Research on Personalized News Recommendation Model Based on Knowledge Graph

Suicheng Li[a], Xiaotian Wang[b]*, Shaoyang Zhang[c], Jiabin Wei[d], Yanying Shang[e]

School of Economics and Management, Xi'an University of Technology, Xi'an, China

[a]lisc@xaut.edu.cn, [b*]496996232@qq.com, [c]2275587224@qq.com, [d]weijiabin95@163.com, [e]505276306@qq.com,

Abstract. By constructing a personalized news recommendation model based on a knowledge graph, the problem of ignoring knowledge correlation in multi-class text information methods has been effectively solved. The study analyzed the characteristics of news texts, treating titles and abstracts as key information components, and using word vectors and knowledge graph feature learning methods for vectorization processing. In terms of model construction, sub-modules such as multi-perspective news features, user click behaviorand click probability prediction were introduced. Through techniques such as convolutional neural networks and attention mechanisms, multi-dimensional information of news was comprehensively considered, achieving more accurate personalized recommendations. The experimental results show that the knowledge graph-based model performs well in AUC, MRR nDCG@5 nDCG@10,Compared to the benchmark model, the evaluation indicators have increased by 3%, 2.6%, 2.5%, and 3%. The ablation experiment has demonstrated the positive effects of knowledge graphs, multi-perspective news features, and attention mechanisms. This study provides new ideas for knowledge graph-based news recommendations, which helps to improve user satisfaction and strengthen the platform's service capabilities.

Keywords: News recommendations; Knowledge graph; Deep learning; Theme model

1 Introduction

With the development of the Internet and media technology, various online platforms have become common communication media and the main channels for people to obtain news. Compared with content recommendations for products, movies, books, and other types of content, online news has the characteristics of fast dissemination and update speed, diverse user interests, etc. On the one hand, users' click behavior is influenced by public opinion and news popularity, while popular and hot news does not have differentiation; On the other hand, the large amount of news content generated every day makes it difficult for users to quickly discover interesting news from the platform. As a tool to help users filter content, in recent years, search engines have

C. Chen et al. (eds.), *Proceedings of the 3rd International Conference on Digital Economy and Computer Application (DECA 2023)*, Atlantis Highlights in Computer Sciences 17,
https://doi.org/10.2991/978-94-6463-304-7_35

gradually improved from being able to passively return relevant pages based on user input terms [1] to being able to actively provide users with direct answers and recommend relevant terms and text information [2]. Therefore, news recommendation (NR) is an effective way to help users filter and filter news content.

In summary, recommendation systems have made some progress and applications in the field of news, but there is still a need to further improve recommendation performance. The main entry points include better news modeling and fully integrating user interests. In news recommendation problems, there is a large amount of network structure data, such as user-level social networks [3], item-level knowledge graphs [4], heterogeneous information networks generated by user news interactions [5]. These data can serve as auxiliary information for news recommendations to enrich news and user representation. Some studies have verified that using graph-type data as auxiliary information can improve the performance of news recommendation systems. Therefore, based on previous research, this type of auxiliary information is combined with news text to introduce a news recommendation system. The learned features are applied to news recommendation problems to improve the accuracy of the recommendation system. This can enhance the ability of news content to provide platform services while enhancing user experience and engagement.

2 Model building

2.1 Problem Description

The main characteristics of news are fast updates, fixed structure, concise sentences, and numerous topics. For users, they may be interested in multiple types of news. Therefore, how to recommend corresponding news to users within the life cycle of news is a key and difficult point in the field of news recommendation. As an important work in the fields of natural language processing and data mining, personalized news recommendation is widely applied to various online platforms. However, due to the characteristics of the news itself, there are still some challenges. However, traditional semantic models cannot discover the hidden knowledge level connections between new news, which limits the recommended news for users to certain topics or themes, Knowledge graph can introduce more potential semantic information;

This article defines the problem as news on a certain online platform, it mainly includes the main parts such as title, abstract, category, subcategory, main text, etc. Each word in the news title and abstract may be associated with a certain entity in the knowledge graph and the relationship between the given entities; For 'user u', given their click history for a certain period, through hierarchical modeling of news text content and user interest modeling, the click probability of each news item in the candidate news set is predicted, and a personalized recommendation list is generated. The mathematical symbols and their definitions used in this chapter are shown in Table 1.

Table 1. Definition of main mathematical notation

Symbol	Define	Symbol	Define
U	List of recommended users	α_t、α_a、α_c、α_{sc}、α_b	Attention weights for news headlines, abstracts, categories, subcategories, and main text
N	User's candidate news set	r_n	News feature representation vector
$\mathrm{e_c}$、$\mathrm{e_{sc}}$	Initialization vectors for news categories and subcategories	u_u	User u's click behavior representation vector
θ_d	Potential Theme Distribution in News Text	y	Click probability
r_t、r_a、r_c、r_{sc}、r_b	Feature vectors for news headlines, abstracts, categories, subcategories, and main text		

2.2 Modeling framework

Based on the above analysis, this article models news and users separately. When learning new feature representation and user feature representation, entity auxiliary information in the knowledge graph is combined with news text, and an attention mechanism is used to give different weights to different features of news and user click history, to better learn news and user representation, and further improve the performance of personalized news recommendation models. And to some extent, it alleviates the problem of cold start of news information and sparse interactive records in news recommendations. The model proposed in this article consists of three main sub-modules: multi-perspective news feature representation, user click behavior representation based on attention mechanism and click probability prediction. The specific structure is shown in Figure 1.

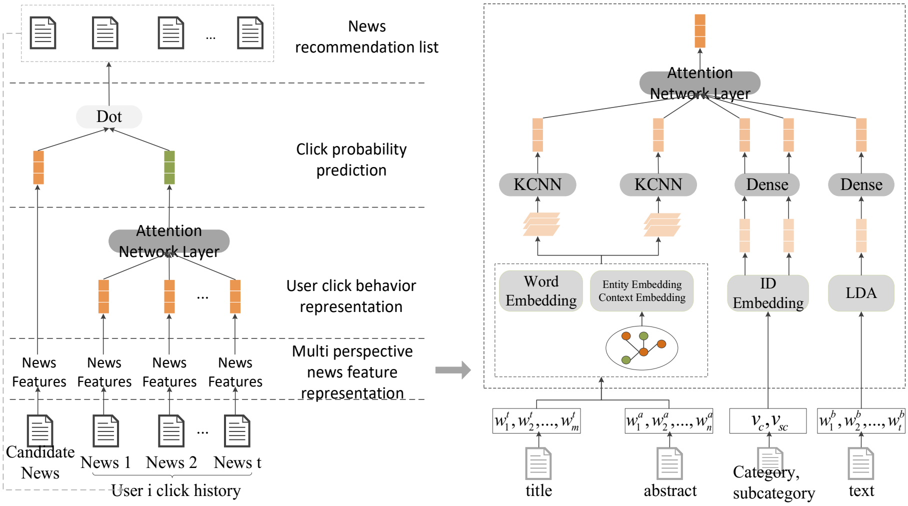

Fig. 1. The Framework of Personalized News Recommendation Model Based on Knowledge Graph

The main function of the multi-perspective news feature representation submodule is to extract title and abstract features based on the word vectors and knowledge graph entity vectors of news titles and abstracts. Then, news features are extracted from the perspectives of categories, subcategories, and text, respectively. Through the attention mechanism at the news level, multi-perspective features are fused to obtain news vector representation based on knowledge graph and text features; The user click behavior representation sub-module is mainly responsible for learning user behavior representation from the news that the user has clicked on. Considering that the historical news that the user clicks on has different levels of interest information, the user-level attention mechanism is used to assign weights to the news that the user reads, thereby determining the user's preference for each news item; The main function of the click probability prediction sub-module is to use the news text features and the user clicks behavior features generated by the first two modules to predict the user's click probability on candidate news. The higher the click probability, the higher the probability of user interest in the candidate news, and then generate the user's TOP-N news recommendation list.

3 Recommendation Algorithm Process and Model Training

3.1 Recommended algorithm process

The specific description of the recommendation algorithm process proposed in this article is as follows:

The first step is to input the different text components of candidate news N and obtain the multi-perspective feature representation vector r_n of the candidate news; Step

2, based on the user click behavior representation process, the user-level attention mechanism is fused with the click history of user U to obtain the user click behavior representation vector u_u; The third step is to use the candidate news feature representation vector and the user clicks behavior representation vector generated in the first two steps, calculate the user's click probability on the candidate news through formula (1), arrange it in descending order according to the size of the y value, and select TOP-N news to generate the user's personalized recommendation list.

$$y = \mathrm{r}_n \mathrm{u}_u \tag{1}$$

3.2 Model training

For the historical click data of users, this article refers to every news provided by the website and clicked by users as a positive sample and news that appears in the same session but is not clicked by users as a negative sample. Due to the highly imbalanced proportion of positive and negative samples in news recommendations, and the higher efficiency of model training through negative sampling technology, this article uses negative sampling to train the model. For each positive sample of the user, randomly sample r negative samples to jointly predict the click probability score of the user on the positive and negative samples, where $\hat{y}^+$ represents the click probability score of the positive sample, $[\hat{y}_1^-, \hat{y}_2^-, \ldots, \hat{y}_r^-]$ and A represents the click probability score of r the negative samples. Firstly, the click probability score of each positive sample is standardized using the $soft\max$ function and then trained using the cross entropy loss function. The calculation process is shown in formulas 2 and 3， $\hat{y}_{i,j}^-$ representing the negative sample corresponding to a positive sample i; P represents the set of all positive sample news.

$$p_i = \frac{\exp(\hat{y}_i^+)}{\exp(\hat{y}_i^+) + \sum_{j=1}^{r} \exp(\hat{y}_{i,j}^-)} \tag{2}$$

$$l = -\sum_{i \in P} \log(p_i) \tag{3}$$

3.3 News recommendation strategy

In the actual business scenarios of online news platforms, the news is usually sorted based on news popularity on a certain interface, and the same number of recommendation lists are generated for all users, or news similar to the current news content is generated on the content interface of a certain news article. In addition, when users refresh their homepage, they will push N news items that they may be interested in based on their characteristics. This article mainly focuses on personalized news recommendations for users in such scenarios, as shown in Figure 2.

For this scenario, the user is mainly generated with a personalized TOP-N news recommendation list that may match their interests based on the size of the click probability value. After representing the features of news from multiple perspectives and user click behavior, the click probability of each candidate news that has not been clicked is predicted for each user to be recommended. The sorted top N news items are sorted in descending order based on the probability value and are used as a push list.

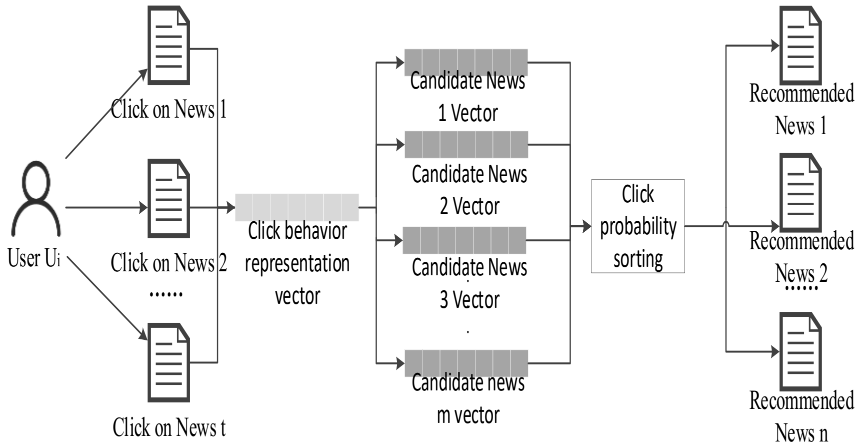

Fig. 2. TOP-N recommendation list generation

4 Case study

This article will construct experiments based on publicly available news datasets to verify the effectiveness of the recommendation model proposed above. Firstly, introduce the experimental environment and data, and conduct statistical analysis on the dataset. Secondly, crawlers are used to obtain the entities and relationships of the knowledge graph subgraph, generate preliminary vector representations of news titles and abstracts, and then extract news and user features using the model from the previous chapter. Recommendation results are generated through click probability prediction, and a certain user has randomly selected for example analysis. Finally, a control experiment was constructed to verify the overall performance of the model. To verify the effectiveness of the knowledge graph, multi-perspective features, and attention mechanism introduction, three sets of ablation experiments were constructed. The experimental results were analyzed based on the evaluation index values of the model to verify the rationality of the model in this paper。

4.1 Parameter Setting and Model Example Analysis

4.1.1 Parameter settings

The word vector, entity vector, news category, and subcategory vector dimensions of news headlines and news abstracts are set to 300 dimensions, 100 dimensions, and

100 dimensions respectively; Set the number of words in the news title and news abstract to 20 and 50 respectively; Set the number of convolutional kernels to 300 and the window size to 2 or 3; Set the number of potential topics in the news body to 100; The user's historical click sequence length is 50; The negative sampling rate is 2; Optimize the model through Adam; To alleviate the overfitting problem of the model, adopt a dropout strategy and set the dropout rate to 0.2.

4.1.2 Case study

To intuitively illustrate the actual effectiveness of the recommended model in this article, a random sample of one user (denoted as u) is selected as the target user, and an example analysis is conducted based on their recommendation results.This article assumes that for a certain user if they click on a news item, they are considered interested in the news. The model first generates feature vectors for each news item through training on the text content of each part of the news and then generates a click behavior representation vector based on the user's click history training. Based on this, the target user's click probability for each candidate news item is calculated. The higher the click probability, the higher the user's liking for the candidate news, and the more likely they are to click on the news. The ranking position should be higher. The news recommendation list generated by the model for the target user u and the actual user clicks are shown in Table 2. In the historical click news records of user u and the corresponding categories and entities of the news, the user's click history clearly shows their points of interest. The model captures this internal correlation, Therefore, predicting users' interest in news containing entities such as "South Carolina" and "Costco" to some extent reflects the rationality of recommendation results.

Table 2. Recommended list for user u

News Headlines	category	entity	Click probability	Click on the label
This was uglier than a brawl. And Myles Garrett deserves suspension for the rest of the year after the helmet attack.	sports	Myles Garrett	0.1777	1
South Carolina teen gets life in prison for deadly elementary school shooting	News	South Carolina、Townville Elementary School shooting	0.1771	0
High tides surge through Venice, locals rush to protect art	weather	Venice、Piazza San Marco	0.1570	0
30 Best Black Friday Deals from Costco	lifestyle	Costco、Black Friday	0.1490	0
Queen Elizabeth Just Rode Horseback at Age 93	lifestyle	Elizabeth II	0.1187	0
Opinion: Colin Kaepernick is about to get what he deserves:	sports	Colin Kaepernick	0.1182	0

a chance				
How much turkey do you need to buy per person?	food and drink	[]	0.1023	0

4.2 Experimental results and analysis

This section mainly analyzes the experimental results, first providing the evaluation indicators used and defining the indicators. Second, to demonstrate the effectiveness of the model, the optimal results of the model and the benchmark model were compared on the same dataset to verify overall performance. Finally, to demonstrate the effectiveness of introducing knowledge graphs, multi-perspective features, and attention mechanisms, ablation experiments were designed for model self-comparison to verify the rationality of the corresponding modules.

4.2.1 Selection of evaluation indicators

After obtaining the experimental results, the model results need to be evaluated through evaluation indicators. Various indicators in the recommendation field can be used to evaluate the effectiveness of the recommendation model, mainly obtained through offline, online experiments, or user surveys [6]. This article mainly uses AUC (Area under curve), MRR (Mean Reciprocal Rank), nDCG@5, and nDCG@10 (Normalized Discounted Cumulative Gain) Four commonly used indicators to evaluate and compare the performance of recommended models.

(1) AUC (Area Under ROC Curve)

AUC is the area of the area below the ROC curve, with a value range $\left[0.5,1\right]$, commonly used for evaluating classification models.

(2) MRR (average reciprocal ranking)

MRR is commonly used to measure the quality of returned results in search scenarios, and can also be used for evaluating recommendation models. This indicator reflects whether the items that the user is interested in are higher in the recommendation ranking list, emphasizing positional relationships and orderliness. The value range is $\left[0,1\right]$, and the closer the value is to 1, the better the effect. The specific calculation method is shown in Formula 4.

$$MRR = \frac{1}{|Q|}\sum_{i=1}^{|Q|}\frac{1}{rank_i} \tag{4}$$

Formula 4, $|Q|$ represents the number of users, and $rank_i$ represents the recommended list for the i-th user, indicating the position of the items that the user accessed in the list.

(3) NDCG (Normalized Accumulated Loss Gain)

In recommendation scenarios, nDCG is mainly used to evaluate the accuracy of sorting lists, taking into account the sorting order of recommendation results. The personalized recommendation model generates a sorted recommendation list for the target user, and the normalized cumulative gain reflects the difference between the sorted list and the user's actual interaction list. The length of the list is often fixed and calculated using nDCG@k K represents the length of the recommendation list, often taken as 5 and 10. The range of indicator values is $\left[0,1\right]$, and the closer the value is to 1, the better the effect [7]. The calculation method is shown in formulas 5, 6, and 7.

$$nDCG@k = \frac{DCG@k}{IDCG@k} \tag{5}$$

$$DCG@k = \sum_{i=1}^{k} \frac{rel_i}{\log_2(i+1)} \tag{6}$$

$$IDCG@k = \sum_{i=1}^{|REL|} \frac{rel_i}{\log_2(i+1)} \tag{7}$$

Formula 5, $DCG@k$ represents the Discounted Cumulative Gain, while $DCG@k$ representing the normalization factor, which is the ideal $DCG@k$ value. The main purpose is to normalize $DCG@k$; In formulas 6 and 7, rel_i represents the level-related value of project location i, and in personalized recommendation models, the value is 0 or 1 [8].

4.2.2 Overall performance verification of the model

To verify the overall performance of the personalized news recommendation model proposed in this article, this section constructs a control experiment on the MIND-small news dataset and calculates the AUC, MRR, and nDCG@5 nDCG@10 Compare the operating results of four evaluation indicators with the model in this article, analyze the indicator results, and then verify the effectiveness of the model[9] [10]. From the experimental results, it can be concluded that NAML, which integrates text content such as title, main text, and category, performs better than DKN, which only utilizes title information, in all four indicators. This indicates the effectiveness and necessity of learning news representation from various text components of news. In addition, compared to the experiments of the other two benchmark models on the same dataset, the model proposed in this paper performs well in AUC, MRR, nDCG@5, and nDCG@10 On average, the four evaluation indicators have improved by 3%, 2.6%, 2.5%, and 3%. The operating results are better than the two benchmark models, indicating that the model in this article improves the performance of personalized news recommendation models[11]. This article aims to fully utilize the textual components of news and combine knowledge graph information to better mine news features. Compared to DKN, this article introduces more text content such as news abstracts, abstract entities, categories, subcategories, and main text to learn news representation. Compared to NAML, this article introduces more news abstracts and knowledge level information. The experimental results indicate that combining a knowledge graph with news text fully utilizes the corresponding knowledge and semantic features of news, thereby improving the model's recommendation effect.The experimental results of different models are shown in Table 3.

Table 3. Experimental results of different models

Model	AUC	MRR	nDCG@5	nDCG@10
DKN	0.6103	0.2782	0.3055	0.3678
NAML	0.6486	0.2949	0.3359	0.3891
proposed method	**0.6593**	**0.3125**	**0.3452**	**0.4086**

5 Conclusion

This article starts from the content of news text, analyzes the various components and characteristics of news, and based on this, vectorizes some text content, including word vector representation, entity vector representation, and entity context vector representation. The experimental results show that the model based on knowledge graph is effective in AUC and MRR nDCG@5 nDCG@10, Compared with the benchmark model, the evaluation indicators have improved by 3%, 2.6%, 2.5%, and 3%, respectively. The ablation experiment has demonstrated the positive effects of knowledge graphs, multi perspective news features, and attention mechanisms. In summary, this study provides a new approach for news recommendation based on knowledge graphs, which helps to improve user satisfaction and enhance the platform's service capabilities.

Acknowledgments

The authors acknowledge financial support from the National Natural Science Foundation of China (Grant No.71872148), the Project of Shaanxi Soft Science Research Program (Grant No. 2022KRM152).

References

1. Ali I , Yadav D , Sharma A K .Question answering system for semantic web: a review[J].International journal of advanced intelligence paradigms, 2022(1/2):22.
2. Jizhou H, Wei Z, Haifeng W, et al. Multi-Task Learning for Entity Recommendation and Document Ranking in Web Search[J]. ACM Transactions on Intelligent Systems and Technology,2020,11(5):24.
3. Zhang K , Liu X , Wang W ,et al.Multi-criteria Recommender System based on Social Relationships and Criteria Preferences[J].Expert Systems with Applications, 2021:114868.DOI:10.1016/j.eswa.2021.114868..
4. Zhou Shuai, Du Yuncheng, Zhang Yangsen. Research progress in personalized news recommendation [J]. Computer Technology and Development, 2023, 33 (02): 1-8.
5. Feng Xiaodong, Hui Kangxin. Social media text topic clustering based on heterogeneous graph neural networks [J]. Data Analysis and Knowledge Discovery, 2022,6 (10): 9-19.
6. Xiang Liang. Practical Application of Recommendation System [M]. Beijing: People's Posts and Telecommunications Publishing House, 2012:23.

7. Fangzhao Wu, Ying Qiao, Jiun-Hung Chen, et al. MIND: A Large-scale Dataset for News Recommendation[C], ACL. 2020:3597-3606.
8. Li Chao, Fu Wei, Ma Ning, et al. Decomposer deep network recommendation algorithm [J]. Small Micro Computer System, 2022, 43 (02): 300-305.
9. Wu H, Huang C, Deng S. Improving aspect-based sentiment analysis with Knowledge-aware Dependency Graph Network[J]. Information Fusion,2023,92:289-299.
10. Park Y, Choi Y, Yun S, et al. Robust Data Augmentation for Neural Machine Translation through EVALNET[J]. Mathematics,2022,11(1):123.
11. Fangzhao Wu, Ying Qiao, Jiun-Hung Chen, et al. MIND: A Large-scale Dataset for News Recommendation[C], ACL. 2020:3597-3606.