Personalized News Recommendations Based on NRMS

Xiao Li, Zhong Xu, Heping Peng(✉), Hongbin Wang, and Qingdan Huang

Guangzhou Power Supply Bureau, Guangdong Power Grid Co. Ltd., Guangzhou, China
papercrane@263.net

Abstract. Personalized news recommendations are a crucial technology for assisting users in discovering news content of their interest and reducing information overload. At the same time, personalized news recommendations allow for effective statistics on news data and will provide assistance in the application of news data across industries and domains. In this regard, accurate modeling of news and users is crucial, while capturing words and the context of news is important for learning news and user representation. Recently, attention models have been widely used and their effectiveness in capturing contextual information is better than that of traditional CNN models. However, the combination of attention models with personalized news recommendation still needs further research. In this article, we have discussed personalized news recommendation using NRMS models and analyzed the application of attention models in personalized news recommendation. We have also proposed several directions for further research to assist researchers in gaining a comprehensive understanding of the implementation of attention models in personalized news recommendation.

Keywords: Component · News recommendation · News modeling · Attention

1 Introduction

A considerable number of readers have shifted from traditional newspapers to online news platforms due to the ease, immediacy, and portability of digital news services [1]. Yet, in the age of information explosion, the vast amount of news content presents significant challenges in processing and effectively utilizing this information. Given the sheer volume of articles being published on a daily basis, sifting through all available news to find content that aligns with a user's interests is virtually impossible [2]. Personalized news recommendation technology addresses this issue by curating content based on individual preferences, ultimately reducing information overload and enhancing the user experience on news platforms. Simultaneously, these personalized recommendations enable efficient data analysis, paving the way for cross-industry and sector-wide applications of news data. In recent years, personalized news recommendation has garnered considerable interest from both academic and industry circles, with significant research efforts devoted to this area.

Refining the feature extraction of news articles and user preferences is crucial in the realm of news recommendation systems [3]. Numerous deep-learning approaches

J. Yen et al. (Eds.): ICBIS 2023, AHCS 14, pp. 137–148, 2023.
https://doi.org/10.2991/978-94-6463-198-2_16

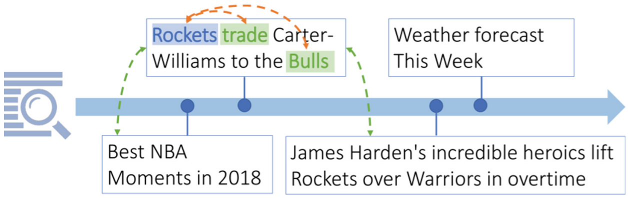


Fig. 1. An example user viewed several news articles, with orange and green dashed lines denoting the connections between words and news, respectively.

have been suggested to address these challenges [4–8]. For instance, Okura et al. [3] introduced a method that employs auto-encoders to derive news features from article bodies, while utilizing gated recurrent units (GRU) to represent users based on their browsing history. However, this approach is hindered by the time-consuming nature of GRUs and their inability to capture word contexts. On the other hand, Wang et al. [9] developed a technique that leverages a convolutional neural network (CNN) with external knowledge integration to extract news features from headlines and model user preferences by examining the similarities between candidate articles and previously browsed content. Despite its merits, this method falls short in capturing long-distance word contexts and fails to account for the relatedness among browsed articles.

The interplay between words in news headlines is crucial for comprehending the content. For instance, as illustrated in Fig. 1, the term “Rockets” is strongly related to “Bulls”. Furthermore, a single word may establish connections with multiple terms, such as “Rockets” exhibiting semantic affinities with “trade”. The relatedness of different news articles viewed by a particular user also plays a role; in Fig. 1, there is a connection between the second news piece and both the first and third articles. Additionally, the significance of individual words in representing news content may vary. In the case of Fig. 1, “NBA” is more informative than “2018”. Similarly, the essentiality of each news article in representing a user may differ; for example, the first three articles convey more meaningful information than the last one.

The attention model effectively utilizes the aforementioned information and addresses the limitations of both GRU and CNN models. Nevertheless, the integration of attention models with personalized news recommendations warrants further investigation. In this paper, we explore personalized news recommendations using the NRMS model, analyze the implementation of attention models in personalized news recommendation systems, and propose several avenues for future research. Our aim is to assist researchers in comprehending the application of attention models within the realm of personalized news recommendations.

2 Related Work

The recommendation of news is a substantial task that falls under the domains of natural language processing and data mining and has numerous practical applications [10, 11]. To achieve effective news recommendation, developing accurate news and user representations is crucial. However, many existing approaches rely on manual feature engineering to learn these representations [12–14]. For instance, Liu et al. [12] utilized a Bayesian model to create interest features and topic categories that were employed as news features. In a similar vein, Son et al. [13] proposed a location-based news recommendation system that incorporated topic and location features extracted from Wikipedia pages to generate news features using their Explicit Localized Semantic Analysis (ELSA) model. And a deep fusion model (DMF) was introduced by Lian et al. [15], which learned news features by integrating multiple handcrafted features, including news categories and title length. Developing these approaches necessitates a significant amount of domain expertise and labor, and they fall short of grasping the contextual and sequential nature of news articles, which are crucial factors in creating precise news and user features.

Lately, recommending news articles using deep learning has gained popularity, with several proposed approaches in this area [3–6, 16, 17]. For instance, Okura et al. [3] proposed a method to learn how to represent news articles by utilizing denoising autoencoders to extract relevant features from the article content, and to represent users based on their browsing history using a GRU network. Wang et al. [16] introduced a CNN network that is capable of incorporating knowledge graph information into its learning process to generate news features based on article headlines, demonstrating knowledge-enhanced news representation. However, these approaches rely on the individual type of news components, which may not be sufficient to accurately represent news and users. In contrast, our approach utilizes various types of news components, such as headlines, article content, and subject classification, through an attentive multi-perspective learning structure, resulting in a more precise depiction of news and users. In-depth experiments on a practical dataset demonstrate that by employing our method, we outperform existing approaches in learning news and user embeddings, leading to superior outcomes in news recommendation.

3 Personalized News Recommendations Core Issues

When a user accesses a news platform in a personalized news recommendation system, the platform initially retrieves a shortlisted set of potential news articles from a vast database of news content. The system then ranks these candidates based on the user's interests, as inferred from their profile. The most relevant K news articles are presented to the user, and the platform records their behavior towards these articles in order to update and maintain the user's profile for future recommendations. To achieve this, the design of a personalized news recommendation system involves several core components, including (1) news feature extraction, (2) user feature extraction, (3) personalized content matching, and (4) model training.

A critical aspect of personalized news recommendation is the development of effective news feature extraction, which enables the system to grasp the essential characteristics of news articles and gain a comprehensive understanding of their content. News feature extraction techniques can be broadly classified into two distinct categories: traditional methods based on handcrafted features and modern methods based on deep learning.

User feature extraction is an essential component of personalized news recommendation systems, as it enables the inference of a user's personal interests in news content.

There are two broad categories of user feature extraction methods in existing news recommendation approaches: traditional methods based on handcrafted features and modern methods based on deep learning.

Interest matching is a crucial component in personalized news recommendation systems that involves aligning candidate news with a user's personal interests to enable personalized ranking and display of news content. Techniques for interest matching can be broadly divided into two types: similarity-based interest matching and reinforcement learning-based interest matching. The former relies on news and user feature extraction to determine the relationship of news to a user's interests, while the latter involves training a reinforcement learning agent to make news recommendations by optimizing a reward function based on the user's behavior.

Personalized news recommendation techniques typically leverage machine learning models for news feature extraction, user feature extraction, and interest matching. Hence, the training of these models is a vital aspect of constructing a reliable and precise news recommendation system. It is imperative to ensure that the models are well-trained using a diverse and representative dataset, to achieve accurate and unbiased recommendations.

4 News Recommendation Base on NRMS

4.1 Introduction to the NRMS Model

The NRMS model is a neural network-driven news recommendation technology that employs multi-headed self-attention to enhance its performance. The core components of the NRMS approach are a news feature extractor and a user feature extractor. The news feature extractor utilizes multi-head self-attention to model cross-impacts among words and learns news features from headlines. The user feature extractor, on the other hand, learns user representations based on their browsing history by employing multi-head self-attention to capture interdependencies. NRMS further improves the quality of news and user representations by utilizing additive attention to select important words and articles. Figure 2 provides a visual representation of the NRMS approach for news recommendation.

4.2 News Feature Extraction

The News Encoder module, as depicted in Fig. 2, is employed to learn news features from headlines and comprises three layers. In the initial layer, known as the word embedding layer, news headlines are transformed from word sequences into low-dimensional embedding vector sequences. This process enables the conversion of each word in the headline into a numerical representation that can be used by the neural network model to identify relationships between words and capture their context. This technique helps to reduce the dimensionality of the input data while preserving the semantic meaning of the words in the headlines. Let us denote a news headline consisting of M words as $[w_1, w_2, \dots, w_M]$. This layer transforms the headline into a sequence of vectors $[e_1, e_2, \dots, e_M]$. The subsequent layer is composed of a multi-headed self-attentive network at the word level, which emphasizes the importance of interactions between words when learning news features. Consequently, multi-headed self-attention is employed to learn the Context-dependent representation of words by capturing these collaborative effects. The learned representation of the i_{th} word, as acquired by the k_{th} attention head, is mathematically computed as:

$$\alpha_{i,j}^{k} = \frac{\exp\left(\boldsymbol{e}_i^T \boldsymbol{Q}_k^w \boldsymbol{e}_j\right)}{\sum_{m=1}^{M} \exp\left(\boldsymbol{e}_i^T \boldsymbol{Q}_k^w \boldsymbol{e}_m\right)} \tag{1}$$

$$h_{(i,k)}^{w} = V_k^w \left(\sum\nolimits_{j=1}^{M} \alpha_{(i,j)}^{k} e_j\right) \tag{2}$$

where $\boldsymbol{Q}_k^w$ and $\boldsymbol{V}_k^w$ are the mapping vectors in the k_{th} self-attention unit, and the relative significance of the interaction between i_{th} and j_{th} words can be quantified by the weight vector $\alpha_{i,j}^{k}$ assigned to it. The multi-attention feature embedding $\boldsymbol{h}_i^w$ of the i_{th} word is merged to form the representations generated by h separate self-attention units, i.e., $\boldsymbol{h}_i^w = \left[\boldsymbol{h}_{i,1}^w; \boldsymbol{h}_{i,2}^w; \ldots; \boldsymbol{h}_{i,h}^w\right]$. The tertiary layer consists of the augmenting word attention network. The usage of different lexical items within a singular news narrative may bear divergent degrees of importance in effectively conveying the intended meaning of that particular story. Consequently, attention mechanisms are employed to pick essential words in news headlines, facilitating the learning of more enlightening news features. The attention weight assigned to the i-th word within a news title, denoted as α_i^w, is determined through a calculation process involving a set of specific algorithms.

$$a_i^w = \boldsymbol{q}_w^T tanh\left(\boldsymbol{V}_w \times \boldsymbol{h}_i^w + \boldsymbol{v}_w\right) \tag{3}$$

$$\alpha_i^w = \frac{\exp\left(a_i^w\right)}{\sum_{j=1}^{M} \exp\left(a_j^w\right)} \tag{4}$$

where $\boldsymbol{V}_w$ and $\boldsymbol{v}_w$ represent the mapping vectors, and $\boldsymbol{q}_w$ is the query vector. The ultimate portrayal of a news article is achieved via a weighted summation of the contextual features attributed to its constituent words, mathematically expressed as:

$$\boldsymbol{r} = \sum_{\boldsymbol{i}=1}^{\boldsymbol{M}} \alpha_i^w \boldsymbol{h}_i^w \tag{5}$$

4.3 User Feature Extraction

The User Encoder module is employed to learn user features based on the news articles they view, and it consists of two layers. In the initial layer features a news-level multi-headed self-attentive network is utilized to extract and integrate relevant features. It is common for news articles accessed by a given user to demonstrate varying degrees of relevance to one another. As a result, multi-headed self-attention is utilized to enrich news features by acquiring their interactions. The learned representation of the i_{th} news, as acquired by the k_{th} attention head, is mathematically computed as:

$$\beta_{i,j}^{k} = \frac{\exp\left(\boldsymbol{r}_i^T \boldsymbol{Q}_k^n \boldsymbol{r}_j\right)}{\sum_{m=1}^{M} \exp\left(\boldsymbol{r}_i^T \boldsymbol{Q}_k^n \boldsymbol{r}_m\right)} \tag{6}$$

$$\boldsymbol{h}_{i,k}^{n} = \boldsymbol{V}_{k}^{n}\left(\sum_{j=1}^{M}\beta_{i,j}^{k}\boldsymbol{r}_{j}\right) \tag{7}$$

where $\boldsymbol{Q}_k^n$ and $\boldsymbol{V}_k^n$ are the mapping vectors in the k_{th} news self-attention unit, and the relative significance of the interaction between j_{th} and k_{th} news can be quantified by the weight vector $\beta_{i,j}^k$ assigned to it. The multi-attention feature embedding of the i_{th} news is merged to form the representations generated by h separate self-attention units, i.e., $\boldsymbol{h}_i^n = \left[\boldsymbol{h}_{i,1}^n; \boldsymbol{h}_{i,2}^n; \ldots; \boldsymbol{h}_{i,h}^n\right]$. The subsequent layer consists of an augmenting news attention network. Different news items may convey varying levels of information in terms of representing users. Therefore, additional attention mechanisms are employed to select important news articles, facilitating the learning of more informative user features. The attention weight assigned to the i_{th} news is determined through a calculation process involving a set of specific algorithms.

$$a_i^n = \boldsymbol{q}_n^T tanh\left(\boldsymbol{V}_n \times \boldsymbol{h}_i^n + \boldsymbol{v}_n\right) \tag{8}$$

$$\alpha_i^n = \frac{\exp\left(a_i^n\right)}{\sum_{j=1}^{N}\exp\left(a_j^n\right)} \tag{9}$$

In this context, $\boldsymbol{V}_n$, $\boldsymbol{v}_n$ and $\boldsymbol{q}_n$ are all vectors within the attention network, and N refers to the total number of browsed news articles. The ultimate representation of a user is derived through a weighted summation of the distinct features pertaining to the news articles that have been browsed by that user. This is mathematically formulated as follows:

$$\boldsymbol{r} = \sum_{\boldsymbol{i}=1}^{\boldsymbol{N}}\alpha_i^n\boldsymbol{h}_i^n \tag{10}$$

4.4 Personalized Content Matching

The click predictor module is utilized as a means to estimate the likelihood of a user engaging with a candidate news article via a click action. Let a denote the representation of D^c prospective news article, with $\boldsymbol{r}^c$ representing its corresponding counterpart. The click probability score, denoted as $\hat{y}$ is mathematically calculated as the inner product of the feature vectors assigned to both the user and the candidate news article, i.e., $\hat{y} = \boldsymbol{u}^T\boldsymbol{r}^c$. While other scoring methods such as perception exist, the dot product has been found to offer prime results and effectiveness.

4.5 Model Training

With the aim of training the NRMS model, negative sampling techniques are utilized. Specifically, this involves randomly selecting a subset of non-relevant news articles to be paired with each training sample in order to simulate a more realistic recommendation

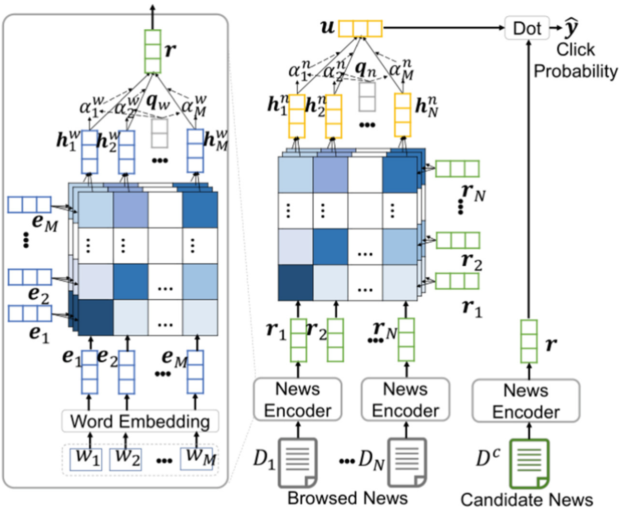


Fig. 2. An explication of the theoretical and practical foundation of the NRMS approach's framework.

scenario. In this study, a positive sample refers to a news article accessed by a given user, while negative samples are randomly drawn from the same impression that were not clicked on by the user. A total of K negative samples are selected for each positive sample. To eliminate the possibility of positional biases, a random shuffle is performed on the order of the selected news articles. For each positive news article, as well as the K corresponding negative articles, click probability scores are assigned and denoted as $\hat{y}^+$ and $\left[\hat{y}_1^-, \hat{y}_2^-, \ldots, \hat{y}_K^-\right]$ respectively. These scores are subsequently normalized through application of the softmax function, resulting in the calculation of the posterior click probability associated with the given positive news article, as expressed by the following formula:

$$p_i = \frac{\exp\left(\hat{y}_i^+\right)}{\exp\left(\hat{y}_i^+\right) + \sum_{j=1}^{K} \exp\left(\hat{y}_{i,j}^-\right)} \tag{11}$$

As part of this study, the task of predicting news click probabilities is reimagined as a pseudo $(K+1) - way$ classification problem. The loss function utilized for model training is defined as the negative log-likelihood of all positive samples, denoted as S. The explicit formulation of the loss function is as follows:

$$\mathcal{L} = -\sum_{i \in S} \log(p_i) \tag{12}$$

4.6 Experiments

To test the efficacy of the NRMS model, we performed experiments using a real-world news recommendation dataset collected from MSN News logs over a one-month period, spanning from December 13, 2018, to January 12, 2019. The logs collected during the final week were reserved for testing, while the rest of the logs were allocated for training. Moreover, we randomly sampled 10% of the training data for validation purposes.

To evaluate the effectiveness of our approach, we conducted a comparative analysis with several baseline methods, namely LibFM, DSSM, Wide&Deep, DeepFM, DFM, DKN, Conv3D, and GRU. The summarized results of these methods are presented in Table 1.

To assess the time efficiency of the NRMS model, we conducted experiments to compare its performance with that of several widely-used news recommendation methods. The comparative results are summarized in Table 2.

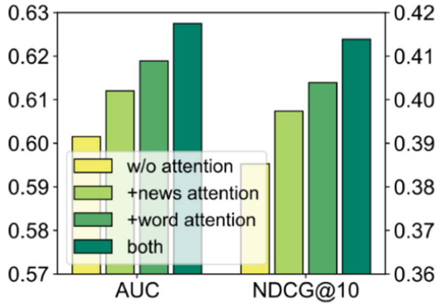
Experimental verification of the efficacy of attention in the NRMS approach: An investigation of word- and news-level attention mechanisms, as illustrated in Fig. 3(a). Our findings highlight the significant advantages of word-level attention, followed by an analysis of the impact of additive and self-attentions on the NRMS approach, as demonstrated in Fig. 3(b). Based on our observations, we concluded that self-attentions hold substantial value.

Table 1. Comparative analysis of method performance: *Significance of improvement at $p < 0.01$.

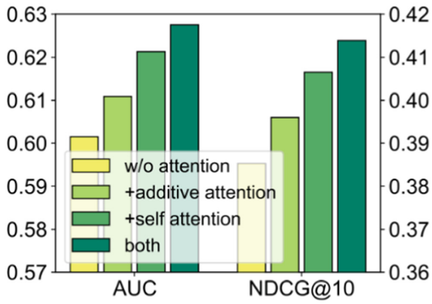
Methods	AUC	MRR	NDCG(@5)	NDCG(@10)
LibFM	0.5661	0.2414	0.2689	0.3552
DSSM	0.5949	0.2675	0.2881	0.3800
Wide&Deep	0.5812	0.2546	0.2765	0.3674
DeepFM	0.5830	0.2570	0.2802	0.3707
DFM	0.5861	0.2609	0.2844	0.3742
DKN	0.6032	0.2744	0.2967	0.3873
Conv-3D	0.6068	0.2788	0.3018	0.3916
GRU	0.6128	0.2826	0.3053	0.3976
NRMS*	0.6290	0.2996	0.3226	0.4160

Table 2. Comparative analysis of NRMS and baseline methods: Parameter count and feature extraction duration for 1 million news and users (*excluding word embeddings).

Methods	# Parameters	News Feature Extraction Duration	User Feature Extraction Duration
DKN	688 K	47.2 s	10.6 min
Conv-3D	590 K	30.6 min	17.6 min
GRU	550 K	60.8 s	148.8 min
NRMS	526 K	39.6s	6.87 min



(a) Analyzing the performance of attention mechanisms across diverse contexts.



(b) Comparing attention mechanisms for their effectiveness in various scenarios.

Fig. 3. Enhancing the performance of attention networks through comparative analysis.

5 Research and Outlook

To further enhance the performance of the NRMS model, several promising avenues could be pursued in future research. First, the current NRMS framework does not take into account the spatial information of words and news, which may provide valuable insights into learning more precise news and user features. To address this limitation, we could explore location encoding techniques, such as positional encoding, to incorporate word position information and news click timestamps within the model. This could potentially lead to a more comprehensive understanding of news articles and user preferences, ultimately improving the recommendation quality. Second, it would be beneficial to examine methods for efficiently integrating multiple sources of news information within the NRMS framework, particularly for longer sequences such as news bodies or extended article summaries. Conventional self-attentive networks may face challenges in terms of efficiency and computational complexity when handling long sequences. To mitigate these issues, we could investigate techniques such as segment-level attention, sliding-window approaches, or sparse attention mechanisms, which have been shown to effectively process lengthy sequences while maintaining computational efficiency. By exploring these potential directions, future work could significantly enhance the capabilities of the NRMS model, leading to more accurate and contextually relevant personalised news recommendations. These advancements would not only benefit users by providing a more tailored news consumption experience but also enable news platforms to better engage their audience and promote high-quality content.

6 Summary

Attention models have demonstrated remarkable capabilities in capturing rich contextual information, making them particularly effective for applications such as personalised news recommendation. By accurately modeling news articles and user preferences, these models can extract more relevant news features and user characteristics, ultimately leading to a tailored and enhanced news consumption experience. In this paper, we delve into personalised news recommendation utilizing the NRMS model as a prime example, providing insights into the practical applications of attention models in this domain. We begin by discussing the fundamentals of attention models, with a focus on their role in the personalised news recommendation landscape. Specifically, we examine the NRMS model, which leverages multi-head self-attention to simultaneously learn multiple representations of news articles and user preferences. This comprehensive approach enables a more nuanced understanding of the intricate relationships between users and news items, contributing to the effectiveness of the model. Subsequently, we analyze the implementation of attention models in personalized news recommendation systems, examining their advantages in handling varying input lengths, capturing long-range dependencies, and automatically learning feature representations. This section also highlights the superior performance of attention models compared to other popular methods, such as collaborative filtering and content-based filtering. Lastly, we outline several promising directions for future research, aimed at enhancing the application of attention models in personalized news recommendation. These avenues include exploring advanced attention mechanisms, incorporating external knowledge sources to enrich the representation of news articles and user preferences, and developing more efficient training and inference techniques to tackle the ever-growing scale of news data. In summary, this paper offers an in-depth analysis of personalized news recommendation with a focus on the NRMS model, showcasing the practical benefits of attention models in this field. By identifying potential areas for further research, we aim to inspire and guide researchers in their pursuit of more sophisticated and effective personalized news recommendation systems.

In our future work, we will apply fast clustering algorithm, such as [18–20], to pre-process news data, due to several benefits below:

1. Data visualization: Cluster analysis can divide data points into different groups, which can be visualized in a way that makes it easier to understand the structure and patterns of the data.
2. Data compression: Cluster analysis can compress data into fewer groups, reducing redundant data and improving data processing efficiency.
3. Data classification: Cluster analysis can classify data into different groups based on different criteria, allowing for more precise classification of data and increasing data reliability.
4. Improved recommendation accuracy: By using cluster analysis to group similar news together, it is possible to more accurately recommend news to users based on their interests, improving the accuracy of news recommendations and user experience.

5. Reduced computation cost: Cluster analysis can reduce computation costs by reducing the dimensionality of the data, improving computation efficiency.

Acknowledgments. Research works in this paper are supported by Guangdong Power Grid Co., Ltd. Guangzhou Power Supply Bureau Science and technology plan project (080037kk52190039/gzhkjxm20190100).

References

1. Banko M, MacKeen B, Ray L. A unified taxonomy of harmful content[C]//Proceedings of the fourth workshop on online abuse and harms. 2020: 125–137.
2. Wu C, Wu F, An M, et al. NPA: neural news recommendation with personalized attention[C]//Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining. 2019: 2576–2584.
3. Okura S, Tagami Y, Ono S, et al. Embedding-based news recommendation for millions of users[C]//Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining. 2017: 1933–1942.
4. Kumar V, Khattar D, Gupta S, et al. Deep Neural Architecture for News Recommendation[C]//CLEF (Working Notes). 2017.
5. Khattar D, Kumar V, Varma V, et al. Weave&rec: A word embedding based 3-d convolutional network for news recommendation[C]//Proceedings of the 27th ACM international conference on information and knowledge management. 2018: 1855–1858.
6. Wu C, Wu F, Liu J, et al. Neural demographic prediction using search query[C]//Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining. 2019: 654–662.
7. Zhu Q, Zhou X, Song Z, et al. Dan: Deep attention neural network for news recommendation[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2019, 33(01): 5973–5980.
8. An M, Wu F, Wu C, et al. Neural news recommendation with long-and short-term user representations[C]//Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. 2019: 336–345.
9. Wu C, Wu F, Liu J, et al. Detecting tweets mentioning drug name and adverse drug reaction with hierarchical tweet representation and multi-head self-attention[C]//Proceedings of the 2018 EMNLP workshop SMM4H: the 3rd social media mining for health applications workshop & shared task. 2018: 34–37.
10. Zheng G, Zhang F, Zheng Z, et al. DRN: A deep reinforcement learning framework for news recommendation[C]//Proceedings of the 2018 world wide web conference. 2018: 167–176.
11. Wu C, Wu F, An M, et al. Neural news recommendation with topic-aware news representation[C]//Proceedings of the 57th Annual meeting of the association for computational linguistics. 2019: 1154–1159.
12. Liu J, Dolan P, Pedersen E R. Personalized news recommendation based on click behavior[C]//Proceedings of the 15th international conference on Intelligent user interfaces. 2010: 31–40.
13. Son J W, Kim A Y, Park S B. A location-based news article recommendation with explicit localized semantic analysis[C]//Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval. 2013: 293–302.

14. Bansal T, Das M, Bhattacharyya C. Content driven user profiling for comment-worthy recommendations of news and blog articles[C]//Proceedings of the 9th ACM Conference on Recommender Systems. 2015: 195–202.
15. Lian J, Zhang F, Xie X, et al. Towards Better Representation Learning for Personalized News Recommendation: a Multi-Channel Deep Fusion Approach[C]//IJCAI. 2018: 3805–3811.
16. Wang H, Zhang F, Xie X, et al. DKN: Deep knowledge-aware network for news recommendation[C]//Proceedings of the 2018 world wide web conference. 2018: 1835–1844.
17. G. Xv, C. Lin, H. Li, J. Su, W. Ye, and Y. Chen, Neutralizing Popularity Bias in Recommendation Models, in Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2022, pp. 2623–2628.
18. Y. Chen, L. Zhou, S. Pei, Z. Yu, Y. Chen, X. Liu, et al., KNN-BLOCK DBSCAN: Fast clustering for large-scale data, IEEE transactions on systems, man, and cybernetics: systems, vol. 51, pp. 3939-3953, 2021.
19. Y. Chen, L. Zhou, N. Bouguila, C. Wang, Y. Chen, and J. Du, BLOCK-DBSCAN: Fast clustering for large scale data, Pattern Recognition, vol. 109, p. 107624, 2021
20. Y. Chen, X. Hu, W. Fan, L. Shen, Z. Zhang, X. Liu, et al., Fast density peak clustering for large scale data based on kNN, Knowledge-Based Systems, vol. 187, p. 104824, 2020.