RESEARCH

A hybrid recommendation model in social media based on deep emotion analysis and multi-source view fusion

Liang Jiang[1,2], Lu Liu[3*], Jingjing Yao[4] and Leilei Shi[1,2]

Abstract

The recommendation system is an effective means to solve the information overload problem that exists in social networks, which is also one of the most common applications of big data technology. Thus, the matrix decomposition recommendation model based on scoring data has been extensively studied and applied in recent years, but the data sparsity problem affects the recommendation quality of the model. To this end, this paper proposes a hybrid recommendation model based on deep emotion analysis and multi-source view fusion which makes a personalized recommendation with user-post interaction ratings, implicit feedback and auxiliary information in a hybrid recommendation system. Specifically, the HITS algorithm is used to process the data set, which can filter out the users and posts with high influence and eliminate most of the low-quality users and posts. Secondly, the calculation method of measuring the similarity of candidate posts and the method of calculating K nearest neighbors are designed, which solves the problem that the text description information of post content in the recommendation system is difficult to mine and utilize. Then, the cooperative training strategy is used to achieve the fusion of two recommended views, which eliminates the data distribution deviation added to the training data pool in the iterative training. Finally, the performance of the DMHR algorithm proposed in this paper is compared with other state-of-art algorithms based on the Twitter dataset. The experimental results show that the DMHR algorithm has significant improvements in score prediction and recommendation performance.

Keywords: Hybrid recommendation system, Emotion analysis, Multi-source view, Fusion

Introduction

With the development of information technology, the data on the Internet has grown exponentially, and how to effectively provide relevant information to users in need is facing great challenges [1–4] in recent years. To this end, various information sharing systems have been spawned, and online social networks are undoubtedly one of the most popular Internet products in the last decade [5–7], which provides the basic conditions for maintaining social relationships, such as discovering users with similar interests and hobbies, and acquiring information and knowledge shared by other users. These features have made online social networks attract a large number of users who generate a large amount of user generated contents since the day they were born. Therefore, how to use these user generated contents to recommend the information users are interested in and how to continuously optimize the recommendation model for improving the recommendation quality have become one of the hot issues of research [8–10].

At present, a large number of recommended algorithms emerge in which collaborative filtering and content-based semantic model are the more popular algorithms in the early development of recommendation systems, which have been greatly developed in the past decade [11, 12]. The recommendation model based on deep learning has gradually become the hot spot of researchers in the face of the remarkable achievements of deep learning technology

* Correspondence: l.liu@leicester.ac.uk
[3]School of Informatics, University of Leicester, Leicester, UK
Full list of author information is available at the end of the article

in many applications of artificial intelligence. Besides, the user rating matrix is still the main data source used by most recommendation systems, but the recommendation based on user reviews, user implicit feedback, and project content information is getting more and more attention [13]. However, the progress made in these aspects of research is not very satisfactory due to the constraints of text mining and user behavior analysis, but they have important potential in solving the recommendation accuracy, cold start and interpretability of the recommendation system. Meanwhile, there are usually more serious data sparse and cold start problems in social networks compared with the traditional recommendation algorithm, which brings great challenges to the research of social recommendation algorithms.

Aiming at the above-mentioned issues, the collaborative filtering algorithm is highly praised by researchers [14, 15]. Its goal is to transform the binary relationship between users and posts into a score prediction problem, and then collaborative filtering or sorting based on users' scores of posts to generate a recommendation list. Furthermore, subsequent research work has found that the recommendation results based on user ratings do not accurately reflect the user's interest preferences due to the constraints of user ratings and the sparseness of the scoring matrix.

In the content-based recommendation, the description text information of the post content is an important recommendation basis [16]. Content-based recommendation can effectively solve the cold start problem, and is not constrained by the score sparsity, which can discover hidden information and has a good user experience. Hence, it receives wide attention in these days. However, the short text natural language description (usually short and fragmented) for the content of the post does not have enough information for the machine to make statistical inferences, which brings great difficulties to the semantic understanding of the post content.

At present, the research of deep learning technology of integrating multi-source heterogeneous data, fusion scoring matrix and review text, and multi-featured collaborative recommendation has become a hot topic [17] [18–20]. Based on the above research, this paper proposes a hybrid recommendation model based on deep emotion analysis and multi-source view fusion (DMHR algorithm), which aims at the balance of user score distribution and the difficulty of multi-recommendation in recommendation system. The multi-source view here is the multidimensional recommendation factor in the recommendation system. And the hybrid recommendation method of this paper combines three recommended views, such as user rating matrix, user review text, and content description information of posts, which is different from the traditional hybrid methods such as weighted fusion and cascading, this paper designs a recommendation algorithm based on collaborative training to achieve the fusion of the behavioral view of user ratings and post content.

The main contribution of this paper is to propose a scoring prediction method based on multi-recommended view fusion of collaborative training, and to explore the integration of auxiliary language information such as user review text in recommendation system by using natural language processing technology based on deep learning. The tasks of this paper are mainly reflected in the following aspects:

(1) A data preprocessing method based on HITS algorithm is introduced, which filters out the users and posts with high influence, so as to eliminate most of the low-quality users and posts and ensure better efficiency in subsequent processing. At the same time, the authority value of the post is obtained and used as the user's initial rating for the post. Besides, a method based on a comprehensive measure of the user's emotional tendency and the original rating level is proposed. The deviation of the user's original score is corrected from the user's real interest preference by mining the emotional tendency of user's reviews. And the perspective pre-filtering method is used to achieve a comprehensive measure of the user's emotional tendency and the original rating level, and provides a more accurate comprehensive scoring data reflecting the user's real interest preference for the post-based collaborative filtering recommendation model.

(2) A method for text information mining based on post content description is proposed. The text information of the content description of the post is mined, the neural network method is used to represent it as a distributed paragraph vector, and the similarity calculation of the content of the post is realized, thereby constructing a recommendation model based on the content of the post. The calculation method of measuring the similarity of candidate posts and the method of calculating K nearest neighbors are designed, which solves the problem that the text description information of post content in the recommendation system is difficult to mine and utilize.

(3) A hybrid recommendation algorithm based on collaborative training is proposed. The cooperative training strategy is used to achieve the fusion of two recommended views, adds a data selection strategy based on confidence estimation and cluster analysis in collaborative training, eliminating the data distribution deviation added to the training data pool in the iterative training. On this basis, the initial recommendation results are filtered and sorted by using the scoring matrix and the similarity of posts output from the collaborative training model, and the final recommendation results are obtained.

The rest of the paper is organized as follows. In section II related work on recommendation algorithms based on collaborative filtering and content description has been discussed. In Section III and IV, a hybrid recommendation model based on deep emotion analysis and multi-source

view fusion is presented. Our experiments have been analyzed and discussed in Section V. The conclusions have been given and our future work has been outlined in the last section.

Related work

The recommendation system is an effective means to solve the information overload problem and is one of the most common applications of big data technology. It utilizes knowledge discovery technology to filter information and products that users are interested in according to their historical information, hobbies and other characteristics, thereby achieving personalized recommendation. In addition, recommendation algorithms based on collaborative filtering and content description are the two most common Recommendation algorithms [21].

Recommendation algorithms based on collaborative filtering

Collaborative filtering recommendation algorithms can often be divided into user rating-based methods and implicit semantic model-based methods. User rating-based methods which use historical scoring data to discover similar users or similar projects, can generate recommendation lists based on similarity. Implicit semantic model-based methods which map the user and the project to a feature vector with some real meaning, can calculate the user's preference for the project by calculating the inner product of the vector. For example, Guo et al. [22] proposed a neural variational collaborative filtering framework for top-k recommendation. The actual effect of the algorithm is improved by incorporating the side information of user and project, and employing a Stochastic Gradient Variational Bayes approach. Yan et al. [23] proposed a stage-wise matrix factorization algorithm by exploiting manifold optimization techniques. Applying this algorithm to the collaborative filtering recommendation model can greatly improve performance and efficiency on large-scale real data. Koren [24] proposed a matrix factorization-based model for recommendation in social rating networks, named SVD++ algorithm, which introduces the trust delivery mechanism into the social recommendation, and better reflects the influence of the social network trust relationship on the recommendation. Although the collaborative filtering algorithm is widely recommended and easy to implement, it has many problems, such as high computing cost, poor scalability and sparse data.

The social network-based recommendation is the extension of Collaborative Filtering Recommendation Algorithm in social networks, which has the characteristics of data diversity, real-time data update and high interaction. Guo et al. [25] proposed a collaborative filtering recommendation algorithm named TDSRec algorithm that integrates the characteristics of social networks. It

obtains the trust and trusted characteristic matrix, and recommends accordingly. It solves the problem of data sparsity in traditional collaborative filtering algorithm to some extent. Forsati et al. [26] proposed a matrix factorization-based model for recommendation in social rating networks, named SocialMF algorithm, which introduces the trust delivery mechanism into the social recommendation, and better reflects the influence of the social network trust relationship on the recommendation. To a certain extent, the social recommendation algorithm has a wide range of applications, such as the huge amount of data, complex data content, complex algorithm implementation, high time complexity, and weak personalized recommendation results.

Recommendation algorithms based on content description

In the content-based recommendation, the content information of the project is an important recommendation basis, and it is also an important way to solve the cold start problem, but this recommendation method is subject to the information acquisition technology [27]. Content-based recommendation is based on the user's favorite project content information to find similar projects for recommendation. The current popular practice is to use the relevant theories, methods and techniques in information retrieval to model the project content information. Zhao et al. [28] proposed a review-based recommendation model by fusing users' internal influence into a matrix factorization to improve the accuracy of rating predictions. User sentimental deviations and the review's reliability are explored to measure their impact on Social Recommendation. McAuley et al. [29] proposed an HFT algorithm which fuses the scoring matrix and the review text during the parameter learning and fitting phases. It models user ratings and user reviews by establishing a link between the topic distribution of the user's reviews and the potential factors of the user or post. Bao et al. [30] proposed TopicMF algorithm which uses non-negative matrix factorization to mine the topic distribution of a single comment. It is considered that the topic distribution reflects user preferences and project characteristics, and maps with user potential factors and project potential factors. Ding et al. [31] proposed a learning algorithm based on the element-wise Alternating Least Squares learner which integrates view data into a recommendation system based on implicit feedback to mine hidden preference information other than primary feedback data such as purchases. However, text content is usually short and fragmented. If historical information is not referred to, it is easy to cause insufficient information for the machine to make statistical inferences, which brings great difficulty to the semantic understanding of the content of the item. The recommended information is also single and the user's interest is limited.

Human emotion expressed in social media plays an increasingly important role in shaping policies and decisions. Emotion analysis on social networks has attracted increasing research attention. In order to improve the accuracy of recommendation, emotion analysis is combined with other factors. Chouchani et al. [18] used information about social influence processes to improve emotion analysis. Phan et al. [19] proposed a new approach based on a feature ensemble model related to tweets containing fuzzy emotion by taking into account elements such as lexical, word-type, semantic, position, and emotion polarity of words. Chung et al. [20] developed a novel framework for dissecting emotion and examining user influence in social media which comprehensively considered emotions, social positions, influence and other factors. However, human emotion is fluctuating, more user history data is needed and multiple recommendation factors are not easy to integrate. Recommended model based on emotion analysis easily restricted by data sparsity and cold start, so it is not easy to obtain good recommendation effect.

Data preprocessing

The data obtained from social networks is disorganized and faces the problem of sparse data and cold start, which requires pre-processing the data to improve the recommendation model. To this end, this paper introduces the HITS algorithm [32] to the recommendation model.

The HITS algorithm is one of the classic algorithms for web search. It finds the authority page and the hub page in the page collection by analyzing the hyperlinks between the pages. These characteristics of the HITS algorithm have attracted many researchers' attention and have been introduced into online social networks. Similarly, the authority value and the hub value are used to represent the influence of users and posts respectively. HITS algorithm is used to process the data set and filter out the users and posts with high influence, so as to eliminate most of the low-quality users and posts, which ensures better efficiency in subsequent processing. At the same time, the authority value of the post is obtained and used as the user's initial rating for the post. The authority value of the post can be represented by the sum of the hub scores of all users who have forward the particular posts:

$$a(p) = \sum h(u) \tag{1}$$

The authority value of the post $a(p)$ is standardized:

$$a(p) = \frac{a(p)}{\sqrt{\sum [a(p)]^2}} \tag{2}$$

The authority value of the post $a(p)$ is obtained by iterating repeatedly until $a(p)$ converges. However, the

initial score obtained in this way does not take into account the time attribute and the actual interest of the user, and the recommended model proposed in this paper overcomes these shortcomings.

Hybrid recommendation model based on deep emotion analysis and multi-source view fusion

In view of the above discussion on the status quo of recommendation model research, this paper proposes a hybrid recommendation model based on deep emotion analysis of user reviews and cooperative fusion of multi-source recommendation views, named DMHR. The process of DMHR hybrid recommendation model is as follow: firstly, The perspective pre-filtering method [33] is used to achieve a comprehensive measure of the user's emotional tendency and the original rating level, and provides a more accurate comprehensive scoring data reflecting the user's real interest preference for the post-based collaborative filtering recommendation model. Simultaneously, the text information of the content description of the post is mined, and the neural network method is used to represent it as a distributed paragraph vector, realizing the similarity calculation of the content of the post, and then a recommendation model based on the content of the post is constructed. Secondly, the cooperative training strategy is used to achieve the fusion of two recommended views, adding a data selection strategy based on confidence estimation and cluster analysis in collaborative training, and eliminating the data distribution deviation added to the training data pool in the iterative training. Finally, On this basis, the initial recommendation results are filtered and sorted by using the scoring matrix and the similarity of posts output from the collaborative training model, and the final recommendation results are obtained. The deviation of the user's original score from the user's real interest preference is corrected by mining the emotional tendency of user's reviews for the next recommendation. The hybrid recommendation model system framework is shown in Fig. 1.

Emotional analysis of user reviews
Distributed vector representation of user review text

Through statistical analysis of the user review text in the recommendation model, it is found that the presentation form is usually a keyword and a short text. Research shows that these short text messages are usually processed differently from long text. The short text has the characteristics of short length and irregular grammar, which makes traditional natural language processing technology powerless in short text analysis. Early analysis and application of short text mainly rely on enumeration or keyword matching, avoiding the semantic understanding of text, while automatic short text understanding usually relies on additional knowledge. In this
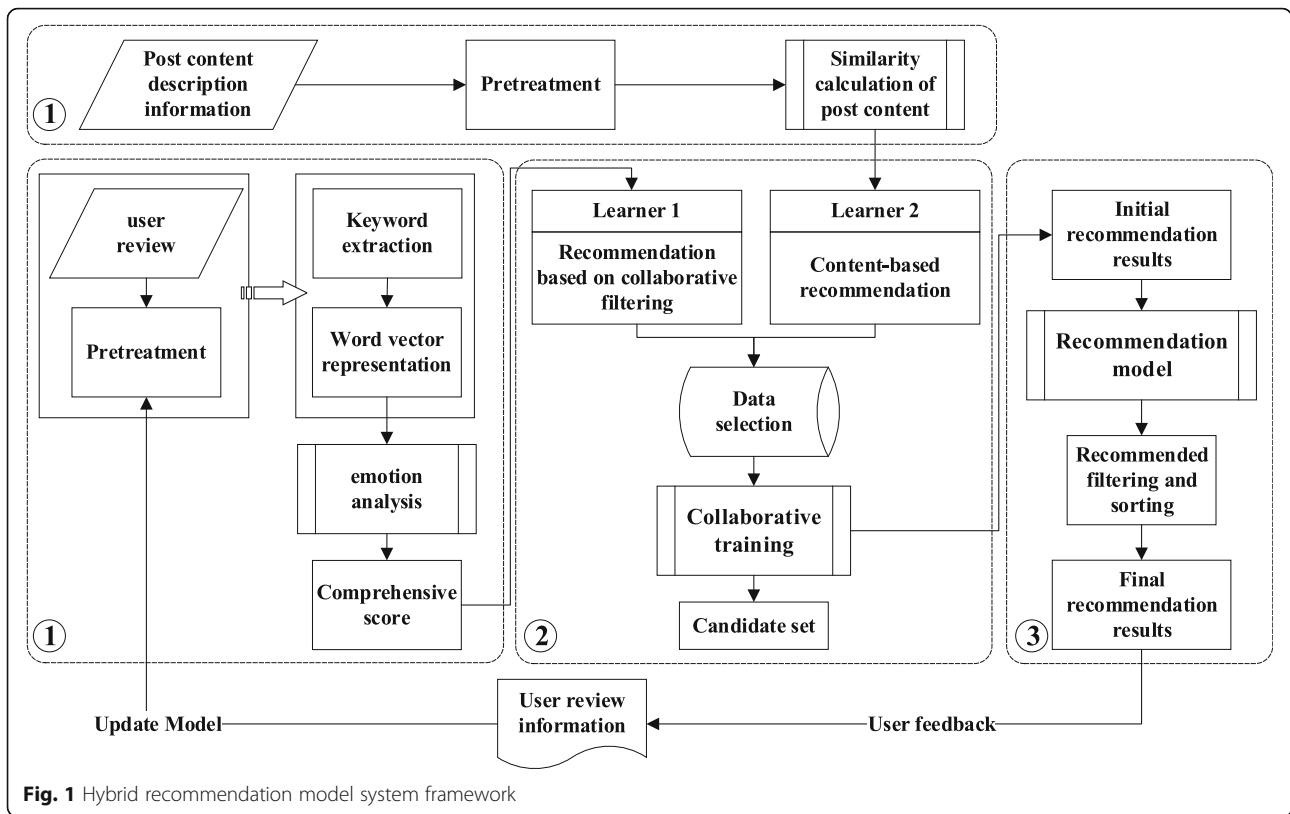
Fig. 1 Hybrid recommendation model system framework

paper, we use the keyword representation method based on word vector to solve the dimension disaster of traditional sparse representation and the problem of unable to express semantic information. At the same time, the association attributes between words are also mined, which improves the accuracy of the semantic representation of keywords.

Word2vec is a predictive model for high-efficiency word nesting learning, including two variants of CBOW model and Skip-Gram model [34]. CBOW predicts the probability of occurrence of a central word through words within the window, while Skip-Gram is based on the probability that the word appears within the window of the central word prediction. Its training goal is to find the vector representation of the words useful for predicting the surrounding words in sentences or documents. If for a given sentence, ω_1, ω_2, ..., ω_T means the words in the sentence, the objective function $g(\omega)$ of Skip-Gram model is to maximize the average logarithmic probability.

$$g(\omega) = \frac{1}{T}\sum\nolimits_{t=1}^{T}\sum\nolimits_{-c \leq j \leq c, j \neq 0} \log p\left(\omega_{t+j}|\omega_t\right) \qquad (3)$$

In the above formula, c denotes the number of training texts, the larger c is, and the higher the accuracy of the model may be. The Skip-Gram model uses the hierarchical Softmax function to define $p(\omega_{t+j}|\omega_t)$. Hierarchy Softmax uses W words as the binary tree representation

of the leaf's output layer. For each node, the relative probability of its sub-nodes is clearly expressed. Random walk algorithm is used to assign the probability of each word.

Word2vec automatically learns syntactic and semantic information from large-scale unlabeled user reviews, enabling the characterization of keywords in user reviews. The use of Word2vec to vectorize the short text information of user reviews is mainly divided into the following two steps:

(i) According to the collected user review text data, using the Skip-Gram or CBOW training word vector model, each word is expressed as a K-dimensional vector real value;

(ii) For the short text of user reviews, Top-N words are extracted to express the emotion of the text based on word segmentation using TF-IDF and other algorithms, and then K-dimensional vector representation of the extracted Top-N words is found from the word vector model.

After obtaining the K-dimension real vector representation of each key word, a common method is to use weighted average method to process the vector of the key word, which is equivalent to the vector representation of the user review text, in order to realize the emotional analysis of the review information. This weighted averaging method ignores the influence of word order

on the affective prediction model. Because word vector representation based on Word2vec is only based on the dimension of words to carry out "semantic analysis", while weighted average processing of word vectors does not have the ability of "semantic analysis" of context. Therefore, this paper constructs an emotional computing model based on word vector and long short-term memory network to realize the emotional analysis of user reviews.

Emotional calculation based on word vector and long short-term memory network

In text information processing, the commonly used method is the Recurrent Neural Network (RNN) [35]. However, RNN can lead to the disappearance of gradient in optimization when dealing with long sequences. To solve this problem, the researchers proposed a threshold (Gated RNN), the most famous of which is the Long Short-Term Memory Network (LSTM) [36]. The research also shows that the neural network with LSTM structure performs better than that with the standard RNN network in many tasks.

LSTM uses a "gate" structure to remove or add information to the cell state. It achieves the purpose of enhancing or forgetting information by adding three "gate" structures of input gate, forgetting gate and output gate in the neuron, so that the weight of the self-loop is changed. The model based on LSTM can effectively avoid the gradient expansion and even disappearance of the RNN network structure by

dynamically changing the accumulation at different times when the parameters are fixed. In the LSTM network structure, the calculation formula of each LSTM unit is as shown in formulas (4) to (9).

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{4}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{5}$$

$$\widetilde{C_t} = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{6}$$

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C_t} \tag{7}$$

$$O_t = \sigma(W_O \cdot [h_{t-1}, x_t] + b_O) \tag{8}$$

$$h_t = O_t * \tanh(C_t) \tag{9}$$

In formulas (4) ~ (9), f_t denotes the forgetting gate, i_t denotes the input gate, O_t denotes the output gate; $\widetilde{C_t}$ denotes the state of the cell at the previous moment, C_t denotes the state of the current cell, and h_{t-1} and h_t respectively represent the previous moment unit output and current unit output.

In this paper, an emotional analysis method based on Word2vec and LSTM is presented as Fig. 2. Firstly, the input of matrix form is coded into the one-dimensional vector by Word2vec to save most useful information; Then, LSTM algorithm is used to train the emotional classification model of user review text, and the grading prediction of user review is realized. At the same time, in order to take account of the interaction of user ratings and review information
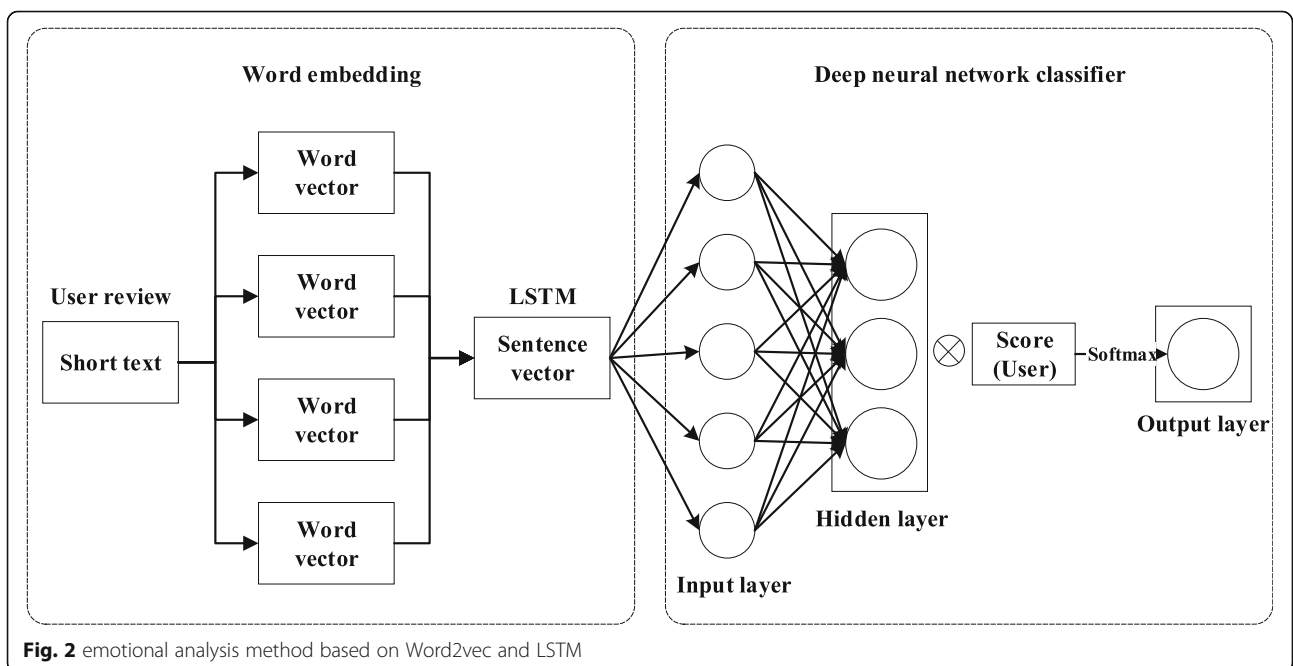

Fig. 2 emotional analysis method based on Word2vec and LSTM

on real emotions, this paper uses the pre-filtering method based on viewpoints and the embedding method based on user ratings to fuse user ratings and emotional prediction ratings respectively. The former uses the LSTM network to get the prediction score, and then weights the sum with the original user score. The method based on user score embedding combines the LSTM network vector with the user rating information, and uses the result as the input of the last layer to directly output the final comprehensive score.

Based on the method of perspective pre-filtering, the emotion analysis of user review text modeling is performed by Word2vec and LSTM, and the emotional tendency score $score_r$ of each user's review on the post is predicted, and the user's original score is weighted and summed to obtain a comprehensive score $score_c$.

$$score_c = \alpha \, score_r + 100(1 - \alpha) \, score_H \qquad (10)$$

In the above formula, $score_r$ represents the user's emotional prediction score for the post review, $score_H$ represents post's authority value in HITS algorithm, due to the limit of the number of data taken, the post's authority value is small. In order to increase its impact on the results, it is expanded by 100 times. α is the balance factor between the two scores.

The method based on user rating embedding is based on the emotional analysis of the user review information, combining the obtained LSTM output vector with the user rating information, then the above result is used as the input to the last layer (fully connected layer) and the final comprehensive emotional score is directly output via the SoftMax activation function.

$$H_i = h_t S\text{core} \otimes (\text{User}_i) \qquad (11)$$

Calculation of similarity based on post content

In the recommendation model, since the natural language description of the post content is short and mostly incomplete, and usually does not follow the grammatical rules. Thus this paper uses the paragraph vector [37] to distribute the short text of the post content description. Paragraph vector is a neural network-based implicit short text comprehension model, which uses a short text vector as "context" to assist in reasoning. In maximal likelihood estimation, text vector is also updated as model parameters. It also adds encoding to the paragraph during the model training process compared with the text vector representation method based on Word2vec. Like ordinary words, paragraph coding is also mapped to a vector (i.e. paragraph coding vector). In the calculation,

paragraph coding vectors and word vectors are accumulated or connected as input of SoftMax in the output layer. The paragraph code remains unchanged during the training of the text description of the post, and semantic information of the entire sentence is integrated every time the word probability is predicted. In the prediction phase, a new paragraph code is assigned to the description text of the post content while keeping the parameters of the word vector and the input layer SoftMax unchanged. Finally, the gradient descent method is used to train the new post description text until it converges, resulting in a low-dimensional vector representation of the post content. The distributed representation of the paragraph vector of the post content is shown below (Fig. 3).

After obtaining the unique d-dimensional distributed vector representation of the post content, the similarity and distance between each two post contents can be obtained by the similarity calculation. This paper uses the cosine formula to measure the similarity between two posts, and uses the Mahala Nobis distance to calculate the distance between the natural language descriptions of the two posts. Assume that the paragraph vectors of the natural language description of the two post contents are represented as $PV_a = (x_{11}, x_{12}, ..., x_{1d})$ and $PV_b = (x_{21}, x_{22}, ..., x_{2d})$, where d denotes the dimensions of two paragraph vectors. Then the similarity and distance between them are defined as follows:

$$\text{sim}(PV_a, PV_b) = \frac{PV_d \bullet PV_d}{\|PV_d\|^2 \bullet \|PV_d\|^2}$$
$$= \frac{\sum_{i=0}^{i=d} x_{1i} x_{2i}}{\sqrt{\sum_{i=0}^{i=d} x_{1i}^2} \sqrt{\sum_{i=0}^{i=d} x_{2i}^2}} \qquad (12)$$

$$dis(PV_a, PV_b) = \sqrt{(PV_a - PV_b)^T S^{-1} (PV_a - PV_b)} \qquad (13)$$

where S is the covariance matrix of eigenvectors PV_a and PV_b.

Multi-source view fusion based on collaborative training

In the construction of the hybrid recommendation model, this paper uses the user comprehensive scoring view to build a post-based collaborative filtering recommendation model; at the same time, a recommendation model based on post content is constructed by using the natural language description view of post content; Finally, the fusion of two
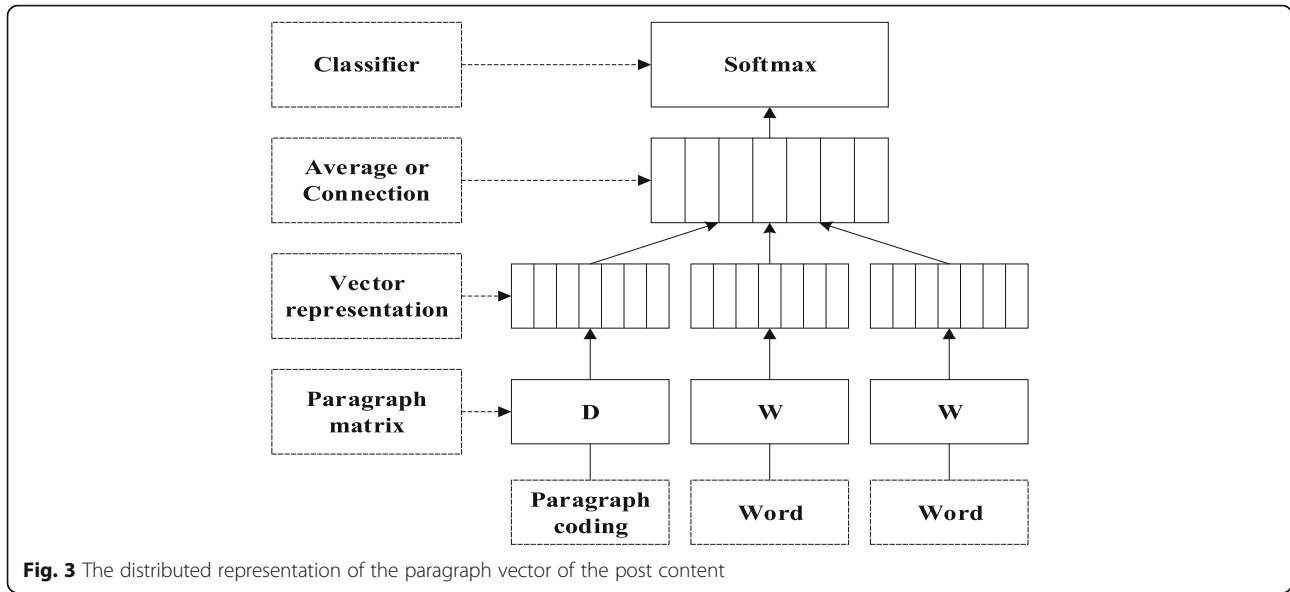
Fig. 3 The distributed representation of the paragraph vector of the post content

recommendation views is realized based on cooperative training strategy. In data selection, data selection algorithm based on confidence estimation and clustering analysis is used to filter the data, and then added to the training data pool of another classifier for the next round of training, so as to iterate.

Hybrid recommendation algorithm based on collaborative training

The hybrid recommendation algorithm based on collaborative training is used to construct the initial scoring matrix based on the user's scoring of the posts. Then the perspective pre-filtering method is used to measure the composite score to update the scoring matrix. Finally, a hybrid recommendation algorithm based on collaborative training is designed in which the scoring matrix is cyclically filled and optimized according to the vector similarity of the comprehensive scoring matrix and the post content description, so as to achieve

recommendation and sorting. Besides the hybrid recommendation algorithm based on collaborative training is shown in the following Fig. 4.

In the recommendation model, the score of user u on post p is recorded as $R_u(p)$ which takes from post's authority value in HITS algorithm; The corresponding scoring matrix is $R_{m \times n}(U, P)$, where the row vector m represents the number of users, and the column vector n represents the number of posts. In the object-based collaborative filtering recommendation model, input the user's original scoring matrix $R_{m \times n}(U, P)$, where $R_u(p) \in [0, 1]$, and the virtual scoring matrix $\overrightarrow{R}_{m \times n}(U, P)$ predicted by the emotion analysis model, where $\overrightarrow{R}_u(p) \in \{0, 1\}$, 0 means that the user's emotion is negative, and 1 means that the user's emotion is positive, output as data set D_{train}. The description of the post-based collaborative filtering recommendation algorithm is as shown in Algorithm 1.
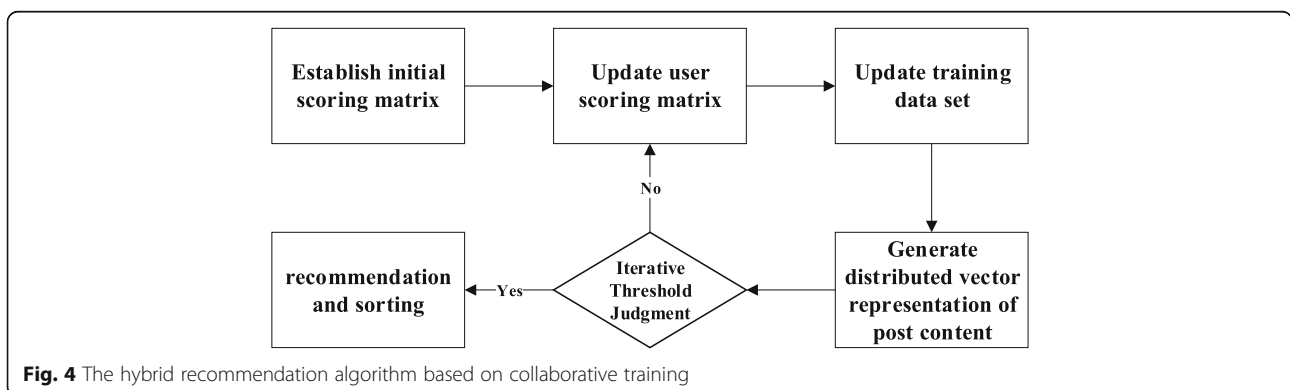

Fig. 4 The hybrid recommendation algorithm based on collaborative training

---

Algorithm 1. Post-based collaborative filtering recommendation algorithm

---

Input: user's scoring matrix for posts $R_{m \times n}(U, P)$, virtual score predicted by emotional computing model $\vec{R}_{m \times n}(U, P)$
Output: training data set based on post collaborative filtering recommendation D_{train}

---

(1) According to user scoring matrix, training data for user u is extracted $D_p = \{R(p)^T | R(p) \in R_{m \times n}(U, P), R_u(p) \neq \emptyset, i \in [1, n]\}$, its category label is $L(p) = R_u(p) \in [0,1]$.
(2) Update training data score $R_u(p)$. User's comprehensive score of posts is calculated by the method of viewpoint pre-filtering. $Time_{cur}$ represents the current time, and $Time_{R_u(p)}$ represents the time of the user's reviews on the

$$R_u(p) = \frac{\alpha \vec{R}_u(p) + 100(1 - \alpha)R_u(p)}{Time_{cur} - Time_{R_u(p)}}$$

(3) Update the training data set, mark the score≥0.7 as a positive class, add it to the data pool $D_p(+)$, and mark the score≤0.3 as a negative class, add it to the data pool $D_p(-)$.
(4) A collaborative filtering recommendation model based on posts is trained. The candidate data $D = \{R(p)^T | R(p) \in R_{m \times n}(U, P), R_u(p) \neq \emptyset\}$ is predicted by the classifier h_1, and the prediction label $L(p)$ is obtained.
$$D'_L \leftarrow \{(D(p), L(p))\}$$
(5) Use data selection algorithm based on confidence estimation and cluster analysis to filter data and return data pre-joined into the training data pool. D_L represents the original data in the second iteration, and D'_L represents the data added in one iteration (the label of the data is the predicted score of the collaborative filtering model).
$$Return\ D_{train} = \{D_L \cup D'_L\}$$

---

description of the recommendation algorithm based on the content of the post is as shown in Algorithm 2.

Algorithm 2. Recommendation algorithm based on post content

Input: User's predicted rating of the post, $D'_L = \{D(p), L(p)\}$, vector representation of the post description $PV(Post)$, training data set, D_{train};
Output: Scoring matrix, $R_{m \times n}(U, P)$;

---

(1) Obtain the user's post collection according to the training data set, D_{train};
(2) Select $2i$ post candidate sets with user score Φ, and calculate the distance and similarity between the candidate posts and the posts in data set D_{pt} respectively;
(3) Select K nearest neighbor posts, $\{D_1, D_2, \ldots, D_K\}$, and calculate the average distance dis_q between post D_p and K nearest neighbors;
(4) Update the value of scoring matrix according to the similarity of post content;
$$FOREACH\ D_i \in D_{2i}\ do$$
$$Score = argmax(a_j | j \in \{0,1\})$$
$$R_u(i) = Score * dis_q$$
(5) Update the score of post p for the $R_u(p)$ value of different users u.
$$L(p)' \leftarrow R(p)$$
(6) Update the score of the training data.
$$D'_L = \{(D(p), L(p))\} \leftarrow \{(D(p), L(p)')\}$$
(7) Data distribution analysis based on confidence estimation and cluster analysis is performed on D'_L, and the filtered data is returned.
$$Return\ R_{m \times n}(U, P)$$

In Algorithm 1, the post-based collaborative filtering recommendation method is used to populate the default value of the user's scoring matrix and update the training data set of user u at the same time. In the emotional classification model, it is generally divided into fine-grained (5-level classification) and coarse-grained (2-level classification). Considering that the accuracy of the 2-level emotional classification model is much higher than that of the 5-level emotional classification model, this paper adopts 2-level emotional classification in the recommendation algorithm. The user's emotions were set to 1 point and 0 point, respectively. Then, the user's emotional scores and original scores were comprehensively measured by means of viewpoint pre-filtering. Finally, the post-based collaborative filtering model is used to predict and fill the scoring matrix, and the data selection algorithm based on confidence estimation and cluster analysis is used to filter the data, and add the incremental data to the training data set of user u.

In the content-based description model, K-nearest neighbor algorithm is used to calculate the distance of content description, and the cosine similarity of posts and the Mahala Nobis distance of K nearest neighbor posts are used to update or fill in the user's score and default value, which is then used in the content-based recommendation model for the next iteration. The

The multiple recommended techniques are mixed within the hybrid recommendation method to compensate for the shortcomings and achieve better recommendations. Different from traditional hybrid recommendation technologies, such as weighted fusion, hybrid recommendation and cascade recommendation, the collaborative training strategy is used in this paper to construct a hybrid model of collaborative filtering recommendation based on posts (Algorithm 1) and content-based recommendation (Algorithm 2). In each iterative training process of the collaborative training model, the calculated comprehensive scoring data is used to train the scoring prediction model to achieve the filling and updating of the scoring matrix. Then, the training model based on the content of the post is trained to be scored according to the updated scoring matrix and the content description information of the post (the posts with the score ≥ 0.7 and the score ≤ 0.3 are respectively placed in the training pool of the post that the user likes and dislikes). The matrix is filled and updated, and it is used as the input of the post-based collaborative filtering recommendation model for the next iteration training. This paper proposes a hybrid recommendation method based on collaborative training compared with weighted fusion hybrid recommendation, which needs to adjust the weight of each recommendation result, the difficulty of ranking hybrid recommendation, and the staged process of cascaded recommendation, which makes full use

of user's scoring information of the post (Post Profile view) and the content description information of the post (meta-data view of the post) in each iteration training to achieve the fusion of the two kinds of recommendation views and a better mixed recommendation effect.

1) *Data selection in collaborative training*

In this paper, a data selection strategy is added to construct the collaborative training model to filter the data to join the training pool. Each grade of the user is specified as a category in the data. The training data in the data pool is tagged data, and the data to be predicted is unlabeled data. In the data selection strategy, not only the confidence score of the sample belongs to a certain category, but also the selected samples are evenly distributed in each cluster, which can avoid the large estimation bias of the selected training data on the Gaussian distribution. A data selection algorithm based on confidence estimation and cluster analysis is described as Algorithm 3.

Algorithm 3. A data selection algorithm based on confidence estimation and cluster analysis

Input: Training sample set *(x,y)*

Output: Selection label $Flag = \{true, false\}$ for sample set *(x, y)*

(1) Initialization:

   Each score level in the training data is taken as a class, and the similarity between the training sample and each class c is calculated as a confidence measure metric of the sample, which is recorded as *f(c;x)*.

(2) Training data partition

   a. K-means clustering algorithm is used to divide candidate data into k sub-clusters $D_1, D_2, \ldots, D_k$.

   b. Calculate the prior probability of each cluster:

$$P(D_k) = (n_k^l + n_k^u)/\sum_k (n_k^l + n_k^u)$$

where n_k^l represents the number of labeled samples in the cluster sub-cluster D_k , n_k^u represents the number of unlabeled samples in the cluster sub-cluster D_k.

   c. Calculate the probability for each category and cluster:

$$P(c|D_k) = \frac{\alpha n_{k,c}^l/n_k^l + (1-\alpha)n_{k,c}^u/n_k^u + \beta}{Z}$$

where $n_{k,c}^l$ represents the number of labeled samples belonging to class c and belonging to cluster sub-cluster D_k, $n_{k,c}^u$ represents the number of unlabeled samples belonging to class c and belonging to cluster sub-cluster D_k, α is the balance factor, β is a constant with a very small value, and Z is a normalization factor that makes $P(c|D_k)$ a function of probability.

(3) Data selection

   a. Select cluster D_k according to probability $P(D_k)$, and select category c according to probability $P(c|D_k)$ in cluster cluster D_k.

   b. Find the intersection of the cluster D_k and the samples of class c, and select the first m samples with the highest confidence estimate (make it *Flag=true*).

   c. Repeat steps a ~ b until the data of all candidate pre-join training pools is filtered.

Experiments

Experiment settings and datasets

The experiment is carried out on a computer with Intel I7 processor and 16GB memory. The datasets selected in this paper cover about 25 million reviews from Twitter from April 2015 to October 2019. The datasets contain the following contents: user information, post information and plain text review information. The specific descriptions of the datasets are shown in the following table (Table 1).

In preprocessing, the HITS algorithm is used to process the data set and filter out the users and posts with high influence, which lays the foundation for subsequent recommendation. The specific results of the HITS algorithm are shown in the following table (Tables 2 and 3).

As can be seen from the above table, the HITS algorithm is used to screen out the top ten most influential users and posts, and the authorities of the posts is used as the user's initial rating for the post to participate in subsequent calculations.

Evaluation measures

In order to evaluate the performance of the proposed algorithm, we choose the classical accuracy index in the recommendation model: mean absolute error (MAE). For a user u and post p in our datasets, r_{up} is the actual score of user u on post p, $\tilde{r}_{up}$ is the predicted score obtained by the algorithm proposed in this paper. T is the number of scores of user u on post p in our datasets. Then the evaluation index MAE in the recommendation model is calculated as follows:

$$\text{MAE} = \frac{\sum_{u,p \in T} |r_{up} - \tilde{r}_{up}|}{T} \qquad (14)$$

The lower the MAE value is, the higher the accuracy of the algorithm prediction is.

Comparative methods

In the experiment of this paper, four more classical recommendation algorithms are selected as the comparison algorithm of the proposed DMHR algorithm. The performance of each algorithm is evaluated by performance indicator MAE. In the case study, the performance of DMHR algorithm in this paper is evaluated by the Top N recommendation of

Table 1 Descriptions of the datasets

Dataset statistics	Quantity
Number of reviews	25,013,682
Number of users	4,621,894
Number of posts	2,184,676
time span	April 2015–October 2019

some specific instances. The four comparison recommendation algorithms are described in detail as follows:

TDSRec algorithm [25]: It is a collaborative filtering recommendation algorithm that integrates the characteristics of social networks. It obtains the trust and trusted characteristic matrix, and recommends accordingly. It solves the problem of data sparsity in traditional collaborative filtering algorithm to some extent.

SocialMF algorithm [26]: It is a matrix factorization-based model for recommendation in social rating networks, which introduces the trust delivery mechanism into the social recommendation, and better reflects the influence of the social network trust relationship on the recommendation.

SVD++ algorithm [24]: It is an improved singular value decomposition (SVD) technique that introduces implicit feedback based on SVD. User's historical browsing data and user's historical rating data are all used as new parameters.

HFT algorithm [29]: It models user ratings and user reviews by establishing a link between the topic distribution of the user's reviews and the potential factors of the user or post.

The influence of parameters
Effect of balance factor α
In the DMHR algorithm proposed in this paper, there is an important parameter α, which reflects the weighting of the original user scores and the emotion analysis virtual scores of the user reviews based on the perspective pre-filtering method. The formula is used to evaluate the emotional tendency of the post:

$$score_c = \alpha\, score_r + 100(1 - \alpha)\, score_H \qquad (15)$$

The larger the value of α is, the greater the weight of the virtual score predicted by the emotion classification model in the comprehensive score. In this experiment, the value of α is set from *0* to *1.0*, and the step size is *0.1*. The experimental results obtained by our datasets are shown in Fig. 5 below.

As can be seen from the data in the above Fig. 5, when $\alpha = 0.7$, the MAE value of the dataset reaches the minimum value. In the perspective pre-filtering method, the value of α represents the weight of the virtual score in the comprehensive score. This shows that the virtual score calculated by the emotion classification model has an important influence on the accuracy of the recommended prediction scoring model. To a certain extent, it also verifies the assumption proposed in this paper that the user's review information can better reflect the user's real interest preferences. In order to reduce the estimation bias of noise data to the score prediction, the weighted synthesis of the original user score and the

user's commentary emotional score based on the perspective pre-filtering method can be used to solve the problem of large deviation between the user's original score and the real interest preference.

The influence of the number of neighboring posts K
In this paper, the collaborative training strategy is used to fuse user score data and post content description information to construct a hybrid recommendation system. In the post content recommendation model, the KNN algorithm is used to calculate the distance of the post content description, and the cosine similarity is used to measure the similarity of the post content description, so as to update or fill in user's score and default value by using the score of K nearest neighbor posts. Finally, experiments have shown that choosing the appropriate K value has an important impact on the final recommendation. In this experiment, the value of K is set from *10* to *100*, and the step size is *10*. The experimental results obtained by our datasets are shown in Fig. 6 below.

As can be seen from the data in the above Fig. 6, the MAE value of the dataset reaches the minimum value when $K = 60$. Subsequently, as the K value continues to increase, the MAE value of the model also increases, indicating that the effect of the recommended model is worse. It is concluded that the recommended effect of the model on the dataset has a greater relationship with the value of the nearest neighbor number K. However, the MAE accuracy of the recommended model is not particularly sensitive to the K value, and the relatively ideal MAE accuracy can be obtained within a certain range when the value of K is large. In this experiment, it is better to choose K in the range of [50, 70]. Therefore, $K = 60$ is selected as the parameter of DMHR algorithm when using KNN algorithm to calculate the content description of similar posts.

Iterations of emotion classification model N
In order to take account of the interaction of user ratings and review information on real emotions, this paper uses the pre-filtering method based on viewpoints and the embedding method based on user ratings to fuse user ratings and emotional prediction ratings respectively. The former uses the LSTM network to get the prediction score, and then weights the sum with the original user score. The LSTM network vector with the user rating information is combined with the method based on user score embedding, and the result is used as the input of the last layer to directly output the final comprehensive score. In order to show the performance of the emotion classification model trained by LSTM algorithm more clearly, we compared the accuracy of the model in different iterations. Set $N = \{1,10,20,30, …, 100\}$ respectively, and the 10-

Table 2 Authorities of the posts

Post	Authority
668,946,589,063,364,608	0.001194922
668,943,987,470,811,136	0.001194922
668,946,589,063,364,608	0.001194922
668,943,987,470,811,136	0.001194922
681,697,568,456,192,001	0.000636248
681,693,469,564,383,232	0.000636248
681,697,568,456,192,001	0.000636248
681,695,337,304,702,976	0.000545355
681,695,337,304,702,976	0.000545355
681,697,928,268,910,593	0.000545355

fold cross-validation method is used to evaluate the dataset. The detailed emotional classification model performance indicators are shown in Fig. 7 below.

As can be seen from the data in the above Fig. 7, in the case of the same parameter settings, the accuracy of the method based on user score embedding reaches the maximum value 92.1% when $N = 20$. With the further increase of the number of iterations, the accuracy of the model fluctuates above 90%, which indicates that the performance of the emotion classification model trained by LSTM algorithm is relatively stable. And when $N = 20$, the emotion classification model can achieve the best results to ensure the effect of subsequent experiments.

Result analysis

In this experiment, MAE is used as the evaluation index to measure the experimental effect of various recommended algorithms. On the same data set, the TDSRec algorithm, the SocialMF algorithm, the SVD++ algorithm, the HFT algorithm and the DMHR algorithm proposed in this paper are used for comparison experiments. In the DMHR algorithm, the comprehensive scoring result of the

Table 3 hubs of the users

user	hub
339,283,603	0.003429355
1,679,619,506	0.003233392
3,693,887,599	0.003135411
3,628,926,974	0.002645503
933,364,430	0.002253576
4,068,440,360	0.001959632
1,000,421,510	0.00186165
3,254,047,099	0.001665687
2,310,175,028	0.001567705
2,168,821,905	0.001567705

perspective pre-filtering method is adopted, and the parameters $\alpha = 0.7$, $K = 60$ when the best result is obtained are set. For each of the other recommended algorithms, the parameters are also set to the parameters at which the best results are obtained. The specific recommendation results are shown below (Fig. 8).

As can be seen from the data in the above figure, overall, the DMHR algorithm proposed in this paper is superior to the other four classic recommendation algorithms in MAE evaluation indicators. The SocialMF algorithm has the worst overall performance, because the algorithm only introduces the trust delivery mechanism into the recommendation model, and cannot achieve good results on the dataset of the social network. The overall performance of the TDSRec algorithm is similar to that of the SVD++ algorithm, and it is significantly improved compared with the SocialMF algorithm. This is mainly because the two algorithms add trust and trust characteristic matrix and user history data to the model respectively, which can improve the performance of the recommended model. It shows that it is feasible to use the auxiliary information such as user reviews to improve the recommendation effect. However, the review information and interest preference are not always positively correlated. The fusion of multiple recommendation views does not always improve the performance of the model. If some unreliable recommendation factors are introduced into the model, it will have a negative effect on the performance of the system.

Based on the above analysis data, the DMHR algorithm proposed in this paper has a significant improvement on the MAE evaluation index compared with the traditional algorithm, which indicates that the prediction accuracy of the recommendation model is related to the real user score and using the perspective pre-filter based method to fuse the virtual score and get the user's comprehensive score can effectively improve the user's scoring accuracy, ultimately affects the recommendation model's scoring prediction accuracy. In addition, the cold start issue is one of the most interesting issues in the recommended scenario, however few records (including ratings and reviews) are considered relevant to the "cold start". The DMHR algorithm proposed in this paper combines post-based collaborative filtering recommendation and post content-based recommendation. The recommendation factor incorporates the emotional tendency of user reviews and the semantic information of natural language description of post content. Moreover, the data preprocessing method is used to process the messy data obtained from social networks and eliminated most of the low-quality users and posts, which ensures better efficiency in subsequent processing. Theoretically, this auxiliary information will help to solve the cold start and sparse data problem to a certain extent.
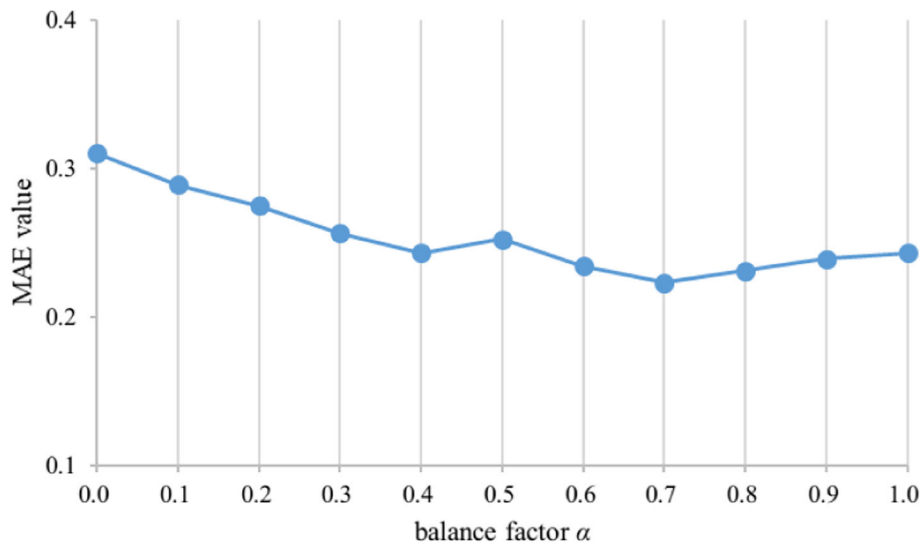
Fig. 5 Effect of balance factor α

Case study

In order to evaluate the performance of the proposed model, this paper uses the leave-one-out method which has been widely used in most literatures [38]. In this section of case study, the way of Top-N recommendation is used to verify the effectiveness of the algorithm. The experiment selects 100 posts that have not been rated by the user and are most similar to the posts that the user likes, as the candidate posts. The 100 posts have been manually sorted based on the user's attention and timeliness of the post. We take this sorting result as a real result and compare it with the recommended results obtained by the various algorithms mentioned above.

The hit rate HR (Hit radio) is used to evaluate the performance of the model.

$$HR@N = \frac{Num@N}{T@N} \qquad (16)$$

where $T@N$ represents the number of candidate test sets, $Num@N$ represents the number of TopN posts obtained by the algorithm mentioned above among the Top N recommended posts of the real result, and ranking is not considered here. The final result is shown in the figure below (Fig. 9).
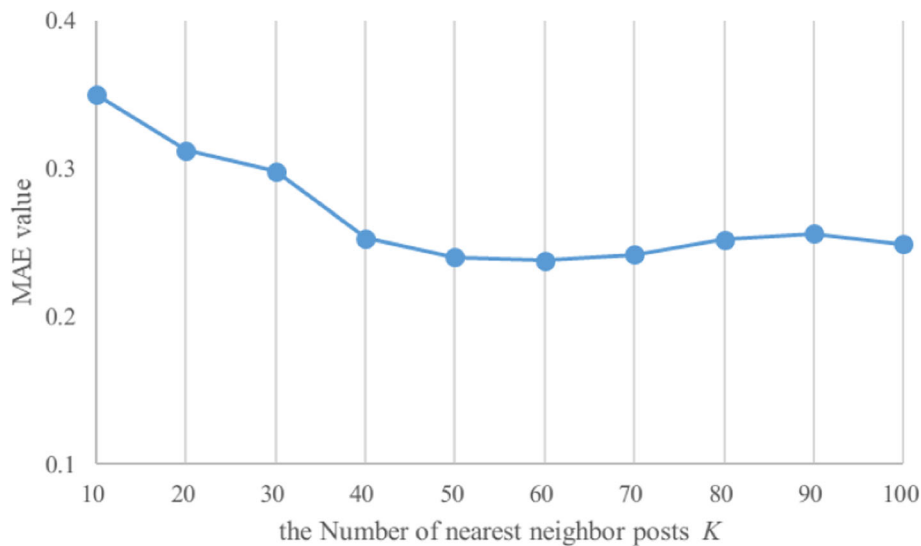

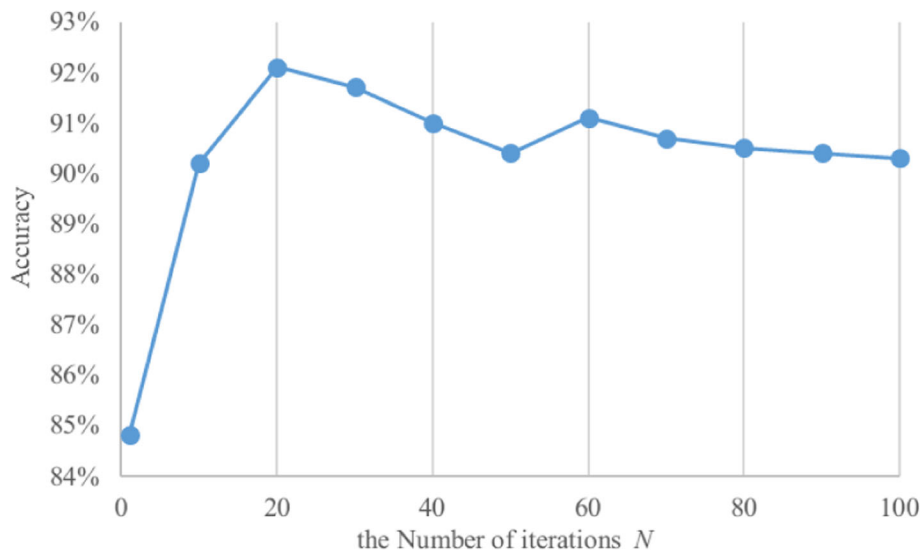**Fig. 6** The influence of the number of neighboring posts K

Fig. 7 Iterations of emotion classification model *N*

It can be seen from the experimental results that the proposed DMHR algorithm and the HFT algorithm achieve better performance than other algorithms in the *HR@N* the evaluation indexes. Since these two algorithms mine user comment information and use the description text information of the post content in the collaborative training model, they can better overcome the cold start problem of the recommendation system. Therefore, they have achieved good results in reflecting the recall rate performance of the recommendation system (*HR@N*). In contrast, other algorithms such as the TDSRec algorithm, the SocialMF algorithm and the SVD++ algorithm only use the traditional collaborative filtering algorithm based on posts and use the topic distribution information to build the model, so they can't get good recommendations. This also proves the feasibility of the idea of building a hybrid recommendation system by integrating multiple recommendation views proposed in this paper. Compared with the HFT algorithm, the DMHR algorithm uses the method of perspective pre-filtering to calculate the user's comprehensive rating of the post, and adds the time dimension to the user's review. The closer the time is, the larger the weighting factor is, the longer the time is, and the smaller the weighting factor is. The time factor of the recommended post is also considered when
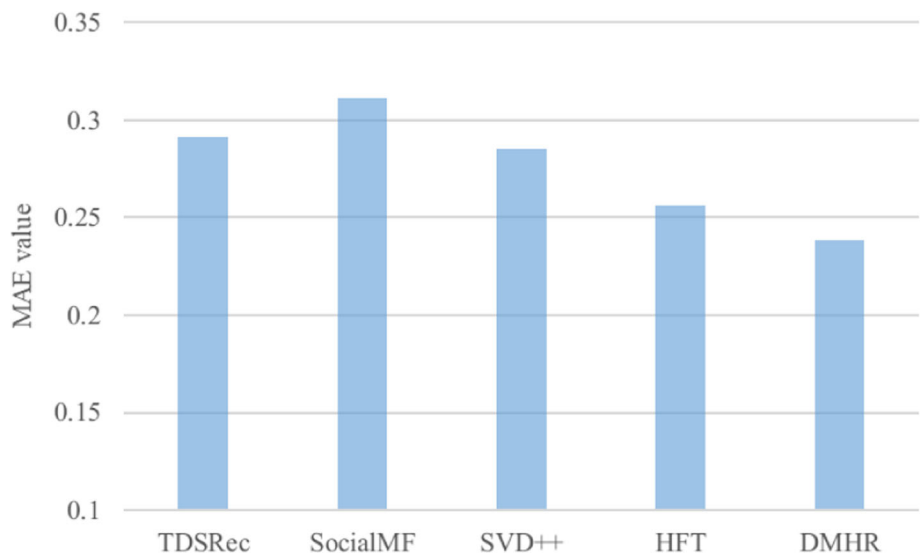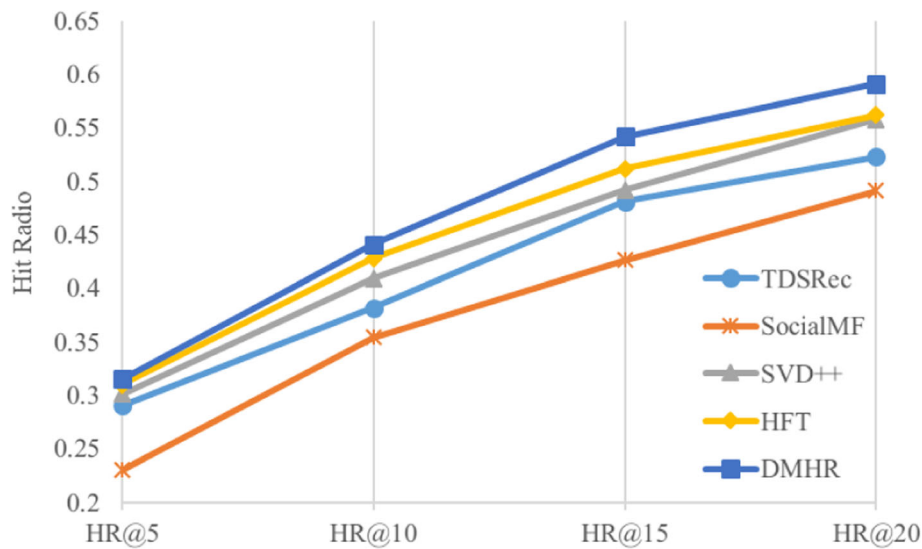

Fig. 8 The MAE value comparison

Fig. 9 Hit radio comparison in the Top-*N* recommendation

the Top-*N* recommendation ranking is performed. Experiments show that considering the time span in the recommendation process has a more important impact on the final recommendation results.

Conclusion and future work

The recommendation system is the most effective tool for solving information overload, and it has received much attention in the current academic and industrial circles. This paper proposes a hybrid recommendation model based on deep emotion analysis and multi-source view fusion. Based on the analysis of user behavior preferences, which focuses on the emotional mining and deep semantic analysis of text information; and the natural language description information of the post content is mined, and combined with the collaborative training strategy in semi-supervised learning, the post-based collaborative filtering recommendation view and the content-based recommendation view are combined to build a hybrid recommendation system. Because the method adopts the collaborative filtering model, it can effectively solve the problem that the user's original score and the real interest preference are deviated in the recommendation system, and the score distribution is extremely uneven. Since the DMHR recommendation algorithm proposed in this paper takes into account the content information of the post, which effectively solves the cold start problem of the recommendation system and improves the recommended recall rate of the recommendation system. Furthermore, in terms of the recommended effect, the experimental results show that the accuracy of the DMHR algorithm proposed in this paper has been significantly improved compared with existing methods, and the problem of cold start has also been solved to some extent.

In the next step, future research can consider the impact of user preferences over time, reviews on text emotions, weights of potential features, and social relationships on recommendations. In addition, the DMHR model can be applied to group recommendation, friend relationship recommendation and other issues in the future work.

Authors' contributions
Liang Jiang, Lu Liu, Jingjing Yao, and Leilei Shi developed the idea of the study, participated in its design and coordination and helped to draft the manuscript. Liang Jiang and Leilei Shi contributed to the acquisition and interpretation of data. Lu Liu provided critical review and substantially revised the manuscript. All authors read and approved the final manuscript.

Authors' information
LIANG JIANG received the B.S. degree from the Nanjing University of Posts and Telecommunications, China, in 2007, and the M.S. degree from Jiangsu University, Zhenjiang, China, in 2011, where he is currently pursuing the Ph.D. degree with the School of Computer Science and Telecommunication Engineering. His research interests include OSNs, computer networks, and network security. LU LIU received the M.S. degree from Brunel University and the Ph.D. degree from the University of Surrey. He is currently a Professor of Distributed Computing with the University of Leicester, U.K. His research interests are in areas of cloud computing, social computing, service-oriented computing, and peer-to-peer computing. Prof. Liu is a fellow of the British Computer Society. JINGJING YAO received the B.E. degree from Jiangsu University, Zhenjiang, China, in 2011, and the D.M. degree from Jiangsu University, Zhenjiang, China, in 2016. Her research interests include complex network, information dissemination. LEILEI SHI received the B.S. degree from Nantong University, Nantong, China, in 2012, and the M.S. degree from Jiangsu University, Zhenjiang, China, in 2015, where he is currently pursuing the Ph.D. degree with the School of Computer Science and Telecommunication Engineering. His research interests include event detection, data mining, social computing, and cloud computing.

Author details
[1]School of Computer Science and Telecommunication Engineering, Jiangsu University, Zhenjiang, China. [2]Jiangsu Key Laboratory of Security Technology for Industrial Cyberspace, Jiangsu University, Zhenjiang, China. [3]School of Informatics, University of Leicester, Leicester, UK. [4]School of Economy and Finance, Jiangsu University, Zhenjiang, China.

References
1. Daud NN, Ab Hamid SH, Saadoon M, Sahran F, Anuar NB (2020) Applications of link prediction in social networks: a review. J Netw Comput Appl 166:102716
2. Yi B et al (2019) Deep matrix factorization with implicit feedback embedding for recommendation system. IEEE Transactions Industrial Informatics 15(8):4591–4601
3. Kant S, Mahara T (2018) Nearest biclusters collaborative filtering framework with fusion. J Comput Sci 25:204–212
4. Salawu S, He Y, Lumsden J (2020) Approaches to automated detection of Cyberbullying: a survey. IEEE Trans Affect Comput 11(1):3–24
5. Wei J, He J, Chen K, Zhou Y, Tang Z (2017) Collaborative filtering and deep learning based recommendation system for cold start items. Expert Syst Appl 69:29–39
6. Nguyen V-D, Sriboonchitta S, Huynh V-N (2017) Using community preference for overcoming sparsity and cold-start problems in collaborative filtering system offering soft ratings. Electron Commer Res Appl 26:101–108
7. Shi L, Liu L, Wu Y, Jiang L, Hardy J (2017) Event detection and user interest discovering in social media data streams. IEEE Access 5:20953–20964
8. Gu K, Fan Y, Di Z (2020) How to predict recommendation lists that users do not like. Physica A: Statistical Mechanics and its Applications 537:122684
9. Yuan W, Wang H, Yu X, Liu N, Li Z (2020) Attention-based context-aware sequential recommendation model. Inf Sci 510:122–134
10. Shi L, Liu L, Wu Y, Jiang L, Panneerselvam J, Crole R (2019) A social sensing model for event detection and user influence discovering in social media data streams. IEEE Transactions on Computational Social Systems:1–10
11. Yu S, Yang M, Qu Q, Shen Y (2019) Contextual-boosted deep neural collaborative filtering model for interpretable recommendation. Expert Syst Appl 136:365–375
12. Liu H et al (2020) Hybrid neural recommendation with joint deep representation learning of ratings and reviews. Neurocomputing 374:77–85
13. Xiao H, Chen Y, Shi X, Xu G (2019) Multi-perspective neural architecture for recommendation system. Neural Netw 118:280–288
14. Rosa RL, Schwartz GM, Ruggiero WV, Rodríguez DZ (2019) A knowledge-based recommendation system that includes sentiment analysis and deep learning. IEEE Transactions on Industrial Informatics 15(4):2124–2135
15. Shi L et al (2019) Human-centric cyber social computing model for hot-event detection and propagation. IEEE Transactions on Computational Social Systems 6(5):1042–1050
16. Sanz-Cruzado J, Castells P, Macdonald C, Ounis I (2020) Effective contact recommendation in social networks by adaptation of information retrieval models. Information Processing & Management 57(5):102285
17. Almaghrabi M, Chetty G (2018) "A Deep Learning Based Collaborative Neural Network Framework for Recommender System," in 2018. International Conference on Machine Learning and Data Engineering (iCMLDE), Los Alamitos, pp 121–127
18. Chouchani N, Abed M (2020) Enhance sentiment analysis on social networks with social influence analytics. J Ambient Intell Humaniz Comput 11(1):139–149
19. Phan HT, Tran VC, Nguyen NT, Hwang D (2020) Improving the performance of sentiment analysis of tweets containing fuzzy sentiment using the feature ensemble model. IEEE Access 8:14630–14641
20. W. Chung and D. Zeng, "Dissecting emotion and user influence in social media communities: An interaction modeling approach" Information Management, 57, 1, 103108, 2020
21. Shi C, Hu B, Zhao WX, Yu PS (2019) Heterogeneous information network embedding for recommendation. IEEE Trans Knowl Data Eng 31(2):357–370
22. Deng X, Zhuang F, Zhu Z (2019) Neural variational collaborative filtering with side information for top-K recommendation. Int J Machine Learning Cybernetics 10(11):3273–3284
23. Yan Y, Tan M, Tsang I, Yang Y, Shi Q, Zhang C (2020) Fast and low memory cost matrix factorization: algorithm, analysis and case study. IEEE Trans Knowl Data Eng 32(2):288–301
24. Koren Y (2008) "Factorization meets the neighborhood: A multifaceted collaborative filtering model," Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, Las Vegas, Nevada, pp 426–434
25. Guo N, Wang B, Hou Y (2018) Collaborative filtering recommendation algorithm based on characteristics of social network. J Front Computer Science and Technology 12(2):208–217
26. Forsati R, Mahdavi M, Shamsfard M, Sarwat M (2014) Matrix factorization with explicit trust and distrust side information for improved social recommendation. ACM Trans Inf Syst 32(4):1–38
27. Feng Y, Zhou P, Wu D, Hu Y (2018) Accurate content push for content-centric social networks: a big data support online learning approach. IEEE Transactions on Emerging Topics in Computational Intelligence 2(6):426–438
28. Zhao G, Lei X, Qian X, Mei T (2019) Exploring Users' internal influence from reviews for social recommendation. IEEE Transactions on Multimedia 21(3):771–781
29. McAuley J, Leskovec J (2013) "Hidden factors and hidden topics: Understanding rating dimensions with review text," Proceedings of the 7th ACM Conference on Recommender Systems, Hong Kong, China, pp 165–172
30. Bao Y, Fang H, Zhang J (2014) TopicMF: simultaneously exploiting ratings and reviews for recommendation. Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence, Québec City, Québec, Canada:2–8
31. Ding J, Yu G, He X et al (2018) Improving Implicit Recommender Systems with View Data. Proceedings of the 27th International Joint Conference on Artificial Intelligence, Stockholm, pp 3343–3349
32. Jiang L, Shi L, Liu L, Yao J, Yuan B, Zheng Y (2019) An efficient evolutionary user interest community discovery model in dynamic social networks for internet of people. IEEE Internet Things J 6(6):9226–9236
33. Pero S, Horvath T (2013) "Opinion-Driven Matrix Factorization for Rating Prediction," Proceedings of the 21st International Conference on User Modeling, Adaptation, and Personalization Rome, Italy, pp 1–13
34. T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2, Lake Tahoe, Nevada, pp. 3111–3119, 2013
35. T. Yu, H. Hui, W. Z. Zhang, and Y. Jia, "Automatic Generation of Review Content in Specific Domain of Social Network Based on RNN," in 2018 IEEE Third International Conference on Data Science in Cyberspace (DSC), Los Alamitos, CA, USA, pp. 601–608, 2018
36. Shi M, Y T, Liu J (2019) Functional and contextual attention-based LSTM for service recommendation in Mashup creation. IEEE Transactions on Parallel and Distributed Systems 30(5):1077–1090
37. Q. Le and T. Mikolov, "Distributed Representations of Sentences and Documents," Proceedings of the 31st International Conference on International Conference on Machine Learning - Volume 32, Beijing, China, pp. 1188–1196, 2014
38. Gantner Z, Rendle s, Freudenthaler C, et al. "MyMedialite A free recommender system library," Proceedings of the 5th ACM Conference on Recommender Systems. Chicago, USA, pp. 305–308, 2011

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.