Crowd Psychological and Emotional Computing Based on PSMU Algorithm

Bei He*
Jiangxi Modern polytechnic College, Nanchang 330095, China
[e-mail: 016087@jxxdxy.edu.cn]
*Corresponding author: Bei He

Received January 23, 2024; revised February 26, 2024; revised July 12, 2024; accepted July 27, 2024; published August 31, 2024

Abstract

The rapid progress of social media allows more people to express their feelings and opinions online. Many data on social media contains people's emotional information, which can be used for people’s psychological analysis and emotional calculation. This research is based on the simplified psychological scale algorithm of multi-theory integration. It aims to accurately analyze people’s psychological emotion. According to the comparative analysis of algorithm performance, the results show that the highest recall rate of the algorithm in this study is 95%, while the highest recall rate of the item response theory algorithm and the social network analysis algorithm is 68% and 87%. The acceleration ratio and data volume of the research algorithm are analyzed. The results show that when 400,000 data are calculated in the Hadoop cluster and there are 8 nodes, the maximum acceleration ratio is 40%. When the data volume is 8GB, the maximum scale ratio of 8 nodes is 43%. Finally, we carried out an empirical analysis on the model that compute the population’s psychological and emotional conditions. During the analysis, the psychological simplification scale algorithm was adopted and multiple theories were taken into account. Then, we collected negative comments and expressions about Japan's discharge of radioactive water in microblog and compared them with the trend derived by the model. The results were consistent. Therefore, this research model has achieved good results in the emotion classification of microblog comments.

Keywords: Credibility, item response theory, population psychological emotion calculation, Psychological reduction scale based on multi-theory fusion algorithm, web crawler.

http://doi.org/10.3837/tiis.2024.08.004 **ISSN : 1976-7277**

1. Introduction

With the wide application of social media, more and more people tend to express their emotion online. Many studies have adopted computer simulations to monitor and predict the movement behavior of crowds and analyze crowd psychology and emotion [1]. Crowd psychological and emotional computing can analyze users' comments and interactions on social media and offers in-depth understanding of people's attitudes and reactions to specific events, products or topics [2]. This can enhance enterprise decision-making and increase efficiency in brand management and market research [3]. The development of social media, which provides vast and diverse information, brings challenges to traditional sentiment analysis techniques. Traditional sentiment analysis techniques rely on text mining and machine learning to infer emotions and attitudes of people [4]. However, the models used are mainly based on the analysis of keywords, syntax, and semantics in a text, not an understanding of human's complex emotion and group behavior when calculating crowd psychology [5]. In this study, a Psychological reduction scale based on multi-theory fusion (PSMU) algorithm was proposed to calculate the psychological emotion of the population. Multi-theory fusion mainly includes classical measurement theory, generalization theory and item response theory. This PSMU algorithm can monitor and interpret of the crowd's psychological state in real time and provide in-depth analysis of crowd's speech and behavior. This study offers a more accurate method to calculate people's psychological and emotional data. This can promote the harmony and stability of society, improve public safety and ensure the sustainable development of the economy. The study consists of four sections. 1. Introduction. 2. The establishment of crowd psychology and emotion model and related research on the Hadoop Streaming framework and the PSMU algorithm. 3. Results analysis of the crowd psychological and emotional computing model that uses PSMU algorithm. 4. Conclusion. The acceleration ratio and data volume of this research algorithm are analyzed. When the number of nodes is 8, the maximum acceleration ratio is 40% when 400,000 data are calculated in the Hadoop cluster; when the data volume is 8GB, the maximum scale ratio of 8 nodes is 43%. The comparison of the changes in social behavior domain (SBD), behavior performance domain (BPD), emotional fluctuation domain (EFD), psychological state domain (PSD) and external environment domain (EED) before and after the simplification of the research algorithm shows that in the field of social behavior, the difference of consistency index (CSI) before and after simplification is the largest 0.4, indicating that the simplified algorithm has a wide application prospect in the field of social behavior analysis and can provide a new method for research in related fields.

2. Literature Review

Many studies have investigated psychological and emotional computing, which is able to enhance the understanding of human emotions and mental health. Deyreh and Asgarian proposed an emotional focus therapy to improve the mental health of mothers of children with specific learning disability. Thirty mothers of children with specific learning disability were selected as samples. Moreover, the experiment used a pre-test, post-test, control group quasi-experimental design and adopted MANCOVA dataset for data analysis. This research method has a significant effect [6]. To test the correlation between people's arts engagement and flourishing in young adults, Bone et al. analyzed data using fixed effects regression and Arellano-Bond methods to control for bidirectional relationships. The study measured flourishing in emotional, psychological, and social wellbeing. It also explored whether changes in the frequency of participation in artistic, musical, or theatrical organizational activities were associated with changes in overall flourishing. After controlling for bidirectional relationships, the study concluded that increased participation in the arts enhanced flourishing years later and improved social wellbeing [7]. Zhang et al. proposed a hybrid neural network approach for fine-grained emotion classification and computing. Their studies aim to address sparse feature selection and emotion classification, and an empirical analysis was conducted on the model. This approach increased the accuracy of emotion classification, exhibited a higher classification performance, reduced the number of iterations and saved computational resources [8]. Guangdi et al. proposed a model based on statistical methods and item reflection theory to simplify multiple psychological scales for children and adolescents in order to make a more accurate assessment of their psychological status [9].

The Hadoop Streaming framework is important to the Hadoop ecosystem as it allows developers to write MapReduce jobs using languages other than Java. Researchers around the world have worked to develop the Hadoop ecosystem. To better manage the data, Goyal and Bedi applied ontology based on H-matching algorithm. Their empirical analysis showed that the proposed algorithm can effectively process parameters in the big data while protecting the data's confidentiality, integrity and authenticity [10]. To overcome problems associated with the existing book collection management and document retrieval systems in universities, He proposed a semantic-based keyword extraction approach that is based on Hadoop algorithm. According to the result of the empirical analysis, the fast-matching method can accurately determine the weight of each keyword and make on-line literature retrieval systems more efficient and accurate. Based on Hadoop technology, the document retrieval systems for universities' online libraries are upgraded [11]. Huang and Wu proposed a Hadoop-based image retrieval system, which combines the bags-of-words approach with SVM classifier to solve the complex computational problems in image matching. According to the empirical analysis, this approach outperforms other latest methods on data processing [12]. In order to effectively utilize the parallel processing capability of cloud platform Hadoop framework, Liu et al. proposed an association rule algorithm for trusted cryptography module (TCM) construction analysis based on Hadoop framework, and compared the algorithm with apriori algorithm. The results show that, the association rule algorithm of TCM composition analysis based on Hadoop framework has better classification effect than apriori algorithm [13].

To sum up, many scholars have carried out extensive research on crowd emotion computing. At the same time. the application of PSMU algorithm and Hadoop Streaming framework has shown significant progress. However, few studies have applied both the PSMU algorithm and the Hadoop Streaming framework to the calculation of crowd emotion. Therefore, this paper has established a model to calculate crowd psychology and emotion based on PSMU algorithm. The new model, which combines PSMU algorithm with Hadoop Streaming framework, overcomes the limitations of the qualitative research and has a strong potential to be widely applied.

3. Research on Crowd Psychological Emotion Test Model Combining Hadoop Streaming Framework and PSMU Algorithm

This study has proposed a method that combines multiple theories with psychological scale simplification. It has built a computer adaptive measurement system to deal with the responses to a simplified mental health scale. This paper starts with the data collection and evaluation methods of psychometric assessments and proposes an efficient scale simplification method. Under the framework of Hadoop Streaming, it has also built a crowd psychological and emotional computing model.

3.1 Research on Psychological and Emotional Measurement Methods Based on PSMU Algorithm

Classical test theory is an important theoretical framework in psychology. It is used to explain and evaluate individual performance in cognitive, emotional, and behavioral aspects [14, 15]. According to classical test theory, the score obtained in the process of measurement is influenced by both the true score of the object and measurement error [16]. Classical test theory includes discrete trend method, correlation coefficient (CCo), Kronbach α coefficient method and item response theory. Among them, the dispersion degree indexes are commonly used in the discrete trend method. They include range, variance, standard deviation, and coefficient of variation. Eq. (1) shows the expression of coefficient of variation.

$$CV = \frac{s}{|\mu|} \times 100\% \tag{1}$$

In Eq. (1), μ represents the average and s represents the standard deviation. As an important components of classical test theory, CCo is a measure of linear correlation between two variables and tests whether the variables change in a similar way. Eq. (2) shows the Pearson CCo when the sum of two variables X and Y is given.

$$\rho_{X,Y} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \sigma_Y} \tag{2}$$

In Eq. (2), $\operatorname{cov}(X,Y)$ is the covariance of the sum of variables X and Y. $\sigma_X \sigma_Y$ is the standard deviation of the sum of the two variables. Kronbach α is commonly used as an indicator of internal consistency reliability. This tool measures the consistency between the various items. Eq. (3) shows the expression.

$$\alpha = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1}^{K} \sigma_{Yi}^2}{\sigma_X^2}\right) \tag{3}$$

In Eq. (3), K is the number of dimensional items and σ_{Yi}^2 is the variance of the score of a certain item. σ_X^2 Represents the variance of the total score for each item. Item Response Theory (IRT) is a theoretical model that evaluates measurement tools, such as tests and questionnaires. In this study, Samejima grade response model was adopted. Eq. (4) shows the expression that measures the score probability of items.

$$P_{ui}(\theta) = \frac{exp\left[-Da_i\left(\theta - b_{ui+1}\right)\right] - exp\left[-Da_i\left(\theta - b_{ui}\right)\right]}{\left\{1 + exp\left[-Da_i\left(\theta - b_{ui}\right)\right]\right\}\left\{1 + exp\left[-Da_i\left(\theta - b_{ui+1}\right)\right]\right\}} \tag{4}$$

In Eq. (4), θ represents the ability; i represents the entry; for variable a_i, i denotes the differentiation degree; u_i is the difficulty level; for b_{ui} i represents the difficulty level. Among them, D is 1.7. Emotion is a complex state of feeling that results in physical and psychological changes that influence thought and behavior. Experts from different fields have proposed various algorithms to quantify human emotion. Classic emotion representation algorithms include web crawler algorithm and left and right entropy algorithm. Web crawler, web spider or search engine bot automatically browses, downloads, and indexes content from the Internet [17, 18]. It aims to learn what every webpage on the web is about, so that the information can be retrieved when it's needed. The timing diagram of the web crawler process is shown in **Fig. 1**.

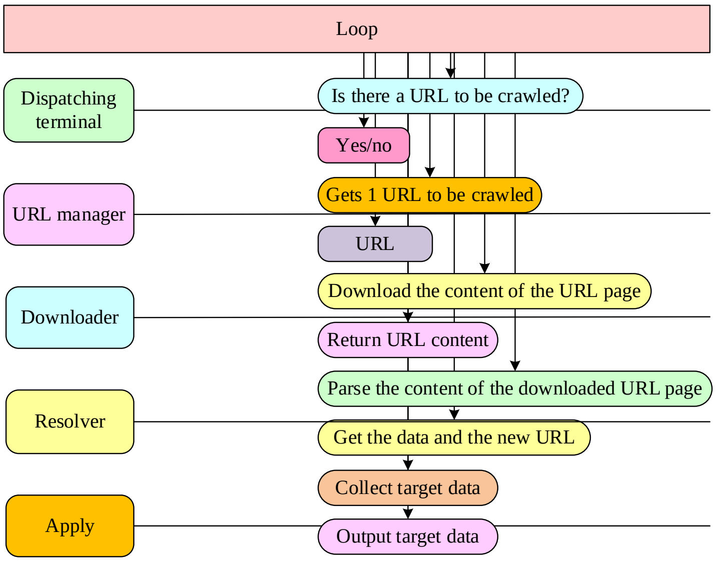


Fig. 1. Time sequence diagram of web crawler process.

Left-right entropy algorithm is often used for sentiment analysis of text data [19, 20]. Based on the frequency of each emotional word that appears in the context, it will select which emotion is expressed most strongly in the text, calculate the probability of different emotion categories, and derive the final classification result. Eq. (5) shows the expression of left-right entropy algorithm:

$$H_l = -\sum_{w_l \in s_l} p(w_l|w)\log_2 p(w_l|w) \tag{5}$$

In Eq. (5), w_l represents the left adjacency set; and s_l denotes the elements in the left adjacency set. To effectively analyze text data, the point-wise mutual information (PMI)-based polarity computation is performed and Eq. (6) shows the expression.

$$PML = \log_2\left(\frac{p(w_i, w_j)}{p(w_i) * p(w_j)}\right) \tag{6}$$

In Eq. (6), variables i and variables j are emotion words. The word “SO” refers to the polarity of emotion words during the analysis. Affective tendencies are represented different values, with positive values indicating positive emotions and negative values representing negative emotions. The expression is shown in Eq. (7).

$$SO(W) = PMI(W, \beta^{+}) - PMI(W, \beta^{-}) \tag{7}$$

In Eq. (7), β^{+} means positive emotional words and β^{-} represents negative emotional words. Association Mining searches for frequent items in the data set and discovers interesting relationships or associations among a set of items. The correlations between current data change series and the historical data change series fall on different set ranges, and the difference is large. Eq. (8) shows the two data change sequences.

$$\begin{cases} x = (x_1, x_2, ..., x_p) \\ y = (y_1, y_2, ..., y_q) \end{cases} \tag{8}$$

In Eq. (8), variables p and q are the dimensions of random variables. Eq. (9) shows the expression of the CCo, in the form of χ_i.

$$\chi_i = \frac{M_k + \min_i \min_k \left| y(k) - x_i(k) \right|}{M_k + \left| y(k) - x_i(k) \right|} \tag{9}$$

In Eq. (9), $x_i(k)$ represents the change sequence of the i parameter. Eq. (10) shows the expression of correlation degree $D(x_i, y)$.

$$D(x_i, y) = \sum_{k=1}^{n} \frac{\omega_k \chi_i(k)}{n} \tag{10}$$

In Eq. (10), the variable k represents the measuring point. A scale is a standardized instrument on which the characteristics are measured. Using a series of questions or statements, it assesses the degree or frequency of the characteristics, attitudes, behaviors, or emotions of the measured object. This study proposed an efficient method to calculate the population combination scale. **Fig. 2** shows the flow chart of the PSMU algorithm.

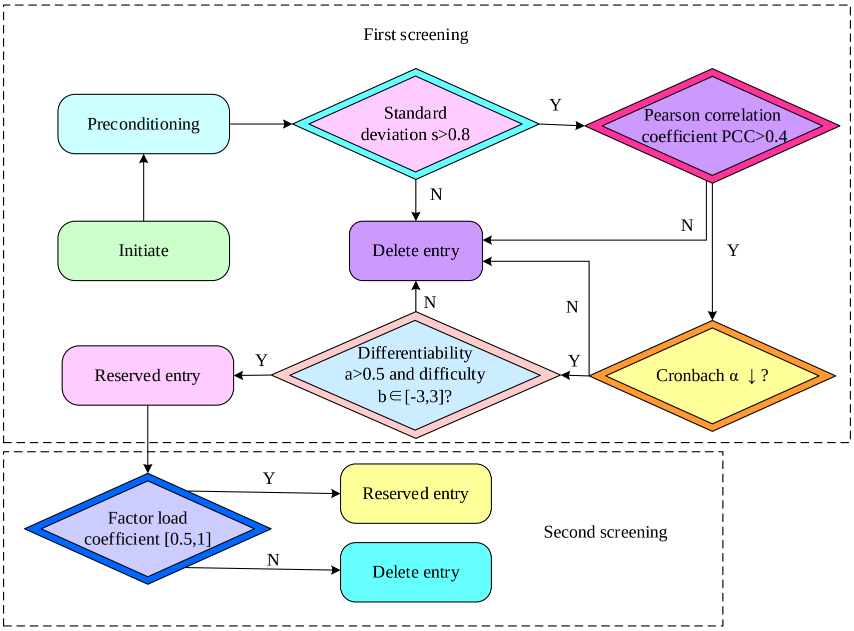


Fig. 2. PSMU algorithm flowchart.

According to **Fig. 2**, the PSMU firstly selects items using the discrete trend method, the CCo method and the Kronbach coefficient method. Next, the first entry filtering is carried out. Based on the results of the first entry filtering, the structural equation model is built and the second entry filtering is performed according to the factor load. Finally, the research model is adjusted and analyzed. **Fig. 3** is the diagram of model adjustment and analysis.

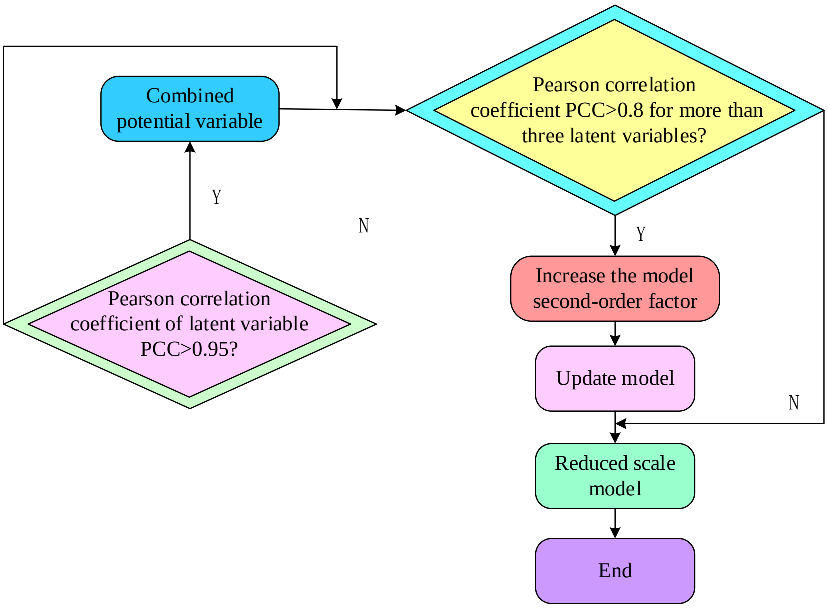


Fig. 3. Model adjustment and analysis diagram.

According to **Fig. 3**, the item structure is optimized and adjusted. Confirmatory factor analysis was adopted to test whether the data fit the hypothesized measurement model, Pearson

CCo of more than three latent variables was used to increase the second-order factor of the model. Finally, a simplified scale model that integrates multiple theories was built.

3.2 Construction of crowd psychological and emotional computing model with Hadoop Streaming framework

Hadoop is an open-source distributed computing framework for processing and storing large-scale data sets. The Hadoop Streaming framework is important to the Hadoop ecosystem as it allows developers to write MapReduce jobs using languages other than Java. The Hadoop Streaming framework enables developers to implement Map and Reduce functions using any programming language. The output of the mapper acts as input for Reducer which performs some sorting and aggregation operation on data and produces the final output. Therefore, the crowd psychological and emotional computing model based on PSMU algorithm will become more widely applicable with the introduction of Hadoop Streaming framework. **Fig. 4** shows the diagram of the Hadoop Streaming framework.

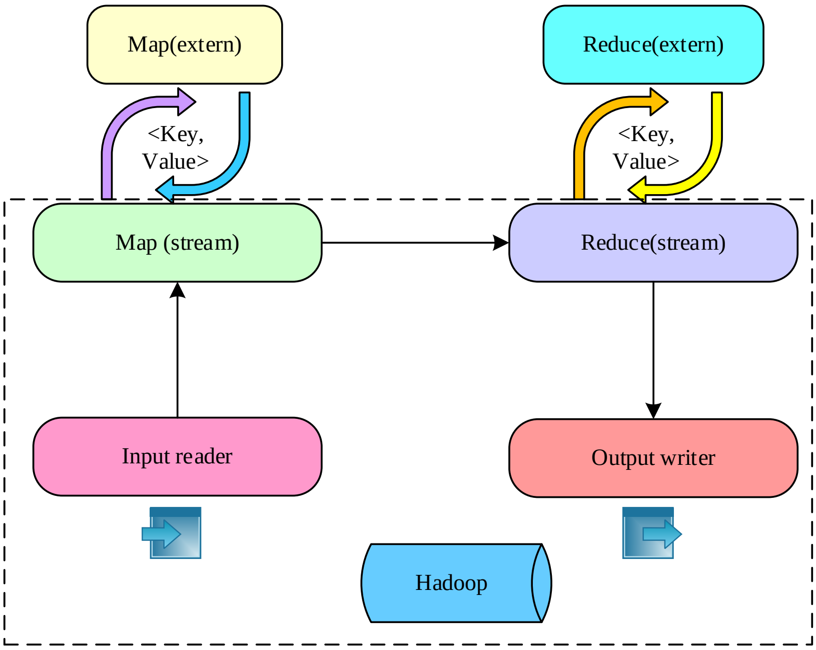


Fig. 4. Diagram of Hadoop Streaming framework.

According to **Fig. 4**, the developer submits the script file to the Hadoop cluster using the Hadoop Streaming command. The Hadoop cluster splits the input data into a series of key-value pairs and passes them to the Mapper. The Mapper processes the input data and generates key-value pairs of intermediate results. Depending on the number of reducers, the Hadoop delivers the intermediate results to the appropriate Reducer after sorting them by the key. Then, Reducer processes the received key-value pairs and produces the final output. The final output will be written to the Hadoop distributed file system or to other specified output location. The acceleration ratio measures to what extent the performance of a computing task in a parallel computing system has been improved. During the practice, the acceleration ratio may be affected by factors such as the load balancing of the parallel task, the communication overhead, and the parallelism degree of the parallel algorithm. As a result, we must take these factors into account when calculating the acceleration ratio in addition to doing experimental verification and performance analysis. The formula for calculating the acceleration ratio is provided in Eq. (11).

$$Speedup(n) = \frac{T(data,1)}{T(data,n)} \tag{11}$$

In Eq. (11), the nodes in the cluster are represented by a variable n; T is the execution time of a single processor; and $T(data,n)$ is the time required for parallel computation of n node. The left-right entropy is used for matrix calculation. Eq. (12) shows the formula of processor computing time T_p consumption.

$$T_p = \frac{n*m*m}{p} \tag{12}$$

In Eq. (12), p is the number of CPU cores, and the window size of left-right entropy n is represented by variables. Eq. (13) shows the efficiency of parallel computing E_p.

$$E_p = \frac{S_p}{p} = \frac{p}{p} = 100\% \tag{13}$$

In Eq. (13), S_p denotes the acceleration of parallel computation. Parallel scalability describes a system's capacity to effectively utilize an increasing number of processors. In the field of computer science, one single computing resource may not be able to process a large volume of data. Therefore, it is necessary to improve the processing capacity of the system through parallel expansion. Eq. (14) shows the formula.

$$Scaleup(data,n) = \frac{T(data,1)}{T(n \times data,n)} \tag{14}$$

In Eq. (14), variable n represents the number of compute nodes. $T(n \times data,n)$ represents the time consumed by multiple data of a compute node. Scale-growing methods can effectively handle the growth of data set. Eq. (15) shows its expression.

$$Sizeup(m) = \frac{T(m \times data,n)}{T(data,n)} \tag{15}$$

In Eq. (15), $T(m \times data,n)$ represents the time consumed by the m multiple data of the computing n node. This study combines characteristics of previous data and Web development technology to build a crowd psychology and emotion computing model under the Hadoop Streaming framework, as shown in **Fig. 5**.

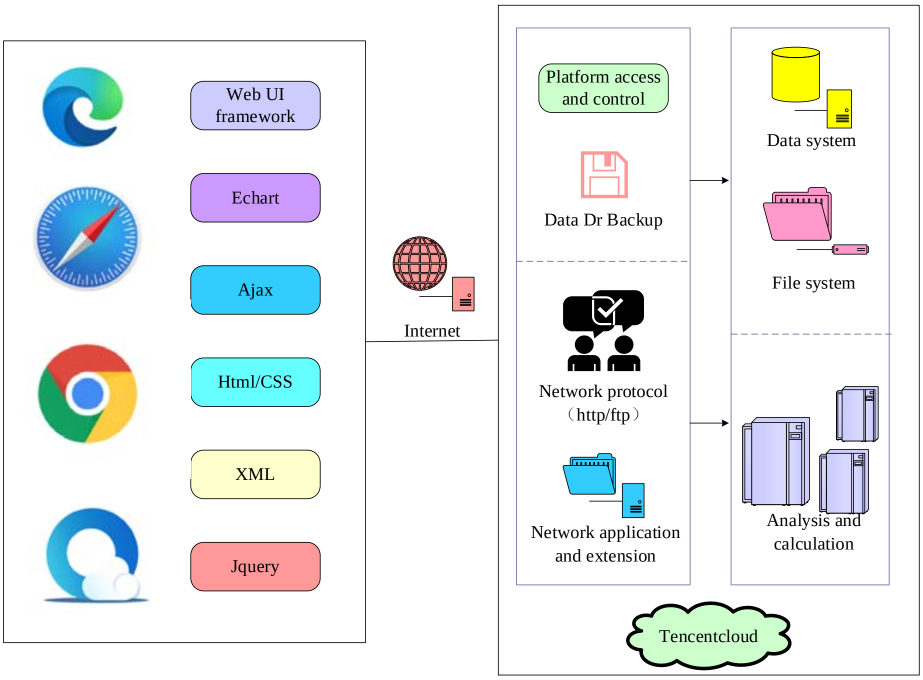


Fig. 5. Crowd psychological and emotional computing model based on Hadoop Streaming framework.

According to **Fig. 5**, during the design of Web front-end, this study aims to create user-friendly pages with a good interactive experience. The pages should offer visualized analysis results and be compatible with various Web browsers. As for the back end of the Web, or the server side, the focus is to efficiently sort out data and support the extension of logic and multiple languages. The research consists of two sections: the establishment of system architecture and function implementation. Based on advanced technology, it incorporates the Hadoop Streaming framework to refresh the data quickly. Lastly, the research model connects to a relational database which stores psychological data and user access history. From the database, it can get a complete psychological data set.

4. Analysis of Crowd Psychological and Emotional Computing Model with PSMU Algorithm

We tested the performance of the algorithm to verify its superiority and adaptability. The assessment of population mental health was mainly from five aspects: psychology, behavior, emotion, interpersonal relationships, and environment. The algorithm was tested on real data for validity and reliability. Finally, the empirical analysis of the research model was conducted.

4.1 Performance Evaluation of Emotion Test Calculation Method Based on PSMU Algorithm

This study selected accuracy rate, recall rate and F1 value as the evaluation indicators. F1 values was adopted to ensure better accuracy and recall. PSMU algorithm, item response theory (IRT) algorithm and social network analysis (SNA) were used to classify emotion on the test set. **Fig. 6** shows the experimental results.

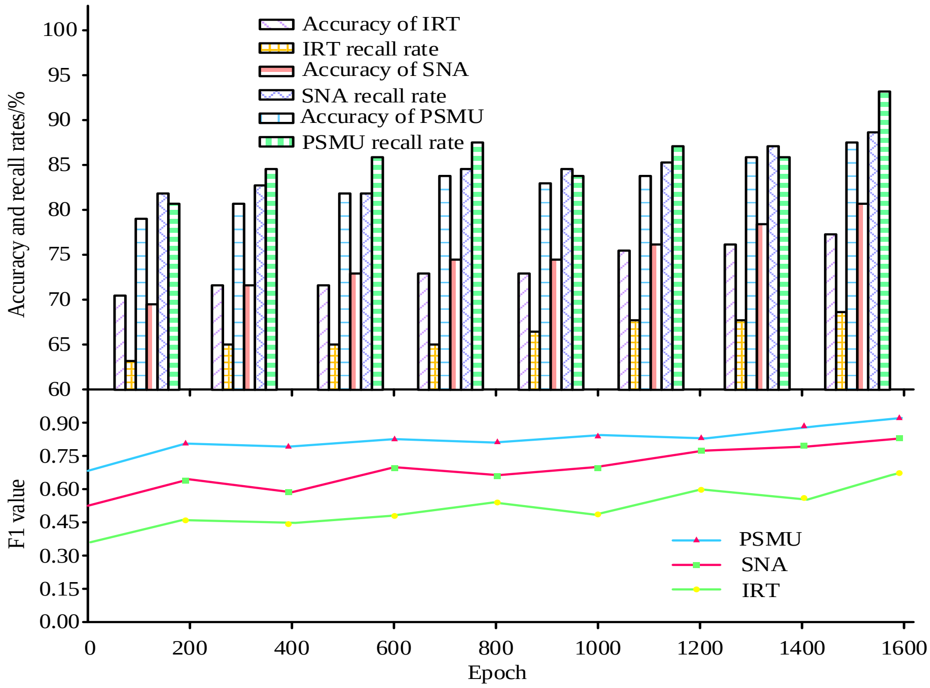


Fig. 6. Accuracy, recall rate, F1 value performance comparison chart.

According to **Fig. 6**, as the number of iterations increased to 1,600, the accuracy rate and recall rate of PSMU algorithm are the highest, which are 89% and 95%, respectively. At the same time, all the three algorithms' F1 line graphs display an upward trend. Among them, the F1 value of the PSMU algorithm fluctuates around 90%; that of the F1 value fluctuates around 75%; and that of the IRT algorithm fluctuates around 45%. In terms of classification, the PSMU algorithm performs the best. After the algorithm is written as an executable program, it needs time and space resources to run. Therefore, to measure the quality of an algorithm, it is generally measured from two dimensions of time and space, namely time complexity and space complexity. Time complexity is a measure of how fast an algorithm is running, while space complexity is a measure of how much extra space an algorithm needs to run. The computational complexity, time complexity or space complexity diagram of the research algorithm is shown in **Fig. 7**.

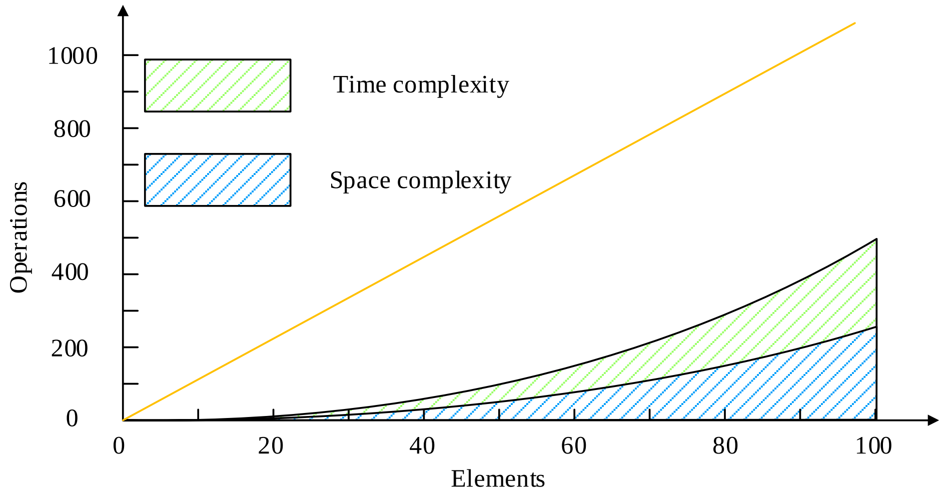


Fig. 7. The computational complexity, time complexity or spatial complexity diagram of the algorithm is studied.

As can be seen from **Fig. 7**, as the amount of data increases, the time complexity curve of the research algorithm is below the linear order of the orange line, which indicates that the algorithm has excellent performance. Similarly, although the space resources required by the research algorithm gradually increase with the increase of the amount of data, its spatial complexity is still below the linear order, which indicates that the algorithm also performs well in terms of spatial efficiency. To verify the parallel performance, the experiment uses two classic evaluation indexes of parallel algorithms and 8 Hadoop cluster nodes to conduct experiments on acceleration ratio and scale ratio. **Fig. 8** shows the experimental results.

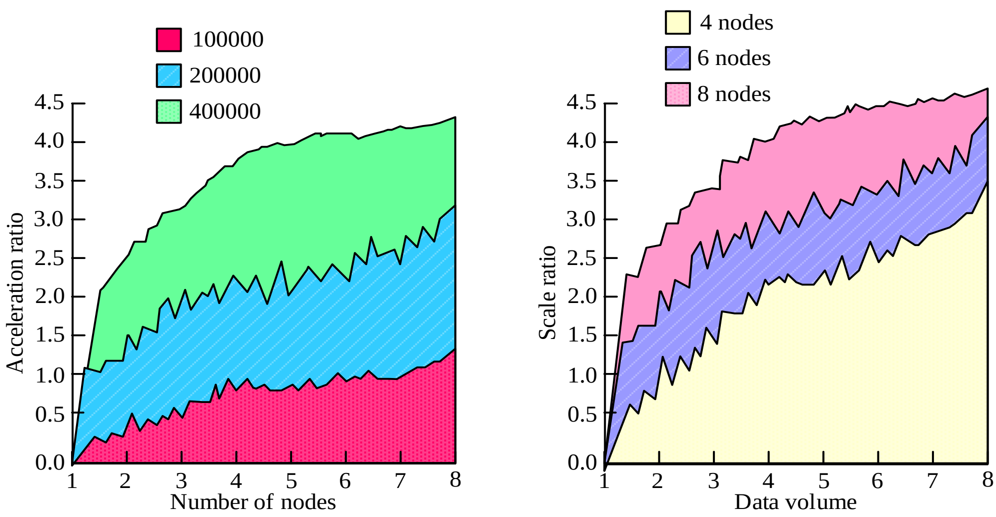


Fig. 8. Acceleration ratio and scale growth diagram.

According to **Fig. 8(a)**, the acceleration ratio curve of the proposed algorithm shows an upward trend as the nodes increase. The highest acceleration ratio is 40% when the number of nodes is 8 and the data calculated in the Hadoop cluster reaches 400,000. The acceleration ratio of the proposed algorithm when the data is in large amount. In **Fig. 8(b)**, the scale ratio curve in this study shows an increasing trend as the data volume increases. The scale ratio of nodes 8 is the highest when the data volume is 8GB, accounting for 43% of the total. To verify the effectiveness of the population combination scale, this study randomly selected 14,034 residents from the city for the experiment and calculated the fitness indicators of 100 residents in terms of SBD, BPD, EFD, PSD and EED. The fitness indicators include: similarity index (SMI), CSI, compatibility index (CPI), perception index (PCI) and adaptability index (API), are shown in **Table 1**.

Table 1. Model fit in each domain

Domain	SMI	CSI	CPI	SRMR	RMSEA	χ2/df
SBD	0.803	0.829	0.839	0.054	0.048	5.789
BPD	0.829	0.938	0.928	0.045	0.069	8.759
EFD	0.925	0.932	0.948	0.036	0.059	6.737
PSD	0.938	0.948	0.938	0.037	0.063	8.393
EED	0.739	0.849	0.839	0.084	0.039	3.085

According to **Table 1**, among the five psychological indicators, the SMI index was lowest (73.9%) in external environment, and highest (93.8%) in mental state. The CSI index was the lowest in social behavior (82.9%) and the highest in psychological state (94.8%). In the field of emotion fluctuation, the standardized residual root mean square value is the lowest (36%). As for external environment, the approximate error of root mean square is the lowest (39%). **Fig. 8** shows the changes in mental state, behavior, emotions, interpersonal relationships, and environmental adaptation before and after the simplification by the proposed algorithm.

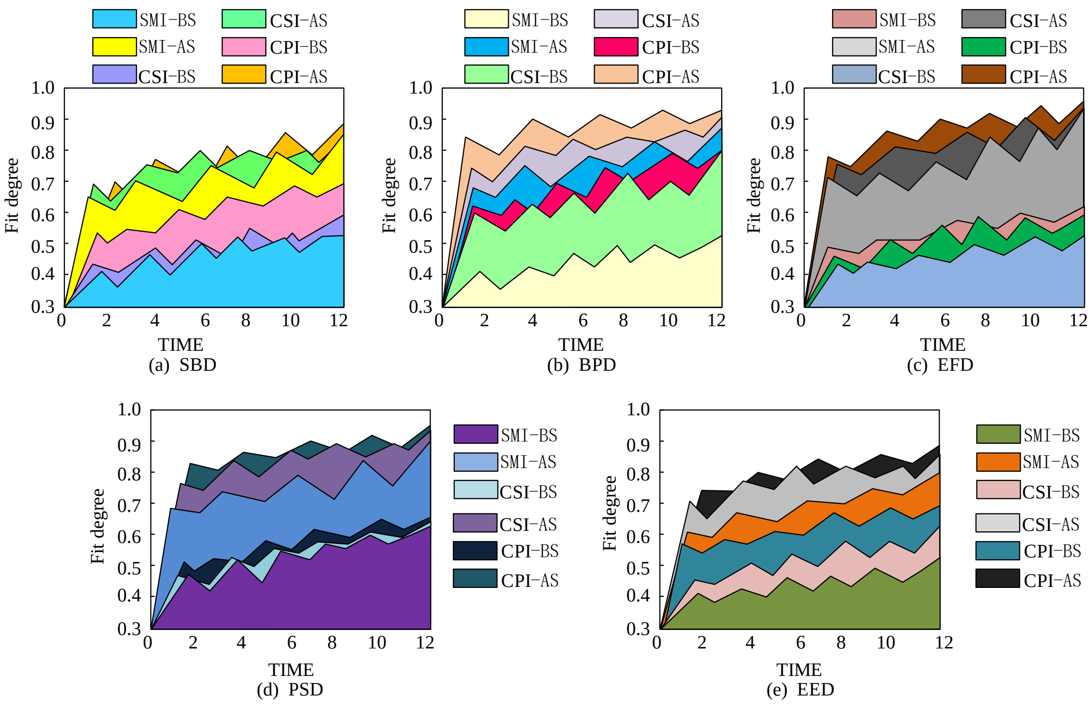


Fig. 9. Changes of model fit index before and after screening.

Before and after the simplification, there is a difference in social behavior's CSI index as great as 0.4, as shown in **Fig. 9(a)**. According to **Fig. 9(b)**, the largest difference of behavioral performance's SMI index is 0.35. In **Fig. 9(c)**, in terms of emotion fluctuation, the largest difference of CPI index before and after simplification is 0.2. According to **Fig. 9(d)**, the difference of psychological states' CSI index before and after simplification is as high as 0.4. According to **Fig. 9(e)**, in the field of external environment, the largest difference of SMI index before and after simplification is 0.4. The indicators above prove that the proposed algorithm has a much better fitness than other algorithms that haven't been simplified.

4.2 Empirical Analysis of PSMU-based Crowd Psychology and Emotion Computing Model

To verify the high efficiency of the crowd's psychology and emotion computing model based on the PSMU algorithm, this study assessed the posts of all Weibo users about Japan's discharge of radioactive water in August 2023. It extracted the content of the posts, counted the number of retweets and comments and conducted visual analysis. It also classified and sorted out the number of negative expressions in daily posts on the microblog about Japan's discharge of radioactive water and used it as a basic standard to measure the change of public negative emotions. **Fig. 10** shows the results.

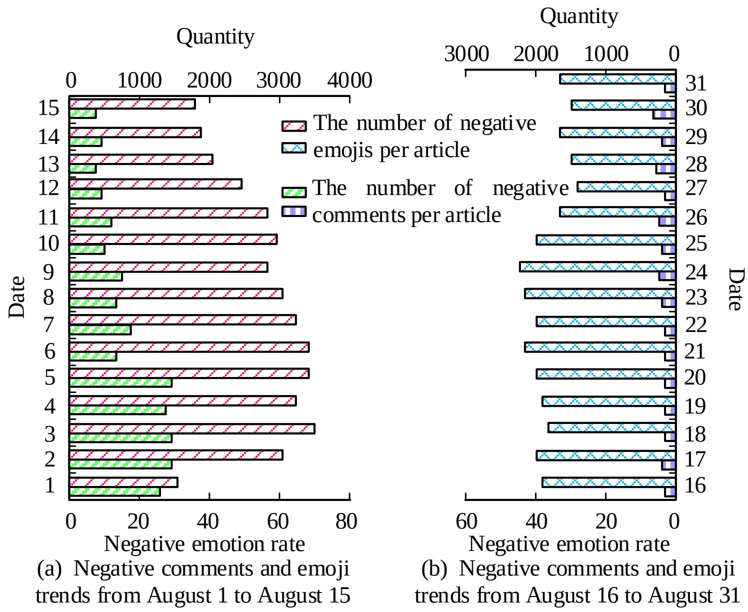
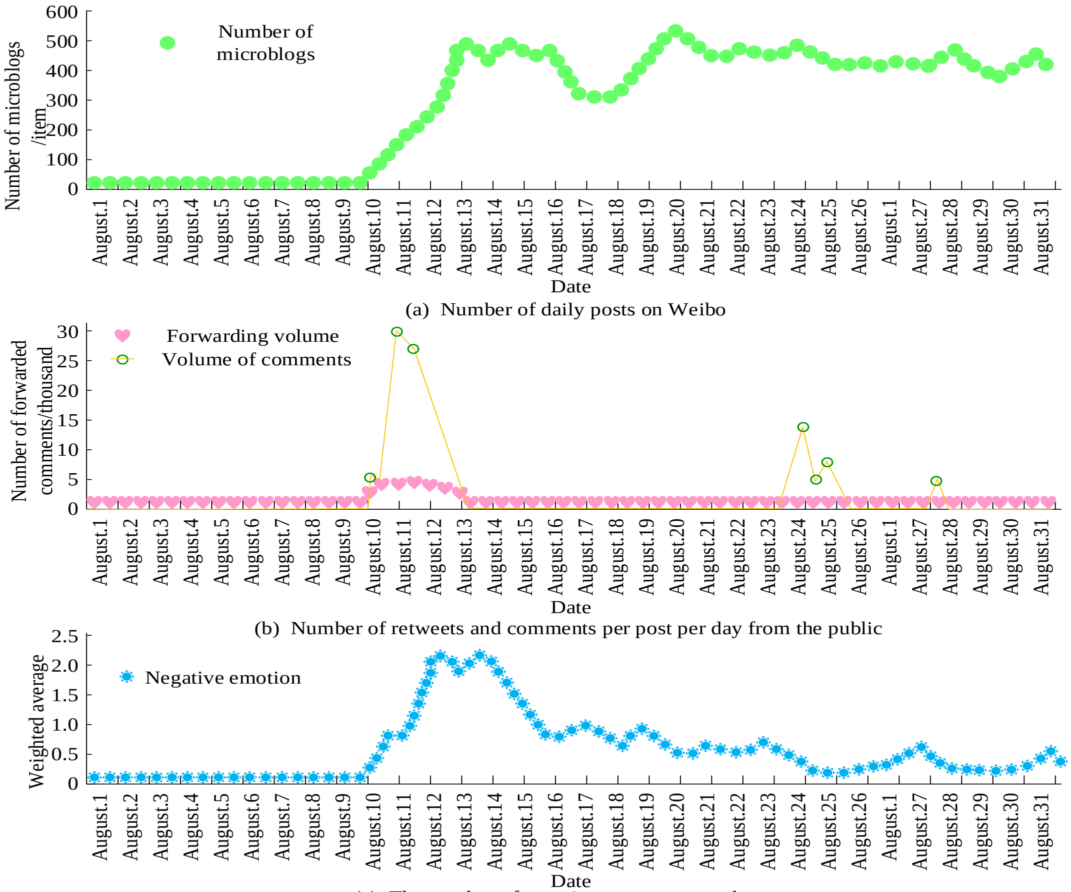


Fig. 10. Negative comments versus negative emoji trends.

From August 1 to August 31, the trend of negative comments about Japan's discharge of radioactive water was consistent with the negative expressions on Weibo comments, as shown in **Fig. 10**. This indicates that the model has achieved good results in the emotion classification of microblog comments. To analyze the trend of the psychology and emotion of the population, the research assessed the posts of all Weibo users about Japan's action to discharge radioactive water. It extracted the content of the posts, and counted the number of reposts and comments, and carried out visual analysis. The results are shown in **Fig. 11**.



(a) Number of daily posts on Weibo

(b) Number of retweets and comments per post per day from the public

(c) The number of negative comments on the post

Fig. 11. The number of tweets, the number of retweets and comments, and the change trend of negative emotions during the study period.

When the daily number of microblog posts about Japan's nuclear pollution discharge surged on August 11, as shown in **Fig. 11(a)**. **Fig. 11(b)** indicated a significant change in the number of daily reposts and comments on Weibo about Japan’s discharge of radioactive water on August 11. According to **Fig. 11(c)**, the level of public negative emotion changes with the number of negative comments per article per day. Besides accurately identifying the negative emotions of the public, the model also presents two different functions: "emotional word search" and "public psychological stress". After selecting dates and clicking the "Download" button, model users can access detailed analysis results, as shown in **Fig. 12**.

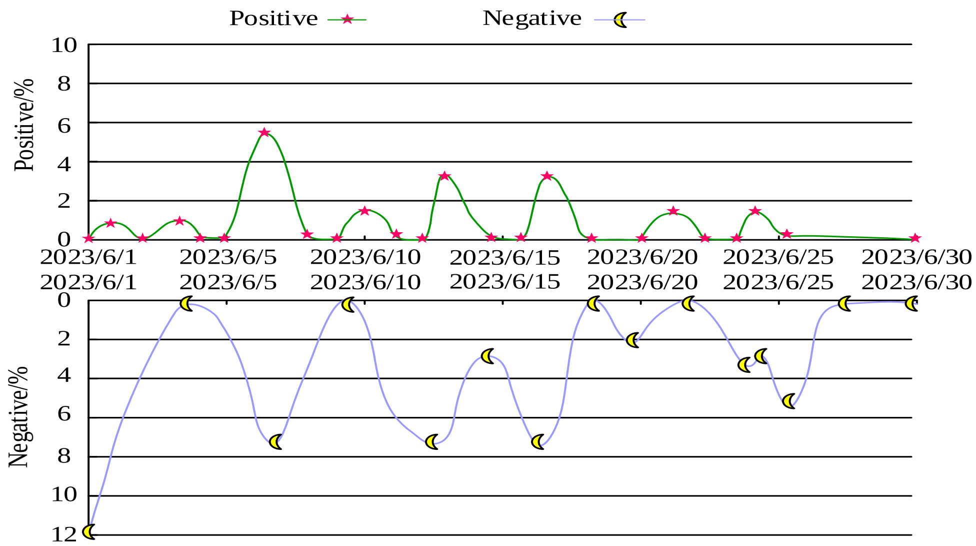


Fig. 12. A visual comparison of positive and negative emotions.

According to **Fig. 12**, before June 5th, users' positive emotions were in a relatively low state, and negative emotions increased sharply. After June 5th, positive emotions gradually increased, and negative emotions gradually decreased. As time went by, the positive emotions and negative emotions shows alternately. This research model has a high application value, as it can accurately recognize users' emotion and identify the crowds' psychology and emotions.

5. Conclusion

Human emotions are complex and the expression is hard to detect. Based on PSMU algorithm, this study proposed a model that calculated human's psychology and emotion. Firstly, it compared the algorithm's performance with others and found that the proposed algorithm had the highest accuracy and recall rate, which were 89% and 95% respectively. Then, it analyzed the acceleration ratio and data volume of the research algorithm and discovered that when the number of nodes was 8, the acceleration ratio of 400,000 data in the Hadoop cluster was 40% being the highest among other algorithms. Besides, when the data volume was 8 GB, the scale ratio of 8 nodes was 43%, also being the highest value than others. Before and after the simplification of the research algorithm, the changes in mental state, behavior, emotion, interpersonal relationships and environmental adaption were compared. Before and after the simplification, the difference of social behavior's CSI index was the largest, which was 0.4. Finally, the study conducted an empirical analysis of the PSMU-based crowd psychology and emotion computing model. It found that the model's results about peoples' responses to Japan's actions to discharge radioactive water was consistent with the negative comments and expressions in microblog. The study assessed the posts of all Weibo users about Japan's discharge of radioactive water from Fukushima. After extracting the content of posts, counting the number of reposts and comments, and performing visual analysis, it discovered that during the release of wastewater from the Fukushima Daiichi nuclear power plant, the level of public negative emotions changed with the number of negative comments per article per day. To sum up, this model is highly accurate in user emotion recognition. This study's limitation is that the amount of data on social media platforms is so huge that it requires the use of efficient algorithms and techniques to process it.

Data Availability Statement

All data generated or analysed during this study are included in this article.

References

[1] N. L. Webster, J. R. Oyebode, C. Jenkins, S. Bicknell, and A. Smythe, "Using technology to support the emotional and social well-being of nurses: A scoping review," *Journal of Advanced Nursing*, vol.76, no.1, pp.109-120, Jan. 2020. Article (CrossRef Link)

[2] V. S. Y. Cheung, J. C. Y. Lo, D. K.W. Chiu, and K. K.W. Ho, "Evaluating social media's communication effectiveness on travel product promotion: Facebook for college students in Hong Kong," *Information Discovery and Delivery*, vol.51, no.1, pp.66-73, Jan. 2023. Article (CrossRef Link)

[3] A. Özdemir, A. Özkan, Z. Günkaya, and M. Banar, "Decision-making for the selection of different leachate treatment/management methods: The ANP and PROMETHEE approaches," *Environmental Science and Pollution Research*, vol.27, no.16, pp.19798-19809, Jun. 2020. Article (CrossRef Link)

[4] D. Gan, X. Ge, and Q Li, "An Optimal Transport-Based Federated Reinforcement Learning Approach for Resource Allocation in Cloud–Edge Collaborative IoT," *IEEE Internet of Things Journal*, vol.11, no.2, pp.2407-2419, Jan. 2024. Article (CrossRef Link)

[5] C. Ma, B. Li, and J. He, "The improved fault location method for flexible direct current grid based on clustering and iterating algorithm," *IET Renewable Power Generation*, vol.15, no.15, pp.3577-3587, Jul. 2021. Article (CrossRef Link)

[6] E. Deyreh, and S. Asgarian, "Effectiveness of Emotion-focused Therapy-based Enrichment Program on Psychological Well-being, Marital Relationship Quality and Marital Satisfaction in Mothers of Children with Specific Learning Disabilities," *Razavi International Journal of Medicine*, vol.9, no.1, pp.63-68, 2021. Article (CrossRef Link)

[7] J. K. Bone, F. Bu, J. K. Sonke, and D. Fancourt, "Longitudinal Associations Between Arts Engagement and Flourishing in Young Adults: A Fixed Effects Analysis of the Panel Study of Income Dynamics," *Affective Science*, vol.4, no.1, pp.131-142, Mar. 2023. Article (CrossRef Link)

[8] W. Zhang, M. Wang, Y. Zhu, J. Wang, and N. Ghei, "A hybrid neural network approach for fine-grained emotion classification and computing," *Journal of Intelligent & Fuzzy Systems*, vol.37, no.3, pp.3081-3091, Oct. 2019. Article (CrossRef Link)

[9] L. Guangdi, L. Yu C., W. Yue, L. Jing Xiang, S. Yong Sheng, Z. Wei, and Z. Le, "Research on Psychological Scales Based on the Multitheory Fusion," *Current Bioinformatics*, vol.15, no.7, pp.741-749, Aug. 2020. Article (CrossRef Link)

[10] S.B. Goyal, and P. Bedi, "Triple DES-WSA: an efficient security algorithm providing integrity with aid of MapReduce process," *International Journal of Spatio-temporal Data Science*, vol.1, no.2, pp.131-148, Jul. 2021. Article (CrossRef Link)

[11] H. He, "Designing a Document Retrieval Method for University Digital Libraries Based on Hadoop Technology," *Journal of Contemporary Educational Research*, vol.5, no.12, pp.82-87, Dec. 2021. Article (CrossRef Link)

[12] Y. F. Huang, and H. Y. Wu, "Image retrieval based on ASIFT features in a Hadoop clustered system," *IET Image Processing*, vol.14, no.1, pp.138-146, Jan. 2020. Article (CrossRef Link)

[13] Y. Liu, X. Zhang, M. Zhang, J. Zhu, L. Wang, and Q. Wu, "A TCM Constitution Analysis Algorithm Based on Association Rules in Hadoop Framework," in *Proc. of 2020 35th Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pp.172-177, Oct. 2020. Article (CrossRef Link)

[14] M. Sharma, K. C. Koenen, C. P. C. Borba, D. R. Williams, and D. K. Deng, "The measurement of war-related trauma amongst internally displaced men and women in South Sudan: Psychometric analysis of the Harvard Trauma Questionnaire," *Journal of Affective Disorders*, vol.304, pp.102-112, May 2022. Article (CrossRef Link)

[15] Y. Yang, and X. Song, "Research on Face Intelligent Perception Technology Integrating Deep Learning under Different Illumination Intensities," *Journal of Computational and Cognitive Engineering*, vol.1, no.1, pp.32-36, Jan. 2022. Article(CrossRefLink)
[16] A. R. N. Tabrizi, and P. Etemad, "Check the grammar of your translation: The usefulness of WhiteSmoke as an automated feedback program," *Digital Scholarship in the Humanities*, vol.36, no.2, pp.411-420, Jun. 2021. Article (CrossRef Link)
[17] R. K. Nelligan, R. S. Hinman, J. Kasza, S. J. C. Crofts, and K. L. Bennell, "Effects of a Self-directed Web-Based Strengthening Exercise and Physical Activity Program Supported by Automated Text Messages for People With Knee Osteoarthritis: A Randomized Clinical Trial," *JAMA Internal Medicine*, vol.181, no 6, pp.776-785, Jun. 2021. Article (CrossRef Link)
[18] S. Yu, B. Wang, and T. Lu, "Sentiment Lexicon Construction Based on Improved Left-Right Entropy Algorithm," *Journal of Donghua University (English Edition)*, vol.39, no.1, pp.65-71, 2022. Article (CrossRef Link)
[19] P. L. dos Santos, and N. Wiener, "By the content of their character? Discrimination, social identity, and observed distributions of income," *The Journal of Mathematical Sociology*, vol.44, no.1, pp.12-41, Jun. 2019. Article (CrossRef Link)
[20] T. Mahmood, and Z. Ali, "Prioritized Muirhead Mean Aggregation Operators under the Complex Single-Valued Neutrosophic Settings and Their Application in Multi-Attribute Decision-Making," *Journal of Computational and Cognitive Engineering*, vol.1, no.2, pp.56-73, 2022. Article (CrossRef Link)

Bei He, She received B.S. and M.S. degrees in School of Psychology from Hunan University of Chinese Medicine in 2009 and Jiangxi Normal University in 2012.
Since Sep 2012, she is working in the Jiangxi Modern Polytechnic College as a psychology teacher.
Her current research interests include Psychological Health Education for college student.