



Construction of the Social Network Information Dissemination Index System Based on CNNs

Weihong Han^{1†}, Linhe Xiao^{1†}, Xiaobo Wu², Daihai He³, Zhen Wang⁴ and Shudong Li^{1*}

¹Cyber Space Institute of Advanced Technology, Guangzhou University, Guangzhou, China, ²School of Computer Science and Cyber Engineering, Guangzhou University, Guangzhou, China, ³Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, China, ⁴School of Artificial Intelligence, Optics and Electronics (IOPEN), Northwestern Polytechnical University, Xi'an, China

OPEN ACCESS

Edited by:

Zhanwei Du,
University of Texas at Austin,
United States

Reviewed by:

Lixiang Li,
Beijing University of Posts and
Telecommunications (BUPT), China
Yan Li,
BP America, United States

*Correspondence:

Shudong Li
lishudong@gzhu.edu.cn

[†]These authors have contributed
equally to this work and share first
authorship

Specialty section:

This article was submitted to
Social Physics,
a section of the journal
Frontiers in Physics

Received: 01 November 2021

Accepted: 28 January 2022

Published: 07 March 2022

Citation:

Han W, Xiao L, Wu X, He D, Wang Z
and Li S (2022) Construction of the
Social Network Information
Dissemination Index System Based
on CNNs.
Front. Phys. 10:807099.
doi: 10.3389/fphy.2022.807099

The information dissemination index system is an effective way to measure the dissemination of public opinion events in social networks. Due to the complexity, variability, and asymmetry of information, the construction of traditional information dissemination index systems demands excessive reliance on manual intervention, has large deviations, and is applied in a limited range. Such shortcomings cannot meet the requirements of constructing an objective, comprehensive, and highly credible index system. Therefore, we propose a method of constructing a multilevel and multigranular information dissemination index system with complex perspectives. In addition, we use the deep learning method of the convolutional neural network to extract the rich convolution features of public opinion events in the information dissemination process. Then, we train the weight, and it forms the corresponding weight of the information dissemination index systems. The experimental results prove that the method we use is superior to other methods and has better performance on the data set of a specific field.

Keywords: convolutional neural network, deep learning, index system, social network, information dissemination

1 INTRODUCTION

With the increasing convenience of Internet social networks, the cost and time of information flow are rapidly decreasing. How to accurately and quantitatively evaluate the dissemination of online public opinion information from the complex and massive amount of online public opinion information is a problem that needs to be solved urgently [1]. But nowadays, the understanding of the evolution mechanism of online public opinion is not deep enough. Many communication laws still need to be explored and summarized. The communication of online information involves a wide range of factors. The role of the constituent elements is interactive promotion and restriction. The main body and role of each stage factors are changing, making it difficult for managers to systematically assess the situation of information dissemination in an all-round way [2]. Therefore, establishing a reasonable and easy-to-use information dissemination evaluation index system, quantifying the information dissemination per unit time, and evaluating and calculating the situation to help managers make accurate decisions have become a current key research topic.

2 RELATED WORK

The research on information dissemination situation assessment mainly includes two aspects: qualitative public opinion situation assessment and quantitative public opinion situation assessment.

2.1 Qualitative Evaluation

In terms of qualitative evaluation, the current research mainly focuses on the evaluation of online public opinion, scope of the event, clustering effect, and tendency of netizens to the event.

Warshaw developed a multilevel regression and post-hierarchical model combining survey and population data so as to be able to detect specific online public opinion on lower-level geographic aggregation and evaluate the spread of public opinion [3]. Gil-Garcia developed two specific dynamic clustering algorithms from the framework proposed by previous scholars for clustering public opinion texts, which can automatically correct the clustering without re-clustering when the massive document data change [4]. Matsumura proposed an impact diffusion model to discover the network opinion leaders on the designated topics on the forum and the spread of network events [5]. Some studies proposed network topology and infectious disease strategies to explore the characteristics of information dissemination in social networks [6,7]. Cha made a detailed comparison of the three indicators for measuring user influence based on Twitter's massive data. Then, based on these three indicators, he also explored the dynamic performance of influence in terms of time and themes and gave a report based on the evaluation method of spreading public opinion of users [8]. Birmingham grabbed large-scale data on possible radical topics from YouTube and used social network analysis and sentiment analysis tools to dig out topics and their emotional tendencies and portray the characteristics of radicals. An analysis method of user emotional orientation in public opinion events [9].

In recent years, topic discovery technology and communication range evaluation technology for new Internet media, especially Twitter short texts, have also received extensive attention from scholars [10]. Li proposed a detection algorithm based on sparse self-encoding to study the mutual influence between nodes in social networks and analyze the information dissemination situation from the perspective of network topology [11]. Han proposed a topic representation model based on user behavior analysis and evaluated the importance of vocabulary in the process of information dissemination [12]. Zhang proposed a method based on vocabulary decomposition to rate the sentiment of Twitter short texts and used machine learning methods to automatically classify the sentiment of new texts to support the qualitative evaluation of user sentiments and responses of specific online public opinion events [13].

2.2 Quantitative Evaluation

In terms of quantitative evaluation of online public opinion situation, the research on online public opinion decision-making can be summarized from three aspects. The first is to design an index system for online public opinion situation decision-making. Based on the analysis of the spatial structure and evolution of online public opinion, corresponding indicators can be designed to represent the situation of online public opinion from different aspects. Designing a systematic, comprehensive, and scientific indicator system is the primary task of making decisions on the online public opinion situation. Identifying key and representative indicators can not only reduce

the complexity of the indicator system but also improve the accuracy of decision-making [14].

Zeng took public opinion risk as the core research object, starting from the source of risk and the internal and external performance of risk, constructed an early warning indicator system for network public opinion emergencies, and introduced the expert evaluation and AHP methods to determine the index weight [15]. Dai designed a three-tier security state decision-making index system from four dimensions: event sensitivity, communication, netizen attitude, and attention to public opinion and explained the statistical methods of each quantitative indicator [16]. Zhu proposed a mechanism to study the role of personal reputation and strategy in social network interactions [17] and also studied the information dissemination phenomenon of social networks in complex networks [18]. Shang regards the public opinion index system as an investment method, and under the guidance of portfolio theory integrates three sets of network public opinion index systems with their own characteristics into a dynamic integrated index system and uses radar chart analysis tools to evaluate the network public opinion. Comprehensive scores can also be used to analyze deep-seated reasons through online public opinion evaluation [19].

Zhang abstracted the key indicator system and logical relationship of the situation evaluation of network public opinion into a Bayesian network, then added some subjective evaluations, and gave a public opinion situation evaluation method based on the Bayesian network model. On the basis of constructing a public opinion index system with three dimensions, [20] Rong proposed a gray prediction method for network public opinion based on the variable-length mechanism of a data sequence. The prediction effect can be iteratively tested to find the optimal data sequence length to achieve the maximum good forecasting effect [21]. He chose the causal loop diagram to represent the mutual influence between the network public opinion indicators and constructed a network public opinion system dynamics model with 38 variables and four core subsystems targeting public opinion popularity [22]. Zhang used the principal component analysis method to refine the primary index system into a few comprehensive indicators without correlation and established an online public opinion early warning model based on the SVM machine learning method [23]. Nie measured and judged the influence of different users on the information dissemination situation in social networks from the perspective of user identity [24]. Huang used the PLSA model to extract the feature word space with no sentimentality, constructed the sentiment word space through word segmentation and TF-IDF function, and then integrated the inclination of all sentiment words based on the HowNet similarity algorithm to support the judgment of the sentiment trend of the network public opinion [25].

Most of the abovementioned scholars use subjective methods to construct indicator system models and use analytic hierarchy process, subjective assignment way, and other weight distribution methods to totally evaluate the situation of social network information dissemination, which have been verified by certain practical applications, but their focus is more limited. For more

TABLE 1 | Dataset.

Dataset	#Hours	#Blogs	#Events
Dataset	6,354	1785034	25

complex social network communication status, as well as the situational assessment of different types of events, the portability is inadequate. In the abovementioned former information dissemination index system construction and evaluation

calculation method, the weight distribution of each fundamental index depends on manual operation, and the value is relatively vague, and sometimes the deviation is large. Considering the real-time nature of social network information dissemination data, it is required to construct a real-time, self-adaptive nonlinear system. The system should include a dynamic index weight learning mechanism and iterate with the constantly updated social network information dissemination situation data to form an absolute index system that can effectually evaluate the information dissemination situation.

TABLE 2 | Information dissemination index system.

Primary index	Secondary index	Tertiary index
Public opinion events	Characteristics of the evolution stage	Propagation time Evolution stage Propagation rate
	Characteristics of post content	Forwarding ratio Include picture ratio Include audio ratio Include video ratio Include topic ratio @ other number
	Characteristics of information dimension	Blog title length Blog character length Blog vectorization feature Blog average TF-IDF
	Characteristics of network structure	Network density Convergence factor Connection strength
Public opinion audience	Characteristics of emotion	Positive affective word frequency Neutral affective word frequency Reverse affective word frequency Emotional intensity rating Positive emotion ratio Neutral emotion ratio Reverse emotion ratio
	Characteristics of user identity	Sex ratio Identity authentication ratio Account registration duration Age distribution of users Education level of users Political distribution User name ratio User head portrait ratio
Media aspect	Characteristics of media	Number of media Total news Proportion of media news reports News likes News forwarding number Media concerning number
	Characteristics of propagation	Total blogs Forwarding number Number of comments Number of likes Number of participating platforms Number of users covered
	Characteristics of distribution	Proportion of original posts Regional coverage Regional aggregation Domain name distribution Proportion of foreign blogs Proportion of domestic blogs

3 CONVOLUTIONAL NEURAL NETWORK

The convolutional neural network (CNN) is currently one of the key research directions in the field of computer vision, based on deep learning. It performs well in applications such as image classification and segmentation, and its powerful feature learning and feature expression capabilities are increasingly being valued by researchers. The convolution operation is a multilayer feedforward neural network model. The network structure is shown in **Figure 1**. Each layer uses a set of convolution kernels separately, which helps extract useful features from locally related data points. In the training process, the CNN learns through the backpropagation algorithm. The objective function optimized by this backpropagation algorithm uses a response-based human-like brain learning mechanism. The CNN imitates the biological neural network, adopting the core weight sharing network structure so that it can adjust the network model magnitude by adjusting the depth and width of the neural network.

3.1 Convolutional Layer

The convolutional neural network model has powerful assumptions about physical images, that is, statistical smoothness and local connection. It can validly reduce the learning complexity of the deep neural network model, making the network connection and weight parameters less, which makes it more than the same scale. The fully connected network is easier to train. It uses the convolution kernel to slide on the image and finally completes the process of calculating the gray value of all image pixels after a series of matrix operations.

3.2 Pooling Layer

The pooling layers can lower the dimensionality of the data by imitating the human visual system and using higher-level features to represent the image. The pooling layer can very effectively reduce the size of the matrix, that is, it can perform set statistical operations on the features of different positions in the local area of the image, thereby alleviating excessive sensitivity of the convolutional layer to image position and reducing the final fully connected layer. Parameters to speed up the calculation speed. The most commonly used pooling methods in practice are Max pooling and Average pooling. In addition to reducing model calculations and reducing information redundancy, they also improve scale invariance and rotation invariance of the model to varying degrees, effectively preventing overfitting. The improvement of various pooling methods also helps better realize feature compression and feature extraction, which greatly reduces the time required for model training.

3.3 Fully Connected Layer

A fully connected layer is consisted of several hidden layers in the CNN and usually appears in the last few layers. Each layer contains multiple neurons, and each neuron is fully connected to the neuron of the next layer, which is used to compare the characteristic structure and the structure designed in the previous section. Through the calculation of the layer and the layer, the feature obtained by the feature map is used as the input of the fully connected layer. The essence of the fully connected layer space is to linearly transform from one feature to another feature

space. In addition, at the end of the CNN, we use different classification functions to calculate the results.

4 MODEL DESIGN AND EXPERIMENT

4.1 Dataset and Environment

We grab a total of 15 from different social platforms including: Weibo, Twitter, and WeChat official accounts; some news websites: Yahoo, Sina, Tencent, and World Wide Web; and various online information dissemination platforms such as forums and blogs. The relevant data of the event and each event is divided on an hourly basis according to the time window, forming a total of 6,354h of data as shown in **Table 1**. We randomly use 10 percent of them as the test set and the remaining as the training set. Each set of data is scored by 10 experts and 100 ordinary users of social networks in accordance with the evaluation criteria, which are used as tags for the data set.

We performed this research in the following environment: CentOS 7.5, Intel(R) Xeon(R) Silver 4210, and Intel(R) Core(TM) i7-8750H CPU.

4.2 Construction of the Index System

The work of this module focuses on the diversified characteristics of the factors involved in the information dissemination situation, and each factor has a different impact on the information dissemination situation at different levels and different granularities. In-depth analysis and mining of influencing information from various perspectives, such as public opinion events, communication media, and public opinion audiences. The characteristics of the attributes of various factors in the communication situation and the law of their influence on the information dissemination situation have constructed a multilevel, multigranular, and multidimensional information dissemination situation indicator system. The three-tier indicators of the information dissemination trend index system are determined through research and use of the Delphi method, and the principal component analysis method is used to determine the main factors affecting the information dissemination trend by analyzing the correlation between different indicators, and finally, the information dissemination trend evaluation index is established, as shown in **Table 2**.

Among them, the first-level index public opinion event is analyzed from the perspective of the public opinion event, including the characteristics of evolution, post content, information dimension, and network structure. It is a class of indicators that describe the state of the public opinion event in the process of dissemination. It is mainly used to judge the communication stage of public opinion so as to analyze the communication trend. In the second-level indicators of the characteristics of the evolution stage of propaganda events, there are three three-level indicators of propagation time span, evolution stage, and propagation rate. In the second-level indicators of post content characteristics, there are three-level indicators such as the proportion of forwarding; the proportion of pictures, audio, video, and topics; and the number of others mentioned. The secondary indicators of the information dimension feature include the length of the post title, the length of the blog post, and the TF-IDF of the blog post, and the

TABLE 3 | Setting the initial model parameters.

Parameters	Setting
Convolution kernel size	2*2
Number of convolution kernels	64*1
Convolution stride	1
Pooling method	Max
Pooling window size	2*2
Pooling stride	2
Padding	SAME
Activation function	Relu
Dropout	0.25
Loss function	Cross-entropy loss
Fully connected layer neurons	128

dissemination situation of the post is measured from the character level. Audience tendency analysis is an indispensable part of public opinion analysis, and tendency analysis also reflects the size, structure, and psychological condition of the audience from another perspective, and is an important component of public opinion dissemination. For the secondary indicators of audience sentiment tendency, it includes the word frequency of positive, neutral, and negative sentiment words as well as the average sentiment intensity of each event and the proportion of positive, neutral, and negative sentiment posts. The spread of posts is measured from the sentiment analysis level. The secondary indicators of user identity characteristics include user information such as the user's gender ratio, age distribution, education level distribution, and political affiliation distribution, as well as account registration time, whether it is identity authentication, and whether the account has a user name and avatar, and other account information.

Public opinion audience is analyzed from the perspective of users participating in a certain public opinion event, including characteristics such as emotional tendency and identity. Audience tendency analysis is an indispensable part of public opinion analysis. Tendency analysis also reflects the size, structure, and psychological status of the audience from another angle and is an important component of public opinion dissemination. Secondary indicators of media participation include the proportion of news reports in posts involving publicity events, number of news media reported, total number of news reports, total number of likes on the news, total number of retweets, and total number of followers of the news media. Leaders play an important intermediary or filtering role in the formation of mass communication effects, and they spread information to audiences to form two-level dissemination of information. An important role in two-level communication is played by the person in the crowd who is first or more exposed to mass media information and disseminates the information that has been reprocessed by himself to others. With the ability to influence the attitudes of others, they intervene in mass communication, speeding it up and expanding its influence. The secondary indicators of dissemination heat include information such as the number of posts, retweets, comments, likes, participating platforms, and users of the propaganda event. These indicators directly reflect the dissemination situation of the propaganda event in social networks.

The media aspect is analyzed from the perspective of the media of public opinion events, including indicators such as media

participation, communication popularity, and regional distribution. It is an important standard for measuring the spread of public opinion events. Secondary indicators of media participation include the proportion of news reports in posts involving publicity events, number of news media reports, total number of news reports, total number of likes, total number of news retweets, and total number of news media followers. The secondary indicators of dissemination heat include information such as the number of posts, retweets, comments, likes, participating platforms, and users of the propaganda event. These indicators directly reflect the dissemination situation of the propaganda event in social networks.

4.3 Construction of the Model Structure

4.3.1 Input

We calculated 25 three-level indicators separately and use the following method to standardize the original statistical data:

$$Z_{ij} = (X_{ij} - \bar{X}_j) / \sigma_j, \sigma_j = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (X_{ij} - \bar{X}_j)^2}. \quad (1)$$

This method first calculates the average value and standard deviation of the original data and then standardizes the data. The data processed in this way conform to the standard normal distribution and conduct the neural network input data. We combine the indicators of each event into a 5*5 two-dimensional matrix as the input of the CNN. The neural network training process is as follows.

4.3.2 Convolution

We used a two-dimensional convolution operation to filter and extract features. The input vector is combined into a 5*5 two-dimensional single-channel grayscale image, and the calculation of the feature on the convolutional layer is as follows:

$$c_{\sigma}^a(x) = f\left(\sum W^{(a,\sigma)} p_i^{(a-1)} + b^{(a,\sigma)}\right), \sigma = 1, 2, \dots, n, \quad (2)$$

$c_{\sigma}^a(x)$ represents the output of the latter layer of features, W represents the weight of the kernel, p represents the feature map of the layer, and f is activation function Relu.

4.3.3 Pooling

In addition to reducing model calculations and reducing information redundancy, pooling operation improves the scale invariance and rotation invariance of the model to varying degrees, effectively preventing overfitting. The improvement of various pooling methods also better realizes feature compression and feature extraction, which greatly reduces the time required for model training. The calculation of features is as follows:

$$s_{\sigma}^a(x) = \max \text{pooling}(c_{\sigma}^{a-1}) + b^{(m,\sigma)}. \quad (3)$$

4.3.4 Output

The convolutional neural network filters the matrix through the aforementioned convolutional layer and pooling layer for feature screening, and then the obtained vector constructs a fully connected layer. After two fully connected layers, finally, the output is analyzed through the Softmax function to obtain the

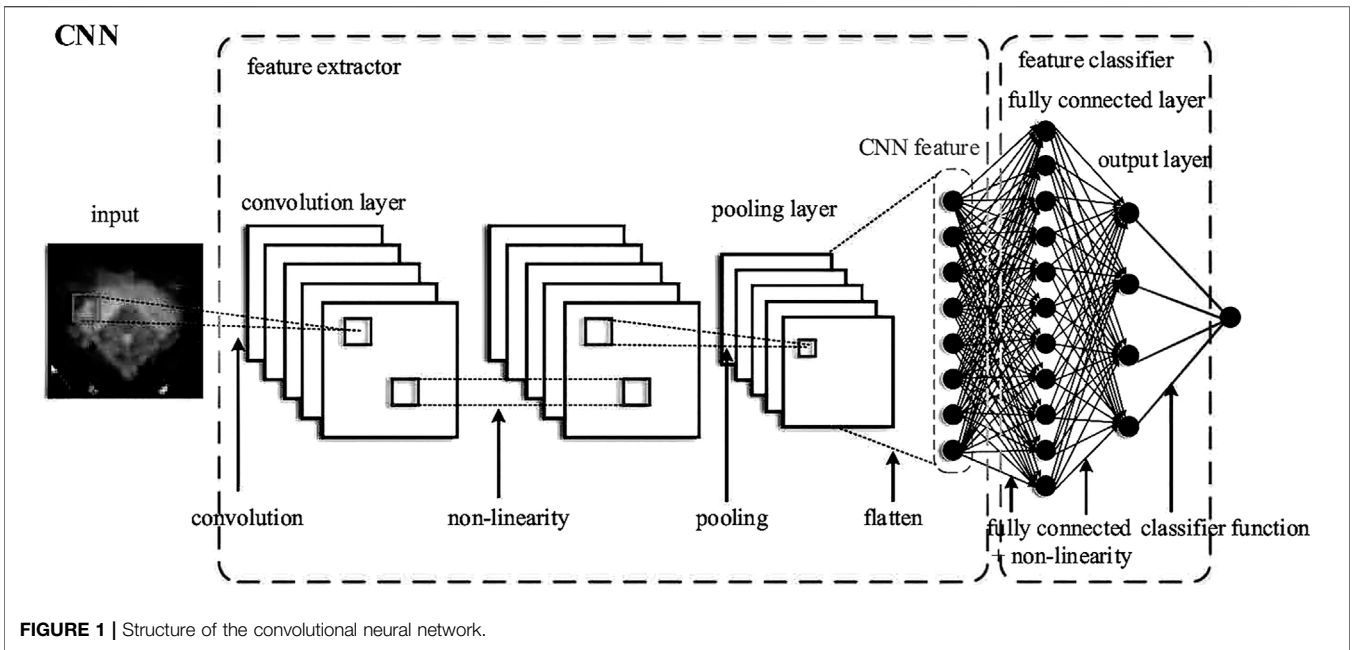


FIGURE 1 | Structure of the convolutional neural network.

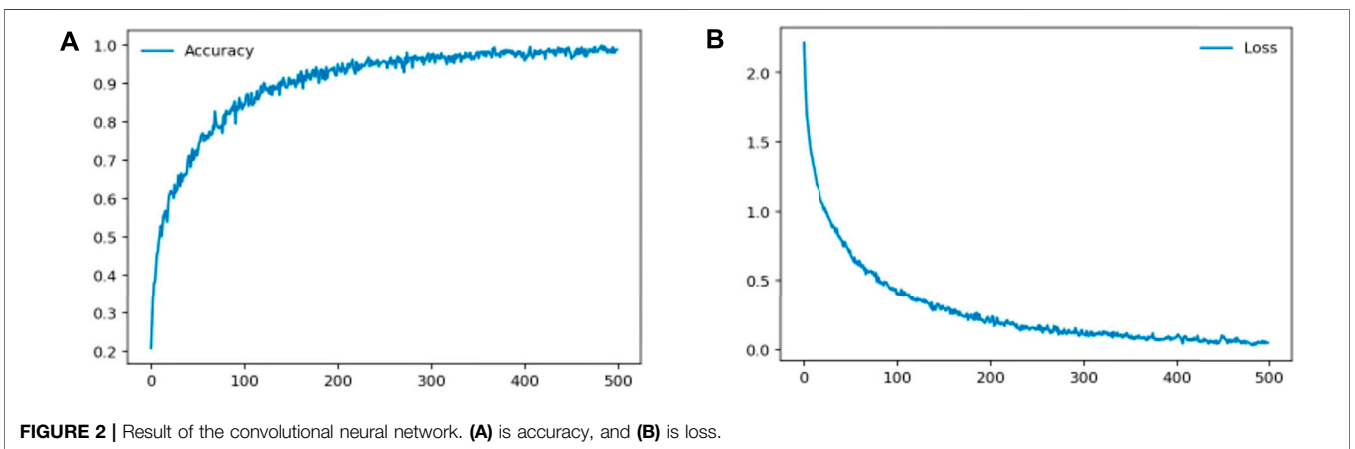


FIGURE 2 | Result of the convolutional neural network. (A) is accuracy, and (B) is loss.

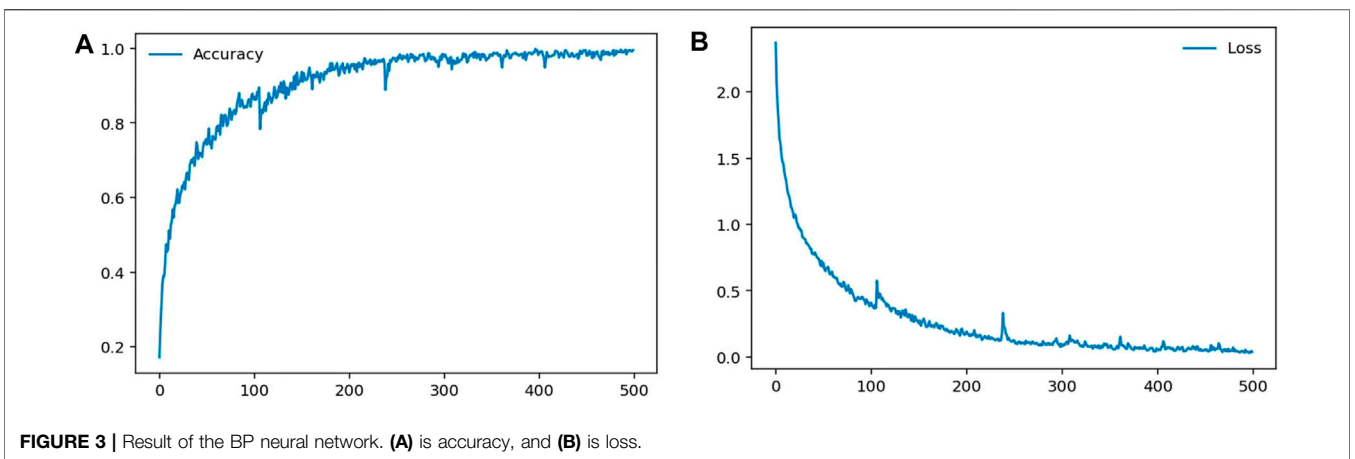


FIGURE 3 | Result of the BP neural network. (A) is accuracy, and (B) is loss.

TABLE 4 | Comparison of regression evaluation indexes.

Model	Accuracy	Recall	Precision	F1-score
LR	0.924	0.916	0.937	0.924
BPNN	0.966	0.950	0.968	0.966
SVM	0.938	0.925	0.958	0.938
CNN	0.986	0.973	0.956	0.986

prediction result, and the offset value and weight value are adjusted at the same time.

4.3.5 Calculation

In the model, we use the Relu function as the activation function of each layer and cross-entropy as the loss function.

$$L = \frac{1}{N} \sum_i = -\frac{1}{N} \sum_i \sum_{c=1}^M y_{ic} \log(p_{ic}). \quad (4)$$

In addition, the Adam algorithm is used for backpropagation optimization. Finally, the model is trained and the Captum interpretability tool is used to obtain the three-level indicator weights. The lower-level index weight is calculated by weighted summation to obtain the upper-level index weight:

$$f(x) = \sum_{i=1}^n w_i x_i. \quad (5)$$

In this way, we calculated the first-level, second-level, and third-level index weights of the index system. When we use the index system, we input the calculated value of the third-level index of an event to be detected. Therefore, we can obtain the corresponding second-level and first-level index values using Eq. 5 and obtain the propagation situation value of the event using model prediction.

4.4 Setting of Model Parameters

Convolutional neural networks can extract features at multiple levels and granularities through convolution and pooling operations and finally learn the propaganda posture features that can really distinguish the propaganda posture. We evaluate the needs and test the propaganda posture at different levels and granularities. The effect of different depth convolutional neural networks and different combinations of data input methods on the performance of the algorithm is to test the effect of convolutional neural networks on the quantitative evaluation of publicity situations. Based on the abovementioned experiments and tests, we have obtained a network structure with excellent experimental results. Two layers of convolutional layer and pooling layer are used, and two layers of fully connected layers, and finally, use the Softmax function is used for output, as shown in Table 3.

4.5 Experiment Results

We preprocessed the data and input it into the constructed convolutional neural network. After training, the results of fitting expert calibration are found to have very good performance and good convergence effect, as shown in Figure 2 and we used the BP neural network which also has a good performance as in Figure 3.

We also used linear regression, SVM, and other methods to calculate and compared the results with the effect of the convolutional neural network. As shown in Table 4, we used the following standards: accuracy, recall, precision, and F1-score to measure the performance of the models. The bold values represent the best effects. We found that the convolutional neural network has more excellent results.

5 CONCLUSION

In summary, the existing methods have several shortcomings, which are mainly concentrated in the following aspects: 1) the calculation of the information dissemination index system relies too much on manual evaluation; 2) the determined index system is relatively limited and one-sided; and 3) the weight value of the index system fluctuates greatly, and sometimes, it cannot accurately reflect the dissemination trend of information dissemination events. In response to these shortcomings, this study put forward a way for constructing an information dissemination index system based on the CNN. We use the convolutional layer for multidimensional and multigranular feature extraction and apply the pooling layer to quickly reduce the size of the information dissemination network and highlight the main features. Through the deep-network structure with several hidden layers of the CNN, we have realized the evolution of simulating expert experience to assess all-round indicators. In addition, it has adaptive features such as self-learning. Through experimental comparison, the calculated results perform better than the other mentioned models. However, this CNN model lacks the best parameter proof, and for specific information dissemination events on different topics, it lacks a more targeted index system construction. This is the direction of our future improvement and research.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

WH and LX contributed equally to this study and share first authorship. SL and LX contributed to the conception and design of the study and performed the experiments. XW and ZW grabbed and analyzed the dataset. SL and LX wrote the first draft of the manuscript.

FUNDING

This research was funded by NSFC (Nos U1803263, 62072131, and 61972106), Grant of Guangdong Recruitment Program of Foreign Experts (2020A1414010081), Grant of (18-163-15-ZD-002-003-10), Guangdong Higher Education Innovation Group, China (No. 2020KCXTD007), and Guangzhou Higher Education Innovation Group, China (No. 202032854).

REFERENCES

- Gao C, Su Z, Liu J, Kurths J. Even central Users Do Not Always Drive Information Diffusion. *Commun ACM* (2019) 62:61–7. doi:10.1145/3224203
- Gao C, Liu J. Network-based Modeling for Characterizing Human Collective Behaviors during Extreme Events. *IEEE Trans Syst Man, Cybernetics: Syst* (2016) 47:171–83. doi:10.1109/tsmc.2016.2608658
- Warshaw C, Rodden J. How Should We Measure District-Level Public Opinion on Individual Issues. *J Polit* (2012) 74:203–19. doi:10.1017/s0022381611001204
- Gil-García R, Pons-Porrata A. Dynamic Hierarchical Algorithms for Document Clustering. *Pattern Recognition Lett* (2010) 31:469–77. doi:10.1016/j.patrec.2009.11.011
- Matsumura N, Ohsawa Y, Ishizuka M. Influence Diffusion Model in Text-Based Communication. *Trans Jpn Soc Artif Intelligence* (2002) 17:259–67. doi:10.1527/tjsai.17.259
- Li S, Zhao D, Wu X, Tian Z, Li A, Wang Z. Functional Immunization of Networks Based on Message Passing. *Appl Maths Comput* (2020) 366:124728. doi:10.1016/j.amc.2019.124728
- Su Z, Gao C, Liu J, Jia T, Wang Z, Kurths J. Emergence of Nonlinear Crossover under Epidemic Dynamics in Heterogeneous Networks. *Phys Rev E* (2020) 102:052311. doi:10.1103/PhysRevE.102.052311
- Cha M, Haddadi H, Benevenuto F, Gummadi K. Measuring User Influence in Twitter: The Million Follower Fallacy. In: *Proceedings of the International AAAI Conference on Web and Social media*, 4 (2010).
- Birmingham A, Conway M, McInerney L, O'Hare N, Smeaton AF. Combining Social Network Analysis and Sentiment Analysis to Explore the Potential for Online Radicalisation. In: 2009 International Conference on Advances in Social Network Analysis and Mining. IEEE (2009). p. 231–6. doi:10.1109/asonam.2009.31
- Li X, Wang Z, Gao C, Shi L. Reasoning Human Emotional Responses from Large-Scale Social and Public media. *Appl Maths Comput* (2017) 310:182–93. doi:10.1016/j.amc.2017.03.031
- Li S, Jiang L, Wu X, Han W, Zhao D, Wang Z. A Weighted Network Community Detection Algorithm Based on Deep Learning. *Appl Maths Comput* (2021) 401:126012. doi:10.1016/j.amc.2021.126012
- Han W, Tian Z, Huang Z, Li S, Jia Y. Topic Representation Model Based on Microblogging Behavior Analysis. *World Wide Web* (2020) 23:3083–97. doi:10.1007/s11280-020-00822-x
- Zhang L, Ghosh R, Dekhil M, Hsu M, Liu B. *Combining Lexicon-Based and Learning-Based Methods for Twitter Sentiment Analysis*. Palo Alto, CA: Hewlett-Packard Laboratories (2011). p. 89. Technical Report HPL-2011.
- Quan Y, Jia Y, Zhou B, Han W, Li S. Repost Prediction Incorporating Time-Sensitive Mutual Influence in Social Networks. *J Comput Sci* (2018) 28:217–27. doi:10.1016/j.jocs.2017.11.015
- Runxi Z. Construction of Early Warning index System for Internet Public Opinion Emergencies. Beijing: Beijing University of Chemical Technology. (2013) (Masters Thesis).
- Yuan D. *Research on My Country's Internet Public Opinion Security Evaluation Index System*. Beijing: Beijing University of Chemical Technology (2008). Ph.D. thesis.
- Zhu P, Wang X, Jia D, Guo Y, Li S, Chu C. Investigating the Co-evolution of Node Reputation and Edge-Strategy in Prisoner's Dilemma Game. *Appl Maths Comput* (2020) 386:125474. doi:10.1016/j.amc.2020.125474
- Zhu P, Wang X, Li S, Guo Y, Wang Z. Investigation of Epidemic Spreading Process on Multiplex Networks by Incorporating Fatal Properties. *Appl Math Comput* (2019) 359:512–24. doi:10.1016/j.amc.2019.02.049
- Runan S, Xiaolong D. Research on the index System of Internet Public Opinion Cases Based on Improved Radar Chart. In: *Beijing: Journal of Beijing University of Posts and Telecommunications (Social Sciences Edition)* (2015). p. 15–20.
- Yiwen Z, Jiayin Q, Binxiang F, Yuxiao L. Research on Internet Public Opinion Early Warning of Unconventional Crisis Events Based on Bayesian Network Modeling. *Libr Inf Work* (2012) 56:76.
- Zizhan R. *Research and Implementation of Online Public Opinion Warning and Pre-plan System*. Beijing: Beijing University of Posts and Telecommunications (2013). Ph.D. thesis.
- Fei H. *Study on Early Warning of Enterprise Network Public Opinion Crisis Based on System Dynamics*. Beijing: Capital University of Economics and Business (2013). Master's thesis.
- Yanyan Z. *Study on Early Warning of Internet Public Opinion Crisis Based on Support Vector Machine*. Nanchang: Jiangxi University of Finance and Economics (2013). Ph.D. thesis, Master's Thesis.
- Nie Y, Jia Y, Li S, Zhu X, Li A, Zhou B. Identifying Users across Social Networks Based on Dynamic Core Interests. *Neurocomputing* (2016) 210: 107–15. doi:10.1016/j.neucom.2015.10.147
- Weidong H, Ping L, Yi D, Hongwei L. Emotional Evolution Analysis of Internet Public Opinion Participants Based on Topic Feature Words. *Intelligence Mag* (2015) 34:117–22.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Han, Xiao, Wu, He, Wang and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.