

## Research Article

# An Assistant Teaching System of English Creative Writing Based on Semantic Mining

Yu Zhang 

Xinyang Vocational and Technical College, Xinyang 464000, China

Correspondence should be addressed to Yu Zhang; zhangyu@xyvtc.edu.cn

Received 13 May 2022; Revised 6 June 2022; Accepted 21 June 2022; Published 8 July 2022

Academic Editor: Wen Zhou

Copyright © 2022 Yu Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To enhance students' English writing ability, English creative writing has become a common course offered by colleges and universities. English reference is an indispensable tool for creative English writing. How to choose appropriate English references is an important guarantee to complete good work. Therefore, this paper proposes a Deep Semantic Mining-based Recommendation (DSMR) algorithm for English writing reference selection and recommendation to assist the completion of high-quality English creative writing works. The model can extract user features and document attributes more accurately by deeply mining semantic information of literature content and user needs, so as to achieve more accurate recommendations. First, the Bidirectional Encoder Representation from Transformers (BERT) pretraining model is adopted to process literature content and user requirement documents. Through in-depth mining of user characteristics and literature attributes, the problems of data sparsity and cold start of items are effectively alleviated. Then, the forward long short-term memory (LSTM) network was used to focus on the changes in user preferences over time, resulting in more accurate recommendations. The experimental results show that the use of heterogeneous information can significantly improve the recommendation performance, and the additional use of user attribute information can also improve the recommendation performance. Compared with other benchmark models, the recommendation quality of this model is greatly improved.

## 1. Introduction

English creative writing is a common course offered by colleges and universities. It uses English as a writing method to express users' thoughts and emotions. In the process of writing, users need to read many excellent English works, so it is important to choose appropriate English references. How to meet the writing needs of different users and provide accurate and personalized references and recommendations for each user in a large number of references is the key of this paper.

The amount of data is exploding, leading to information overload and making it difficult for users to find what they are interested in. To improve user experience, the recommendation system has been applied to music, film, advertising, and other recommended scenes [1, 2]. The Collaborative filtering (CF) based recommendation system is widely used because it can effectively capture user preferences and is easy to implement in a variety of

scenarios without feature extraction in a content-based recommendation system. However, CF-based recommendations have problems of data sparsity and cold start [3]. To solve these problems, a hybrid recommendation system is put forward. It leverages multiple recommendation techniques to overcome the limitations of a single recommendation approach, exploring various types of supporting information, such as item attributes, item reviews, and the user's social network.

At first, researchers tried to use comment text in topic modeling [4], which achieved higher prediction accuracy than the model using only score data. However, this approach only focuses on thematic clues and ignores semantic content. Comments are usually expressed as word bags and context information is ignored [5], thus limiting the further improvement of prediction accuracy. In recent years, many studies have begun to combine deep learning with review text, and many excellent algorithms have been proposed. This results in a more accurate recommendation than the

topic-based modeling approach. Literature [6] connects multiple comments into a long document and uses Convolutional Neural Network (CNN) to learn useful features from the comment text. However, document-based modeling indiscriminately connects all comments to the same document without distinguishing the different importance of different comments, which is not conducive to extracting effective features [7]. Therefore, researchers began to use review-based modeling for each comment separately, and finally aggregating the features of each comment into a general feature. Literature [8] is based on review modeling and uses attention mechanism to distinguish the importance of different reviews, which achieves higher recommendation accuracy than the model based on document modeling.

To sum up, we note the limitations of much of the current work. (1) It is still a lot of models using CNN to extract the characteristics of users and items in the review, can capture the local characteristics, for a long sequence of text feature extraction, it effectively to a certain extent, limits the accuracy of recommendation. (2) Based on the comment model, much of the work does not take into account the changes in users' interests and preferences over time [9], but rather the same views from the past. (3) None of the excellent models mentioned above that use review text to improve recommendation accuracy emphasizes the use of article description documents as well as the use of review text. Item description documents contain a comprehensive introduction of item attributes, which plays an important role in alleviating the cold start of items. (4) For training data, existing methods do not consider the large difference in the number of different points. The scores of 4 and 5 are a large proportion. The training results are unfair to the data with low scores, which is easy to cause overfitting and poor robustness of the model.

For solving the problems, the article propose a recommendation model based on deep semantic mining, and design an English creative writing-assisted teaching system based on this model. This paper mainly completed the following work.

- (1) Use the pretrained BERT [10] model (Bert\_base\_uncase) to process the comment text instead of CNN. It overcomes the weakness that CNN can only extract local features, can more accurately capture the semantic meaning of words in different contexts, and measure the contribution of different comments to user characteristics. In addition, the forward long short term memory (LSTM) model was used to learn the user's interest transfer over time, which improved the recommendation accuracy. Many models have a Recurrent Neural Network (RNN) to process data. But for our model, the semantic information has been learned by BERT, and we only expect LSTM to learn the change in user interest over time. Since only existing comments can affect future comments, future comments cannot affect existing comments, and backward LSTM is not effective in the transfer of learning interest and will only increase the complexity of the model, so we do not adopt it.
- (2) Introduce English literature description documents and user demand description documents into the model. This can help us better describe the features of English literature and improve the accuracy of prediction. In addition, when new English literature lacks comments, English literature description documents can alleviate the problem of the cold start of English literature.
- (3) For the experimental data, we randomly sampled the comment data of the five score values from 1 to 5 as 1:1:1:1:1 to ensure that the data amount of each score value is equal, so as to reduce the overfitting combination and improve the robustness of the model.
- (4) Comparative experiments on datasets show that, compared with other models, our recommendation model DSMR based on deep semantic mining has high prediction scoring accuracy and significantly improved recommendation performance.

The main objective of this research paper is to assist users in creative writing in English by providing accurate and personalized reference recommendations for each user among a large number of references. To achieve that goal, this paper mainly does the following work. First, we extract the semantic information of literature content and user requirements and extract user features and literature attribute features more accurately. Second, we use BERT pre-training model to process the document content and user requirements, deep mining user features and document attributes, and solve the problem of data sparsity and item cold start. Third, the forward LSTM is used to focus on the changes of user preferences over time to make the recommendations more accurate.

This model has the following advantages.

- (1) English literature description and user demand are used as reference data for the recommendation model to improve the recommendation quality.
- (2) Equalize the comment data with different scores to improve the robustness of the model.

This paper mainly consists of five parts, including the first introduction, the second state of the art, the third methodology, the fourth experiment and analysis, and the fifth conclusion.

## 2. State of the Art

In recent years, the success of deep learning in fields has brought the recommendation community to notice this powerful tool. Scholars began to explore the use of deep learning methods to improve some of the weaknesses of the current recommendation system, such as sparse data, cold start, poor scalability, and other problems [11]. Data sparsity refers to that under the condition of huge data volume and sparse data, first of all, it is difficult to find the existence of the nearest neighbour user set, and second, the cost of computing the similarity is also high. At the same time,

information is often lost, leading to the reduction of the recommendation effect. A cold start is when a project first appears, and there is absolutely no user review of it in detail, so there is no way to predict ratings and recommendations for the project. At the same time, the accuracy of new items is poor because users have few comments when they appear. The reason for poor scalability is that as the number of users and items in the recommendation system continues to increase, the amount of computation of the collaborative filtering recommendation algorithm will also increase, leading to the gradual decline of system performance and thus affecting user experience. In particular, the emergence of CNN and RNN has achieved great success in many Natural Language Processing (NLP) tasks. Therefore, people began to try to use deep learning methods, such as DeepCoNN and D-ATTN [12], etc., to mine user preferences and characteristics of products in review texts, so as to directly apply them to predict scores. DeepCoNN consists of two parallel neural networks based on CNN, learning the implicit representation of users and objects respectively. By connecting the two parts at the top of the network to learn the interaction, the effectiveness of the comment text in alleviating the sparse problem is proved.

The key of the attention mechanism [13] is to learn a weight to mark the degree of importance, which has been widely used in natural language processing since it was proposed. The most advanced results have been achieved in machine translation, reading comprehension, speech recognition, and other fields [14]. Therefore, the attention mechanism attracted the attention of the recommendation field and began to be used in review-based recommendation algorithms [15]. Literature [16] uses attentional mechanisms to learn the usefulness of different reviews, better model users and items, predict item ratings, and generate explanations. Different from the attention mechanism at D-ATTN word level, the attention mechanism at the comment level is adopted in literature [17]. Literature [18] puts forward a new learning scheme based on Pointers, which enables users to carry out deep text interactions with objects and achieves good results.

The development of NLP has greatly promoted the application of review text in the field of recommendation. The pretraining language model [19] has developed rapidly since it was proposed, producing many excellent methods, such as features-based ELMo [20] and fine-tuning-based Open AI GPT [21]. However, these language models are one-way in nature, limiting the ability of pretrained representation. Therefore, literature [22] proposed a bidirectional pretraining model BERT, which uses an Encoder in the transformer to read the whole text at one time, so that the model can learn based on both sides of words, so as to grasp the meaning of words expressed in sentences more accurately. Therefore, BERT has a natural bidirectional and strong generalization ability, which provides a good foundation for downstream tasks.

Data preprocessing technology [23] is to process data information in advance, so as to improve the accuracy of data mining. For example, in keyword retrieval, data preprocessing can sort the information resources in the

database to improve retrieval accuracy and efficiency. The technology generally goes through data review, data screening, data sorting, etc., to achieve the effect of enhancing the efficiency of data information processing. The working principle of preprocessing technology generally includes data cleaning, integration, transformation, reduction and other technical processing to improve the accuracy of data retrieval in the later period. Data cleaning is carried out by filling missing values, identifying outliers, and correcting inconsistencies in data. Data integration needs to consider many problems, such as redundancy. The commonly used redundancy analysis methods include Pearson product distance coefficient, Chi-square test, numerical attribute covariance, and so on. Data transformation transforms data into a form suitable for learning, including data smoothing, aggregation, generalization, normalization, etc. The data reduction technique is used to obtain the reduced representation of the dataset, which greatly reduces the size of the dataset from dimension to quantity while approaching the integrity of the original data.

### 3. Methodology

*3.1. Overall Framework of Auxiliary Teaching System.* Figure 1 is the overall structure of the English creative writing assistant teaching system. The system consists of the bottom database module, the middle recommendation algorithm module, and the top user demand module. The database stores literature characteristic data and user demand data. Recommendation algorithm for English literature recommendation based on semantic mining. The user requirements module is used for front-end interaction. Its running process is shown in Figure 2.

#### 3.2. DSMR Model

*3.2.1. Model Framework.* Every User of English creative writing will browse many English references and comment on many English references, so we can use reviews as an indication of user preference. But for the user, the description of the reference is just as important. Because users only choose to browse the reference and see the comments received by the reference if they are attracted by its description. In addition, for a new English reference, there is little or no browsing and evaluation, and the reference description provides rich information on literature attributes, which helps to solve the problem of reference cold start. Many models, when using text for modeling, only make use of the comment text and do not pay attention to the reference description document. We thought this would lose some important information, so we added the reference descriptions to the model to get more accurate predictions.

DSMR uses the BERT pretraining model to process text data and distinguish the importance of different reviews, thus helping us to more accurately predict a user's rating of an English reference. The structure of the DSMR model is shown in Figure 3. The model is divided into two parallel parts. One is the user module and the other is the documentation module. In the User module, enter description

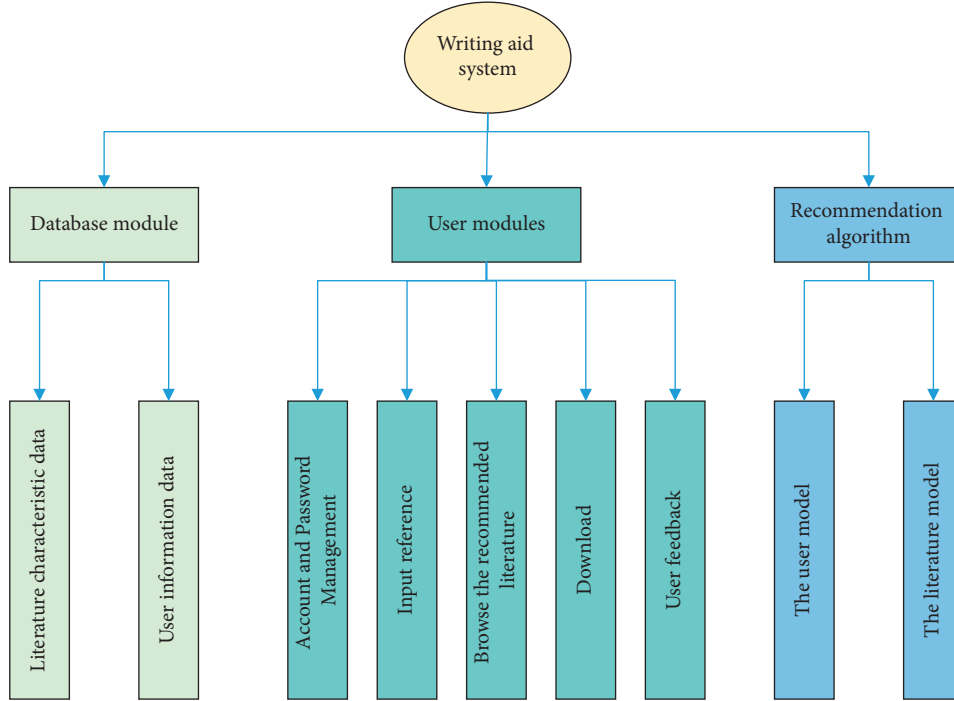


FIGURE 1: The overall structure of the English creative writing assistant teaching system.

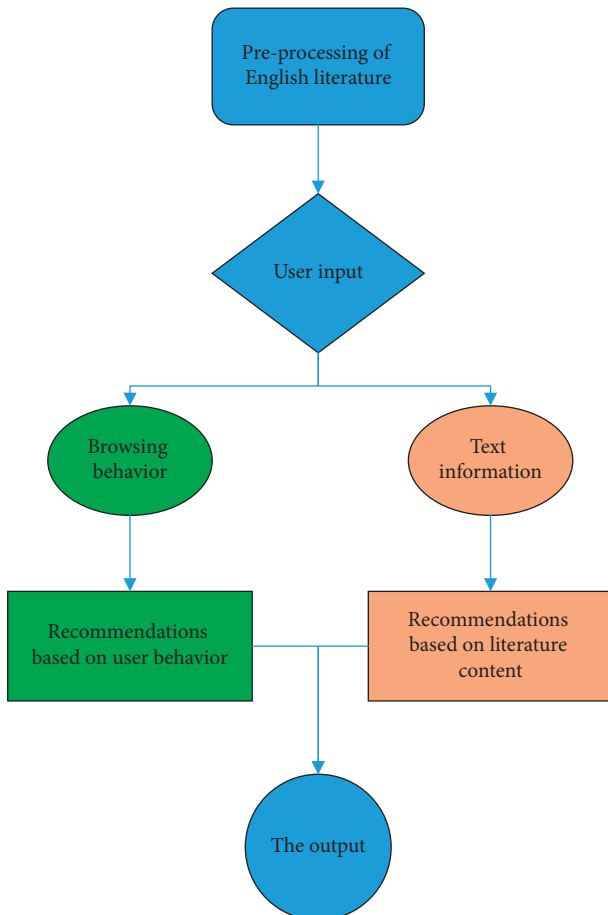


FIGURE 2: Operation process of English creative writing assistant teaching system.

documents for all references reviewed for that user and all comments received for each reference. In the literature module, enter all comments received for this reference and a description of this reference. Finally, the results of the two modules are dotted to get the user's prediction score of the reference. Since the user module is similar to the literature module in structure, this paper takes the user module as an example to introduce our model in detail.

### 3.2.2. Implementation Details

(1) *Encoder*. For a user  $u$ , all English references he has reviewed are represented by  $\hat{X}_{px}$  ( $x = 1, 2, \dots, t$ ). Pass  $\hat{X}_{px}$  into the Item\_encoder module. The specific structure of Item\_encoder is shown in the left box in Figure 4. Where  $\oplus$  means the sum. In the Item\_encoder module, the description document  $D_x$  of the document  $\hat{X}_{px}$  and all comment  $\hat{R}_{xy}$  ( $y = 1, 2, \dots, w$ ) received by document  $\hat{X}_{px}$  are passed into BERT. Our comparison model uses CNN for comment text processing and can only establish short-distance dependence on input sequences. However, self-attention in the transformer can process variable length information sequences by dynamically generating weights of different connections, which can realize parallelization and improve training speed.

The word vector representation of the reference description document  $D_x$  was obtained after BERT pretraining. BERT (Bidirectional Encoder Representations from Transformers) is a language representation model. Its main model structure is a stack of transformer's encoder. It is a 2-stage framework for pretraining, and for fine-tuning on each

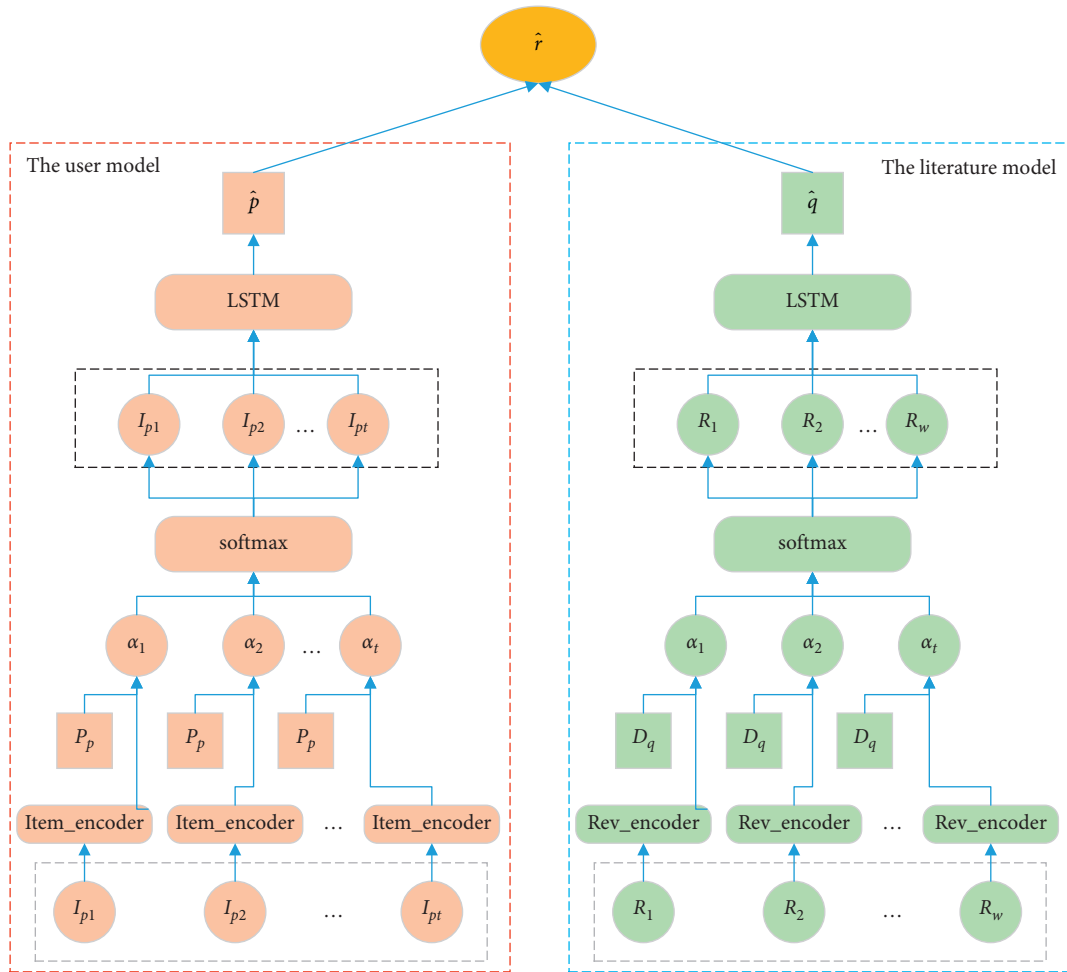


FIGURE 3: DSMR model overall framework.

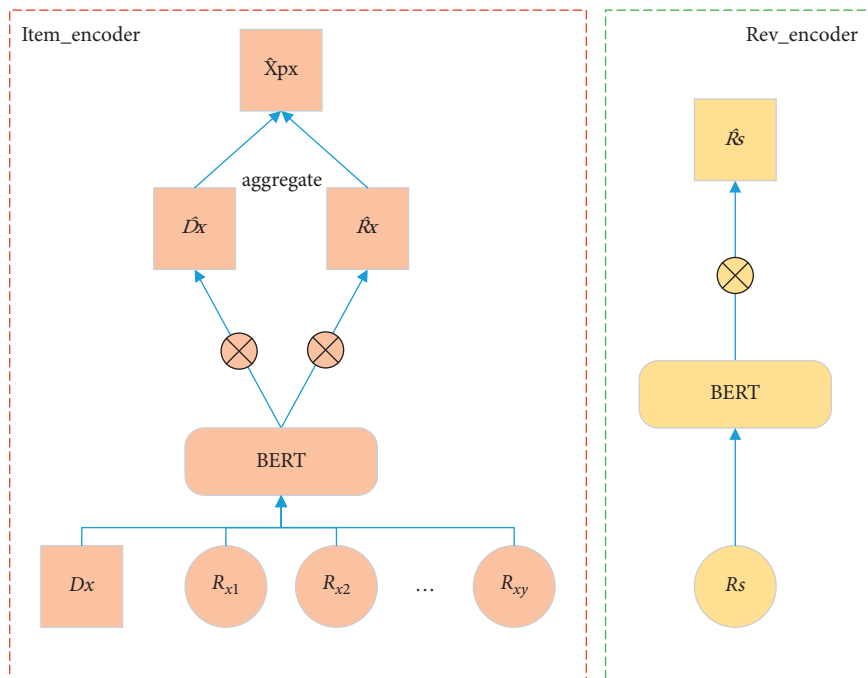


FIGURE 4: The coding section.

specific task.  $\widehat{D}_x$  is obtained by adding the vectors of each word. The implicit representation of each comment was obtained after BERT pretraining. The sum of the implicit representations yields  $\widehat{R}_x$ . The literature embedding vector  $\widehat{X}_{px}$  is obtained by combining  $\widehat{D}_x$  and  $\widehat{R}_x$ .  $\widehat{X}_{px}$  describes the characteristics of reference  $x$ . The formula is as follows:

$$\begin{aligned}\widehat{D}_x &= \text{BERT}(D_x), \\ \widehat{R}_x &= \text{sum}(\text{BERT}(R_{x1}, R_{x2}, \dots, R_{xw})), \\ \widehat{X}_{px} &= \widehat{D}_x \odot \widehat{R}_x,\end{aligned}\quad (1)$$

where,  $\odot$  represents the splicing of two vectors.

For reference  $q$ , all comments received by it are represented by  $R_s$  ( $s=1, 2, \dots, w$ ). Comments are implicitly expressed as  $\widehat{R}_s$  after passing through BERT model, as shown in rev\_encoder on the right side of Figure 4.

(2) *LSTM*. We use word embedding to represent user ID as a user embedding vector  $P_p$  ( $p=1, 2, \dots, d$ ),  $d$  is the total number of users. Mapping  $P_p$  to the same space as the literature embedding vector  $\widehat{X}_{px}$  for dot product operation, the correlation degree  $\alpha_x$  between user  $p$  and literature  $x$  features is obtained. The higher the value of  $\alpha_x$  is, the higher the correlation degree is, and the more interested users are in the literature.

$$\alpha_x = P_p \cdot \widehat{X}_{px}, x = 1, 2, \dots, t, \quad (2)$$

$\alpha_x$  ( $x=1, 2, \dots, t$ ) was normalized by Softmax, and the normalized  $\alpha_x$  was multiplied by  $\widehat{X}_{px}$  to obtain the contribution degree of each literature to user characteristics. Finally,  $\widehat{X}_{px}$  is sent into LSTM to learn user interest transfer over time, and the output vector  $\widehat{p}$  of the user model is obtained. Softmax is a normalized exponential function. It is an extended application of the binary classification function sigmoid to multi-classification, to present the results of multi-classification in the form of probabilities.

$$\widehat{p} = \text{LSTM}(\text{softmax}(\alpha_x) \cdot \widehat{X}_{px}), x = 1, 2, \dots, t. \quad (3)$$

Similarly, we represent the description document of literature  $q$  as  $Dq$ , and map  $Dq$  to the same space with the review embedding vector  $\widehat{R}_s$  to obtain the output vector  $\widehat{q}$  of the literature model.

(3) *Score Prediction*. The final prediction score  $\widehat{r}$  is obtained by the dot product of the user model's output vector  $\widehat{p}$  with the literature model's output vector  $\widehat{q}$ .

$$\widehat{r} = \widehat{p} \cdot \widehat{q}. \quad (4)$$

(4) *Model Learning*. The goal of DSMR model is actually to improve the accuracy of score prediction, which is equivalent to a regression problem. For regression problems, the most commonly used objective function is the squared loss function. In the training set sample  $W$ , the predicted score of user  $p$  for reference  $x$  is  $\widehat{r}_{px}$  and the real score is  $r_{px}$ , so the objective function can be expressed as:

$$L = \sum_{p,x \in W} (\widehat{r}_{px} - r_{px})^2. \quad (5)$$

Our task is to minimize the target function. We choose Adam optimization algorithm to optimize the objective function because Adam uses momentum and adaptive learning rate to accelerate the convergence speed, which is suitable for problems with large data volume and only requires a small amount of memory.

## 4. Result Analysis and Discussion

4.1. *The Data Set*. To better evaluate the model proposed in this paper, a dataset containing interactive and user-assisted information is necessary. This paper collects 600 English novels from the website as an experimental dataset, which includes 10,020 anonymous ratings (value range from 1 to 5) generated by 6,020 users on about 600 English novels. The user attribute information includes age level, gender, grade and major.

In the process of processing the dataset, we consider that although there are 5 scores ranging from 1 to 5, 5 and 4 still account for the majority of the scores. Almost all the models that have been proposed do not take this situation into account. We don't think this is fair to a score of 1 or 2. It overfits the training results. In this way, the data of each score value in the dataset are of equal amount, and the results are more objective and the model is more robust. The extraction results are shown in Table 1.

4.2. *Comparison Model*. The proposed DMSR is compared with the recommended model in the following literature, and the hyper-parameter settings of the comparison model are the same as those in the original text, except as specifically indicated. Table 2 shows the comparison algorithm.

4.3. *Experimental Settings*. The BERT pretraining model used by us has an initial learning rate of 0.01 for Bert\_base\_uncase, Review-DSMR, and DSMR models trained by Google, and then the NoamOpt optimizer is used for dynamic adjustment. The loss rate is set to [0.05, 0.1, 0.3, 0.5], the batch size is set to [3, 5, 8, 16, 32], and the number of potential factors is set to [32, 64, 128, 256].

The ratio of the training set, validation set, and test set was 3 : 1 : 1. Each experiment was repeated 3 times and the average performance value was taken.

Models were evaluated in two experimental scenarios: (1) in the click-through rate (CTR) prediction, trained models were used to predict each interaction in the test set, followed by Precision, Recall, and F1 to evaluate the CTR prediction. (2) in Top@K recommendation, the trained model is used to select  $K$  items with the highest predicted click probability for each user in the test set, and then select Precision@K, Recall@K and F1@K to evaluate the recommendation set, where  $K=1, 2, 5, 10, 20, 50, 100$ .

4.4. *Evaluation Indicators*. In CTR prediction, AUC, Precision and compromise accuracy, and recall rate score (F1)

TABLE 1: Dataset preprocessing results.

Score	Comment number	Number of English novels	The number of users
1	5000	40	200
2	5000	40	200
3	5000	40	200
4	5000	40	200
5	5000	40	200

TABLE 2: The comparison algorithms.

Number	Algorithm	Theory	Operation principle
1	Literature [24]	Matrix decomposition	The rating matrix is used as input, the inner product of the low-rank matrix of users and items is used to represent the rating, and the objective function is minimized by the alternate least squares (ALS) technique
2	Literature [25]	Probability matrix decomposition	Only score data is used for collaborative filtering, and Gaussian distribution is introduced to model the underlying factors of users and items
3	Literature [26]	Parallel neural network	An additional sharing layer on top of the two neural networks connects the two parallel networks so that the learned user and project potential factors can interactively predict ratings. This model proves that the sparsity problem can be effectively alleviated by using comment text

are used as evaluation indexes. In Top@K recommendation, Precision, Recall and compromised accuracy and Recall score (F1) are used as evaluation indexes. AUC considers the sorting quality of samples, which is closely related to the sorting error. Their calculation formula is as follows:

$$AUC = \frac{\sum_{x \in U} \text{rank}_x - (|U| * (|U| + 1) / 2)}{|U| * |T|},$$

$$\text{Precision} = \frac{\sum_{p \in P} |R(p) \cap N(p)|}{\sum_{p \in P} |R(p)|},$$

$$\text{Recall} = \frac{\sum_{p \in P} |R(p) \cap N(p)|}{\sum_{p \in P} |N(p)|},$$

$$F1 = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}.$$
(6)

where,  $R(p)$  is the recommendation list made to users based on their behaviours in the training set, and  $N(p)$  is the behaviour list of users in the test set. Rank is the sample sort position, and it starts at 1.  $|U|$  is positive sample number,  $|T|$  is negative sample number. When  $|U| > |T|$ ,  $1 \geq AUC > 0.5$ .  $|U| = |T|$ ,  $AUC = 0.5$ .  $|U| < |T|$ ,  $0 \leq AUC < 0.5$ .

**4.5. Results and Analysis.** This section describes the comparison results between the different models and the proposed model (DMSR) in this paper. Table 3 and Figure 5 show the prediction results of CTR of the model.

As can be seen from Table 3 and Figure 5, the proposed model has achieved good performance in all indicators of CTR. Compared with the best-performing model literature [25], the proposed model improved by 1.52% on AUC, 1.82% on Precision, and 1.88% on F1. Compared with the reference [24], the proposed model improves 2.41% in AUC, 2.84% in Precision and 2.9% in F1. Compared with reference [26], the proposed model improves 1.84% in AUC, 2.23% in

TABLE 3: AUC, precision and F1 results predicted by CTR.

Models	AUC	Precision	F1
Literature [24]	0.8926	0.8137	0.8161
Literature [25]	0.9015	0.8239	0.8263
Literature [26]	0.8983	0.8198	0.8218
Proposed (DMSR)	0.9167	0.8421	0.8451

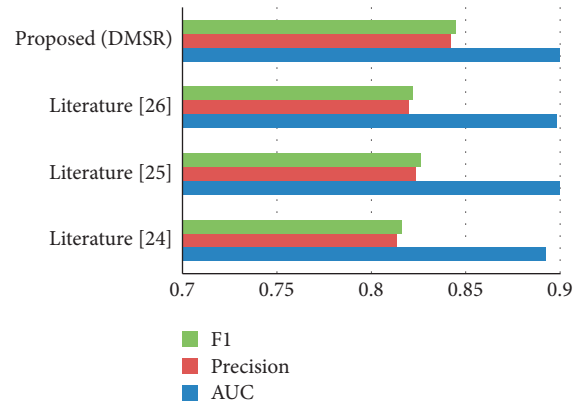


FIGURE 5: Results of AUC, precision and F1.

Precision, and 2.33% in F1. The model in this paper achieves better recommendation results mainly for the following reasons.

- (1) The model proposed in this paper adopts BERT pretraining model to process the document content and user requirement documents. It digs into user characteristics and document attributes, effectively alleviates the problem of sparse data, and improves the accuracy of recommendations.

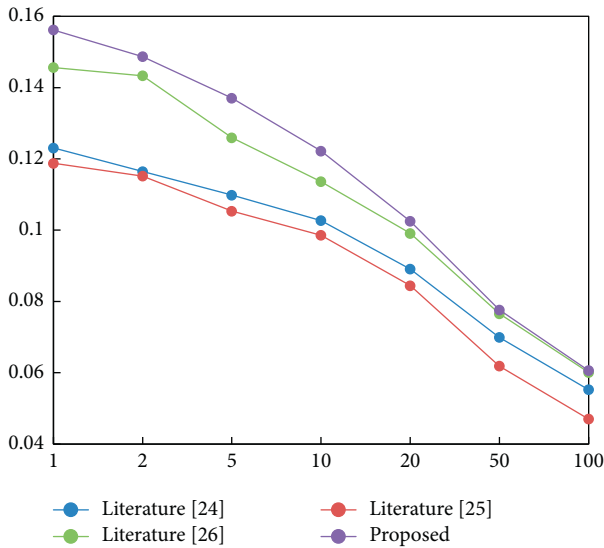


FIGURE 6: Accuracy in Top@K.

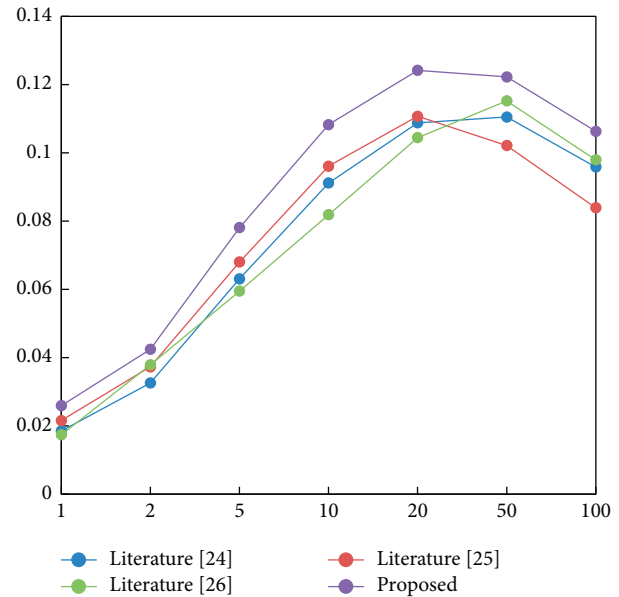


FIGURE 8: F1 in Top@K.

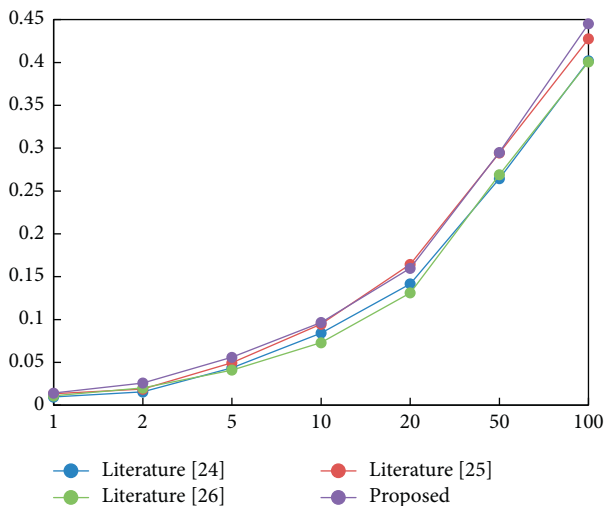


FIGURE 7: Recall in Top@K.

- (2) The model uses forward LSTM to pay attention to the changes in user preferences over time, thus generating more accurate recommendations.

Figures 6–8 show the accuracy rate, recall rate, and F1 line graph recommended by Top@K respectively. As you can see from these figures, the proposed model also achieves good performance on Precision@K, Recall@K, and F1@K. When  $K = 10$ , reference [26] showed the best performance in accuracy, and the proposed model improved by 7.09% compared with it. Literature [25] has the best performance in recall rate and F1. The proposed model has a 2.68% improvement in recall rate and a 12.88% improvement in F1. Compared with the reference [24], the proposed model improved 18.63% in accuracy, 0.69% in recall rate, and 12.52% in F1.

In CTR prediction and top-K recommendation, the proposed model achieved good performance in all indicators. In the comparison model, reference [24] performed well in CTR recommendation scenarios, but not in top-K recommendation scenarios. The proposed model does well in both recommended scenarios. This is because reference [24], reference [25], reference [26], and the proposed model all combine rich heterogeneous information in English literature. However, reference [24], reference [25] and reference [26] only integrate the auxiliary information and its relationship to the content of the literature. Only the proposed model combines their respective auxiliary information and its relationship at both the user end and the literature end. There are two potential reasons for the performance improvement of the proposed model in this paper: the first reason is the use of more information sources. The second reason is the use of structured descriptions to model heterogeneous information. The proposed model shows that it is feasible to fuse auxiliary information of users on the user side and auxiliary information of documents on the document side. The use of heterogeneous information can significantly improve the recommendation performance, and the additional use of user attribute information can also improve the recommendation performance.

## 5. Conclusion

There are two problems in the traditional collaborative filtering recommendation algorithm: cold start and personalized recommendation. This paper conducts in-depth research on these two problems and obtains the optimized experimental results after the improved scheme through relevant experiments. Based on the literature content and user needs, this paper proposes a deep semantic mining recommendation model DSMR, which can predict the score



more accurately. It uses BERT pretraining model to learn the accurate semantics of words in context information. LSTM is also used to learn the internal relationship between contents, explore the change of user preferences over time, and introduce literature content documents to alleviate the problem of cold start. Experimental results show that the improved algorithm in this paper can optimize the cold start and personalized recommendation problems of the recommendation algorithm, so as to improve the data processing ability in the big data environment and give users a better user experience.

For future studies, we will focus on the recommended interpretability issues. The interpretability of recommendation is also an important aspect to improve the effect of recommendation. Persuasive and appropriate reasons for recommendation will improve the trust of users. How to use review text to generate recommendation reasons is a promising research direction in the future.

## Data Availability

The labeled dataset used to support the findings of this study are available from the author upon request.

## Conflicts of Interest

The author declares no conflicts of interest.

## References

- [1] M. Aivazoglou, A. O. Roussos, D. Margaritis, C. Vassilakis, S. Loannidis, and J. Polakis, "A fine-grained social network recommender system," *Social Network Analysis and Mining*, vol. 10, no. 1, pp. 1–18, 2020.
- [2] H. Tahmasebi, R. Ravanmehr, and R. Mohamadrezai, "Social movie recommender system based on deep autoencoder network using Twitter data," *Neural Computing & Applications*, vol. 33, no. 5, pp. 1607–1623, 2021.
- [3] T. Anwar, V. Uma, M. Hussain, and M. Pantula, "Collaborative filtering and kNN based recommendation to overcome cold start and sparsity issues: a comparative analysis," *Multimedia Tools and Applications*, pp. 1–19, 2022.
- [4] T. R. Hannigan, R. F. J. Haans, K. Vakili et al., "Topic modeling in management research: rendering new theory from textual data," *The Academy of Management Annals*, vol. 13, no. 2, pp. 586–632, 2019.
- [5] L. Yue, W. Chen, X. Li, W. Zuo, and M. Yin, "A survey of sentiment analysis in social media," *Knowledge and Information Systems*, vol. 60, no. 2, pp. 617–663, 2019.
- [6] U. D. Gandhi, P. Malarvizhi Kumar, G. Chandra Babu, and G. Karthick, "Sentiment analysis on twitter data by using convolutional neural network (CNN) and long short term memory (LSTM)," *Wireless Personal Communications*, pp. 1–10, 2021.
- [7] J. Devezas and S. Nunes, "A review of graph-based models for entity-oriented search," *SN Computer Science*, vol. 2, no. 6, pp. 1–36, 2021.
- [8] S. Xing, Q. Liu, X. Wang, X. Zhao, and T. Li, "A hierarchical attention model for rating prediction by leveraging user and product reviews," *Neurocomputing*, vol. 332, pp. 417–427, 2019.
- [9] C. Gao, W. Lei, X. He, M. de Rijke, and T.-S. Chua, "Advances and challenges in conversational recommender systems: a survey," *AI Open*, vol. 2, pp. 100–126, 2021.
- [10] J. Carrión and V. Serrano, "Revisión sistemática de literatura: características y funcionamiento respecto a los modelos BERT y SQuAD," *CEDAMAZ*, vol. 11, no. 1, pp. 79–86, 2021.
- [11] A. Da'u and N. Salim, "Recommendation system based on deep learning methods: a systematic review and new directions," *Artificial Intelligence Review*, vol. 53, no. 4, pp. 2709–2748, 2020.
- [12] Z. Y. Khan, Z. Niu, and S. Sandiwarno, "Deep learning techniques for rating prediction: a survey of the state-of-the-art," *Artificial Intelligence Review*, vol. 54, no. 1, pp. 95–135, 2021.
- [13] Z. Niu, G. Zhong, and H. Yu, "A review on the attention mechanism of deep learning," *Neurocomputing*, vol. 452, pp. 48–62, 2021.
- [14] L. N. Vieira, M. O'Hagan, and C. O'Sullivan, "Understanding the societal impacts of machine translation: a critical review of the literature on medical and legal use cases," *Information, Communication & Society*, vol. 24, no. 11, pp. 1515–1532, 2021.
- [15] Y. Du, L. Wang, Z. Peng, and W. Guo, "Review-based hierarchical attention cooperative neural networks for recommendation," *Neurocomputing*, vol. 447, pp. 38–47, 2021.
- [16] H. Liu, Y. Wang, Q. Peng et al., "Hybrid neural recommendation with joint deep representation learning of ratings and reviews," *Neurocomputing*, vol. 374, pp. 77–85, 2020.
- [17] A. Da'u and N. Salim, "Sentiment-aware deep recommender system with neural attention networks," *IEEE Access*, vol. 7, pp. 45472–45484, 2019.
- [18] J. Li, A. Sun, J. Han, and C. Li, "A survey on deep learning for named entity recognition," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 1, pp. 50–70, 2020.
- [19] Y. Cui, W. Che, T. Liu, B. Qin, and Z. Yang, "Pre-training with whole word masking for Chinese BERT," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 3504–3514, 2021.
- [20] Y. Qiao, X. Zhu, and H. Gong, "BERT-Kcr: prediction of lysine crotonylation sites by a transfer learning method with pre-trained BERT models," *Bioinformatics*, vol. 38, no. 3, pp. 648–654, 2022.
- [21] M. Zhang and J. Li, "A commentary of GPT-3 in MIT technology review 2021," *Fundamental Research*, vol. 1, no. 6, pp. 831–833, 2021.
- [22] J. Hou, X. Li, H. Yao, H. Sun, T. Mai, and R. Zhu, "BERT-based Chinese relation extraction for public security," *IEEE Access*, vol. 8, pp. 132367–132375, 2020.
- [23] S. Rajeswari and K. Suthendran, "C5.0: advanced Decision Tree (ADT) classification model for agricultural data analysis on cloud," *Computers and Electronics in Agriculture*, vol. 156, pp. 530–539, 2019.
- [24] N. Singh, L. Ma, H. Yang, and E. Solomonik, "Comparison of accuracy and scalability of gauss–Newton and alternating least squares for CANDECOM/PARAFAC decomposition," *SIAM Journal on Scientific Computing*, vol. 43, no. 4, pp. C290–C311, 2021.
- [25] C. Feng, J. Liang, P. Song, and Z. Wang, "A fusion collaborative filtering method for sparse data in recommender systems," *Information Sciences*, vol. 521, pp. 365–379, 2020.
- [26] C. E. Lawson, J. M. Martí, T. Radivojevic et al., "Machine learning for metabolic engineering: a review," *Metabolic Engineering*, vol. 63, pp. 34–60, 2021.