

Research Article

Research on the Prediction of Popularity of News Dissemination Public Opinion Based on Data Mining

Jing Tian , Huayin Fan, and Zengwen Hou

Journalism & Communications College, Jilin Normal University, Changchun 130012, Jilin, China

Correspondence should be addressed to Jing Tian; tj80@jlnu.edu.cn

Received 9 November 2021; Accepted 22 December 2021; Published 29 March 2022

Academic Editor: Akshi Kumar

Copyright © 2022 Jing Tian et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The development of news public opinion presents the characteristics of dynamic changes, and its life cycle is generally relatively short. For news public opinion to be welcomed by everyone, that is, to become hot news, it must be able to spread to a large number of readers in a short time. And, some of its characteristic attributes must satisfy the interests of most users and arouse users' desire to read. Therefore, it is particularly important to extract and study these characteristic attributes that determine the popularity of news public opinion and finally establish a model to describe the relationship between the popularity of news public opinion and these characteristic attributes. Based on data mining, this article mainly studies the popularity of news public opinion from two aspects. First is the sampling and attribute feature extraction of news. Then, considering the nonlinear relationship between the features, an improved principal component analysis method is proposed to analyze the correlation of the features. This can select important features from many irrelevant features and effectively reduce the original high-dimensional features. Second, the application of neural network in the prediction of news public opinion is studied. BP is an efficient data mining method. Considering BP network has some shortcomings. This work uses an improved particle swarm optimization to optimize the initial parameters for BP network, which can compensate for the defects of BP network. After that, the BP network with optimized parameters is used to establish a prediction model for the popularity of news public opinion. The experimental results prove that the neural network model proposed can accurately predict the popularity of news public opinion.

1. Introduction

The term news was first born in the Tang Dynasty and refers to the time that was recently heard or what happened recently in society. The meaning of the word news has not changed much after more than a thousand years of development. However, the dissemination of news has undergone earth-shaking changes after more than a thousand years of historical changes. Traditional news dissemination, from small-scale oral transmission to newspapers, to wireless broadcasting, and to TV broadcasting, blooms in many ways and has various channels. Nowadays, the news is presented on everyone's mobile terminal, and traditional news dissemination methods have gradually withdrawn from the market [1–5].

With the rapid development of the Internet, the Internet has become a part of people's lives. People use mobile

phones to receive information and view news anytime and anywhere on the Internet. Compared with traditional news dissemination, the Internet not only brings more traffic but also shortens the generation, creation, and dissemination time of news exponentially. The data shows that more than 2,600,000 news are shared every minute on Facebook, the average number of online videos shared every minute on LINE is 8,333, and the number of news links shared every minute on Twitter is close to 300,000. Interconnection has detonated the information age, and news hotspots have been generated all the time. Traditional news dissemination is restricted by channels, and the audience is limited. Even if it is a hot topic, its influence is also limited. But nowadays, the coverage of breaking news is so wide that it is unimaginable [6–10].

In the Internet age, news and public opinions cover all aspects of our lives. The real-time, uncertain, and rapid

spread of the Internet make the characteristics of online news public opinion show the characteristics of the new era such as randomness, suddenness, and concealment. In such an era, everyone on the Internet can become a publisher and disseminator of information, as well as a creator of news and public opinion. Of course, news and public opinion are not all good, and the domestic Internet environment is not a pure land, which is full of negative, distorted, extreme, inciting violence and division factors. Popular news, especially breaking news, as a powerful means of communication, is not only conducive to the rapid dissemination of information but also conducive to the transfer of value. Under such a background, predictive analysis of the popularity of news and public opinion is beneficial to the healthy development of the Internet and curbs negative information such as rumors and violence, at the same time, producing news or advertisements that can arouse people's general goodwill. It has great value for many Internet publishers, portal websites, and Internet advertising companies. The development of news public opinion presents the characteristics of dynamic changes, and its life cycle is generally relatively short. If a piece of news is to be welcomed by everyone, it becomes hot news. It must be able to spread to a large number of readers in a relatively short period of time, and some of its characteristic attributes must satisfy the interests of most users and arouse users' desire to read. Therefore, it is particularly important to extract and study these characteristic attributes that determine the popularity of online news, analyze their interrelationships, and finally, establish a model to describe the relationship between the popularity of news public opinion and these characteristic attributes [11–15].

The contributions are as follows: (1) This work proposes an improved principal component analysis method to analyze the correlation of features. In this way, important features can be selected from many irrelevant features, effectively reducing the dimensionality of the original high-dimensional features. (2) This work uses an improved particle swarm optimization to optimize the initial parameters for the BP network. It can make up for the shortcomings of BP and use the BP network with optimized parameters to establish a prediction model for the popularity of news and public opinion.

2. Related Work

Compared with traditional news and advertising industries, Internet news, with its low-cost, fast-spreading, and wide-ranging features, had a huge impact on the traditional news industry and had shown initial advantages. Due to the real-time characteristics of news and the guiding characteristics of public opinion, accurately predicting the popularity of a news manuscript for the roles of journalists, content providers, advertisers, etc., had become a new research hotspot. However, predicting the popularity of news articles on the Internet was a challenging task. First of all, the content of these articles was only available online. The structure of the social network that spreads news was different, which would affect the probability that users can see this article. In addition, the coverage of the content of the article was also a

factor that affects the prediction of the results. For example, articles that were closely related to the general public on a global scale are more popular than articles that only targeted a small number of social groups.

For this challenging work, a lot of research had been devoted to it, and a lot of questions had been raised. Literature [16] put forward the problem of tracking the spread of online topics. Literature [17] proposed detecting the public opinion orientation of online communities. The literature [18] put forward the field of prediction of the popularity of social media. There were many different ways of expressing the concept of popularity. For example, our most common classic method was the number of clicks on a piece of news, but this information was of little significance to researchers. Because in fact, due to the existence of web crawlers and search engines, it was difficult for us to estimate the number of times, an article was actually requested by users, which would directly lead to the inaccuracy of the original data. In addition, there were some other methods based on the user's own participation, such as comments, voting, and sharing via email or WeChat. Predicting the popularity of news articles was a difficult task, and a lot of research results had been accumulated over the years. Literature [19] proposed a model to predict the popularity of news published by science and technology information websites. This method believed that the popularity of articles will gradually increase over time. Literature [20] proposed two algorithms for predicting video popularity and news popularity and achieved good prediction results. Literature [21] proposed a linear logarithm method, which can meet the forecasting demand in a specific range. Literature [22] proposed a different way of thinking. He established a new prediction model based on the social influence of news and the characteristics of the publishing platform. The disadvantage of this type of algorithm was that it could only be applied to predictive analysis in a specific field but failed to analyze articles on different topics. Literature [23] proposed a new forecasting idea, which focused on the prediction of the popularity of certain news that would continue to be paid attention to after a certain period of time. Literature [24] proposed a method. Using the number of news sharing as a popular indicator, a clustering algorithm was used to divide different articles into 4 categories according to their popularity. Literature [25] proposed a new idea to predict the popularity of online news and ranked it according to the popularity. The disadvantage of this type of method was that it could not give an accurate prediction value.

Driven by the Internet industry chain, Internet + had become an industrialized concept, and network news was quickly spread through the Internet as a carrier. Data showed that more than 26,000 news were shared per minute on Facebook, the average number of online videos shared per minute on LINE was 8333, and the number of news links shared per minute on Twitter was close to 3,000. Clarifying how popular the public was for a piece of news has become a new branch of research.

At present, many models for tracking and forecasting online news based on time series had been proposed. These models assumed that the evaluation of popularity was time-

sensitive. The number of users of YouTube videos fluctuated continuously within 24 hours, and the number of their shares was also constantly changing. The literature [26] aimed at analyzing the life cycle of online news; this research used a threshold value to compare and calculate the number of news shares in 7 days. At the same time, the feature extraction method was another subtle branch in the field of social news research. It mainly focused on the news itself and removed some possible influence factors, such as time, comment, and emotional color. The reason why time was eliminated was that the number of news sharing gradually increases with the passage of time. But its impact on users would gradually become zero after a certain point in time. In addition, the results of this research field were diversified because of the existence of various media forms, including videos, music records, and plain text information [27].

3. Method

3.1. Improved Principal Component Analysis

3.1.1. Algorithm Background. When predicting the popularity of news public opinion, it is necessary to extract the characteristics of the news. Under normal circumstances, these features are high-dimensional, and processing these features is more complicated. Therefore, it is necessary to reduce the dimensionality of the features before making subsequent predictions. Principal component analysis has become a classic and effective statistical dimensionality reduction method. It has been successfully applied to linearly distributed data for a variety of practical problems and has also achieved remarkable results. At the same time, the algorithm for further expansion of principal component analysis has also been effectively developed.

First of all, through research, we can know that sparse principal component analysis is based on principal component analysis, which mainly transforms the problem of solving the principal component load into a ridge regression problem. A regularization penalty term is added to the essence of such an optimization regression problem to compare the load of each principal component. This promotes a load of some relatively weak variables to approach zero, achieving a sparse effect. At present, there are generally two types of penalty items in sparse principal component analysis: L_1 penalty items and TL_1 penalty items. The L_1 penalty term was originally proposed by the literature [28] when the L_1 penalty term was combined with the ridge regression term in the optimization problem transformed by principal component analysis. It can constitute an elastic net penalty item. The penalty characteristic of the elastic net penalty item is between the L_1 penalty item and the L_2 penalty item. The TL_1 penalty term is based on the literature [29] with a certain transformation in the form of the L_1 penalty term, and new parameters are added to the numerator and denominator to control its characteristics between the L_0 penalty term and the L_1 penalty term. However, the essence of sparse principal component analysis is still aimed at processing linear data, and the processing of nonlinear data has not yet achieved the desired effect.

In view of the situation that nonlinear data cannot be handled well, kernel principal component analysis can solve this problem well. For high-dimensional feature space, this method uses the kernel function to project sample points from the original space with a linear distribution. This method and PCA have been combined naturally. This obtains sample points that are linearly distributed in high-dimensional space and finally uses principal component analysis in this space. Through the addition of the kernel function, the principal component analysis problem of nonlinear data is successfully solved. At the same time, it avoids the difficulty of not being able to obtain an explicit projection function by constructing a kernel matrix. This successfully transformed the way of analyzing the data covariance matrix into the way of analyzing the kernel matrix.

In this section, this article will focus on giving another extension method based on principal component analysis. This method is also an improved method based on the thinking of other expansion methods of principal component analysis in this article. The sparse kernel principal component analysis based on the TL_1 penalty term is based on the thinking of sparse penalty and kernel principal component analysis. After learning sparse principal component analysis and kernel principal component analysis, it is logical to think of merging the two to get a sparse kernel principal component analysis in a high-dimensional space. In this way, the linear transformation of nonlinear data can be realized, and the sparse principal components can also be obtained in the feature space. This highlights the important variables and weakens the influence of some variables on the overall result, which has a simple and robust effect. Therefore, based on the method research on the principal component analysis, this paper proposes a sparse kernel principal component analysis.

3.1.2. Improvement Strategy. The sparse kernel principal component analysis with TL_1 penalty term combines the characteristics of the kernel principal component analysis as well as the sparse penalty, which can be implemented on nonlinearly distributed data. Its idea is similar to the sparse principal component analysis. What is considered in the sparse principal component analysis is to compare and sparse the linearly distributed data subjected to principal component analysis in the original space. When the covariance matrix is used to decompose to obtain the load, it is converted into an optimization problem, and then a penalty term is added to approach the partial load to zero. Therefore, through similar considerations, the original variables are transformed in the nuclear principal component analysis. At the same time, the eigenvalue decomposition of the sample point covariance matrix in the high-dimensional eigenspace is required. However, since the projection function cannot be obtained explicitly, it is finally transformed into the decomposition of the kernel matrix. Therefore, when the kernel matrix is eigen-decomposed, it can be transformed into an optimization problem. Then, we add the TL_1 penalty term to this optimization problem, so that the new coordinates of the sample points in the feature space are sparse.

The characteristic equation for solving the kernel principal component analysis is

$$Kv_i = f_i v_i, \quad (1)$$

where K is the kernel matrix and f_i is the eigenvalue corresponding to the eigenvector v_i .

Then, the above equation can be written as an optimized regression model:

$$\hat{v}_i = \arg \min_{v_i} K - f_i v_i v_i^{T2}. \quad (2)$$

Assuming $v_i = f_i v_i$ and $\alpha_i = v_i$, then equation (2) can be written as

$$(\hat{\alpha}_i, \hat{v}_i) = \arg \min K - v_i \alpha_i^{T2}. \quad (3)$$

So far, this paper has transformed the principal component loading problem in nuclear principal component analysis into a regression optimization model. Therefore, the TL_1 penalty term can be added to the above optimization problem, so that v_i can be sparse. The model based on TL_1 penalty term sparse kernel principal component analysis is obtained as

$$\hat{v}_i = \arg \min_{v_i} K - v_i \alpha_i^{T2} + \lambda_1 \frac{(b+1)v_{i1}}{b+v_{i1}}, \quad (4)$$

where λ_1 and b are the parameters that need to be customized.

So far, this paper has given a sparse kernel principal component analysis model based on the TL_1 penalty term. It can be seen from this model that its idea is similar to sparse principal component analysis. The biggest difference is that sparse principal component analysis is aimed at the covariance matrix of the data. The sparse kernel principal component analysis based on the TL_1 penalty term is aimed at the mapped kernel matrix. From the perspective of the threshold iteration method, the threshold iteration algorithm for solving the sparse kernel principal component analysis based on the TL_1 penalty term is as follows:

Step 1: normalize the input matrix to get the output matrix X . Step 2: use the kernel method to process the matrix X to obtain the kernel matrix K in the high-dimensional space. Step 3: use general principal component analysis on the matrix K to get the load matrix V . Step 4: let A' be the first k -dimensional column vector of the load matrix V ; that is, $A' = V[:, 1:k]$. Step 5: let the initial coefficient of the i -th principal component be x_i , given reasonable parameter values. Step 6: compare the size relationship between the parameters, and then perform the corresponding processing. Step 7: repeat Step 4 and Step 5; stop iterating when the conditions are met.

Therefore, this paper presents the basic model and algorithm of sparse kernel principal component analysis with TL_1 penalty terms. It can be seen from the above steps that the penalty functions used are all TL_1 penalty items. The process is basically similar to the solution of TL_1 penalty

term sparse principal component analysis. However, the most essential difference between the sparse kernel principal component analysis with TL_1 penalty term and the sparse principal component analysis with the TL_1 penalty term in the solution process is whether the original data has undergone a nonlinear transformation, that is to say, whether the data is mapped to the high-dimensional space. If it has undergone nonlinear processing, it is necessary to decompose and iterate the kernel matrix. Without this step, you only need to iteratively decompose the covariance matrix.

In general, an iterative algorithm is a more direct algorithm for solving optimization models. Iterative algorithms do not need to give a specific algebraic solution. Instead, it reaches a local optimal or global optimal solution by going through multiple cycles and iterations in a certain direction until it converges. However, there are many aspects that need attention in iterative algorithms. First, the selection of parameters in the iterative algorithm plays an important role and may even affect the selection of the final optimal solution. Therefore, a suitable method is needed to help the selection of parameters. Secondly, because the iterative algorithm achieves the purpose of solving by means of continuous loop iterations, so it is deeply affected by the number of iterations and accuracy. There may be a relatively large amount of calculation, especially when iterating all the elements in each vector of a matrix. Therefore, it can determine the appropriate threshold and number of iterations.

3.2. Improved BP Network

3.2.1. Improved Particle Swarm Optimization. After using the proposed principal component analysis method to reduce the dimensionality of the features, this paper uses the BP network to model these features to predict the popularity of news public opinion.

Prediction accuracy is low due to random selection of initial threshold and weight parameters and the BP network's tendency to slip into the local minimum state. This results in poor performance. BP neural network model predictions can be enhanced by using an upgraded particle swarm technique to optimize the starting parameters. Model predictions can be made using a more accurate BP model that has been optimized for use in this study.

Particle swarm optimization (PSO) is an algorithm based on swarms that is extremely clever. The algorithm is easy to develop, is precise, and quickly converges to the best solution. It can effectively optimize various functions and is suitable for application in multiobjective constrained optimization problems. And it has better efficiency and practicability in solving various problems. At present, it has achieved good application results in constrained optimization, function optimization, various engineering design problems, and other fields.

Based on population dynamics, particle swarm optimization (PSO) is a population-based approach for solving optimization problems. One of its most fundamental aspects involves the creation of an initial group of particles in space using a randomized search process. Iteration is used to find

the best solution for these particles, which begins with a random solution. There are three ways to express the attributes of a particle in an optimization problem: position, speed, and the fitness value. The fitness function can be used to calculate the particle's fitness value, which can be used to determine the particle's advantages and disadvantages. During each repetition, the particle updates itself by observing two points of departure. The individual extreme value is the one being sought by the particle, whereas the global extreme value is the one being sought by the entire group at the same time. Particles modify their speed and position by following these two extreme values, which each particle follows.

The basic principle of PSO mainly includes the initialization of particle swarm, the calculation for particle fitness, the update of particle speed and position, and the determination of termination conditions. The specific steps are as follows:

Step 1: initialize the particle swarm. Step 2: calculate the objective function value, namely, fitness. Step 3: calculate fitness after the location update. The fitness is analyzed and compared with its individual extreme value P_{best} . If $fitness > P_{best}$, replace P_{best} with fitness. Step 4: compare fitness with the global extreme value G_{best} . If $fitness > G_{best}$, replace G_{best} with fitness. Step 5: update the speed and position. Step 6: judgment of termination conditions. If the algorithm progresses until the end condition is met, the iterative calculation process is stopped, and the result is output at this time. Otherwise, the algorithm returns to Step 2.

A particle swarm algorithm can better solve combinatorial optimization problems. But this algorithm also has some shortcomings. Algorithm search accuracy is not very great and it is easy to get caught in the trap of local minimum solutions when using it. A modified particle swarm method using inertia weight has been developed to address these problems; in this way, the algorithm's search performance can be improved and it can avoid getting stuck in a local optimal scenario.

Among the various parameters of PSO, inertia weight w is an important parameter. This parameter affects the local optimal ability as well as the global optimal ability, and its value can reflect the speed of the particle's flight. If, in the algorithm parameter setting, the value for inertia weight w increases, the flying speed will increase, which can enhance the global search capability. On the contrary, if the value of w is reduced, the flying speed of the particles will slow down, which is beneficial to improve the local searchability. Therefore, it is necessary to select an accurate inertia weight w , so that the particles have balanced searchability, and effectively improve the exploration ability of the algorithm.

At present, many scholars have proposed improved methods for inertial weights, such as random weighting algorithms and fuzzy inertial weighting algorithms. One of the more common algorithms uses the method of linearly decreasing weights. Aiming at the characteristic that the PSO algorithm is easy to mature prematurely, literature [30] proposed a linear decreasing weight method. By introducing

a linear transformation weight, the inertia weight decreases linearly from large to small. The formula for change with the number of iterations is as follows:

$$w = \frac{w_{\max} - (w_{\max} - w_{\min})t}{t_{\max}}, \quad (5)$$

where w_{\max} is the maximum inertia weight, w_{\min} is the minimum inertia weight, t represents the current iteration, and t_{\max} is the maximum iteration.

Through analysis, although the introduction of linear inertia weight can make the PSO algorithm adjust the particle's optimization ability, this algorithm has some shortcomings. First of all, in the early stage of operation, due to the large value of the inertial weight w parameter, it is conducive to the global search of particles. If the optimal point is detected in the early stage of the algorithm, it is hoped that the algorithm will quickly converge to its optimal point. However, the linear decreasing characteristic of the inertia weight w slows down the convergence speed of the algorithm. Secondly, in the late running of the algorithm, as w decreases, the global search ability gradually decreases, which is easy to fall into the local optimal situation at this time.

Aiming at the shortcomings, an improved method of nonlinear decrease of inertia weight is proposed on the basis of the algorithm to improve the shortcomings. This algorithm is described as

$$w = w_{\max} - (w_{\max} - w_{\min}) \times \tan\left(\frac{\pi t}{4t_{\max}}\right). \quad (6)$$

In the iterative process of the algorithm, when t is small, the global search capability can be guaranteed. Then, with the increase of t , local search ability can be guaranteed, so the algorithm has strong balanced searchability. Through the above-mentioned improved method, the nonlinear dynamic inertia weights better realize the balance between the global search and the local search. This improves the performance in convergence speed and global optimization.

3.2.2. Improvement Strategy for BP. BP network is an algorithm based on gradient descent, as shown in Figure 1. As a result, the training effect of the network model will be negatively impacted by its inability to search globally. This cannot provide a more effective training method. Global optimization and simultaneous search are two features of the particle swarm algorithm. This has the potential to enhance global convergence. Neural networks can be improved by using this technique in the neural network algorithm. The particle swarm algorithm's global optimization ability is merged with the neural network's advantages by comparing their properties. It is possible to improve the learning and training speed of the network by establishing a BP neural network model using the particle swarm algorithm and optimizing the initial weights and threshold parameters of the BP network.

PSO may be used to optimize the initial parameters of a BP neural network model by integrating the particle swarm

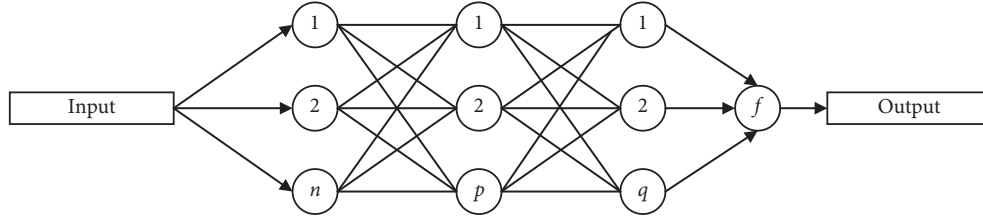


FIGURE 1: Structure of the BP network.

method into the model. Iterative particle optimization can be used to find the best parameters for the BP network. When it comes to the PSO algorithm, the BP neural network weights and thresholds are the flying particles of dimension two. In order to find the most optimal particles, you must first select the best initial parameters for the BP network. Additionally, iterative particles constantly adjust their velocity, position, and initial characteristics in space. Searching for the ideal particle is done in accordance with the function's optimization aim so that the particle's optimal parameter can be determined. The fitness value of a particle is the difference between the projected output of the BP model and the expected output, which is the optimization objective function. The fitness function is

$$f = \frac{1}{n} \sum_{i=1}^n |y_i - p_i|, \quad (7)$$

where n is the number of samples, y_i is the actual output, and p_i is the predicted output.

Through the continuous update for the position, velocity, weight, and threshold, we determine the parameter value when the particle obtains the optimal fitness, assign the parameter value, and use the parameter optimized model for learning and training.

The principle of using the PSO algorithm to optimize the BP network algorithm mainly includes three parts: determining BP network structure, the PSO algorithm optimizing BP network parameters, and BP network training and prediction. The specific steps are as follows. (1) First, initialize the parameters for the model, including the number of network layers and learning rate of BP network, the number of populations of particle swarm algorithm, inertia weight, and the number of iterations. (2) Calculate and compare the fitness value. The output error of BP network model training is used as the fitness objective function to obtain the fitness value. (3) According to the fitness function, the improved particle swarm algorithm is used to iteratively search the space and update speed and position. (4) Determination of the termination condition of particle population update: when the algorithm reaches the maximum iterations, the loop iterative calculation process is stopped, the optimal fitness value is determined, and the optimal initial parameters are obtained. (5) Assign the optimized parameters to the BP network as initial weight and threshold parameters of the BP model. Combined with the sample data of news and public opinion, the optimized BP model is used to train until the entire model meets the convergence accuracy. (6) Use the trained BP model to predict the popularity of news public opinion.

4. Experiment and Discussion

4.1. Dataset. This article self-made two data sets DS1 and DS2. DS1 comes from online news published by Mashable Publishing House, and a total of 1389 news data are selected. DS2 comes from news published by Facebook, and a total of 1473 news data are selected. The specific distribution of the training set and test set is shown in Table 1. For each piece of news data, a 60-dimensional feature is extracted, and the dimensionality is reduced to a 20-dimensional feature vector through an improved principal component analysis algorithm. Each piece of news is divided into two categories: popular and unpopular. In this work, precision, recall, and F1 score are utilized to evaluate the performance of prediction.

4.2. Evaluation of Model Convergence. In the BP network, whether the model converges is an important indicator for evaluating network performance. If the model fails to converge, subsequent predictions are meaningless. Therefore, this article first compares the training loss of the proposed network on two data sets. Experimental results are illustrated in Figure 2.

As the training progresses, the loss of the network gradually decreases. And at the 40th epoch, the loss is basically not decreasing, indicating that the network has reached a state of convergence. Besides the training loss, this work also compares the test performance on two datasets. Experimental results are illustrated in Figure 3.

Similarly, when the training iteration is 40 epochs, the performance of the network on the test set tends to converge, which can obtain 0.91 precision, 0.82 recall, and 0.86 F1 on DS1, 0.93 precision, 0.85 recall, and 0.88 F1 on DS2. In summary, the designed network can finally converge and make stable and efficient predictions.

4.3. Evaluation on Improved PCA. In this work, an improved PCA is proposed to reduce the dimensionality of features. To prove that this method can effectively improve the performance, the performance when using the improved PCA (IPCA) is compared with that of the improved BP (IBP) network alone. The experimental results are illustrated in Figure 4.

Combining the IPCA method with the BP network can effectively improve the performance. On the DS1 dataset, compared with not using the IPCA method, the performance improvement obtained is 6%, 5%, and 5% on three indexes. On the DS2 dataset, compared with not using the

TABLE 1: The division of training set and test set.

Dataset	Training set	Test set
DS1	1041	348
DS2	1104	369

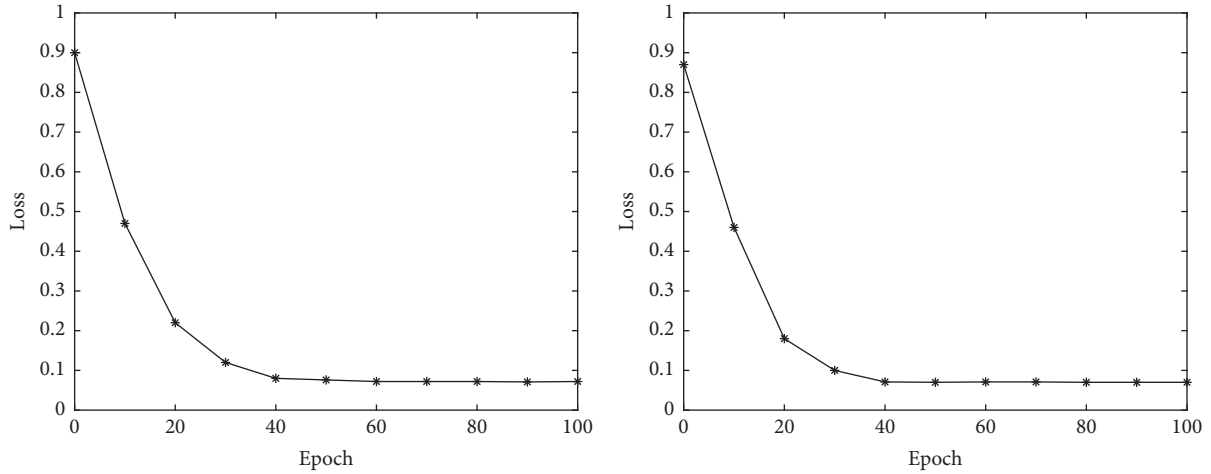


FIGURE 2: The training loss on DS1 and DS2.

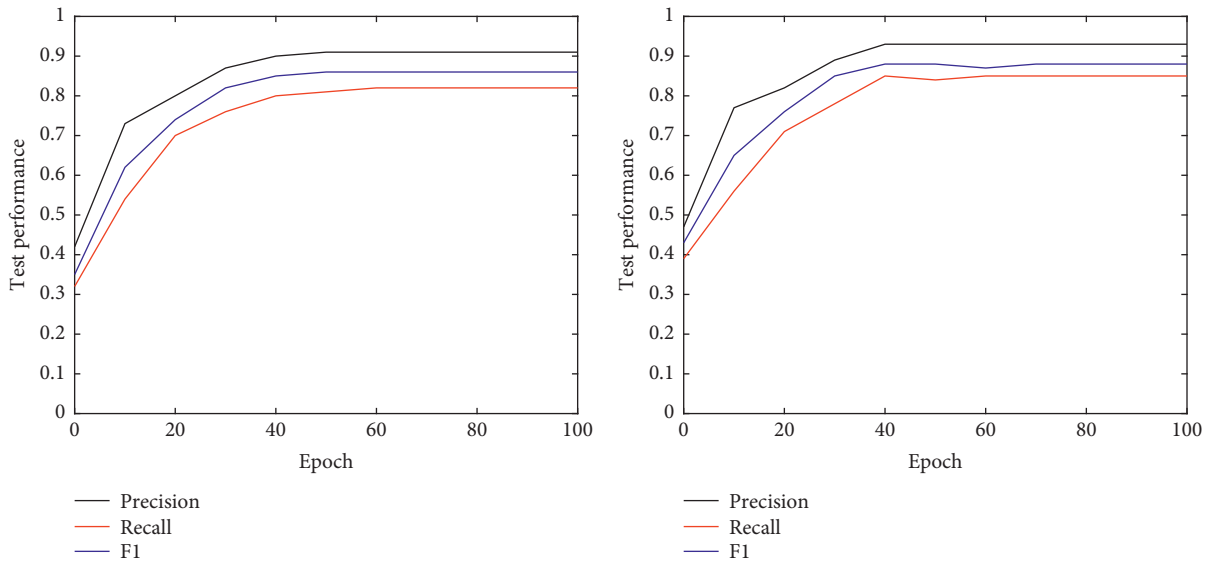


FIGURE 3: The test performance on DS1 and DS2.

IPCA method, the performance improvement obtained is 7%, 6%, and 5% on three indexes. To further illustrate the superiority of the IPCA method, it is compared with the traditional PCA. Experimental results are illustrated in Table 2.

Obviously, compared with the PCA algorithm, the IPCA method can further improve the performance of the BP network. This proves the reliability and effectiveness of the IPCA method.

4.4. Evaluation of Improved PSO. In this paper, an improved PSO is proposed to optimize the BP network. To prove that this method can effectively improve the performance of the BP network, the performance when using the improved PSO (IPSO) is compared with that of the BP network alone. The experimental results are illustrated in Figure 5.

It can be seen that combining the IPSO method with the BP network can improve performance. On the DS1 dataset, compared with not using the BP method, the performance

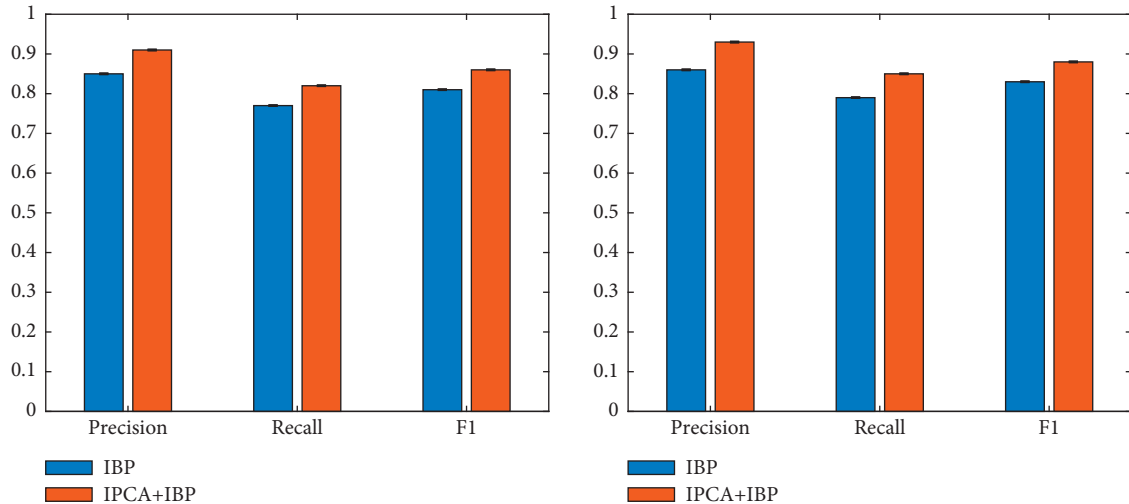


FIGURE 4: Evaluation of improved PCA.

TABLE 2: Comparison of IPCA and PCA.

Method	DS1			DS2		
	Precision	Recall	<i>F1</i>	Precision	Recall	<i>F1</i>
PCA + IBP	0.89	0.80	0.83	0.90	0.82	0.84
IPCA + IBP	0.91	0.82	0.86	0.93	0.85	0.88

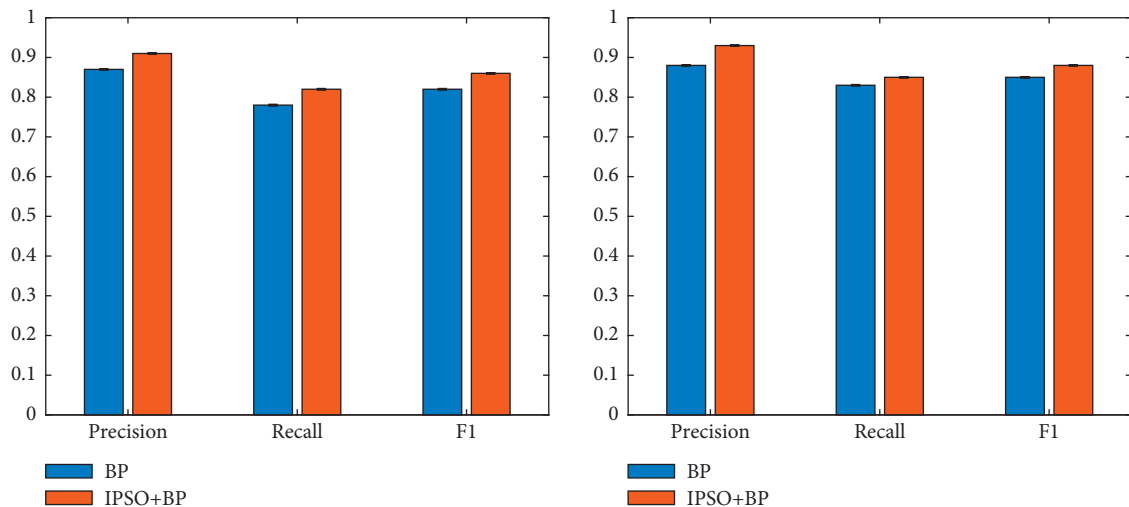


FIGURE 5: Evaluation of improved PSO.

improvement obtained is 4%, 4%, and 4% on three indexes. On the DS2 dataset, compared with not using the IPSO method, the performance improvement obtained is 5%, 2%, and 3% on three indexes. To further illustrate the superiority of the IPSO method, it is compared with the traditional PSO method. Experimental results are illustrated in Table 3.

Obviously, compared with PSO, the IPSO method can further improve the performance of the BP network. This proves the reliability and effectiveness of the IPSO method in this paper.

TABLE 3: Comparison of IPCA and PCA.

Method	DS1			DS2		
	Precision	Recall	<i>F1</i>	Precision	Recall	<i>F1</i>
PSO + BP	0.88	0.80	0.84	0.91	0.84	0.85
IPSO + BP	0.91	0.82	0.86	0.93	0.85	0.88

4.5. Comparison to Other Methods. To illustrate the effectiveness of our strategy, this article compares our method with other methods. Experimental results are illustrated in Table 4.

TABLE 4: Comparison to other methods.

Method	DS1			DS2		
	Precision	Recall	F1	Precision	Recall	F1
DT	0.75	0.67	0.71	0.79	0.69	0.73
SVM	0.81	0.72	0.75	0.81	0.78	0.75
BP	0.85	0.77	0.79	0.86	0.83	0.82
Ours	0.91	0.82	0.86	0.93	0.85	0.88

Obviously, compared with other methods, our method can obtain the best performance improvement. This demonstrates the correctness of the method in this article.

5. Conclusion

With the development of the Internet, news and public opinion have exploded. It is a very important subject to extract and study the characteristic attributes that determine the popularity of news public opinion and establish a network model to predict the popularity of news public opinion. The main work of this paper has the following two contents. The first is to realize the sampling of news and the extraction of attribute features. Then, considering the nonlinear relationship between features, an improved principal component analysis method is proposed to analyze the correlation of features. In this way, important features can be selected from many irrelevant features, effectively reducing the original high-dimensional features. Second, neural networks are examined for their ability to anticipate public opinion on news stories. By using BP, we can mine data more effectively. The BP neural network, on the other hand, has a number of limitations and uses an improved approach for optimizing the initial parameters of the BP neural network, which successfully compensates for the BP neural network's flaws. Finally, a prediction model for the popularity of news public opinion is built using the BP network with optimal parameters. The new neural network model described in this paper can accurately predict the popularity of news public opinion, as demonstrated by experimental findings.

Data Availability

The datasets used are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Jilin Provincial Planning Fund Office of Philosophy and Social Science, under projects "Research on Promoting the Four Forces in the Communication of Local News in the Reporting of Major Public Events" (Project no. 2020WB6) and "Research on the Development of the "Internet Plus" Cultural Industry in Jilin" (Project no. 2019wt20).

References

- [1] T. C. Serm, P. Blanchfield, and K. Su, "Mobile newspaper development framework: guidelines for newspaper companies for creating usable mobile news portals," in *Proceedings of the 2006 International Conference on Computing & Informatics (ICOCI 2006)*, pp. 1–8, IEEE, Kuala Lumpur, Malaysia, June 2006.
- [2] W. Wang, "News development in the 5G network era based on machine learning and FPGA," *Microprocessors and Microsystems*, Article ID 103391, 2020.
- [3] B. H. Shah, M. Adnan, and I. Sultana, "Development news in print media of Pakistan: the case of Urdu and English press," *SJESR*, vol. 3, no. 2, pp. 244–251, 2020.
- [4] C. A. Watson, "Information literacy in a fake/false news world: an overview of the characteristics of fake news and its historical development," *International Journal of Legal Information*, vol. 46, no. 2, pp. 93–96, 2018.
- [5] B. P. Putra, K. Ansari, and I. Pramuniati, "The development of news text-based problems teaching materials in 12th grade students of vocational high school at TR sinar husni of 2019/2020," *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal*, vol. 3, no. 1, pp. 293–305, 2020.
- [6] J. M. Cleaton, C. Viboud, L. Simonsen, A. M. Hurtado, and G. Chowell, "Characterizing ebola transmission patterns based on internet news reports," *Clinical Infectious Diseases*, vol. 62, no. 1, pp. 24–31, 2016.
- [7] Y. Zhang, W. Song, D. Shen, and W. Zhang, "Market reaction to internet news: information diffusion and price pressure," *Economic Modelling*, vol. 56, pp. 43–49, 2016.
- [8] J. W. Ayers, B. M. Althouse, M. Dredze, E. C. Leas, and S. M. Noar, "News and internet searches about human immunodeficiency virus after charlie sheen's disclosure," *JAMA Internal Medicine*, vol. 176, no. 4, p. 552, 2016.
- [9] J. Pei, K. Zhong, and J. Li, "ECNN: evaluating a cluster-neural network model for city innovation capability," *Neural Computing & Applications*, pp. 1–13, 2021.
- [10] A. C. Micu and I. Pentina, "Examining search as opposed to experience goods when investigating synergies of internet news articles and banner ads," *Internet Research*, vol. 25, no. 3, pp. 378–398, 2015.
- [11] Y. Huang, Z. Yu, J. Guo, Z. Yu, and Y. Xian, "Legal public opinion news abstractive summarization by incorporating topic information," *International Journal of Machine Learning and Cybernetics*, vol. 11, no. 9, pp. 2039–2050, 2020.
- [12] C. Yasai, "The formation of Chinese online public opinion on the 7.23 wenzhou train crash: through the analysis of sina news site, sina weibo, and tianya BBS," *Journal of Mass Communication Studies*, vol. 86, pp. 123–142, 2015.
- [13] M. Yousaf, W. Li, and Z. Yousuf, "The news media exposure and polarization of public opinion: the limits of social functions of the agenda-setting role of mass media," *Global Media Journal*, vol. 13, no. 1, pp. 36–51, 2020.
- [14] O. C. Dewi and R. M. Amalia, "Construction of public opinion in online news on the issue of nurses' expulsion from boarding house: a critical discourse analysis," *Teknosastik*, vol. 19, no. 1, p. 48, 2021.
- [15] I.-H. Hong, "Analysis of public opinion on sports by the 4th industrial revolution using newspaper big data," *Korean Journal of Sport Management*, vol. 25, no. 3, pp. 1–9, 2020.
- [16] J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in *Proceedings of the ACM SIGKDD International Conference on Knowledge*

- Discovery and Data Mining*, pp. 497–506, Paris, France, June 2009.
- [17] L. A. Adamic and N. Glance, “The political blogosphere and the 2004 US election: divided they blog,” in *Proceedings of the International Workshop on Link Discovery*, pp. 36–43, Chicago, IL, USA, August 2005.
 - [18] M. Tsagkias, W. Weerkamp, and M. De Rijke, “Predicting the volume of comments on online news stories,” in *Proceedings of the ACM Conference on Information and Knowledge Management*, pp. 1765–1768, Beijing, China, 2009.
 - [19] A. Kaltenbrunner, V. Gómez, and V. López, “Description and prediction of slashdot activity,” in *Proceedings of the Latin American Web Conference*, pp. 57–66, Santiago, Chile, 2007.
 - [20] G. Szabo and B. A. Huberman, “Predicting the popularity of online content,” *Communications of the ACM*, vol. 53, no. 8, pp. 80–88, 2010.
 - [21] M. Tsagkias, M. De Rijke, and W. Weerkamp, “Linking online news and social media,” in *Proceedings of the ACM International Conference on Web Search and Data Mining*, pp. 565–574, Hong Kong, China, 2011.
 - [22] K. Lerman and R. Ghosh, “Information contagion: an empirical study of the spread of news on digg and twitter social networks,” *Computer ENCE*, vol. 52, pp. 166–176, 2010.
 - [23] J. G. Lee, S. Moon, and K. Salamatian, “An approach to model and predict the popularity of online contents with explanatory factors,” in *Proceedings of the International Conference on Web Intelligence and Intelligent Agent Technology*, pp. 623–630, Toronto, Canada, August 2010.
 - [24] R. Bandari, S. Asur, and B. Huberman, “The pulse of news in social media: forecasting popularity,” in *Proceedings of the AAAI Conference on Web and Social Media*, pp. 12–17, Dublin, Ireland, 2012.
 - [25] A. Tatar, P. Antoniadis, M. D. d. Amorim, and S. Fdida, “From popularity prediction to ranking online news,” *Social Network Analysis and Mining*, vol. 4, no. 1, p. 174, 2014.
 - [26] M. Tsagkias, W. Weerkamp, and M. De Rijke, “News comments: exploring, modeling, and online prediction,” *Lecture Notes in Computer Science*, pp. 191–203, 2010.
 - [27] M. D. Buhmann, *Radial Basis Functions: Theory and Implementations*, Cambridge University Press, Cambridge, UK, 2003.
 - [28] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society: Series B*, vol. 58, no. 1, pp. 267–288, 1996.
 - [29] M. Nikolova, “Local strong homogeneity of a regularized estimator,” *SIAM Journal on Applied Mathematics*, vol. 61, no. 2, pp. 633–658, 2000.
 - [30] Y. Shi and R. Eberhart, “A modified particle swarm optimizer,” in *Proceedings of the IEEE International Conference on Evolutionary Computation*, pp. 69–73, Anchorage, AK, USA, 1998.