Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2023 SCPE. Volume 24, Issues 4, pp. 999–1010, DOI 10.12694/scpe.v24i4.2479

USER SENTIMENT ANALYSIS METHODS FOR ELDERLY SOCIAL MEDIA NETWORKS

SHUAIZHI SHEN*

Abstract. For the analysis of user sentiment in social media networks for the elderly population, emotional sentences are first extracted to classify movie reviews. Afterwards, social network data of the elderly population based on user search behavior is analyzed. The movie reviews of elderly social media users are analyzed for rating prediction. The research results indicate that the accuracy of sentiment classification results is in descending order of Dirichlet, maximum entropy, and support vector machine. The highest classification accuracy of the three algorithms is 87.1%, 86.9%, and 86.5%, respectively. The classification accuracy of the first level classifiers of Dirichlet, maximum entropy, and support vector machine are 90.7%, 88.7%, and 87.4%, respectively. The classification accuracy of the second level classifier is 86.7%, 83.7%, and 80.4%, respectively. The predictive analysis results of the research method are superior to those generated by using Slope One. The method proposed in the study can promote emotional analysis of film review texts, improving the analysis accuracy.

Key words: Elderly population; Social media networks; Emotional analysis; Film reviews; Personalized recommendations

1. Introduction. In recent years, the rapid development of the Internet has made communication between people more and more convenient. With the emergence of various new multimedia social platforms, communication around these platforms is also increasing day by day. Every day, tens of thousands of users express their opinions on these platforms [1]. In textual information, there are abundant vocabulary with emotional tendencies. These words can well reflect the user's emotional state at a certain moment [2]. The comment text of movies belongs to the common comment information on multimedia social media platforms, which contains the emotions and opinions of social media network users. Moreover, by analyzing the same specific movie, different users can better understand their emotional tendencies and analyze their emotions. At the same time, analyzing the emotions in movie reviews will help clarify everyone's overall view of the movie, thereby promoting adjustments in movie promotion and scheduling. There is currently a lot of research on film criticism, but in the traditional field of criticism, it is difficult to break through emotional analysis [3]. Personalized movie recommendations are generally based on proactive recommendations, which use software to conduct a series of data on users who have operated on the platform, including useful historical footprints, personal information, and operations, in order to infer user preferences. After organizing user data, it is then analyzed and combined with user preferences, Furthermore, targeted movie recommendations can be made to users. Due to the fact that on multimedia social media platforms, movie review information not only affects audience choices and viewing decisions to a certain extent, but also helps producers obtain feedback information from audiences in a timely manner after watching a movie, allowing for targeted promotion. In order to analyze the emotions of the elderly towards movies, make targeted recommendations based on emotions, and expand the film market for the elderly, this study intends to use comment text information from multimedia social platforms to conduct research on film recommendation and rating prediction methods. The article is mainly divided into five parts. The first part is an introduction, which mainly introduces the research background and purpose. The second part is a literature review, mainly summarizing the current research situation of different scholars at home and abroad. The third part is the research method, the fourth part is the result analysis, and the fifth part is the conclusion.

2. Related Work. Multimedia social platforms are gradually becoming text databases for viewpoint and sentiment analysis. Higher requirements have also been put forward for the mining and analysis of public opinion information. Chakraborty K and other scholars proposed a detection method based on social network user similarity selection to analyze the emotions in social media data. Community-based user data technology and

*School of Media and Art Design, Guilin University of Aerospace Technology, Guilin, 541004, China (PHDshen2023@163.com)

emotional evaluation methods have been studied and classified. The research results are conducive to analyzing the emotions of social media data and promoting accurate understanding of social media expression content [4]. Abid et al. designed an emotional analysis method based on bidirectional recursive convolutional neural network to solve the low accuracy of emotional data analysis in multi-source social media. The research results show that compared with conventional sentiment analysis methods, the accuracy of this method is greatly improved, reaching 89.67% [5]. The traditional Convolutional neural network cannot accurately capture the semantic features in emotion analysis. Therefore, researchers such as Alam proposed a domain specific distributed word representation method based on social media text resources. Then, the convolutional neural network is expanded. The results show that this method can reduce the convolution dimension and expand the size of Receptive field. It can combine different inflation rates to obtain long-term contextual semantic information and reduce computational costs [6]. Ducange et al. designed a decision support system based on machine learning text classification to improve the accuracy of online social media sentiment analysis. By combining sentiment analysis engines, user emotions in comments are analyzed across sources. The research results indicate that the accuracy of sentiment analysis using this method is as high as 90%, which is beneficial for analyzing users' emotions and making improvements [7]. Vasishtha S and other scholars proposed an unsupervised analysis system based on fuzzy rules for sentiment analysis of posts on social media. Fuzzy rules are used to calculate and analyze the emotions in posts. The research results indicate that this method can be applied to all datasets. Compared to traditional methods, it has higher performance [8].

Social media has become a new channel for product opportunities. Therefore, Jeong et al. proposed an opportunity recognition method based on social media data. Through theme modeling and sentiment analysis, product opportunities are mined and identified. Emotional analysis is used to evaluate product satisfaction. The research results indicate that this method can promote users' understanding of product emotions and recommend products based on customer needs [9]. The mixing of multiple languages increases the difficulty of recognition. Therefore, Bansal N and other scholars designed a language recognition system with mixed data sets and logistic regression classifiers. The results show that the accuracy of Logistic regression is the highest, 86.63. This method helps to improve the recognition accuracy of mixed languages [10].

Based on the above research, it can be seen that machine learning methods have certain effects in emotional analysis of social media data. However, there is limited analysis of user emotions in social media networks for the elderly. Given the increasing popularity of movie reviews on social media, in order to analyze the emotions of elderly users, strengthen movie recommendations for elderly users, and expand the video market for the elderly, research will combine machine learning and other methods to analyze the user emotions of movie reviews.

3. User sentiment analysis of social media networks targeting the elderly population. In order to analyze the emotional analysis of elderly social media network users towards movies, we aim to expand the film market for the elderly population through emotional analysis. The study first analyzes social network data based on user search behavior, then extracts emotional sentences from comments of the elderly population, classifies movie comments, and finally analyzes movie comments of elderly social media users before conducting rating predictions.

3.1. Analysis of Social Network Data for the Elderly Based on User Search Behavior. Currently, with the rapid development of social media, a large amount of information has emerged on social media platforms. Analyzing these data can provide more valuable information. So, research based on cloud technology has become a very important way to obtain information. By exploring different search patterns, different search results are revealed from different perspectives. Then the retrieval quality and retrieval behavior are analyzed [11]. An algorithm that combines user search behavior has been proposed. Based on the user's search logs, their relevant behaviors are extracted. Analyzing user behavior patterns can increase the volume of search information services for users and find keywords. This operation can provide user behavior data for future rating predictions, improving rating accuracy.

As a stable distributed computing architecture, Hadoop's core is HDFS and MapRecurce. The most important aspect of the Hadoop architecture is the implementation of distributed storage for the underlying framework through HDFS. Distributed programming support for parallel tasks can be achieved through MapReduce. q and C represent search behavior and click through volume, respectively. Users generally ignore the return

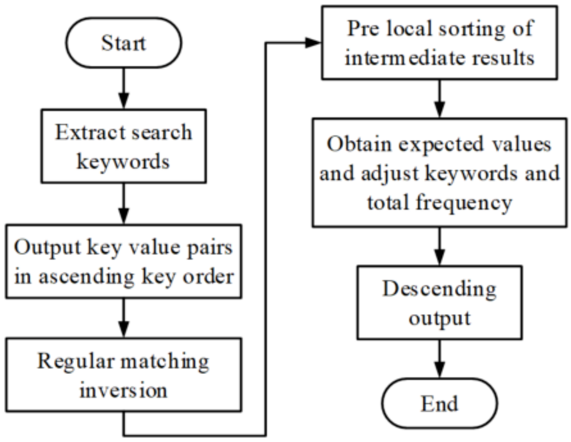


Fig. 3.1: The statistical process of users' search keywords

result URL, which greatly affects their search behavior. This defect can be remedied through equation 3.1.

$$U_q = \sum_{i=1}^{n} c(A,q) * \text{click}(A,q) \tag{3.1}$$

In equation 3.1, $c(A,q)$ represents the click through volume of the balance factor click(A,q) of webpage A. The higher the U_q value, the more popular the page is. When a user performs a search activity, if they believe there are similarities in the search activity, they will stay for a period of time. This behavior does not affect users' satisfaction with search activities. For this purpose, equation 3.2 is used to represent the weight of the search time.

$$\text{Time}(A,q) = \frac{t_i}{\sum_{i=1}^{n} t_i} \tag{3.2}$$

In equation 3.2, t_i is the set of words q that users find when spending time browsing websites. In the cloud computing search process, there is a correlation between pages i and j, but there is a significant difference in their weights [12]. When performing N iterations, in $[0,t]$, the page construction matrix clicked by the user is C_{NxN}. $C_{i,j}$ represents the number of clicks on i and j. If $C_{i,j}$ and $C_{j,k}$ are greater than 0, there is a relationship between I, j, and k, as shown in equation 3.3.

$$K(A,T_i) = \lambda(\text{ID}_A, \text{ID}_{T_i}) \tag{3.3}$$

In equation 3.3, $K(A,T_i)$ represents the correlation between A and T_i. $\lambda(\text{ID}_A$ is used to describe the relevant values found for two-page IDs. When users search for their desired content, their search behavior characteristics and browsed information will be recorded in the log directory. Therefore, the main research content of the article is the data analysis of logs. User log analysis includes searching for key words and user time periods. The specific implementation steps are shown in Figure 3.1.

From Figure 3.1, it can be seen that the input data of the Map function is a segment in the log text. The first MapReduce calculation process analysis found that the text content in the first column is a string. The second column is the number of times this text is presented. During the second MapReduce calculation, a set interface is used to pass the parameters to the user. Then the task is controlled by the main control program [13]. However, when conducting searches, users choose web pages based on a large number of similar topics. The imbalance is inevitable. Therefore, in addition to simple connections, an implicit connection also needs to be considered. Therefore, the PR expression for calculating the conventional page ranking and page

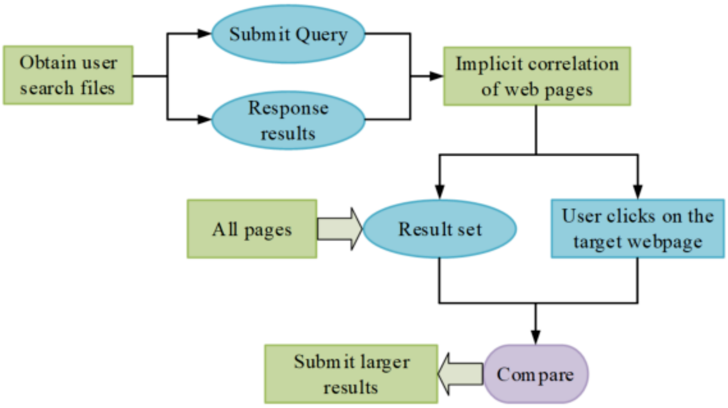


Fig. 3.2: Detailed comparison process for real-time feedback

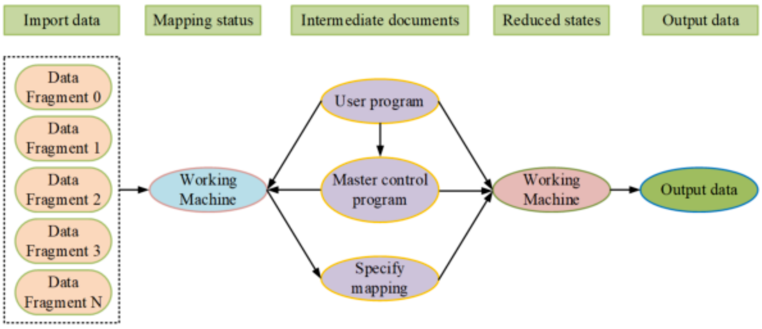


Fig. 3.3: Map/Reduce model structure

X is shown in equation 3.4

$$PR(X) = \sum_{(X,T_i)\in R} \left(\frac{PR(T_i)\left(\delta_1 f(X,T_i) + \delta_2 T(X,q) + \delta_3 k(X,T_i)\right)}{\sum_{k=1}^{M} click(T_i, X)} \right) \tag{3.4}$$

In equation 3.4, the parameters δ_1, δ_2 and δ_3 all represent the influencing factors. $\delta_1 + \delta_2 + \delta_3 = 1$. E is used to describe the total number of web pages. d represents the damping factor. (T_i, X) represents the number of clicks on web pages T_i and X. The higher the click through volume, the greater its relevance. During user retrieval using cloud computing, a specific example of real-time feedback comparison can be obtained. This process is shown in Figure 3.2. From Figure 3.2, it can be seen that the user selects the target webpage after receiving a set of results from the search engine. The relevant ID number can be obtained. In addition, based on the implicit value of association, the result set is compared with the implicit degree of association. The close correlation of web pages is fed back to users as new search results [14]. The Hadoop distributed computing platform can analyze and mine data from multiple aspects to infer user search preferences. Based on this, search results are correlated with content to infer users' emotional tendencies. The Map/Reduce model is shown in Figure 3.3.

From Figure 3.3, it can be seen that the Map/Reduce model mainly consists of five parts. They are input data, mapping state, intermediate document, reduced state, and output data. Hadoop cloud computing is used to count user query records. Due to the large number of logs queried, Mapreduce needs to be calculated in parallel. Through this operation, the number of keyword queries, the ranking of URL feedback results, and

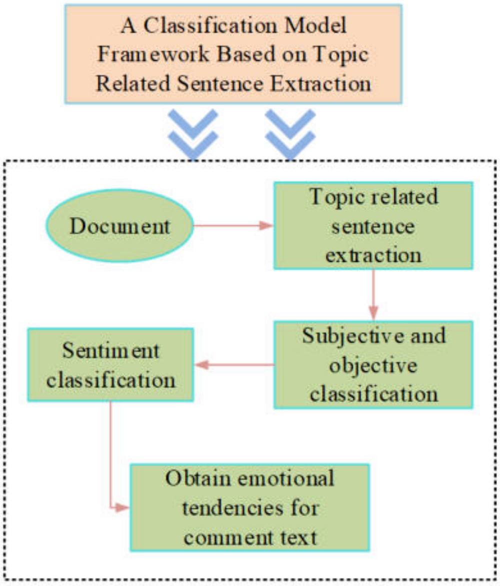


Fig. 3.4: A Classification Model Framework Based on Topic Related Sentence Extraction

the number of users can be calculated. Based on the user's query patterns and preferences, a summary of association algorithms is conducted. User search topics can improve user search efficiency.

3.2. Classification of movie reviews based on sentiment sentence extraction. In movie reviews, text information often includes multiple movie themes. In addition, other movie theme reviews also include some emotional words [15]. Therefore, when classifying words based on emotions, words and sentences that are not related to the movie theme should be eliminated. Given the divergent nature of film review themes, a method for emotion classification is proposed based on this research. That is to extract emotional theme related sentences from movie review texts. The model framework for sentiment classification is shown in Figure 3.4. From Figure 3.4, it can be seen that the extraction method is mainly divided into three stages. Firstly, sentences related to the topic are found in the text. Then, sentences related to the theme are analyzed to determine the subjectivity and objectivity of the text. Finally, the theme sentences related to subjectivity can be found in the text, which is the theme sentiment sentences. The sentiment analysis of text is an sentiment classifier constructed based on machine learning methods. Then the entire comment text's emotions are classified. Finally, the emotional orientation of the text is obtained [16]. Sentences related to the topic are divided into subjective and objective categories. Objectivity sentences are eliminated. Afterwards, sentiment classification is performed on the topic sentiment sentences. Machine learning methods are used to conduct emotional analysis on comment texts and assess their emotional tendencies. Common machine learning methods include support vector machine (SVM), maximum entropy, Latent Dirichlet Allocation (LDA) and other methods. SVM is a supervised learning model, which can analyze data, recognize high-dimensional patterns, classify emotions, Logistic regression, etc. The SVM model performs well in classification performance. The samples are divided into two types. A sample set is a collection of sample points $\{(X_1, C_1), (X_2, C_2), \ldots, (X_n, C_n)\}$. In the sample point set, C_j represents the category, with a value range of [-1,1]. A high-dimensional space vector is represented by a number X_i. The mathematical expression of the classification hyperplane is shown in equation 3.5.

$$w \cdot x - b = 0 \tag{3.5}$$

The two are parallel to the optimal hyperplane and closest to the support vector. SVM modeling method is used to find the optimal hyperplane, so as to transform the problem into the Quadratic programming

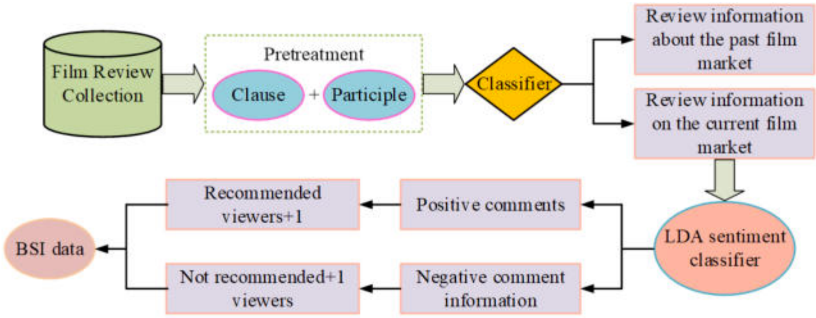


Fig. 3.5: Step diagram of user sentiment prediction

optimization [17]. Based on Lagrand's Mean value theorem, the hyperplane can be described by equation 3.6.

$$f(x)=\sum_{i=1}^{n}(a_i c_1 x_i^T)+b \tag{3.6}$$

The maximum entropy can predict the results of probability distribution. That is to say, the prediction result is to try to choose a uniformly distributed result. x represents the feature vector of the sample. $y$represents the category of the sample. The eigenvector values of sample x in $p(y|x)$ represent probability. The expression for maximum entropy is shown in equation 3.7

$$H(p)=-\sum_{(x,y)} p(y\mid x)\log(y\mid x) \tag{3.7}$$

Combined with adjustment function, known information can be presented relatively easily. So, the classification based on the maximum entropy model evolves into the optimization of feature functions. This feature function has constraints. The expression for obtaining the maximum entropy probability is shown in equation 3.8.

$$f_i(x,y)=\begin{cases}1 & \text{if "meet the conditions"}\\ x & \text{otherwise}\end{cases} \tag{3.8}$$

The optimized maximum entropy expression is shown in equation 3.9.

$$P(y\mid x)=\frac{\exp\left(\sum_i \lambda_i f_i(x,y)\right)}{Z(x)} \tag{3.9}$$

In equation 3.9, $Z(x)$represents the factor. λ is used to describe the weight value of feature i. The maximum entropy model does not need to consider whether each feature value is related, nor does it require independent feature assumptions. However, based on attribute vectors, different attributes are randomly selected according to their differences [18]. Therefore, when studying multiple categories, the maximum entropy model is the most effective. On the basis of early emotional analysis, a prediction model that can automatically identify users' future emotional tendencies towards movies has been constructed. In movie reviews, comments on future movie emotions also include current movie sentiment reviews. Therefore, when conducting emotional analysis of the text, this factor should be included. Based on the previous emotional analysis, a new emotional prediction model has been constructed. The implementation steps for user sentiment prediction are described in Figure 3.5.

From Figure 3.5, the model has two classifiers. Therefore, it can be divided into two stages to analyze the emotions of words. The sentences recognized by the first level classifier are sentiment classified by the second level classifier. This classification will generate two types of surprises and disappointments. Among them, BSI data is a type of data in statistical economics. The expression is shown in equation 3.10.

$$\begin{cases}BSI(1)=\frac{TJ}{TJ+BTJ}\times 100\%\\ BSI(2)=\frac{BTJ}{TJ+BTJ}\times 100\%\end{cases} \tag{3.10}$$

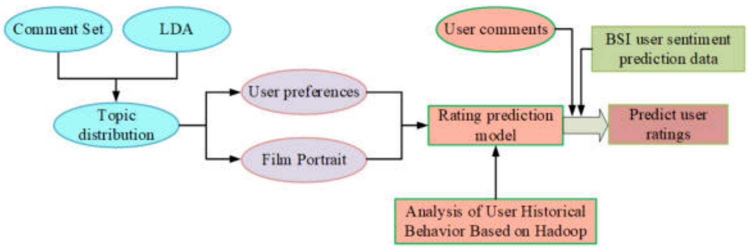


Fig. 3.6: Rating prediction process based on user comments

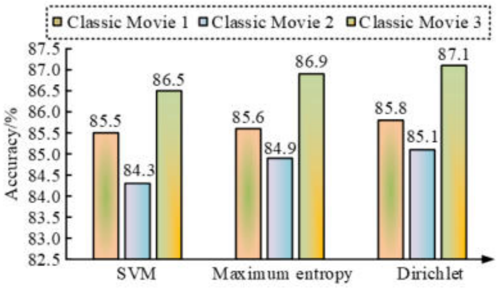
In equation 3.10, TJ and BTJ represent the number of recommended users and the number of non-recommended users, respectively. The first classifier is a classifier based on temporal order. The second classifier is the potential Dirichlet classifier. The accuracy of both classifiers has a significant impact on the accuracy of BSI. A secondary classifier is constructed in the study. It is used as an influential factor in Sentiment analysis.

3.3. Rating prediction based on film review analysis. With the development of the Internet, recommendation technology is increasingly widely used in multimedia social media. The recommendation system can accurately predict users' preferences and help them find the most suitable product for themselves. The current research focus is on how to accurately judge user preferences and product categories, thereby improving the accuracy of recommendations. A new emotion prediction model is constructed by utilizing words and symbols with emotional meanings such as comments' text information, user ratings, and emoticons. On this basis, a sentiment prediction model based on evaluation is proposed. This model not only considers the application of evaluation methods, but also needs to refine the evaluation text to mine information such as user preferences and movie classification [19]. First, through the topic model, the potential topic distribution in the comment information is determined. Then, through the Logistic regression model, each potential word that has an impact on the generation of comment information and scoring is searched out. The internal correlation between possible topics and actual comment information and ratings is determined. After understanding the preferences and search behavior of elderly users, a new method is proposed to accurately grade and predict text information. Then, the scores of the scoring prediction results can be sorted in descending order. Traditional recommendation models focus on the impact of evaluation results on elderly users. However, traditional models have cold start issues [20, 15]. After extracting the scoring features in the comment text, feature analysis can be performed to associate user preferences with product categories. This can alleviate the cold start issue during the recommendation process. In addition, the information in the comments can also be presented to elderly users by selecting the most representative comment information based on the potential topic distribution. The specific implementation process of a rating prediction model based on user comments is shown in Figure 3.6.

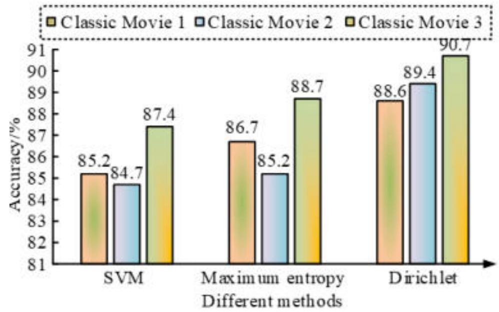
From Figure 3.6, the movie portrait module and rating prediction module are the two functional modules that the model must have. User preferences, movie portraits, and Hadoop based user behavior analysis are organically combined. Based on the regression model, a user rating prediction model is constructed. After training the model, the trained model is used to predict ratings for movies that have not been rated by users.

4. User emotions analysis in social media networks. The movie reviews of elderly social media network users are combined with machine learning, first level classifiers, and second level classifiers for sentiment analysis. According to movie review analysis, rating prediction is achieved.

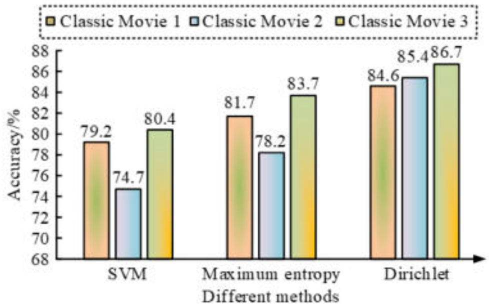
4.1. Emotional Analysis Results of Film Review Text. Based on machine learning for sentiment analysis, various methods are used to analyze emotions in various movies, including SVMs, maximum entropy, and Dirichlet. Then, based on this, a new model is established to analyze the text to determine the classification accuracy of different machine learning methods. In the experiment, three methods are still used, including SVM, maximum entropy and Dirichlet. The first level classifier is based on time series. The differentiated feature values are used as evaluation indicators for the classifier. In the experiment, the 5-fold crossover method is used to validate in the first level classifier and the second level classifier, respectively. The classification accuracy



(a) The accuracy of emotional classification in movies



(b) The classification results of the first level classifier on different feature items



(c) The classification results of a two-level classifier on different feature items

Fig. 4.1: Classification accuracy and sentiment analysis results of the first and second level classifiers

and sentiment analysis results of the first and second level classifiers are shown in Figure 4.1.

From Figure 4.3a, under machine learning, the accuracy of emotion classification results is in descending order of Dirichlet, maximum entropy, and support vector machine. The highest classification accuracy of the three algorithms is 87.1%, 86.9%, and 86.5%, respectively. From Figures 4.3b and 4.1c, in the sentiment analysis of the first and second level classifiers, the three classifiers have good classification performance for three different types of features. Especially when both unary and binary grammars are used simultaneously, their classification performance is better than the first two. The classification accuracy of the first level classifiers of Dirichlet, maximum entropy, and support vector machine are 90.7%, 88.7%, and 87.4%, respectively. The classification accuracy of the second level classifier is 86.7%, 83.7%, and 80.4%, respectively. From the classification results of the three classifiers, Dirichlet has better classification performance than the first two classifiers. Therefore, the method of analyzing the text of Dirichlet's film reviews will be integrated into the research method for the next step of emotional analysis.In the first and second level classifiers, the lack of distinction between emotions results in the presence of both positive and negative emotion words in the comment text. In addition, the mixing of different types of movies increases the difficulty of emotion classification. Thus, the classification results are not satisfactory. Therefore, the emotional analysis of the film review text continues. In the experiment, 5 PC servers are used to build a Hadoop distributed computing platform. Figure 4.1 shows the comparison results of nodes and accuracy based on different data volumes.

From Figure 4.2, the overall effectiveness of the research method is superior to the other two traditional methods. Hadoop is used to analyze the search behavior of elderly users, which can obtain relevant information needed by elderly users and be used for subsequent emotional analysis. In addition, cloud computing models will help better analyze user data. It can also effectively compensate for the shortcomings in the existing

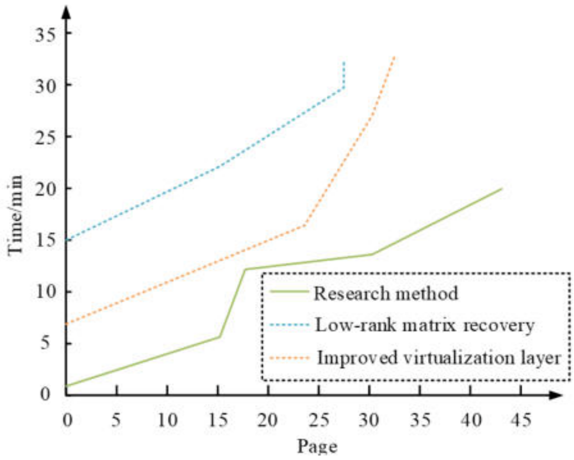


Fig. 4.2: Comparison results of three algorithms based on different page views

Table 4.1: Statistical results of user data on social media networks for the elderly population

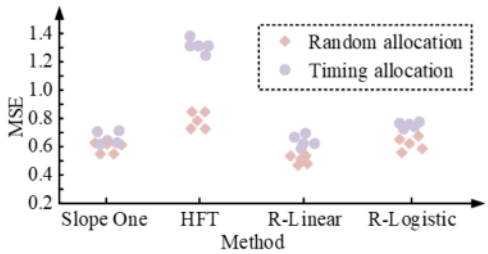
Number	Film Category	Number of Users/People	Number of Movies/Films	Comment/Article
(1)	Action	13,171	229	101,719
(2)	Comedy	17,018	119	170,175
(3)	Affectional film	12,843	151	58,419
(4)	Science fiction film	8,657	120	86,562
(5)	Suspense film	2,353	61	23,522
(6)	Horror film	5,452	70	54,519
(7)	Total	59,494	750	494,916

Hadoop distributed computing architecture.

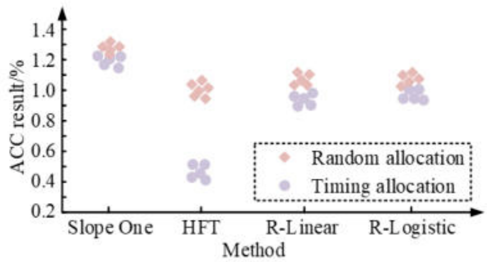
4.2. Scoring prediction results based on movie review analysis. The experiment collected a large amount of comment information from middle-aged and elderly social media network users in the data center, including 59494 elderly users, 750 movies, and 494916 comments. Among them, movies mainly include six types of domestic and foreign films: action films, comedy films, romance films, science fiction films, suspense films, and horror films. The user data statistics results are shown in Table 4.1.

The dataset of movie reviews was randomly divided into a training set and a testing set in an 8:2 ratio. The theme content and ratings in movie reviews are linked by HFT. The HFT source code is downloaded from the website. The collaborative filtering (CF) method means that users with the same preferences can choose the same movie. The subjective decision-making of traditional CF associated users is used to filter comment text. Slope One is a widely used commodity Collaborative filtering method. This method has the characteristics of simplicity and high efficiency, which is implemented by the open-source tool My Media Lite3.10. In this chapter's experiment, four methods, Slope One, HFT, R-Linear, and R-Logistic, are used as test systems. 5 tests are conducted to demonstrate the effectiveness of the system. The Mean squared error (MSE) and accuracy (ACC) results of the four methods are shown in Figure 4.3.

From Figure 4.3, the results of MSE are a comparison of the prediction accuracy of the scores. In actual scores, the size of the score is generally between 1 and 5. However, the scores obtained through Slope One, HFT, and linear regression are all decimal. Therefore, the evaluation values containing decimals were rounded to meet the calculation requirements of ACC. Among them, random allocation and temporal allocation have the highest MSE results on HFT, with 0.78 and 1.32, respectively. Random allocation and temporal allocation

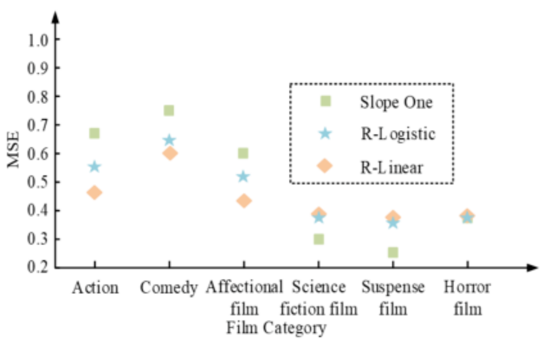


(a) MSE Results

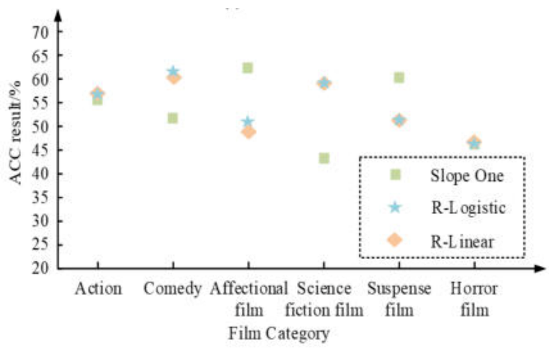


(b) ACC Results

Fig. 4.3: MSE and ACC results of four methods



(a) MSE Results



(b) ACC Results

Fig. 4.4: Prediction Results of Six Types of Movies Using Slope One, R-Linear, and R-Logistic Methods From Figure 4.4, the linear regression method has achieved good results in all six movie categories.

achieved the highest ACC results on Slope One, with 56.5% and 55.4%, respectively. Slope One, R-Linear, and R-Logistic methods are used to predict action, comedy, romance, science fiction, suspense, and horror films. The results are shown in Figure 4.4.

However, when the data suddenly changes, the results obtained by using the Slope One method to predict six categories have significant fluctuations. From the experimental results, it can be seen that the predictive analysis results of the research method are superior to the results generated by using Slope One. In the case of sparse data, the research method can still have good stability and small fluctuations, with good MSE and ACC results. When the recommendation system recommends a movie to users, they will first browse the review information of the movie before they can understand the general content of the movie. The selection of user representative comments is predicted. The results are shown in Figure 4.5.

From Figure 4.5, a video recommendation model has been established through the evaluation of viewers and text analysis of their comments. Finding the comment topic in the movie's comment text to obtain the topic distribution of the comment text can project both user information and movie information into the same space. The above process can obtain user preferences and movie portraits to establish a regression model. The theme distribution and prediction are linked to recommend movies to users. Combining predictive methods, representative evaluations were selected as support for this model. From the experimental data, the scoring prediction model has shown good performance and special adaptability to data sparsity.

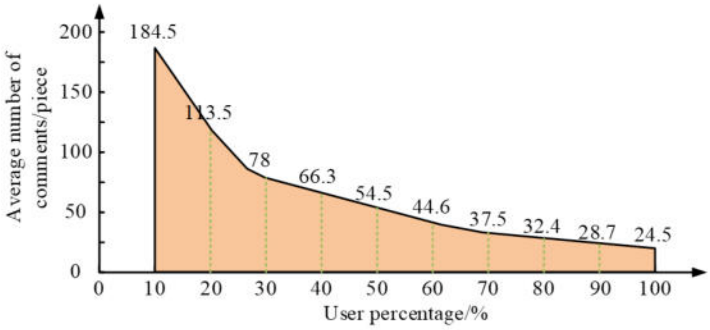


Fig. 4.5: Prediction of user representative comments selection

5. Conclusion. With the development of the Internet, there are more and more comment texts that integrate user emotions. There are more emotional commentary texts in the film field. To strengthen the user sentiment analysis of social media networks for the elderly population, movie reviews are analyzed through user search behavior. The results indicate that by analyzing the logs of elderly users, the behavioral patterns of users can be discovered. The accuracy of sentiment classification results is ranked from high to low by Dirichlet, maximum entropy, and SVM. The highest classification accuracy of the three algorithms is 87.1%, 86.9%, and 86.5%, respectively. The overall effectiveness of the research method is superior to the other two traditional methods. The predictive analysis results of the research method are superior to those generated by using Slope One. In the case of sparse data, the research method can still have good stability and small fluctuations, with good MSE and ACC results. Therefore, this study will have certain guiding significance and practical value. In the future, research will be conducted on automatic classification of subjective and objective texts.

Funding. The research is supported by: 2021 Guangxi Philosophy and Social Science Planning Research Project. Approval No:21FXW013.

REFERENCES

[1] Derakhshan, A. & Beigy, H. Sentiment analysis on stock social media for stock price movement prediction. *Engineering Applications Of Artificial Intelligence.* **85** pp. 569-578 (2019)
[2] Babvey, P., Borrelli, D., Lipizzi, C. & Galaxies, R. Digital Fingerprints of Discussions on Social Media. *IEEE Transactions On Computational Social Systems.* **8**, 294-307 (2021)
[3] Liu, H., Chatterjee, I., Zhou, M., Lu, X. & Abusorrah A. Aspect-Based Sentiment Analysis:, A. Survey of Deep Learning Methods. *IEEE Transactions On Computational Social Systems.* **7**, 1358-1375 (2020)
[4] Chakraborty, K., Bhattacharyya, S. & Bag, R. Survey of Sentiment Analysis from Social Media Data. *IEEE Transactions On Computational Social Systems.* **7**, 450-464 (2020)
[5] Abid, F., Li, C. & Alam, M. Multi-source social media data sentiment analysis using bidirectional recurrent convolutional neural networks. *Computer Communications.* **157**, 102-115 (2020)
[6] Alam, M., Abid, F., Cong, G. & Yunrong, L. Social media sentiment analysis through parallel dilated convolutional neural network for smart city applications. *Computer Communications.* **154**, 129-137 (2020)
[7] Ducange, P., Fazzolari, M., Petrocchi, M. & Vecchio, M. An effective Decision Support System for social media listening based on cross-source sentiment analysis models. *Engineering Applications Of Artificial Intelligence.* **78** pp. 71-85 (2019)
[8] Vashishtha, S. & Susan, S. Fuzzy Rule based Unsupervised Sentiment Analysis from Social Media Posts. *Expert Systems With Applications.* **138**, 1-11283 (2019)
[9] Jeong, B., Yoon, J. & Lee, J. Social media mining for product planning: A product opportunity mining approach based on topic modeling and sentiment analysis. *International Journal Of Information Management.* **48** pp. 280-290 (2019)
[10] Bansal, N., Goyal, V. & Rani, S. Experimenting Language Identification for Sentiment Analysis of English Punjabi Code Mixed Social Media Text. *International Journal Of E-Adoption (IJEA).* **12**, 52-62 (2020)
[11] Rathore, P., Soni, J., Prabakar, N., Palaniswami, M. & Santi, P. Identifying Groups of Fake Reviewers Using a Semisupervised Approach. *IEEE Transactions On Computational Social Systems.* **8**, 1369-1378 (2021)
[12] Amina, B. & Azim T. Scancpeclens:, A. FrameWork for Automatic Lexicon Generation and Sentiment Analysis of MicroBlogging Data on China Pakistan Economic Corridor. *IEEE Access.* **7**, 33876-13388 (2019)

[13] Liu, H., Burnap, P., Alorainy, W. & Williams, M. Fuzzy Approach to Text Classification with Two-Stage Training for Ambiguous Instances. *IEEE Transactions On Computational Social Systems*. **6**, 227-240 (2019)
[14] Lu, R., Tsao, H., Lin, H., Ma, Y. & Chuang, C. Sentiment Analysis of Brand Personality Positioning Through Text Mining. *Journal Of Information Technology Research*. **12**, 93-103 (2019)
[15] Nazir, A., Rao, Y., Wu, L. & And, S. and Challenges of Aspect-based Sentiment Analysis: A Comprehensive Survey. *IEEE Transactions On Affective Computing*. **13**, 845-863 (2022)
[16] Ibrahim, N. & Wang, X. Decoding the sentiment dynamics of online retailing customers: Time series analysis of social media. *Computers In Human Behavior*. **96** pp. 32-45 (2019)
[17] Shi, L., Liu, L., Wu, Y., Jiang, L. & Panneerselvam, J. Human-Centric Cyber Social Computing Model for Hot-Event Detection and Propagation. *IEEE Transactions On Computational Social Systems*. **6**, 1042-1050 (2019)
[18] Wang, B., Wang, E., Zhu, Z., Sun, Y., Tao, Y. & Wang, W. An explainable sentiment prediction model based on the portraits of users sharing representative opinions in social sensors. *International Journal Of Distributed Sensor Networks*. **17**, 3323-3330 (2021)
[19] Liu, Y., Jiang, C. & Zhao, H. Assessing product competitive advantages from the perspective of customers by mining user-generated content on social media. *Decision Support Systems*. **3079**, 1-11307 (2019)
[20] Khaled, A., Ouchani, S. & Chohra, C. Recommendations-based on semantic analysis of social networks in learning environments. *Computers In Human Behavior*. **101** pp. 435-449 (2019)
[21] Nimrah, S. & Saifullah, S. Context-Free Word Importance Scores for Attacking Neural Networks. *Journal Of Computational And Cognitive Engineering*. **1**, 187-192 (2022)

Edited by: Rajni Mohana
Special issue on: Sentiment Analysis and Affective computing in Multimedia Data on Social Network
Received: Aug 1, 2023
Accepted: Oct 16, 2023