**Journal of Organizational and End User Computing**
Volume 35 • Issue 3

# The Social Media Big Data Analysis for Demand Forecasting in the Context of Globalization:
## Development and Case Implementation of Innovative Frameworks

Yifang Gao, College of Art Design, Shanghai Jian Qiao University, China
Junwei Wang, School of Media and Law, NingboTech University, China
Zhi Li, School of Media and Law, NingboTech University, China
Zengjun Peng, School of Media and Law, NingboTech University, China*

## ABSTRACT

This paper aims to analyze the predictive effect of artificial intelligence on user demand in big data social media and to provide suggestions for developing enterprise innovation frameworks and implementing marketing strategies. In response to the inconsistency between the supply of enterprise products and services and market demand, deep learning algorithms have been introduced using social media big data analysis. This algorithm has been improved to construct a user demand prediction model in social media big data based on bidirectional long short-term memory (BiLSTM) fused with Word2Vec. The model uses data acquisition and pre-processing, Word2Vec algorithm to vectorization the data information, and BiLSTM network to model and train the sequence. Finally, the model is evaluated as an example.

## KEYWORDS

Big Data, Deep Learning, Innovation Framework, Marketing Strategy, Social Media

## INTRODUCTION

With the continuous deepening of globalization, social media has become an integral part of people's daily life. The rapid increase in user activity and information exchange on social media platforms can provide companies with a larger market and more opportunities to obtain user data. In this increasingly competitive market environment, these data can help companies better understand market demand and consumer behavior, enhance their competitiveness and economic efficiency, and increase the market share (Meng et al., 2022; Rajković et al., 2021). However, these methods suffer from information lag

DOI: 10.4018/JOEUC.325217
*Corresponding Author

and limitations, often making it difficult to make timely and accurate predictions of rapidly changing market demands. To better predict consumer demand and behavior, companies need to master and analyze a large amount of data (Chen et al., 2021). With the popularization of the Internet and social media, the application of big data is becoming increasingly widespread. Social media has become a highly potential source of data. Big data analysis techniques are used to analyze user behavior and speech in social media. Enterprises can better understand consumers' needs and preferences and adjust product and service strategies (Lei et al., 2022; Tajpour et al., 2023). Therefore, how to use social media big data to predict user needs accurately has become the focus of many scholars in related fields.

With the rise and popularity of social media, more and more people are starting to express their opinions and needs through social media. Using big data on social media for demand forecasting has become a new approach, and many scholars have conducted relevant research. Feng et al. (2021) studied environmental governance startups and pointed out that the information sharing of venture capital networks and social media positively correlated with investment performance. In social media, the degree of enterprise risk information sharing was increasing. Based on the native big data of the Internet of Things, Chen and Du (2022) elaborated on the blocking process of the evolution of Online Public Sentient (OPS) and utilized artificial intelligence (AI) technology and big data to simulate and predict the development of OPS. The results indicated that Adam's optimized Long Short-Term Memory (LSTM) neural network model could predict the heat of OPS in dynamic evolution with high prediction accuracy. Dai (2022) improved the existing CNN and applied it to social media for predicting user demand in the stock market. The results showed that the proposed improved convolutional neural network had significant advantages in prediction accuracy based on specific data from the Chinese stock market for empirical analysis. Liu and Chen (2023) proposed a new potential feature topic model and introduced a time series model to establish a topic evolution network. The high-frequency words from three periods in social media were compared and analyzed. The results showed that the topic model found the evolution law of AI domain knowledge structure. The research of these scholars shows that the digital transformation of enterprises is the strategic demand of enterprise development at present and that it plays an important role in promoting economic benefits. However, there are relatively few relevant studies on user demand prediction in enterprise products from the perspective of social media big data, such as on the specific influencing factors of the development direction of social media big data.

Under the globalized economy, this paper aims to develop an innovative social media big data analysis framework for predicting market demand and consumer behavior; it has extremely practical significance for developing enterprise innovation frameworks and optimizing marketing strategies. The innovation addresses the inconsistency between the supply of enterprise products and services and market demand. Deep learning algorithms are introduced based on social media big data analysis, and a user demand prediction model in social media big data is constructed using Bidirectional Long Short-Term Memory (BiLSTM) fused with Word2Vec. Ultimately, the data on social media is analyzed and predicted to provide more valuable marketing recommendations for optimizing enterprise products and services.

The overall organizational structure is divided into four sections. Section 1 is an introduction explaining the background related to the development of social media under the rapid development of globalization. Additionally, this section summarizes and analyzes the current related research and gaps and puts forward innovations and contributions. Section 2 models social media big data. Deep learning is introduced, and a user demand prediction model based on BiLSTM fused with Word2Vec in social media big data is constructed. Section 3 discusses the experimental design and provides the performance evaluation. The model's performance is evaluated and compared with different schemes for discussion and analysis. Section 4 is the conclusion, briefly describing the results and explaining this research's limitations and prospects.

## METHOD

### Analysis of Social Media Big Data

Compared to traditional methods, social media big data has the advantages of instant information, a wide range, a large data volume, and multiple types of data, which can more comprehensively reflect the changing trends of market demand (Korcsmáros & Csinger, 2022; Sakas et al., 2022). However, the use of social media big data for demand forecasting still faces a series of challenges. The characteristics and problems of big data social media are shown in Figure 1.

In Figure 1, it is shown that the quality of information on social media varies, and there is a significant amount of noise and error. Humans need to use effective algorithms for data cleaning and pre-processing. Secondly, the scale of big data social media is enormous, and extracting effective information from it and analyzing it is also challenging.

In addition, how to combine big data social media with traditional sales data to understand changes in consumer demand and achieve more accurate predictions of enterprise marketing strategies and product planning is also an important issue. In order to provide more precise market demand prediction for enterprises, big data enterprise is extracted through data collection, pre-processing, feature extraction, and modelling prediction (Ye & Chen, 2021). The information extraction process in social media big data is shown in Figure 2.

In Figure 2, crawler technology collects information related to enterprise products on social media. Secondly, this information is analyzed and processed, extracting features and keywords pertaining to demand prediction. Then, the machine learning model models and predicts consumers' needs and behavioral trends based on these features and keywords. Finally, the predicted results are fed back to the enterprise to develop more accurate marketing strategies and product planning (Abkenar et al., 2021; Tao et al., 2020; Zhou et al., 2020). Among them, machine learning algorithms extract information from big data from social media, which is the core of modelling and prediction.

### Machine Learning Applied to Information Prediction Analysis of Social Media Big Data

Machine learning applied to information prediction analysis of social media big data is a method that utilizes AI technology to analyze and predict information trends and behaviors on social media. Massive social media data is processed and analyzed. Machine learning algorithms can explore the laws and patterns within them and apply them to future enterprise product planning prediction and analysis of marketing strategy behavior.

**Figure 1. Characteristics and issues of social media big data**

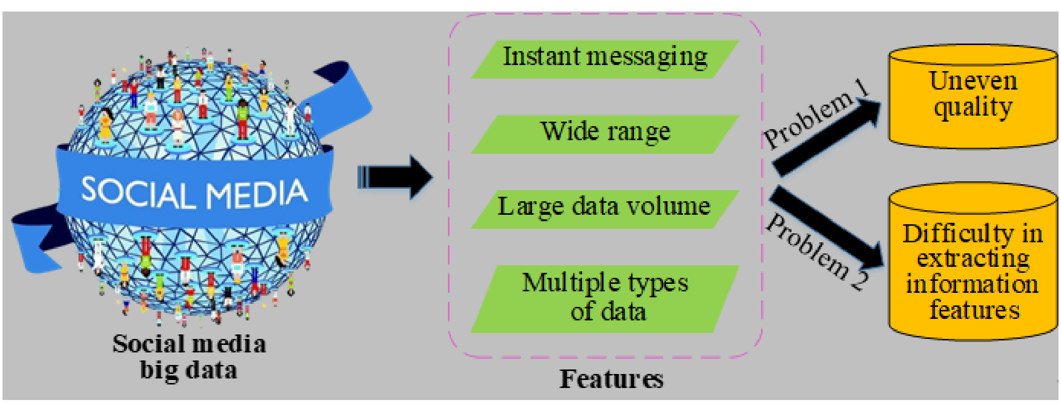
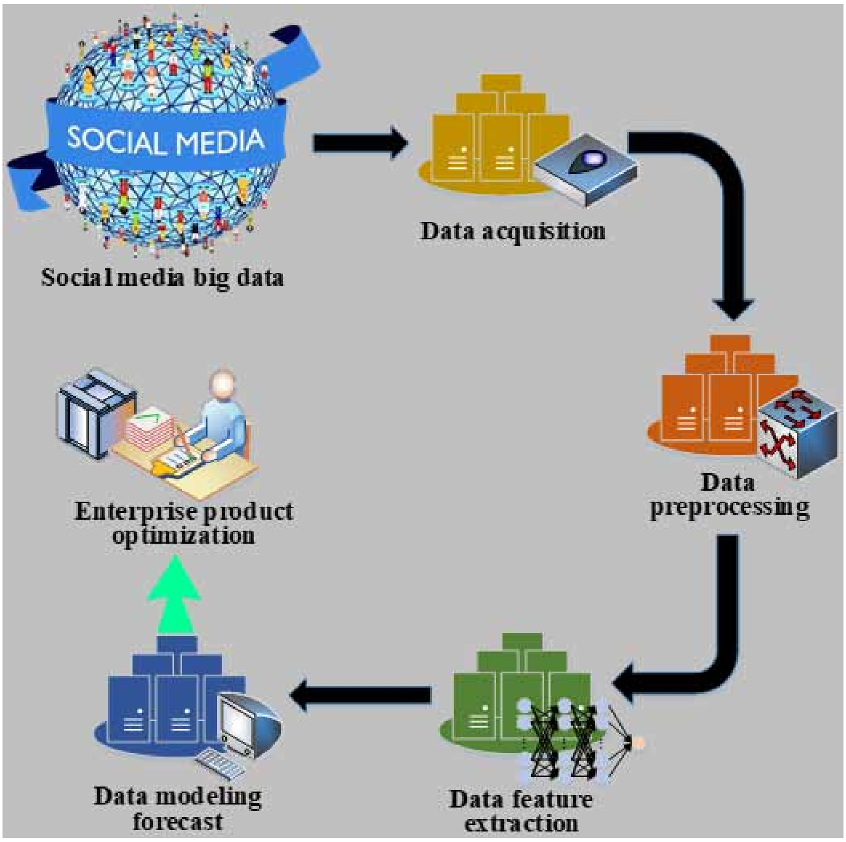


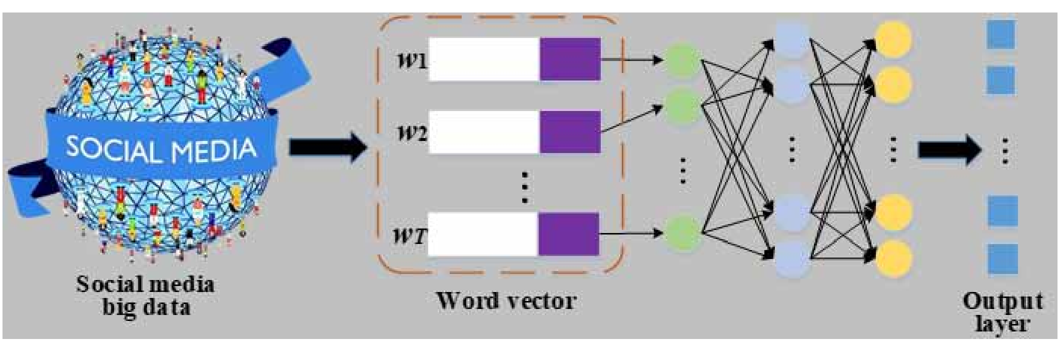
Figure 2. Information extraction process in social media big data




When extracting information features from big data of social media, relevant evaluation feedback texts are used to classify user demand features, which can directly learn low-dimensional word vectors. The Word2Vec algorithm predicts similar texts around the text (Liu et al., 2022). The process of extracting information from big social media data using the Word2Vec algorithm is shown in Figure 3.

The log probability of any text surrounding a central word is maximized. The text vocabulary $w_1, w_2, \cdots, w_T$ that need to be trained are given. The objective function is shown in Eq. (1):

Figure 3. Word2Vec algorithm for extracting information from social media big data



$$J\left(\theta\right)=\frac{1}{T}\sum\nolimits_{t=1}^{T}\sum\nolimits_{-c\leq j\leq c,j\neq 0}\log p\left(w_{t+j}\mid w_{t}\right) \tag{1}$$

In Eq. (1), *c* refers to the length of the window. The larger the window length, the more text vocabulary is trained and the higher the accuracy. $p\left(w_{t+j}\mid w_{t}\right)$ is shown in Eq. (2):

$$p\left(w_{O}\mid w_{I}\right)=\frac{\exp\left(v'^{T}_{w_O}\ v_{w_I}\right)}{\sum\nolimits_{w=1}^{W}\exp\left(v'^{T}_{w_O}\ v_{w_I}\right)} \tag{2}$$

In Eq. (2), $v$ and $v'$ represent the "input" and "output" vectors of the distribution w. Therefore, each w has two vector representations. W refers to the number of words in the vocabulary. The gradient descent method maximizes the objective function and obtains the final output vector.

In the Word2Vec algorithm, each vector describes the word's position in space. Therefore, spatial distance can represent the similarity between two words, meaning the similarity in grammar and semantics between two words (Feng & Chen, 2022). In the embedding space, the relationship between similarity can be determined using vector subtraction for category judgment, as shown in Eq. (3):

$$X_{apple}-X_{apples}\approx X_{car}-X_{cars}\approx X_{family}-X_{familys} \tag{3}$$

The Word2Vec algorithm can be used to analyze text and connect similar words. Adding an embedding layer between the input and the hidden layers of the original neural network can make the neural network more accurate. The Word2Vec algorithm can better represent word vectors in big data social media, reducing the learning cost of neural networks for user demand-related feedback texts in social media.

Usually, the information in social media big data is real-time, and understanding the temporal characteristics of data information is essential. This paper uses IOB2 to label the set. In this tag set, 'B' refers to the first word in a labelled unit. '*r*' refers to the word belonging to an intermediate or closing term in a labelling unit, representing the continuation. 'O' indicates that the word does not belong to any annotation unit. Therefore, the sequence label set is introduced, and the semantic recognition problem of text can ultimately be transformed into a sequence label problem. This paper formalizes the semantic recognition problem of text frames as a sentence, $S=w_{1},w_{2},\cdots,w_{n}$; the target word $w_{t}$ in the sentence and the frame $f_{w_t}$ to which it belongs are given. The sequence label set IOB2 is used to label each word in the sentence appropriately; $t_{i}$. $t_{i}\in\left\{B,I,O\right\}$ represents the semantic role boundary label corresponding to the $i$-th word in sentence *S*. A marker sequence $T=t_{1},t_{2},\cdots,t_{n}$ is obtained. The semantic role recognition task is transformed into Eq. (4):

$$T^{*}=\arg\max P\left(T=t_{1},t_{2},\cdots,t_{n}\mid S=w_{1},w_{2},\cdots,w_{n},f_{w_t}\right) \tag{4}$$

In Eq. (4), $T^{*}$ refers to a proper sequence that can restore semantic role information.

BiLSTM, as an improved algorithm for recurrent neural networks (RNN), has become the mainstream model in this field. The standard RNN recursion execution is shown in Eq. (5):

$$h_t = H\left(W\left[h_{t-1}, x_t\right] + b\right) \tag{5}$$

The creation of hidden layer output is referred to as $h_t$ in Eq. (5). *H* is a nonlinear function, such as the straightforward tan*h* function or a chain of more intricate transformations. *W* is for weight. In an RNN, in addition to $x_t$, the output $h_{t-1}$ of the preceding hidden layer is also included in the input of the hidden layer at time *t*. The information $h_{t-1}$ of the words before the *t*-th word in a sentence is used to forecast how the ruling will process the emotional information in the text.

The standard RNN also has two issues: Firstly, if the gradient transfer is too long, capturing long-distance dependencies in the sequence will be difficult. Secondly, the neural network may experience gradient vanishing or exploding when processing long sequences. The standard RNN has been improved. The most widely used structure is LSTM (Behera et al., 2021). Standard RNN or LSTM, however, can record information only about past sequences. Future information is occasionally required in sequence tagging, given the complexity of the sentences in the Chinese language. The information behind the current word (on the right) may be needed (Ali et al., 2021; Wadud et al., 2022). The forward and backward dimensions are used to analyze the periodicity of historical sequence data, namely BiLSTM. The application framework of BiLSTM in social media big data is shown in Figure 4.

In Figure 4, the fundamental design of the BiLSTM is to model the sequence from front to back (ahead) and from back to front (reverse) using two LSTMs in the hidden layer, respectively.

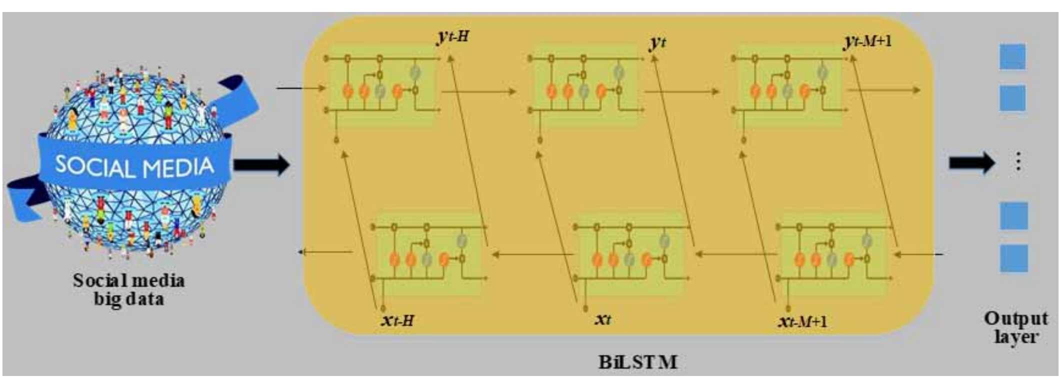
In Figure 4, the basic framework of BiLSTM is to use two LSTMs in the hidden layer to model the sequence from front to back (forward) and from back to front (backward), respectively. Then, their outputs are connected (Tripathi, 2021). The calculation of RNN from front to back is shown in Eq. (6):

$$h_t^l = H\left(W^l \cdot \left[h_{t-1}^l, x_t\right] + b^l\right) \tag{6}$$

The calculation of RNN from back to front is shown in Eq. (7):

$$h_t^r = H\left(W^r \cdot \left[h_{t-1}^r, x_t\right] + b^r\right) h_t = \left[h_t^l : h_t^r\right] \tag{7}$$

**Figure 4. Application process of BiLSTM in social media big data**



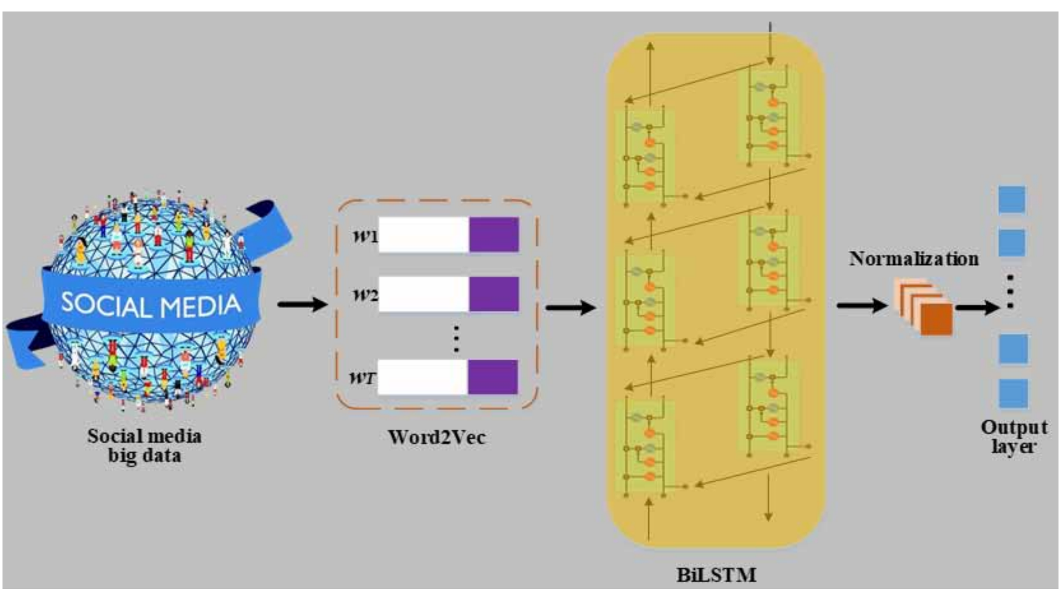
At time $t$, $h_t = \left[h_t^l : h_t^r\right]$ means that the beginning $h_t^l$ and $h_t^r$ end of two vectors will be concatenated. The cyclic units in BiRNN are swapped out for LSTM structures in BiLSTM. The two advantages are integrated to recognize information features and user needs in text sentences. Therefore, this paper combines BiLSTM and Word2Vec algorithms. This method can not only capture the semantic relationship between social media big data but also sequentially model the needs issued by users to achieve the effect of effective identification of user needs in social media big data.

### Construction of User Demand Prediction Model in Social Media Big Data Based on BiLSTM Fusion With Word2Vec

When designing their business strategy, enterprises must clarify their services or products to meet the needs of target users. Therefore, the key to a company's success in influencing the local market is whether the products it launches can meet market demand and whether the strategy is scientifically reasonable. This paper combines two algorithms, BiLSTM and Word2Vec, and uses the BiLSTM network to model the sequence of user requirements. Meanwhile, Word2Vec technology converts text into vector representations, allowing network parameters in deep networks to be updated and learned and capturing the semantic relationships between social media and big data. User needs are understood to provide brand enterprises with intelligent business intelligence and help them make marketing and decision plans. The user demand prediction model in social media big data based on BiLSTM fusion with Word2Vec is shown in Figure 5.

In Figure 5, the user demand prediction model is subjected to data pre-processing. User data on social media is cleaned and pre-processed, including with noise removal, word segmentation, and the removal of stop words. The pre-processed data trains the Word2Vec model to convert each word into a vector representation, thereby better capturing the semantic relationships between texts. The BiLSTM network models user requirements in sequence, transforming user requirements into vector sequences. This network is used to model sequences and capture the relationships between them. Finally, the vector output by Word2Vec and BiLSTM are fused to extract feature vectors for subsequent classification or regression tasks. When training and optimizing the model, the annotated

**Figure 5. Prediction model for user needs in social media big data based on BiLSTM fusion with Word2Vec**



training data is used to train the model. Optimization algorithms such as random gradient descent optimize the model and improve prediction accuracy.

During the training process, user needs are specifically rated. The loss function is shown in Eq. (8).

$$L\left(y,y\right)=\frac{1}{2}\left(y-y\right)^{2} \tag{8}$$

The mini-batch method optimizes the loss function when the user demand is predicted. When the mini-batch method is used to reduce the value of the loss function, the learning rate needs to be dynamically adjusted. At the initial stage of training, a significant learning rate is required to speed up the decline of the loss function. However, in the later stage, it is necessary to reduce the value of the learning rate to avoid not reaching the optimal value for a long time. This paper uses the "Poly" learning rate adjustment method with polynomial decay (Khan et al., 2022), as shown in Eq. (9):

$$init\_lr\times\left(1-\frac{epoch}{\max\_epoch}\right)^{power} \tag{9}$$

In Eq. (9), the initial learning rate $init\_lr$ is 0.0005 (or 5e$^{-4}$). The power is set to 0.9.

In the process of data information normalization, the normalization standards used in this paper are the minimum and maximum value normalization methods.

$$P'=\left(P-P_{\min}\right)/\left(P_{\max}-P_{\min}\right) \tag{10}$$

In Eq. (10), $P$ refers to the original data value at the current time. $P_{\min}$ and $P_{\max}$ refer to the minimum and maximum values of all data, respectively. $P'$ refers to the normalization result of user source data in all social media requests.

The actual output projected data must be returned to its original size using a normalization reduction function after normalization is finished in the data input stage. Eq. (11) illustrates the reduction functions for the minimum and maximum standardization procedures.

$$P_{\mathrm{Pr}e}=P'_{\mathrm{Pr}e}/\left(P_{\max}-P_{\min}\right)+P_{\min} \tag{11}$$

In Eq. (11), $P'_{\mathrm{Pr}e}$ refers to the predicted output value of the unreduced BiLSTM neural network. $P_{\mathrm{Pr}e}$ refers to the expected value of social media big data after restoration. Weighted Cross Entropy (WCE) is employed as a cost function throughout the training process to enhance model training (Chen & Du, 2022a; Kaladevi & Thyagarajah, 2021).

## Experimental and Performance Evaluation

Through the experimental design and data collection methods, this paper can verify the performance of the proposed user demand prediction model based on BiLSTM fused Word2Vec in social media big data.

This paper uses Matlab simulation software to simulate it. The data obtained in the experiment all come from the group comments about Midea on the Weibo APP platform. Among them, 6,121 texts of 25 frames are selected as corpus. 3×2 cross-validation is adopted for experiments. The corpus is divided into four parts of the same size, two of which are selected as the training set; the remaining two are used as the verification set, and 2-fold cross-validation is implemented, so there are six sets of results in total. Additionally, when dividing the corpus, the example sentences under each framework are evenly divided into each part according to the lexical units, ensuring the balance of the framework and lexical units of the training and verification sets.

The model uses the mini-batch gradient descent algorithm to update model parameters (Ali, El-Sappagh, et al., 2021). The batch size is set to 10. Each iteration uses 10 sentences to update the parameters. The learning rate attenuation coefficient is 0.05. Due to 3×2 cross-validation, the training set is half of the corpus, and the training corpus is relatively small. In order to prevent overfitting, the model training uses the dropout method. The value of the dropout rate is 0.5. Training is performed for 80 iterations.

Among them, the specific simulation experiment environment configuration is mainly prepared from two aspects of hardware and software. In the software, the operating system is Linux 64bit, the Python version is Python 3.6.1, and the development platform is PyCharm. In the hardware, the CPU is Intel core i7-7700@4.2GHz 8-core, the memory is Kingston ddr4 2400MHz 16G, and the GPU is Nvidia GeForce 1060 8G. In the Python environment, related third-party libraries are installed, including NumPy, Pandas, sci-kit-learn, and TensorFlow. Pip or conda are allowed to be used for installation.

Based on Accuracy, Precision, Recall, and F1 values, LSTM (Al Salem et al., 2021), BiLSTM (Zhang et al., 2022), RNN (Hamayel & Owda, 2021), and Liu & Chen (2023) are evaluated. Accuracy, Precision, Recall, and F1 value base calculations are shown in Eqs. (12)-(15):

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \tag{12}$$

$$\Pr ecision = \frac{TP}{TP + FP} \tag{13}$$

$$\operatorname{Re} call = \frac{TP}{TP + FN} \tag{14}$$

$$F1 = \frac{2 \Pr ecision \cdot \operatorname{Re} call}{\Pr ecision + \operatorname{Re} call} \tag{15}$$

In Eqs. (12)-(15), *TP* is the number of positive samples predicted to be positive. *FP* is the number of negative samples predicted to be positive. *FN* is the number of positive samples predicted as negative. *TN* is the number of negative samples predicted to be negative. Accuracy (*ACC*) is used to measure the overall classification accuracy: the proportion of predicted samples. Recall (*Rec*) measures the coverage of positive samples: the proportion of correctly classified positive samples to the total number of positive samples. Accuracy (*Pre*) represents the ratio of examples classified as positive to positive. The F-measure is the weighted harmonic mean of Precision and Recall.

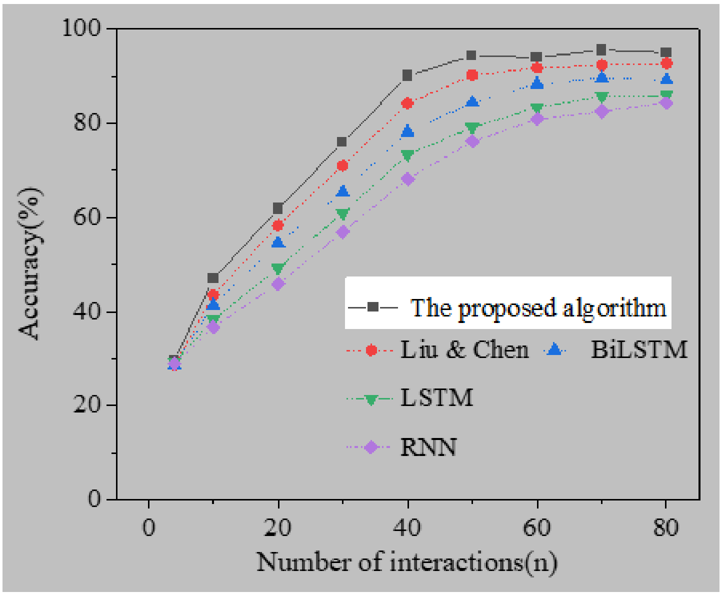
## RESULTS AND DISCUSSION

### Performance Analysis of Different Algorithm Models

The constructed algorithm is analysed from the perspectives of Accuracy, Precision, Recall, and F1 values, respectively, with LSTM, BiLSTM, RNN, and Liu and Chen (2023), as shown in Figures 6-9.

In Figure 6, as the number of iterations increases, the accuracy of user demand prediction shows a trend of first increasing and then stabilizing. The accuracy of the model algorithm in this paper is significantly better than that of other algorithms. When the number of iterations is 80, the accuracy reaches 95.12%. The accuracy of the model algorithm proposed by Liu and Chen (2023) is only 92.80%, and the prediction accuracy of the RNN algorithm is 84.37%. The accuracy of each algorithm is arranged in order from large to small as the proposed algorithm > the model algorithm proposed

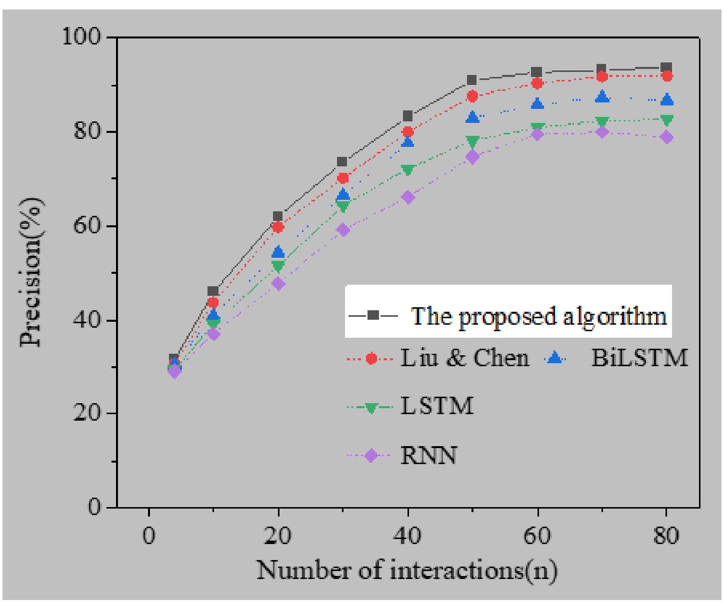
Figure 6. Accuracy results of user demand prediction in social media big data based on different algorithms




by Liu and Chen (2023) > BiLSTM > LSTM > RNN. Therefore, the constructed user demand prediction model based on BiLSTM fusion of Word2Vec in social media big data can make more accurate predictions of user demand in social media big data as the number of iterations increases.

In Figure 7, as the number of iterations increases, the prediction precision of user demand shows a trend of first increasing and then stabilizing. This paper's precision of the model algorithm is significantly better than other models'. When the number of iterations is 80, the precision reaches

Figure 7. Precision results of user demand prediction in social media big data based on different algorithms



93.81%. The precision of the model algorithm proposed by Liu and Chen (2023) is only 92.00%. The prediction precision of the RNN algorithm is 78.95%. The precision of each algorithm is arranged in order from large to small as the proposed algorithm > the model algorithm proposed by Liu and Chen (2023) > BiLSTM > LSTM > RNN. Therefore, the constructed BiLSTM fusion Word2Vec algorithm model can make more accurate predictions of user needs in big data social media.

In Figure 8, as the number of iterations increases, the predicted Recall of user demand shows a trend of first increasing and then stabilizing. This paper's Recall of the model algorithm is significantly better than others models'. When the number of iterations is 80, the Recall reaches 98.07%. The Recall of the model algorithm proposed by Liu and Chen (2023) is only 85.57%. The predicted Recall of the RNN algorithm is 72.60%. The Recall of each algorithm is arranged in order from large to small as the proposed algorithm > the model algorithm proposed by Liu and Chen (2023) > BiLSTM > LSTM > RNN.

In Figure 9, as the number of iterations increases, the predicted F1 value of user demand shows a trend of first increasing and then stabilizing. This paper's F1 value of the model algorithm is significantly better than other models'. When the number of iterations is 80, the F1 value reaches 85.97%. The F1 value of the model algorithm proposed by Liu and Chen (2023) is only 84.57%. The predicted F1 value of the RNN algorithm is 75.63%. The F1 value of each algorithm is arranged in order from large to small as the proposed algorithm > the model algorithm proposed by Liu and Chen (2023) > BiLSTM > LSTM > RNN. Therefore, the constructed BiLSTM fusion Word2Vec algorithm model can make more accurate predictions of user needs in big data social media.

## Discussion

This paper predicts user needs in social media big data. The results show that the proposed social media big data model algorithm based on BiLSTM fused with Word2Vec can achieve an accuracy of 95.12% when predicting user needs in social media. The accuracy of the model algorithm proposed by Liu and Chen (2023) is only 92.80%, and the prediction accuracy of the RNN algorithm is 84.37%, which is significantly better than the existing algorithms. This model can make more accurate predictions of

**Figure 8. Recall results of user demand prediction in social media big data based on different algorithms**

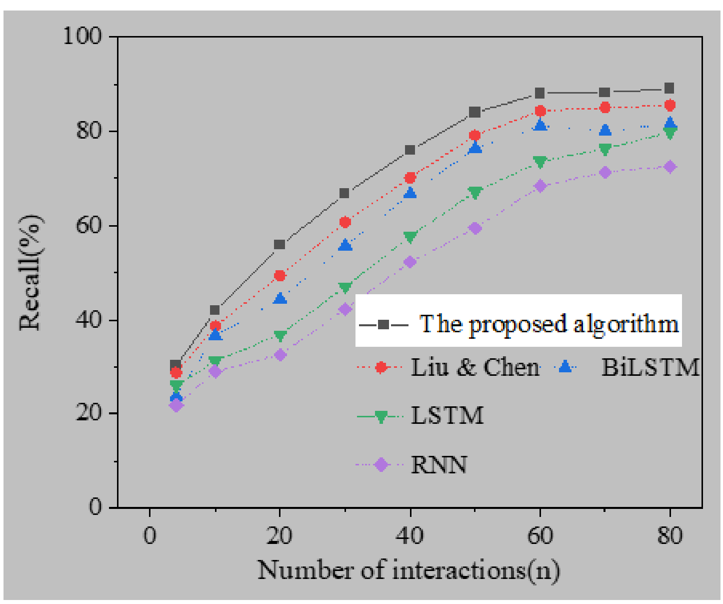
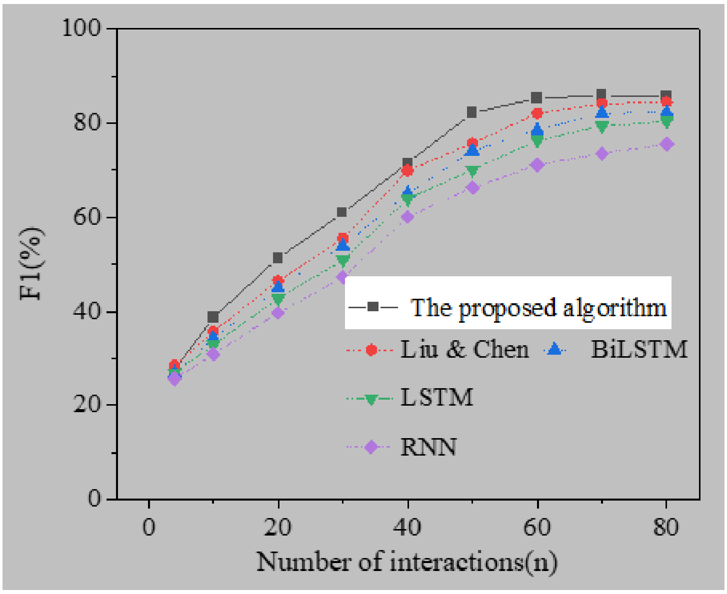


**Figure 9. F1 results of user demand prediction in social media big data based on different algorithms**




user needs in social media. Based on the experimental results of this model, marketing optimization strategies for products and services in enterprises in the new media environment are proposed.

Firstly, the demand-shaping ability of enterprise products has been enhanced. Enterprises can increase their demand-shaping capacity by developing strategic Weibo and spreading and influencing the purchasing decisions of others through powerful means such as opinion leaders and electronic word-of-mouth. Businesses can promote their products through influencers on social media. These individuals can post product information on social media, influencing other users' purchasing decisions. Secondly, the marketing of social media has been launched. Enterprises can assist in efficient marketing through homepage settings, fan headline discount privileges, personalized lottery conditions, multi-dimensional data analysis, card, coupon marketing, etc. (Dwivedi et al., 2022). Finally, a virtual community for sustainable enterprise product experiences is established (Wang & Dai, 2022). In virtual communities, users' understanding of enterprise products is deepened, effectively promoting their knowledge-sharing behavior and sustainable product experience development.

The proposed user demand prediction model based on BiLSTM fused with Word2Vec in social media big data has performed well in experiments and can more accurately predict user needs in social media. Additionally, the marketing optimization strategy of products and services in enterprises under the new media environment is proposed. These strategies have certain practical application value for the healthy and sustainable development of the social media field.

## CONCLUSION

In response to the inconsistency between the supply of enterprise products and services and market demand, deep learning has been introduced to construct a user demand prediction model in social media big data based on BiLSTM fused with Word2Vec. Finally, the model is evaluated as an example. Compared with other existing model algorithms, the prediction accuracy of this paper's model for user demand in social media big data reaches 95.12%. The prediction accuracy is improved by at least 2.31%. Additionally, combined with the experimental results, this paper proposes an optimization plan for the marketing strategy of products and services in enterprises under the new media environment.

Enterprises can enhance product demand-shaping capabilities, carry out social media marketing, and establish sustainable virtual communities for enterprise product experience. These strategies are consistent with the research results of relevant scholars and have certain feasibility and practical application value in practice. Therefore, this paper can make a more accurate prediction of user needs in social media big data and provide an experimental basis for optimizing marketing strategies for products and services in enterprises in the new media environment. However, some shortcomings still exist. In the instance evaluation, the data collection object is the original Weibo, and the number of times it is forwarded is not considered. Therefore, future research will use multiple keywords for search, including acquiring diverse content such as forwarding Weibo. This approach enables the model to extract more comprehensive user needs and conduct more in-depth research on optimizing the enterprise's marketing strategy.

## ACKNOWLEDGMENT

This work was supported by Project of Ningbo Yangming Culture Creative Change and Communication Base in 2021: The Study of the Creative Change Path form Yangming Culture from the Perspective of Urban Cultural Symbols. This work was supported by the Zhejiang Federation of Humanities and Social Sciences Circles in China or Zhejiang Federation of Humanities and Social Sciences in China (ZFHSS): The Cultural Conflict in Media Ethics Under the Background of Globalization and Enlightenment to China's International Communication (No. 17NDJC032Z).

## REFERENCES

Abkenar, S. B., Kashani, M. H., Mahdipour, E., & Jameii, S. M. (2021). Big data analytics meets social media: A systematic review of techniques, open issues, and future directions. *Telematics and Informatics*, *57*, 101517. doi:10.1016/j.tele.2020.101517 PMID:34887614

Al-Sarem, M., Alsaeedi, A., Saeed, F., Boulila, W., & AmeerBakhsh, O. (2021). A novel hybrid deep learning model for detecting COVID-19-related rumors on social media based on LSTM and concatenated parallel CNNs. *Applied Sciences (Basel, Switzerland)*, *11*(17), 7940. doi:10.3390/app11177940

Ali, F., Ali, A., Imran, M., Naqvi, R. A., Siddiqi, M. H., & Kwak, K. S. (2021). Traffic accident detection and condition analysis based on social networking data. *Accident; Analysis and Prevention*, *151*, 105973. doi:10.1016/j.aap.2021.105973 PMID:33461071

Ali, F., El-Sappagh, S., Islam, S. R., Ali, A., Attique, M., Imran, M., & Kwak, K. S. (2021). An intelligent healthcare monitoring framework using wearable sensors and social networking data. *Future Generation Computer Systems*, *114*, 23–43. doi:10.1016/j.future.2020.07.047

Behera, R. K., Jena, M., Rath, S. K., & Misra, S. (2021). Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data. *Information Processing & Management*, *58*(1), 102435. doi:10.1016/j.ipm.2020.102435

Chen, M., & Du, W. (2022a). Dynamic relationship network and international management of enterprise supply chain by particle swarm optimization algorithm under deep learning. *Expert Systems: International Journal of Knowledge Engineering and Neural Networks*. Advance online publication. doi:10.1111/exsy.13081

Chen, M., & Du, W. (2022b). The predicting public sentiment evolution on public emergencies under deep learning and internet of things. *The Journal of Supercomputing*, *79*(6), 6452–6470. doi:10.1007/s11227-022-04900-x

Chen, T., Peng, L., Yang, J., & Cong, G. (2021). Analysis of user needs on downloading behavior of English vocabulary APPs based on data mining for online comments. *Mathematics*, *9*(12), 1341. doi:10.3390/math9121341

Christian, H., Suhartono, D., Chowanda, A., & Zamli, K. Z. (2021). Text based personality prediction from multiple social media data sources using pre-trained language model and model averaging. *Journal of Big Data*, *8*(1), 1–20. doi:10.1186/s40537-021-00459-1 PMID:33425651

Dai, W. (2022). Application of improved convolution neural network in financial forecasting. *Journal of Organizational and End User Computing*, *34*(3), 1–16. doi:10.4018/JOEUC.289222

Dwivedi, Y. K., Shareef, M. A., Akram, M. S., Bhatti, Z. A., & Rana, N. P. (2022). Examining the effects of enterprise social media on operational and social performance during environmental disruption. *Technological Forecasting and Social Change*, *175*, 121364. doi:10.1016/j.techfore.2021.121364

Feng, B., Sun, K., Zhong, Z., & Chen, M. (2021). The internal connection analysis of information sharing and investment performance in the venture capital network community. *International Journal of Environmental Research and Public Health*, *18*(22), 11943. doi:10.3390/ijerph182211943 PMID:34831699

Feng, Z., & Chen, M. (2022). Platformance-based cross-border import retail e-commerce service quality evaluation using an artificial neural network analysis. *Journal of Global Information Management*, *30*(11), 1–17. doi:10.4018/JGIM.306271

Hamayel, M. J., & Owda, A. Y. (2021). A novel cryptocurrency price prediction model using GRU, LSTM and bi-LSTM machine learning algorithms. *AI, 2*(4), 477-496.

Kaladevi, P., & Thyagarajah, K. (2021). Integrated CNN-and LSTM-DNN-based sentiment analysis over big social data for opinion mining [retracted]. *Behaviour & Information Technology, 40*(9), xi-xix.

Khan, L., Amjad, A., Afaq, K. M., & Chang, H. T. (2022). Deep sentiment analysis using CNN-LSTM architecture of English and Roman Urdu text shared in social media. *Applied Sciences (Basel, Switzerland)*, *12*(5), 2694.

Korcsmáros, E., & Csinger, B. (2022). Sustainable competitiveness in the case of SMEs—Opportunities provided by social media in an international comparison. *Sustainability (Basel)*, *14*(19), 12505. doi:10.3390/su141912505

Lei, Z., Gong, G., Wang, T., & Li, W. (2022). Accounting information quality, financing constraints, and company innovation investment efficiency by big data analysis. *Journal of Organizational and End User Computing*, *34*(3), 1–21. doi:10.4018/JOEUC.292525

Liu, Y., & Chen, M. (2023). The knowledge structure and development trend in artificial intelligence based on latent feature topic model. *IEEE Transactions on Engineering Management*, 1–12. Advance online publication. doi:10.1109/TEM.2022.3232178

Liu, Y., Xu, C., Kuai, X., Deng, H., Wang, K., & Luo, Q. (2022). Analysis of the causes of inferiority feelings based on social media data with Word2Vec. *Scientific Reports*, *12*(1), 5218. doi:10.1038/s41598-022-09075-2 PMID:35338206

Meng, T., Li, Q., Dong, Z., & Zhao, F. (2022). Research on the risk of social stability of enterprise credit supervision mechanism based on big data. *Journal of Organizational and End User Computing*, *34*(3), 1–16. doi:10.4018/JOEUC.289223

Rajković, B., Đurić, I., Zarić, V., & Glauben, T. (2021). Gaining trust in the digital age: The potential of social media for increasing the competitiveness of small and medium enterprises. *Sustainability (Basel)*, *13*(4), 1884. doi:10.3390/su13041884

Sakas, D. P., Giannakopoulos, N. T., Terzi, M. C., Kamperos, I. D. G., Nasiopoulos, D. K., Reklitis, D. P., & Kanellos, N. (2022). Social media strategy processes for centralized payment network firms after a war crisis outset. *Processes (Basel, Switzerland)*, *10*(10), 1995. doi:10.3390/pr10101995

Tajpour, M., Hosseini, E., Ratten, V., Bahman-Zangi, B., & Soleymanian, S. M. (2023). The role of entrepreneurial thinking mediated by social media on the sustainability of small and medium-sized enterprises in Iran. *Sustainability (Basel)*, *15*(5), 4518. doi:10.3390/su15054518

Tao, D., Yang, P., & Feng, H. (2020). Utilization of text mining as a big data analysis tool for food science and nutrition. *Comprehensive Reviews in Food Science and Food Safety*, *19*(2), 875–894. doi:10.1111/1541-4337.12540 PMID:33325182

Tripathi, M. (2021). Sentiment analysis of Nepali covid19 tweets using nb svm and LSTM. *Journal of Artificial Intelligence*, *3*(3), 151–168.

Wadud, M. A. H., Kabir, M. M., Mridha, M. F., Ali, M. A., Hamid, M. A., & Monowar, M. M. (2022). How can we manage offensive text in social media-a text classification approach using LSTM-BOOST. *International Journal of Information Management Data Insights*, *2*(2), 100095. doi:10.1016/j.jjimei.2022.100095

Wang, J., & Dai, Y. (2022). The agglomeration mechanism of network emerging e-commerce industry based on social science. *Journal of Organizational and End User Computing*, *34*(3), 1–16. doi:10.4018/JOEUC.291561

Ye, S., & Chen, M. (2021). Leveraging team expertise location awareness in improving team improvisation: A dynamic knowledge integration perspective. *Psychology Research and Behavior Management*, *14*, 2135–2146. doi:10.2147/PRBM.S341685 PMID:34984035

Zhang, H., Fan, L., Chen, M., & Qiu, C. (2022). The impact of SIPOC on process reengineering and sustainability of enterprise procurement management in e-commerce environments using deep learning. *Journal of Organizational and End User Computing*, *34*(8), 1–17. doi:10.4018/JOEUC.306270

Zhou, X., Liang, W., Kevin, I., Wang, K., & Yang, L. T. (2020). Deep correlation mining based on hierarchical hybrid networks for heterogeneous big data recommendations. *IEEE Transactions on Computational Social Systems*, *8*(1), 171–178. doi:10.1109/TCSS.2020.2987846