Short Video Recommendation Method Based on Sentiment Analysis and K-means++

Rong Hu*, Wei Yue
School of Visual Arts, Hunan Mass Media Vocational and Technical College, Changsha, 410100, China

***Abstract*—With the explosive growth of short video content, effectively recommending videos that interest users has become a major challenge. In this study, a short video recommendation model based on barrage sentiment analysis and improved K-means++ was raised to address the interest matching problem in short video recommendation systems. The model uses sentiment vectors to represent bullet content, clusters short videos through sentiment similarity calculation, and studies the use of clustering density to eliminate abnormal sample points during the clustering process. The study validated the effectiveness of the raised model through simulation experiments. The outcomes denoted that when the historical data size increased to 7000, the model's prediction accuracy could reach 0.81, recall rate was 0.822, and F1 value was 0.832. Compared with the current four mainstream recommendation algorithms, this model showed advantages in clustering time and complexity, with clustering time reduced to 8.2 seconds, demonstrating the efficiency of the model in raising recommendation efficiency and accuracy. In summary, the model proposed in the study has high recommendation accuracy in short video recommendation systems and meets the real-time demands of short video recommendation, which can effectively raise the quality of short video recommendations.**

***Keywords*—Short videos; barrage; sentiment analysis; K-means++; recommendation; cluster density**

I. Introduction

As the quick advancement of the digital age, short video content has experienced explosive growth, providing rich leisure and entertainment methods and information sources for online users. Meanwhile, the rise in the number of short videos has also brought a challenge, which is how to quickly and accurately find content that attracts users in the vast video library. This problem has made the research of short video recommendation systems a focus of attention. In the past, short video recommendation systems relied on traditional collaborative filtering algorithms or content recommendation algorithms. Although these algorithms can achieve content recommendation to a certain extent, there are still several core issues that have not been resolved. Firstly, the accuracy of video recommendations is insufficient to accurately match the diverse and personalized needs of users. There are limitations in exploring the preferences of users, which can easily overlook the sentiment needs of users [1-2]. Shao P et al. believe that the collaborative filtering algorithm has biased prediction results due to the user's sensitive attributes, so the correlation of sensitive attributes is reduced in the prediction rule, and then a fairer recommendation model is proposed, and the performance of this model is verified in the real data set [3]. Wu B improved the traditional filtering method to solve the problems of low accuracy and low user interest. In the traditional method, data such as periodic update and trust are introduced to match the characteristics of news data with users' preferences, and finally complete the recommendation. This method has been proved by experiments to have a good performance in news recommendation [4].

In terms of emotional response to video content, the recommendation system can only judge based on video labels, and it is difficult to extract the emotional tendency of video, and lacks an effective mechanism to capture users' emotional response to video content, resulting in deviations between the recommendation results and the actual needs of users. Souza MLF et al. proposed a multi-channel bullet-screen text emotion analysis model, which first characterized the dynamic features of bullet-screen text, then encoded the sentences, and finally used dynamic routing to obtain the relationship features between local text and global text. This model achieves a good recognition effect in the data set of this paper [5].

The main research issues are how to effectively identify the emotional information of short video through bullet-screen emotion analysis, how to extract and quantify the emotional tendency of users to optimize the performance of the recommendation system, and how to improve the clustering algorithm to improve the accuracy of short video recommendation. The research goal is to build an emotion dictionary and emotion analysis model based on bullet screen data, and develop a short video recommendation system based on emotion analysis combined with emotion analysis results.

Therefore, this paper proposes a short video classification model based on emotion analysis of bullet screen and realizes the recognition of users' emotional preferences by constructing emotion dictionary. Then, the improved K-Means ++ algorithm is introduced to perform cluster analysis of video categories, to realize personalized video recommendation based on users' emotional preferences.

The research aims to construct a short video recommendation model that can accurately match user interests and sentiment needs by integrating barrage sentiment analysis and improving the K-means++. This model is expected to raise the accuracy and user satisfaction of the recommendation system. The research innovation lies in combining barrage sentiment analysis with improved K-means++ to introduce sentiment recognition mechanism into short video recommendation systems. By improving the K-means++, the selection of initial clustering centers is optimized.

The research is composed of six sections. Section II summarizes the research achievements of domestic and foreign scholars on sentiment analysis and video recommendation methods, and analyzes the shortcomings of current research. The Section III is to extract short video barrage comments for sentiment analysis, construct a short video sentiment classification model, and then optimize and improve the clustering method of K-means using clustering density. Section IV is to carry out simulation experiments on the proposed model to determine the optimal parameters of the model, and verify the effectiveness and progressiveness of the research method through comparative experiments. Discussion and conclusion is given in Section V and Section VI respectively.

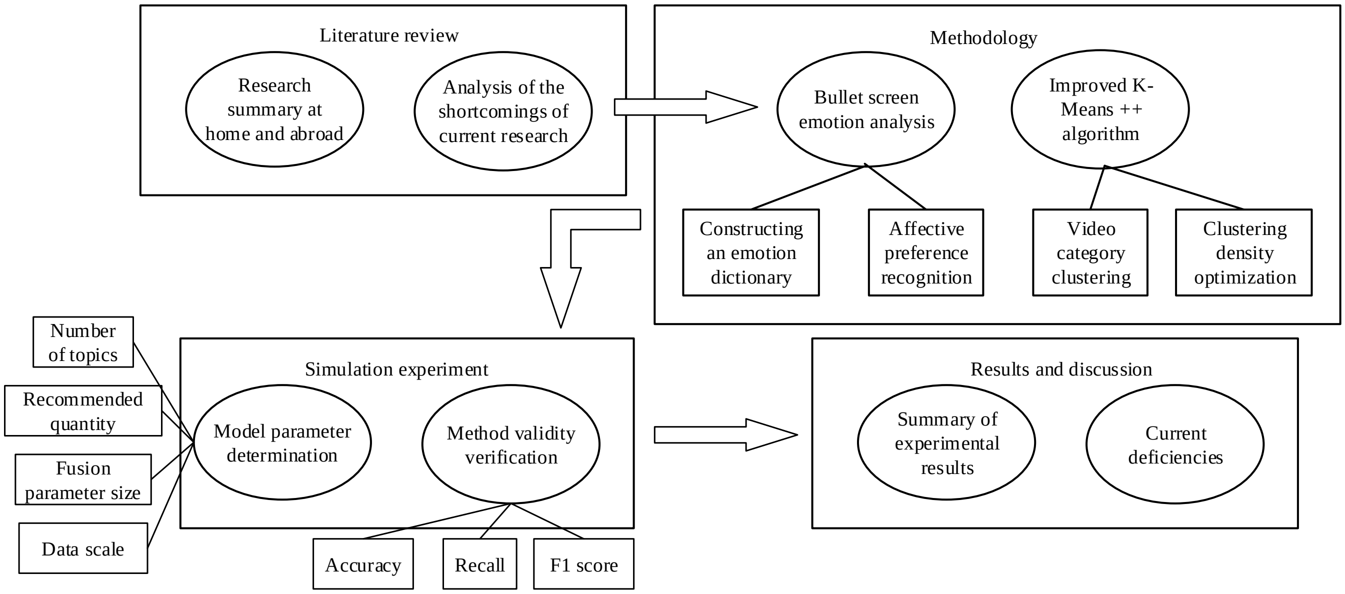


Fig. 1. Work roadmap.

II. Related Works

With the development of the information age, short video browsing has become the main popular way of leisure. Although there is currently a massive amount of short videos, video recommendation content is hard to meet the sentiment demands of users. Therefore, many scholars all over the world have organized extensive research on user sentiment analysis and recommendation algorithms. Alwehaibi et al. put forward an optimized sentiment classification for dialect Arabic short texts at the document level based on deep learning. The proposed model was trained and tested on a dataset consisting of modern standard Arabic and dialect Arabic corpora, and the findings indicated a great upgrading in the classification accuracy of Arabic texts, ranging from 88% to 69.7% [6]. Jiang W et al. raised a hybrid classification model that integrates algorithms such as K-means++, convolutional neural networks, and long short-term memory networks. The raised model was applied to balanced and imbalanced corpora, and the comparison outcomes denoted that the proposed model outperformed commonly used models in text sentiment classification [7]. Imran AS et al. used two Generative Adversarial Network (GAN) models, CatGAN and SentiGAN, to synthesize text for balancing highly imbalanced datasets. Special emphasis was placed on the diversity of samples synthesized to fill minority groups. The experiment findings on highly imbalanced datasets indicated that the effect of the model on the dataset was greatly raised after balancing the sentiment classification task with synthetic data [8]. Edara DC proposed a new sentiment analysis model with a distributed framework of long short-term memory neural networks, and evaluated the effect of the raised framework. The effect of each text mining and classification method on three datasets was assessed and they were compared with each other. The outcomes indicated that the proposed method performed better than other methods in accuracy and execution time [9].

Iwendi C et al. proposed an item-based recommendation system for personalized product recommendation problems and used a machine learning model to rate the recommended items. The system was tested for performance on the Yelp dataset, with an accuracy of 79%, an mean absolute error of 21%, a recall rate of 80%, and an F1 value of 79%. The results indicated that this method improved the accuracy of product recommendation [10]. Park J et al. used text mining methods to analyze the usefulness and consistency of comments, and assessed the effectiveness of the raised method. The experiment findings expressed that the usefulness and consistency of comments could improve the performance of personalized recommendation services and increase customer satisfaction [11]. Wang Y et al. constructed an Application Programming Interface (API) recommendation method with sequence awareness and designed new metrics to assess the method's ability to prioritize API usage. The experiment outcomes illustrated that compared with the baseline, the raised method not only realized better results on commonly used indicators, but also outperformed the baseline method on the newly proposed sequence indicators [12]. Liu W et al. brought temporal contextual information into typical collaborative filtering algorithms and used a popularity penalty function to weight the similarity between recommended and historical short videos. User context was also introduced into traditional collaborative filtering recommendation algorithms, taking into account user context information during the recommendation generation stage. Finally, the accuracy and diversity of this method were demonstrated through case analysis [13].

To sum up, text sentiment analysis and recommendation algorithm have some research achievements at present. Collaborative filtering and content recommendation algorithms have been applied in video recommendation systems, and have improved the performance of the system to some extent. Sentiment analysis technology has achieved good results in the field of text analysis and product recommendation. Compared with the existing methods, the main difference is that the proposed method introduces user emotion into the recommendation system for recognition. Because most of the existing short video recommendation systems do not make full use of the user's emotional response during the viewing process; There are few researches on the application of sentiment analysis in short video recommendation, and the existing methods mainly focus on text or product recommendation. It is difficult for collaborative filtering and content recommendation algorithms to guarantee high accuracy recommendation results in the face of massive and rapidly updated short video content. Algorithms often ignore users' emotional tendencies and focus only on historical behavioral data. Therefore, this paper proposes a bullet-screen-based emotion analysis method to identify users' emotional responses in the process of watching videos, build a special emotion dictionary, and improve the accuracy and comprehensiveness of emotion analysis. Fig. 1 shows work roadmap.

III. Construction of a Short Video Recommendation Model Based on Barrage Sentiment Analysis and Improved K-Means++

Aiming at the problem of low quality short video recommendations that do not meet the interests and sentiment needs of users, a barrage-based short video sentiment analysis is proposed. By matching video sentiment similarity with user sentiment preferences, K-means++ is used for short video category classification, and finally video recommendation is completed.

A. *Short Video Sentiment Analysis Based on Barrage Comments*

At present, short videos have high traffic in the Internet, but a large number of self-made short videos lack user ratings and other measures, resulting in low accuracy of short video recommendation. With the emergence of "barrage", it has gained the love of netizens in long videos. "Barrage" is a text comment method based on the video timeline and can be displayed in the video, which has social and sentiment characteristics [14]. With the application of "barrage" in short videos, short videos can identify user interests and hobbies through sentiment analysis of bullet comments. The main reason for the analysis of bullet screen is that users tend to use emotional words when expressing their opinions and emotions, which provides basic data for emotion analysis and makes it easier to obtain users' emotional needs. Therefore, a short video recommendation model based on sentiment analysis and K-means++ (SVRSA-K-means++) was raised in this study. The main structure of the research and construction model is shown in Fig. 2.

In Fig. 2, the main structure of the SVRSA-K-means++model includes two parts. The first part is sentiment analysis of short videos based on barrage comments. The second part is short video recommendation based on the improved K-means++. From Fig. 2, to calculate the sentiment similarity of the video, the first step is to extract the barrage from the short video. The study uses web crawlers to crawl bullet comments in short videos using Extensible Markup Language. Web crawlers are scripts or programs that automatically obtain the required resources from the network based on a unified resource locator, and are the most critical technical means in current search engine crawling systems. The general structure of web crawlers is shown in Fig. 3.

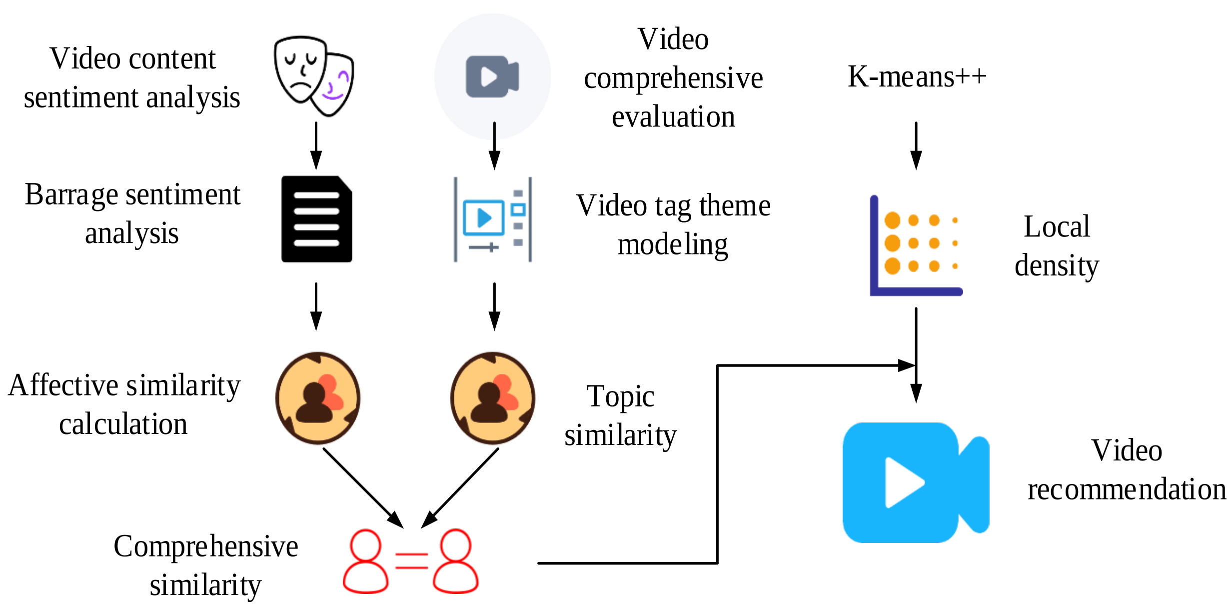


Fig. 2. Schematic diagram of SVRSA-K-means++ model structure.

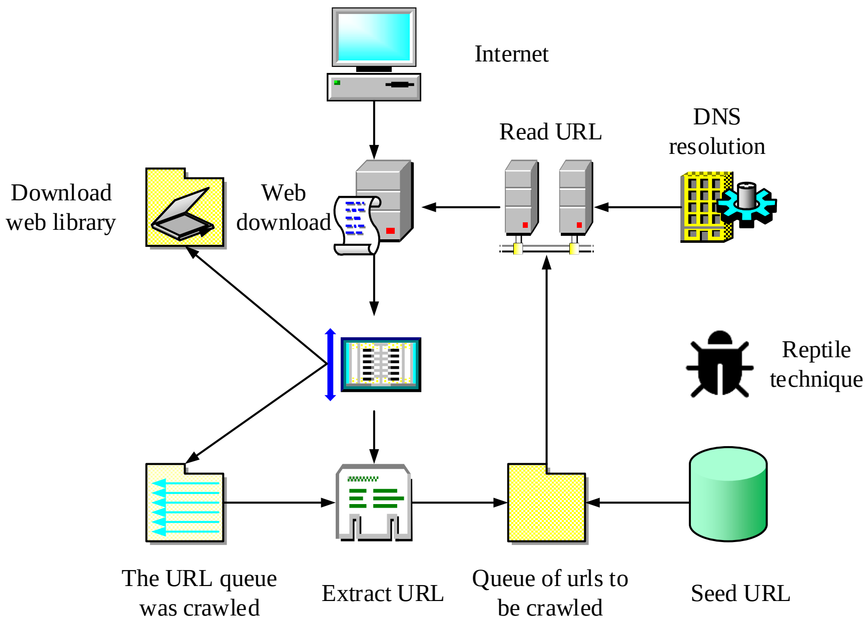


Fig. 3. General structure of web crawlers.

In Fig. 3, the first step is to determine the partial unified resource locator and seed URL based on the business scenario and data acquisition purpose. The second step is to read the URL. The third step is to use a Domain Name System (DNS) to resolve URLs. Finally, the webpage downloader reads the corresponding address and performs the operation, which can determine whether the connection has been accessed by the crawler program. Therefore, the study utilizes Python to crawl bullet comments from Bilibili video networks and remove reread bullet comments. Finally, a barrage comment text sentiment dictionary is constructed to segment barrage comment documents. Emotion dictionary can effectively classify and quantify users' emotional responses. It is built based on a large number of emotion words and artificial annotations, and contains a variety of emotion categories, which can effectively cover the main emotions expressed by users in the bullet screen. Through natural language processing technology and human intervention, the sentiment dictionary can be continuously updated and optimized to maintain adaptability to the user's language habits. If the sentiment dictionary is not comprehensive or accurate, the sentiment analysis results will be distorted, which will affect the video similarity calculation and recommendation effect. Research constructs a 7-dimensional sentiment vector based on an sentiment dictionary, which includes joy, anger, sadness, joy, fear, evil, and shock. A set of sentiment words is represented through a 7-dimensional vector, and the corresponding aspect of sentiments corresponding to sentiment words is assigned weights in the corresponding vector. After preprocessing the barrage comments, each barrage comment can be regarded as a set of several words. By adding and normalizing the sentiment vectors of all barrage comment words in the video, the sentiment vectors of each barrage comment can be obtained, as shown in Formula (1) [15].

$$E_d = \frac{1}{M}\sum_{i \in d} E_{W_i} \tag{1}$$

In Formula (1), E_d means the barrage sentiment vector. E_{W_i} represents the sentiment vector of the i th sentiment word in the barrage, and M represents the maximum sum of the sentiment word vectors. Formula (1) For each bullet screen, it is necessary to perform word segmentation processing first, and then find the corresponding emotion word and its vector according to the emotion dictionary. The emotion vector of the barrage is obtained by adding and normalizing the vectors of these emotion words. If the processed barrage does not match any words in the sentiment dictionary, the barrage will not be calculated. The average of the sentiment vectors is taken for all bullet comments in the video to obtain the sentiment vector as shown in Formula (2).

$$E_V = \frac{1}{n}\sum_{k=1}^{n} E_{d_k} \tag{2}$$

In Formula (2), n means the amount of barrage comments in the video, and E_{d_k} represents the sentiment vector of bullet comments. Formula (2) can obtain the overall emotion vector of the video by calculating the average of the emotion vector of all the bullets in the video. After calculating the video sentiment vector E_V, the sentiment similarity between two videos can be calculated using cosine similarity, as shown in Formula (3).

$$sim_{E_V}(E_{V_i}, E_{V_j}) = \frac{\sum_{k=1}^{7} e_{V_i}^k \times e_{V_j}^k}{\sqrt{\sum_{k=1}^{7} (e_{V_i}^k)^2} \times \sqrt{\sum_{k=1}^{7} (e_{V_j}^k)^2}} \tag{3}$$

In Formula (3), $sim_{E_V}(E_{V_i}, E_{V_j})$ represents the sentiment vectors of different videos, and $e_{V_i}^k$ represents the sentiment index in the sentiment vectors of videos. Formula (3) can quantify the emotional similarity of two videos by calculating the cosine similarity of the emotion vector of the two videos. The cosine similarity value is between -1 and 1, and the larger the value, the more similar the emotions of the two videos are. Regarding the processing of video classification labels, the study calculates the topic similarity between videos using Formula (4) [16].

$$sim_{T_V}(T_{V_i}, T_{V_j}) = \frac{\sum_{k=1}^{7} t_{V_i}^k \times t_{V_j}^k}{\sqrt{\sum_{k=1}^{7} (t_{V_i}^k)^2 \times \sqrt{\sum_{k=1}^{7} (t_{V_j}^k)^2}}} \tag{4}$$

In Formula (4), T indicates the distribution of topics, and T indicates the weight of video topics. Formula (4) By modeling and analyzing the theme of the video, the theme distribution of the video can be obtained, and the theme similarity of the two videos can be calculated by using these distributions. After calculating the sentiment similarity and topic similarity of the video, the comprehensive similarity of the video is obtained by weighted sum. The comprehensive similarity enables the model to find a balance between emotion and content, improving the personalization and relevance of recommendations. If it can not effectively improve the relevance of recommendation, it will directly affect the user satisfaction and the practicability of the system. The comprehensive similarity is calculated as shown in Formula (5).

$$sim_V(V_i, V_j) = \alpha sim_{E_V}(E_{V_i}, E_{V_j}) + (1-\alpha) sim_{T_V}(T_{V_i}, T_{V_j}) \tag{5}$$

In Formula (5), $sim_V(V_i, V_j)$ represents the comprehensive similarity between videos, and α represents the fusion weight coefficient. Formula (5) is used to calculate the comprehensive similarity of the video, and the affective similarity and topic similarity are weighted and summed. When the α value is 1, the comprehensive similarity of the video is equal to the sentiment similarity of the video, and the video lacks video labels. When the α value is 0, the comprehensive similarity of the video is equal to the theme similarity of the video, and the video lacks barrage comments. The user's preference for videos is obtained by using the user's historical viewing video set and the similarity between the video in the set and the target video to obtain Formula (6) [17].

$$prefer = \frac{\sum_{i \in H_u} sim_V(V_i, V_k)}{|H_u|} \tag{6}$$

In Formula (6), H_u represents the number of historical watched video collections. Formula (6) By calculating the average comprehensive similarity between each video and the target video in the user's historical viewing video set, the user's preference for the target video can be obtained. Through the above sentiment analysis and topic analysis, short videos are divided into similarity and sentiment, and finally K-means algorithm is used to classify and recommend data samples.

B. Short Video Classification Model Based on Density Improved K-Means++

The current K-means algorithm has two important shortcomings, namely the selection of K values and initial clustering centers. In response to the shortcomings of the K-means clustering algorithm, this study optimized it by reducing the iterations in the clustering and the amount of data in the clustering process, resulting in the K-means++[18-19]. The K-means++ selects the initial cluster center by calculating the shortest cluster between each sample and the existing cluster center. The K-means++ algorithm can avoid the local optimal problem common in the traditional K-means algorithm by optimizing the initial cluster center selection. By improving the initial center selection strategy, the algorithm can better deal with complex and high-dimensional data and improve the clustering effect. Although it improves clustering accuracy, the effect still needs to be improved. The study first preprocesses the data, assuming two distance thresholds of φ_1 and φ_2 in the sample dataset, and then selects samples from the dataset to get the Euclidean distance d between the remaining samples and the selected samples. If d is less than φ_1, it will add the data to the latest dataset. If d is less than φ_2, it will remove the sample from the dataset. The mean distance of all sample data in the sample dataset is shown in Formula (7).

$$MeanDis(D) = \frac{2}{n(n-1)} \sum_{i=1}^{n} \sum_{j=j+1}^{n} d(d_i, d_j) \tag{7}$$

The density of data objects in the sample dataset is set to $\rho(i)$, and its calculation is denoted in Formula (8).

$$\begin{cases} \rho(i) = \sum_{j=1}^{n} f(d_{ij} - MeanDis(D)) \\ f(x) = \begin{cases} 1 & x < 0 \\ 0 & x \geq 0 \end{cases} \end{cases} \tag{8}$$

According to Formula (8), samples can form a cluster, and the average distance between samples in the cluster is defined as $a(i)$, which is expressed as Formula (9).

$$a(i) = \frac{2}{\rho(i)[\rho(i)-1]} \sum_{i=1}^{\rho(i)} \sum_{j=j+1}^{\rho(j)} d(x_i, x_j) \tag{9}$$

The study assumes that there are video i and video j in a certain cluster, and the local density between the two is located as j, and the local density is compared with the distance between the samples. If the $s(i)$ between video i and video j is the maximum, then $\max\{d(i,j)\}$ is used to represent the maximum distance between the two. If $\rho(j) > \rho(i)$, $s(i)$ is defined as the minimum distance and represented by $\min\{d(i,j)\}$, so the expression for $s(i)$ is shown in Formula (10).

$$s(i)=\begin{cases}\min\limits_{\rho(j)>\rho(i)}\{d(i,j)\}, \exists j, \rho(j)>\rho(i) \\ \max\{d(i,j)\}, !\exists, \rho(j)>\rho(i)\end{cases} \tag{10}$$

Formulas (7) to (10) can effectively remove outliers and select more reasonable initial clustering centers through density calculation and preprocessing of data, thus reducing the number of iterations of K-means algorithm and improving the accuracy and efficiency of clustering. Through the above data preprocessing, research can remove outliers from initial clustering data to improve the performance of clustering algorithms. The specific process of data preprocessing is indicated in Fig. 4.

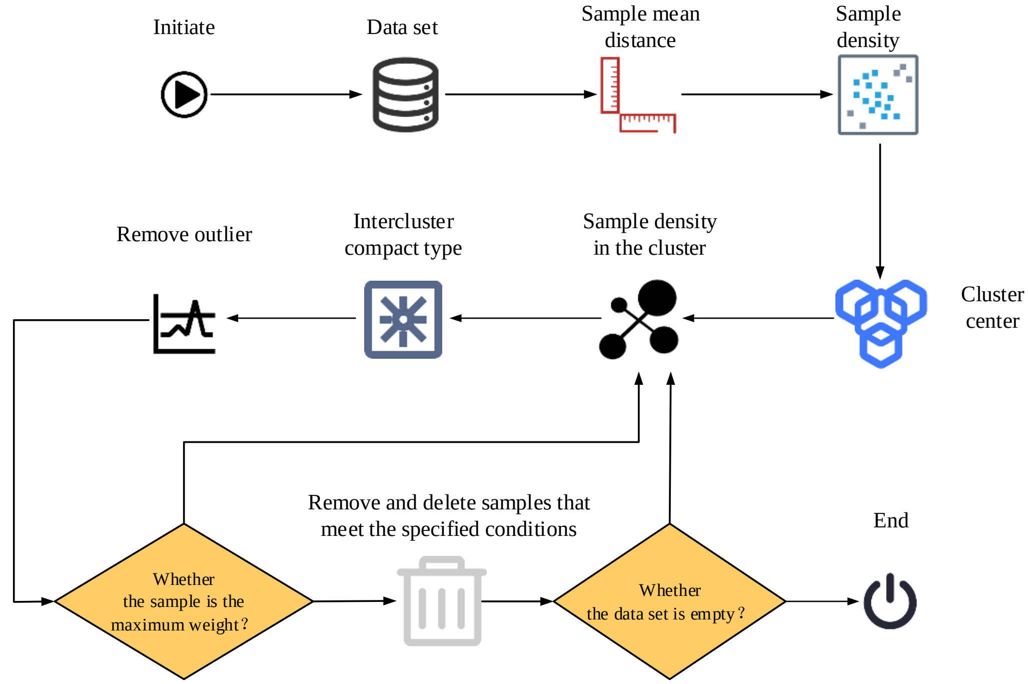


Fig. 4. Data preprocessing process.

In Fig. 4, the sample dataset is first established, and then the average distance and density of the samples are measured to determine the clustering center. Secondly, it needs to calculate the sample cluster density and inter cluster compactness. Then it needs to delete the outliers and check if the dataset is empty. If there are still samples in the dataset, it will continue to calculate cluster density and compactness. If the dataset is empty, it will stop the algorithm. By studying the process shown in Fig. 4, the processed dataset, cluster values, and initial cluster center set can be obtained. Due to the significant impact of selecting the initial clustering center on the clustering effect, research has been conducted to remove low-density points by calculating the density of data points. At the same time, it needs to calculate the mini distance between the data points and other points with higher density, in order to distinguish between common points and maximum density points in the cluster, and remove outliers below the average density, thereby further optimizing the selection of initial cluster centers. The clustering dataset is set as $G=\{g_1,g_2,..,g_n\}$ and the initial cluster center set as $C=\{C_1,C_2,..,C_n\}$, and the Euclidean distance between the data is obtained using Formula (11).

$$d(g_i,C_i)=\sqrt{\sum_{g=1}^{n}(g_{id}-C_{id})^2} \tag{11}$$

In Formula (11), $C=\{C_1,C_2,..,C_n\}$ and SS are the coordinates of the dataset samples and the initial cluster center in two-dimensional coordinates, respectively. Study calculates of the centroid points of each cluster using Formula (12), and determines the relationship between the change in the centroid points of the cluster and the initial cluster center points of that class.

$$x(C_i)=\frac{1}{|C_i|}\sum_{a_i\in C_i}a_i \tag{12}$$

In Formula (12), $|C_i|$ means the amount of data objects at the initial cluster center, and a_i represents the centroid point of the cluster. The calculation of the centroid point is shown in Formula (13).

$$\omega_r=\frac{1}{|C_i|}\sum_{\omega_i\in C_{r,i}}g_i-\frac{1}{|C_{i-1}|}\sum_{\omega_i\in C_{r-1,i}}C_i \tag{13}$$

In Formula (13), r refers to the iterations of the algorithm, and ω_r denotes the variable at the cluster center point. Formulas (11) to (13) can determine the distance relationship between the data point and the cluster center by calculating the Euclidean distance. Calculating the centroid helps determine the central location of each cluster. The final clustering result can be obtained by updating the centroid points until the change of centroid points satisfies the set conditions. Research is conducted to determine whether the variable of the cluster centroid meets the condition of being less than the initial setting based on ω_r. If it meets the condition, it is added to the feature set and deleted from the dataset. Finally, it will traverse all center points and update them, calculate the centroid of each cluster whose change in center points is greater than the set value, and use it as the new cluster center. The above steps are repeated until the final clustering result cluster is got, as shown in Fig. 5.

In Fig. 5, if the grid density of the sample is less than the threshold obtained by the maximum weight method calculation, the sample is removed. After removing the outliers, an initial cluster center can be generated. In traditional algorithms, these centers are generated randomly. This method divides the data of each dimension into K segments, and uses the average of each segment as the coordinate of the corresponding initial cluster center in that dimension.

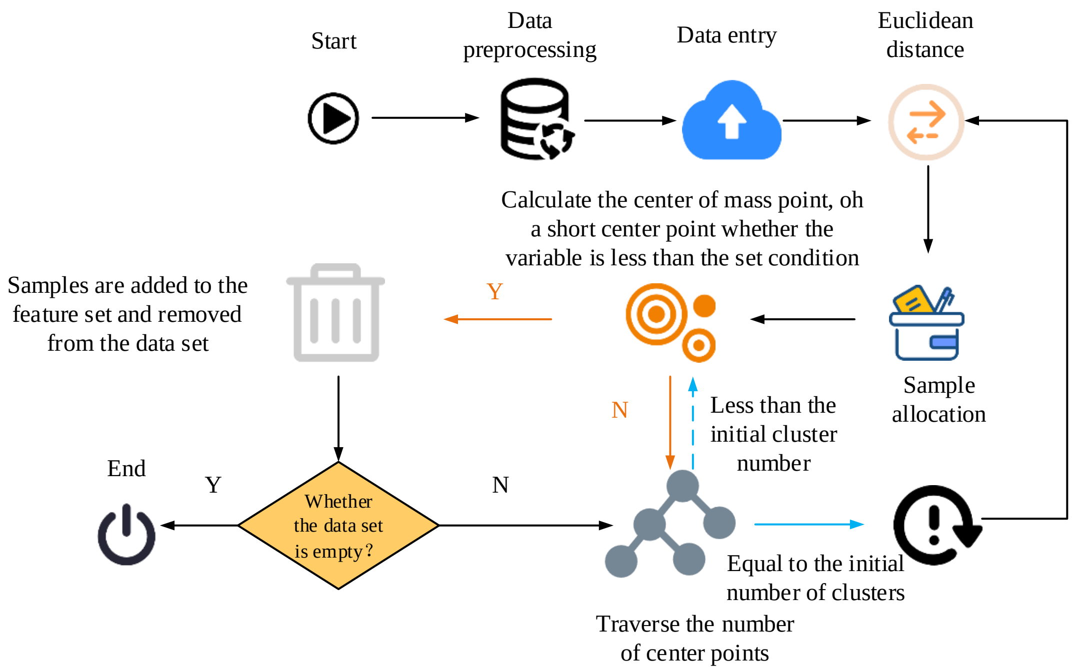


Fig. 5. K-means++ flowchart based on density improvement.

IV. Analysis of Short Video Recommendation Performance Based on Sentiment Analysis and K-Means++

The study validated the effect of the raised method. Firstly, the optimal parameters of the model were determined through simulation experiments. The experiment adjusted the number of short video topics, the number of recommendations, the size of fusion parameters and historical record data. Next, the clustering performance of the model was analyzed. Finally, through comparative experiments, the research model was analyzed based on evaluation indicators such as iteration number, clustering time, and complexity comparison.

A. *Algorithm Parameter Determination and Clustering Analysis*

The study used a focused crawler to crawl all relevant user interaction data of 3205 videos from the short video channel of the Bilibili video website, and used this data as an experimental dataset. The dataset included 3485992 bullet comments and involves 1652144 users. After data preprocessing, the barrage was deduplicated, sparse historical users were deleted, and some abnormal data videos were removed, leaving 2752 videos. There were 1071 active users, and each user had an average of about 60 historical viewing records. The experimental hardware environment was an Intel Core i5-7400 processor with 16 GB of memory. The software environment was Windows 10 x64 operating system, and the code was Python 3.7.0. The experimental parameter values involved in researching algorithms are denoted in Table I.

TABLE I. EXPERIMENTAL DATA PARAMETER VALUES

Argument	Value
Short video topic	10,20,30,40,50,60,70,80,90
Recommended quantity	10,20,30,40,50,60
Fusion parameter	0-1.0
Historical data size	2000,3000,4000,5000,6000,7000

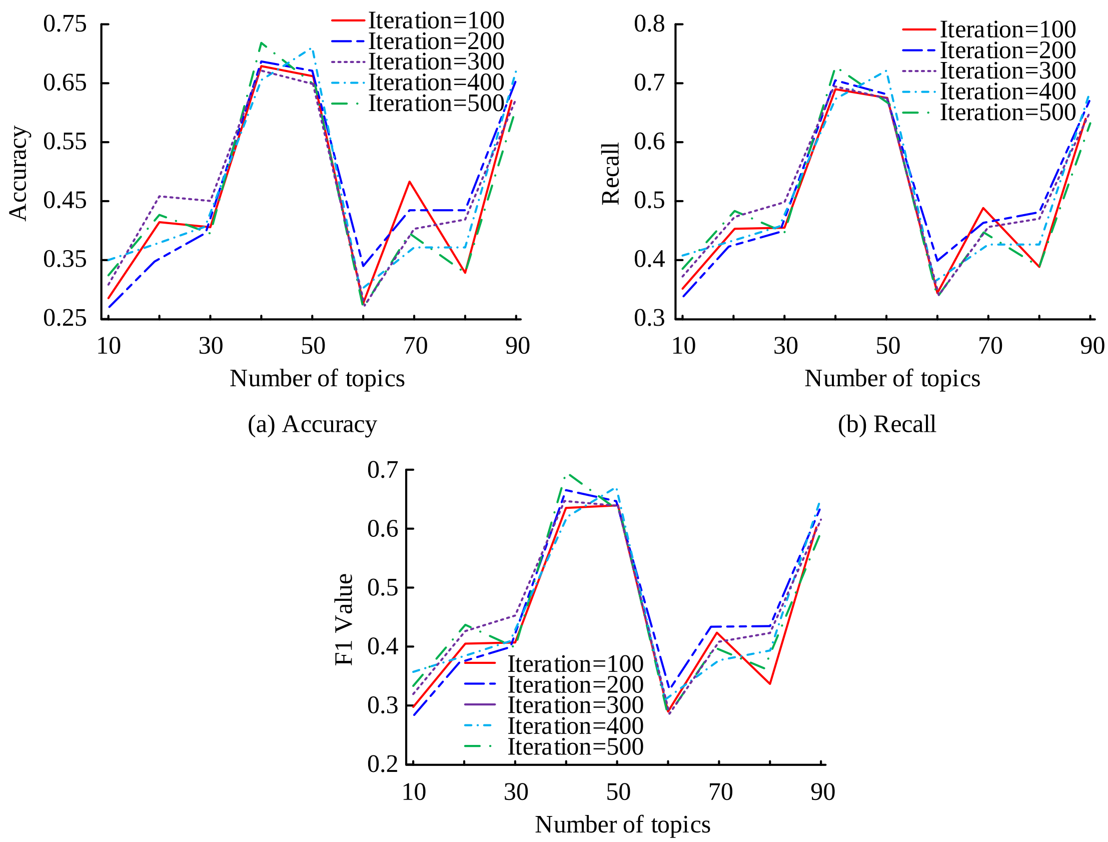
In the construction of the short video theme sentiment model, the research assumed a historical record size of 2000, a recommendation quantity of 10, and a fusion parameter value of 0. The simulation findings of the model are expressed in Fig. 6. Fig. 6 (a) showcases the accuracy outcomes of the model. Fig. 6 (b) showcases the recall rate results of the model. Fig. 6 (c) showcases the F1 value outcomes of the model. In Fig. 6, as the amount of short video themes increased, all three evaluation indicators first gradually increased and then suddenly decreased. When the amount of topics was 40 or 50, the model's prediction accuracy reached around 0.66, the recall rate reached around 0.70, and the F1 value was 0.71. From the perspective of iteration times, it can be seen that the impact of iteration times on the model was relatively small. Therefore, in the research and construction model, the number of theme sentiments parameter was 40 and the number of iterations was 500.

After determining the model parameters, further analysis was conducted on the impact of historical data size and different fusion parameters on the performance of the model. The outcomes are indicated in Fig. 7. Fig. 7 (a) showcases the impact of the number of historical records on the accuracy of the model. When the amount of historical records was 2000, the accuracy of the model in predicting short videos was 0.250. When the amount of historical records was 7000, the model's accuracy in predicting short videos was 0.81. Fig. 7 (b) showcases the impact of the amount of historical records on the model's recall rate. When the number of historical records was 2000, the model's accuracy in predicting short videos was 0.225. When the amount of historical records was 7000, the model's accuracy in predicting short videos was 0.822. Fig. 7 (c) showcases the impact of the number of historical records on the F1 value of the model. When the amount of historical records was 2000, the accuracy of the model in predicting short videos was 0.314. When the amount of historical records was 7000, the model's accuracy in predicting short videos was 0.832. Meanwhile, under the same amount of historical records, the more recommendations there were, the higher the effectiveness of the model. Therefore, the historical record data size of the model was set to 7000.

The evaluation results of different fusion parameters are denoted in Fig. 8. Fig. 8 (a) showcases the accuracy of the model under different fusion parameters. Fig. 8 (b) showcases the recall rate of the model under different fusion parameters. Fig. 8 (c) showcases the F1 values of the model under various fusion parameters. The results showed that there was not much difference in model performance when the fusion parameter values ranged 0.2-0.5. When the fusion parameter value was 0.3, the model accuracy reached 0.825, the recall rate was 0.805, and the F1 value was 0.812. Therefore, when the fusion parameter value of the model was 0.3, the model performed best.

B. Analysis of Short Video Recommendation Performance Based on Barrage Screen Sentiment Analysis

Research was conducted to prove the clustering and recommendation performance of the proposed model through instance validation using crawled datasets. The experiment first conducted clustering validation on three sentimental categories, and the outcomes are indicated in Fig. 9. Fig. 9 (a) showcases the data distribution before noise point removal, and Fig. 9 (b) showcases the data distribution after noise point removal. After comparison, the outliers around the sentiment clustering have been removed, indicating that the proposed method had good ability to block outliers.

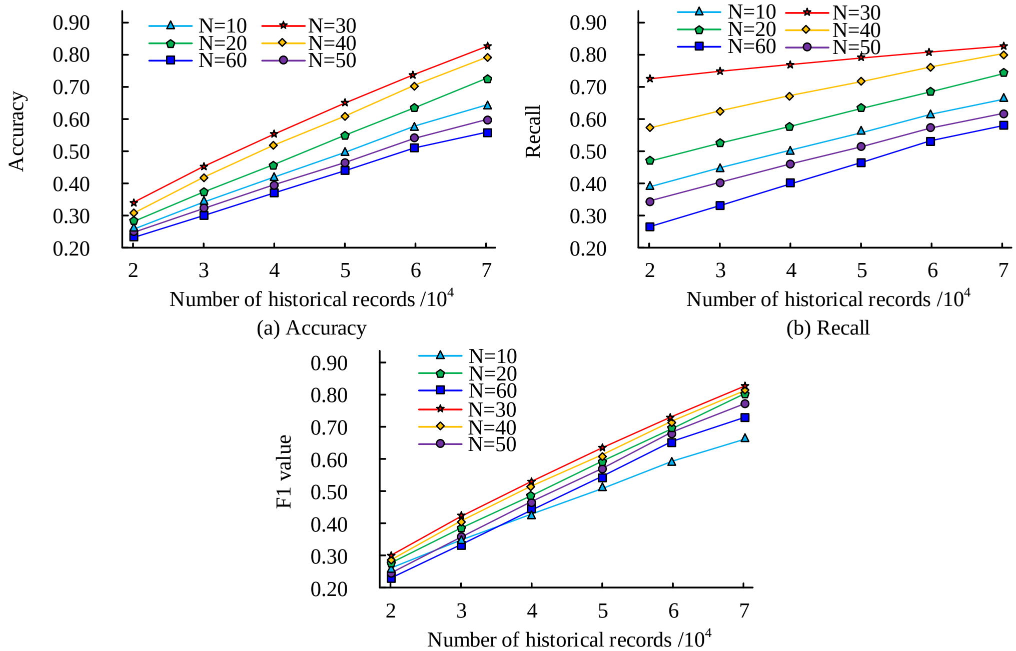


(a) Accuracy

(b) Recall

(c) F1 Value

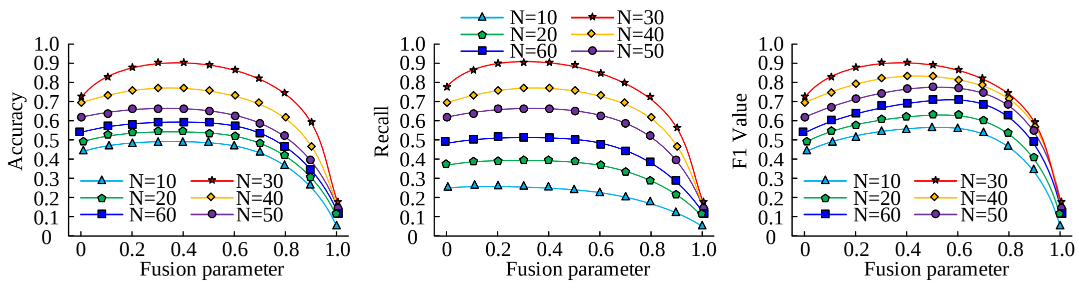
Fig. 6. Evaluation metrics for different number of topics and iterations.



(a) Accuracy (b) Recall

(c) F1 value

Fig. 7. Impact of historical records on model accuracy.



(a) Accuracy (b) Recall (c) F1 Value

Fig. 8. Evaluation results of different fusion parameters.

After removing the noise points from the dataset, the study validated the selection and clustering effect of the improved K-means++ clustering center. The selection and clustering process of the de-clustering center are shown in Fig. 10. Fig. 10 (a) -10 (h) represent the clustering center selection and clustering process of the algorithm. The findings indicated that the improved K-means++, after selecting the initial cluster center, continuously calculated and iterated to change the position of the cluster center, thereby making the cluster centers as far apart as possible and approaching the center points of different clusters, thereby improving the clustering effect.

To assess the progressiveness of the proposed method, the research conducted a comparative analysis with the current four mainstream algorithms, including Item-based Collaborative Filtering (ItembasedCF), Tag-based Latent Dirichlet Allocation (Tag-LDA) algorithm, Unifying LDA andRatings Collaborative Filtering (ULR-itemCF), and Danmaku-Related Collaborative Filtering and Topic model-based Recommendation, DRCFT) [20-21]. The study used algorithm clustering time and complexity as evaluation indicators, and the clustering time comparison outcomes of the five algorithms are expressed in Fig 11.

In Fig. 11, the clustering time of the five algorithms was directly proportional to the amount of samples in the dataset. The more samples in the dataset, the longer the clustering time of the algorithms. From the perspective of the same amount of samples in the dataset, when the sample size was 7000, there was a significant difference in clustering time among the five algorithms. The itembasedCF clustering time was 11.2 seconds. Tag-LDA clustering took 9.8 seconds. ULR-itemCF clustering took 10.4 seconds. The clustering time of DRCFT was 8.7 seconds. The clustering time of the research method was 8.2 seconds. The comparison findings of the complexity of the five algorithms are expressed in Fig. 12.

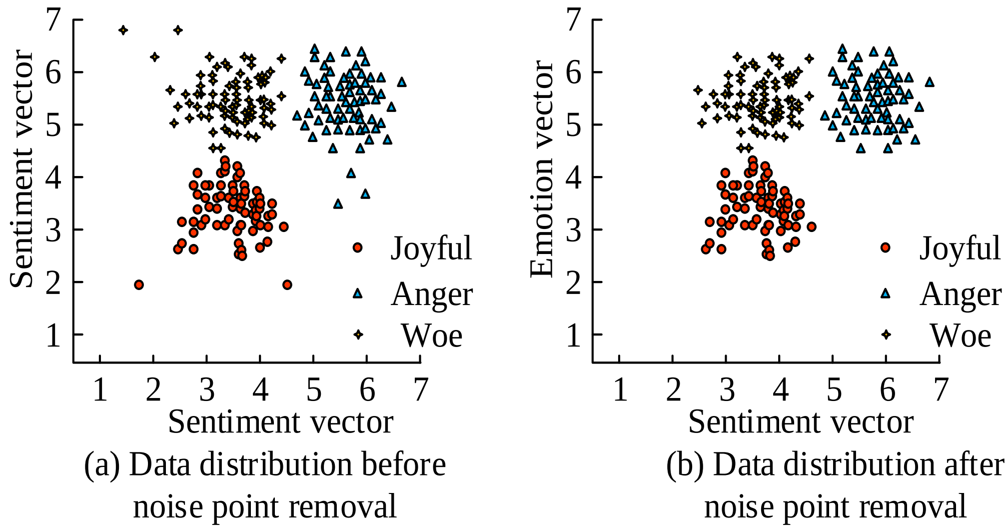


Fig. 9. Data distribution of sentiment categories.

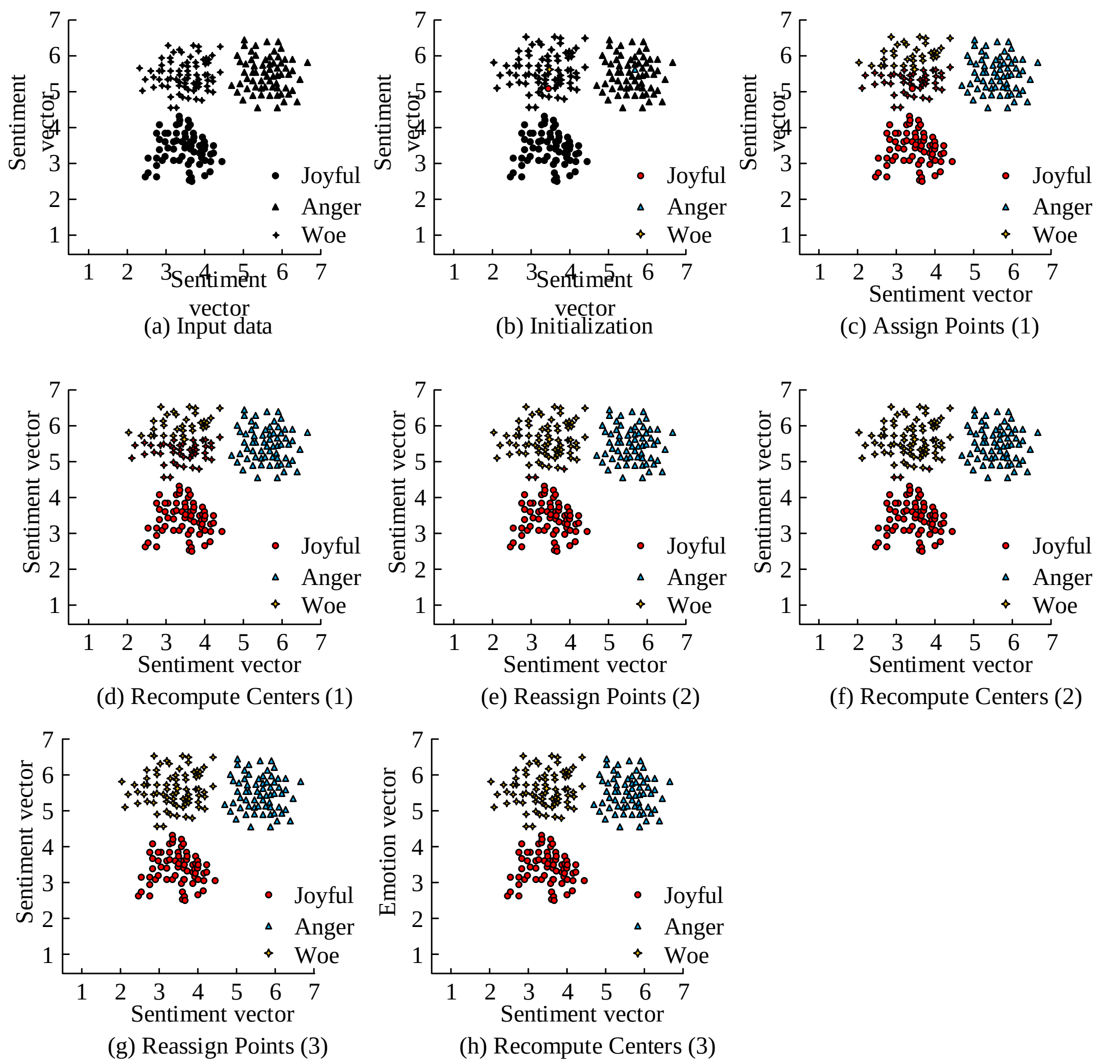


Fig. 10. Model cluster center selection and clustering process.

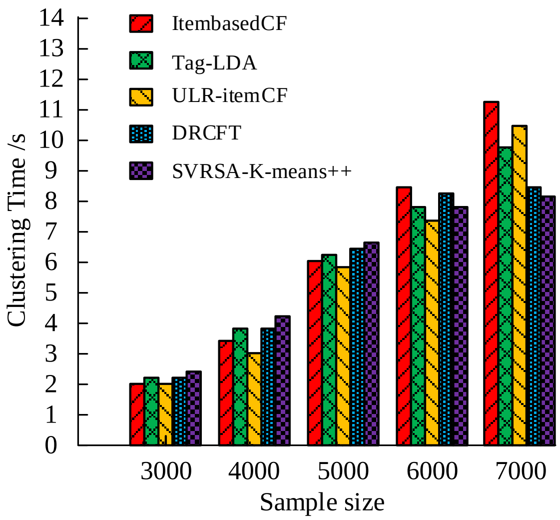


Fig. 11. Clustering time results of different algorithms.

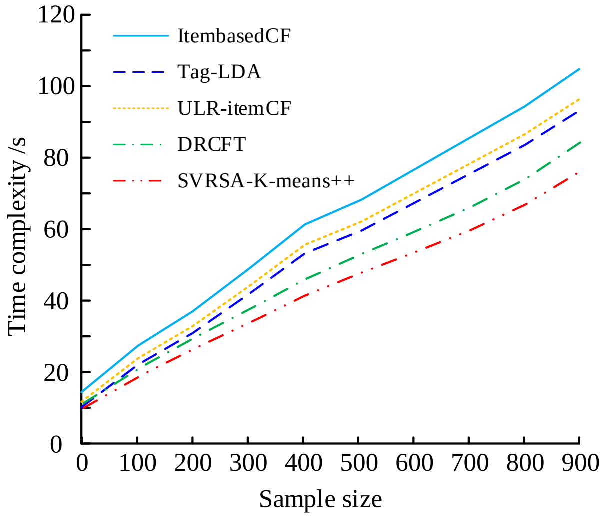


Fig. 12. Comparison of complexity of different algorithms.

The outcomes in Fig. 12 denoted that the time complexity of ItembasedCF was 105, Tag-LDA was 92, and ULR-itemCF was 97. The clustering time for DRCFT was 84. The clustering time of the research method was 78. Based on the above results, the proposed method had good clustering performance and consumed less clustering time. The algorithm itself had low complexity and could effectively meet the practical application requirements of short video recommendation. The results show that the limitations of the current method are low recommendation accuracy, lack of sentiment analysis, poor clustering effect and high computational complexity. In the face of massive and rapidly updated short video content, it is difficult for traditional methods to guarantee high accuracy of recommendation results. The lack of sentiment analysis makes the recommendation system lack the performance of individuation and relevance, and it is difficult to improve user satisfaction. Long clustering time and low computing efficiency degrade the performance of the system in high concurrency environment and make it difficult to provide timely recommendation service.

V. Discussion

Simulation experiments were conducted to verify the model. With the support of 7000 historical records, the accuracy rate of the recommended model based on SVRSA-K-Means ++ reached 81.0%, the recall rate was 82.2%, and the F1 score was 83.2%. In the clustering time detection, the recommended model only takes 8.2s in the environment of 7000 samples, which is reduced by 3s compared with ItembasedCF. In the time complexity detection, the recommended model has a time complexity of 78, which shows lower time complexity and higher efficiency than other mainstream algorithms. The experimental results show that emotion feature recognition model plays a key role in optimizing the clustering effect of K-Means ++ algorithm. Through the similarity calculation of emotion vector, the model can assign similar videos to the same category more accurately, thus improving the clustering effect and the accuracy of recommendation. The results show that the introduction of emotion analysis can reflect the emotional response of users in the process of watching videos, which makes the recommendation results more in line with the actual needs of users. At the same time, the experimental results are consistent with the purpose of improving the accuracy of video recommendation.

VI. Conclusion

Aiming at the requirements of short video recommendation systems, this study analyzed the accuracy and efficiency of video recommendation and proposed a short video recommendation model based on SVRSA-K-means++. The model crawled and analyzed a large number of short videos and their barrage data on the Bilibili website, constructed an emotion based barrage analysis framework, and performed sentiment annotation and category classification on the videos. The model is verified by simulation experiments, and the proposed method has certain reliability in the short video recommendation system, can effectively meet the emotional needs of users, and provides a new idea for the development and optimization of the future video recommendation system. The specific contribution of the research is to improve users' viewing experience by building a more accurate short video recommendation system. The personalized recommendation system can accurately identify the emotional preferences of users and recommend the content they are interested in, reducing the time and energy of users looking for videos they are interested in. At the same time, the model helps to disseminate high-quality, culturally valuable and educational content more widely to users, promoting cultural exchange and diversity. Although research has obtained certain outcomes in the short video recommendation, there are still some shortcomings, such as the accuracy of sentiment analysis being limited by the comprehensiveness and accuracy of sentiment dictionaries. Future research directions will focus on more accurately identifying and classifying the emotional attributes of video content, exploring more dimensions of user data analysis and fusion. Moreover, in addition to the emotion analysis of bullet screen text, future research can combine audio, images, user facial expressions and voice emotions in video for multimodal emotion analysis. By integrating various emotional features, users' emotional responses can be captured more comprehensively and accurately, providing more possibilities for further improving the performance of the short video recommendation system.

ACKNOWLEDGMENT

The research is supported by: This paper is the research result of the 2023 Hunan Provincial Natural Science Foundation project "Research on ISMAS Model Agriculture, rural areas and farmers Short term Video Frequency Communication and Application Innovation in the Perspective of Rural Revitalization" (No.2023JJ60174).

REFERENCES

[1] Zhou K, Yang C, Li L, Cong M, Song L, Jiang P, Su J. A folksonomy-based collaborative filtering method for crowdsourcing knowledge-sharing communities. Kybernetes: The International Journal of Systems & Cybernetics, 2023, 52(1):328-343.

[2] Wang S, Fan Y, Yu C, Jin S, Paul T A, Carlos F, Stroe D. A novel collaborative multiscale weighting factor″daptive Kalman filtering method for the time(arying whole‐‐‐ife`ycle state of charge estimation of lithium¯¨¯on batteries). International Journal of Energy Research, 2022, 46(6):7704-7721.

[3] Shao P, Wu L, Zhang K, Lian D, Hong R, Li Y, et al. Average User-Side Counterfactual Fairness for Collaborative Filtering. ACM Transactions on Information Systems, 2024, 42(5): 1-26.

[4] Wu B. Study on news recommendation of social media platform based on improved collaborative filtering. International Journal of Web Based Communities, 2024, 20(1-2): 27-37.

[5] Zhang M, Wang S, Yuan K. Sentiment Analysis of Barrage Text Based on ALBERT and Multi-channel Capsule Network. Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery: Proceedings of the ICNC-FSKD 2021 17. Springer International Publishing, 2022: 89(1):718-726.

[6] Alwehaibi A, Bikdash M, Albogmi M, et al. A study of the performance of embedding methods for Arabic short-text sentiment analysis using deep learning approaches. Journal of King Saud University-Computer and Information Sciences, 2022, 34(8): 6140-6149.

[7] Jiang W, Zhou K, Xiong C, Du G, Ou C, Zhang J. KSCB: A novel unsupervised method for text sentiment analysis. Applied Intelligence, 2023, 53(1): 301-311.

[8] Imran A S, Yang R, Kastrati Z, Daudpota S M, Shaikh S. The impact of synthetic text generation for sentiment analysis using GAN based models. Egyptian Informatics Journal, 2022, 23(3): 547-557.

[9] Edara D C, Vanukuri L P, Sistla V, et al. Sentiment analysis and text categorization of cancer medical records with LSTM. Journal of Ambient Intelligence and Humanized Computing, 2023, 14(5): 5309-5325.

[10] Iwendi C, Ibeke E, Eggoni H. Pointer-Based Item-to-Item Collaborative Filtering Recommendation System Using a Machine Learning Model. International Journal of Information Technology & Decision Making, 2022, 21(1):463-484.

[11] Park J, Li X, Li Q, Kim J. Impact on recommendation performance of online review helpfulness and consistency. Data Technologies and Applications, 2023, 57(2):199-221.

[12] Wang Y, Zhou Y, Chen T, Zhang J, Yang W, Huang Z. Sequence-Aware API Recommendation Based on Collaborative Filtering. International Journal of Software Engineering and Knowledge Engineering, 2022, 32(8):1203-1228.

[13] Liu W, Wan H, Yan B. Short Video Recommendation Algorithm Incorporating Temporal Contextual Information and User Context. CMES-Computer Modeling in Engineering & Sciences, 2023, 135(1): 239-258.

[14] Pal S, Roy A, Shivakumara P, Pal U. Adapting a Swin Transformer for License Plate Number and Text Detection in Drone Images. Artificial Intelligence and Applications, 2023, 1(3), 145-154.

[15] Bhosle K, Musande V. Evaluation of Deep Learning CNN Model for Recognition of Devanagari Digit. Artif. Intell. Appl., 2023, 1(2):114-118.

[16] Gheisari M, Hamidpour H, Liu Y, Saedi P, Raza A, Jalili A, Rokhsati H, Amin R. Data Mining Techniques for Web Mining: A Survey. Artificial Intelligence and Applications, 2023, 1(1): 3-10.

[17] Pandey K K, Shukla D. NDPD: an improved initial centroid method of partitional clustering for big data mining.Journal of Advances in Management Research, 2023, 20(1):1-34.

[18] Soleymani F, Vasighi M. Efficient portfolio construction by means of CVaR and k - means++ clustering analysis: Evidence from the NYSE. International Journal of Finance & Economics, 2022, 27(3): 3679-3693.

[19] Zhang R, Lu S, Wang X, Yu H, Liu Z. A multi-model fusion soft measurement method for cement clinker f-CaO content based on K-means++ and EMD-MKRVM. Transactions of the Institute of Measurement and Control, 2023, 45(2): 287-301.

[20] Horasan F. Latent Semantic Indexing-Based Hybrid Collaborative Filtering for Recommender Systems. Arabian Journal for Science and Engineering, 2022, 47(8): 10639-10653.

[21] Singh P K, Sinha S, Choudhury P. An improved item-based collaborative filtering using a modified Bhattacharyya coefficient and user‐user similarity as weight. Knowledge and Information Systems, 2022, 64(3): 665-701.