with predator effect. International Journal of Energy Research, 46(5), 5847–5863. doi:10.1002/er.7526.

[22] Bai, J., Tian, M., & Li, J. (2022). Control System of Liquid Fertilizer Variable-Rate Fertilization Based on Beetle Antennae Search Algorithm. Processes, 10(2), 357. doi:10.3390/pr10020357.

[23] Umbarkar, A. J., Moon, L. R., & Sheth, P. D. (2018). Comparative Study of CEC'2013 Problem Using Dual Population Genetic Algorithm. International Journal of Information Engineering and Electronic Business, 10(5), 40–45. doi:10.5815/ijieeb.2018.05.06.

[24] Mahmoudabadi, A., Kanaani, M., & Ghazimahalleh, F. P. (2022). Modifying Hidden Layer in Neural Network Models to Improve Prediction Accuracy: A Combined Model for Estimating Stock Price. HighTech and Innovation Journal, 3(1), 45-55. doi:10.28991/HIJ-2022-03-01-05.

[25] Li, C., Bai, L., Liu, W., Yao, L., & Waller, S. T. (2022). Graph Neural Network for Robust Public Transit Demand Prediction. IEEE Transactions on Intelligent Transportation Systems, 23(5), 4086–4098. doi:10.1109/TITS.2020.3041234.

[26] Cheng, Y., Wang, C., Wu, J., Zhu, H., & Lee, C. K. M. (2022). Multi-dimensional recurrent neural network for remaining useful life prediction under variable operating conditions and multiple fault modes. Applied Soft Computing, 118, 108507. doi:10.1016/j.asoc.2022.108507.