SentiRec: Sentiment Diversity-aware Neural News Recommendation

Chuhan Wu[†] **Fangzhao Wu**[‡] **Tao Qi**[†] **Yongfeng Huang**[†]
[†]Department of Electronic Engineering & BNRist, Tsinghua University, Beijing 100084, China
[‡]Microsoft Research Asia, Beijing 100080, China
{wuchuhan15,wufangzhao,taoqi.qt}@gmail.com
yfhuang@tsinghua.edu.cn

Abstract

Personalized news recommendation is important for online news services. Many news recommendation methods recommend news based on their relevance to users' historical browsed news, and the recommended news usually have similar sentiment with browsed news. However, if browsed news is dominated by certain kinds of sentiment, the model may intensively recommend news with the same sentiment orientation, making it difficult for users to receive diverse opinions and news events. In this paper, we propose a sentiment diversity-aware neural news recommendation approach, which can recommend news with more diverse sentiment. In our approach, we propose a sentiment-aware news encoder, which is jointly trained with an auxiliary sentiment prediction task, to learn sentiment-aware news representations. We learn user representations from browsed news representations, and compute click scores based on user and candidate news representations. In addition, we propose a sentiment diversity regularization method to penalize the model by combining the overall sentiment orientation of browsed news as well as the click and sentiment scores of candidate news. Extensive experiments on real-world dataset show that our approach can effectively improve the sentiment diversity in news recommendation without performance sacrifice.

1 Introduction

Online news websites such as Google news[1] have gained huge popularity for consuming digital news (Das et al., 2007). However, it is difficult for users to find their interested news information due to the huge volume of news emerging every day (Okura et al., 2017). Thus, personalized news recommendation is important for news websites to target user interest and alleviate information overload (Wu et al., 2019a).

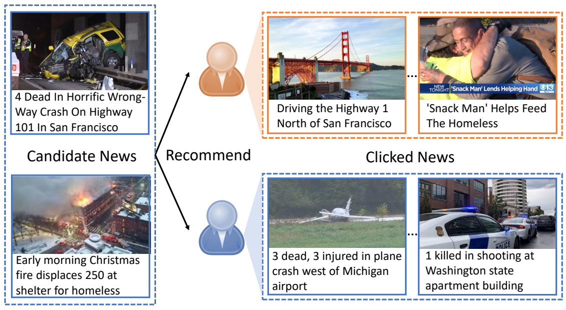


Figure 1: Several news browsed by two users of MSN News and the candidate news recommended to them.

Many existing news recommendation methods rank candidate news based on their relevance to the interests of users inferred from their historical browsed news (Okura et al., 2017; Wu et al., 2019c). For example, Okura et al. (2017) proposed to learn news representations from news texts via autoencoders, and learn user representations from browsed news using a gated recurrent unit (GRU) network. They ranked candidate news based on the inner product of the user representation and candidate news representation. Wu et al. (2019c) proposed to learn news and user representations using multi-head self-attention networks. They ranked news based on the click scores computed by the dot product between news and user representations. The news articles recommended by these methods are usually similar to those previously browsed by a user in many aspects, such as content and sentiment. For example, in Fig. 1 the two candidate news articles are recommended to both users. The first user browses a news about the highway in San Francisco and a news about a person helping the homeless, which has inherent relatedness with the content of the candidate news. The second user browses several news about deadly accidents and

[1]https://news.google.com/

Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing, pages 44–53
December 4 - 7, 2020.

crime, which has the same sentiment orientation as the candidate news. However, like the recommendations for the second user in Fig. 1, if a user mainly browses news articles that have a certain kind of sentiment (e.g., negative sentiment), many existing methods may intensively recommend news with the same sentiment orientation, which is not beneficial for this user to receive diverse opinions and news events that convey other sentiments.

In this paper, we propose a sentiment diversity-aware news recommendation approach named *SentiRec*, which can improve the sentiment diversity of news recommendation by considering the sentiment orientation of candidate and browsed news. In our approach, we propose a sentiment-aware news encoder, which is jointly trained with an auxiliary news sentiment prediction task, to incorporate sentiment information into news modeling and generate sentiment-aware news representations. We learn user representations from the representations of browsed news, and compute the click scores of candidate news based on their relevance to the user representations. In addition, to enhance the sentiment diversity of news recommendation, we propose a sentiment diversity regularization method to penalize our model during model training, which is based on the overall sentiment orientation of browsed news as well as the sentiment scores and click scores of candidate news. We conduct extensive experiments on a real-world benchmark dataset, and the results show that our approach can achieve better sentiment diversity and recommendation accuracy than many baseline methods.

The contributions of this paper are summarized as follows:

- To the best of our knowledge, this is the first work that explores to improve the sentiment diversity of news recommendation.
- We propose a sentiment-aware news encoder that incorporates an auxiliary news sentiment prediction task to encode sentiment-aware news representations.
- We propose a sentiment diversity regularization method to encourage the model to recommend news with diverse sentiment from the browsed news.
- Extensive experiments on real-world benchmark dataset verify that our approach can recommend news with diverse sentiment without performance loss.

2 Related Work

News recommendation is an important technique for online news websites to provide personalized news reading services (Zheng et al., 2018). A core problem in news recommendation is building accurate representations of news and users and further ranking candidate news according to news and user representations (Okura et al., 2017). In many news recommendation methods, news ranking is based on the representations of news and users built by manual feature engineering (Liu et al., 2010; Capelle et al., 2012; Son et al., 2013; Karkali et al., 2013; Garcin et al., 2013; Bansal et al., 2015; Ren et al., 2015; Chen et al., 2017; Zihayat et al., 2019). For example, Liu et al. (2010) proposed to use topic categories and interest features generated by a Bayesian model to build news and user representations. They ranked candidate news based on the product of a content-based score computed from news representations and a filter-based score computed by collaborative filtering. Son et al. (2013) proposed an Explicit Localized Semantic Analysis (ELSA) model for location-based news recommendation. They proposed to represent news and users by extracting topic and location features from Wikipedia pages, and ranked news based on the cosine distance between the representations of news and user. Lian et al. (2018) proposed to use various handcrafted features to represent news and users, such as title length, news categories, user profiles and features extracted from user behavior histories. They ranked candidate news based on the click scores computed by a neural factorization machine. However, these methods rely on manual feature engineering to build news and user representations, which usually necessitate massive expertise. In addition, handcrafted features may not be optimal in representing news content and user interest.

In recent years, several news recommendation methods based on deep learning techniques are proposed (Okura et al., 2017; Khattar et al., 2018; Wang et al., 2018; Wu et al., 2019a; An et al., 2019; Wu et al., 2019b,c; Ge et al., 2020). For example, Okura et al. (2017) proposed to learn first news representations from news bodies using autoencoders, and then learn representations of users from their clicked news with a GRU network. Candidate news are ranked based on the click scores computed by the dot products between news and user representations. Wang et al. (2018) proposed to learn news representations from news titles and their entities via a knowledge-aware CNN network, and learn user representations from clicked news with a candidate-aware attention network. They ranked candidate news based on the click scores computed from the concatenation of news and user representations via a feed-forward neural network. Wu et al. (2019c) proposed to learn news and user representations with a combination of multi-head self-attention and additive attention networks. They also used dot product to compute click scores for news ranking. These methods tend to recommend news articles which are similar with the news users previously browsed (Lin et al., 2014). Thus, these methods may recommend news with similar sentiment orientation with those previously browsed by users, which is not beneficial for users to receive diverse news information. Different from these methods, our approach can effectively recommend news with diverse sentiment to users by incorporating sentiment information into news modeling via a sentiment-aware news encoder and regularizing the model based on the sentiment orientation of browsed and candidate news.

3 Our Approach

In this section, we first present the formal definitions of the problem explored in this paper, then introduce the details of our sentiment diversity-aware news recommendation (SentiRec) approach.

3.1 Problem Definition

The problem studied in this paper is defined as follows. Given a user u with her news browsing history $H = [D_1, D_2, ..., D_N]$ and a set of candidate news[2] $C = [D^c_1, D^c_2, ..., D^c_P]$ (N and P respectively denote the number of browsed news and candidate news), the goal of the news recommendation model is to predict the personalized click scores $[\hat{y}_1, \hat{y}_2, ..., \hat{y}_P]$ of these candidate news, which are further used for ranking and display. We denote the sentiment labels of the browsed news and candidate news as $[s_1, s_2, ..., s_N]$ and $[s^c_1, s^c_2, ..., s^c_P]$, respectively. In this paper we assume the sentiment labels are real values from -1 to 1, which indicate the sentiment polarity of news articles. We denote the overall sentiment orientation of browsed news as s. The sentiment diversity is defined as the differences between the sentiment orientation of recommended news and the overall sentiment of browsed news.[3] The sentiment diversity of the news ranking results C' for the user u is measured by a function $d = f(C', s)$. The recommendation diversity is better if more top ranked news in C' have the different sentiment orientation with s.

3.2 News Recommendation Framework

In this section, we introduce the general news recommendation framework of our *SentiRec* approach, as shown in Fig. 2. There are three core components in this framework for news recommendation, i.e., sentiment-aware (SA) news encoder, user encoder, and click predictor. The sentiment-aware news encoder aims to learn representations of news articles from their texts, where their sentiments are taken into consideration. We apply the sentiment-aware news encoder to the browsed news $[D_1, D_2, ..., D_N]$ and the candidate news D^c to encode their sentiment-aware representations, which are respectively denoted as $[\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N]$ and $\mathbf{r}^c$. The user encoder aims to learn representations of users from the sentiment-aware representations of their browsed news. Motivated by (Vaswani et al., 2017), we use Transformer to capture the relatedness between browsed news and learn a unified representation $\mathbf{u}$ for each user. The click predictor aims to compute the personalized click scores of candidate news by measuring the relevance between user and candidate news representations. Following many previous works (Okura et al., 2017; Wu et al., 2019b), we use dot product to implement the click predictor, and the click score $\hat{y}$ is predicted by $\hat{y} = \mathbf{u}^\top \mathbf{r}^c$.

3.3 Sentiment-Aware News Encoder

In this section, we introduce the details of the sentiment-aware news encoders in our *SentiRec* approach. Its architecture is shown in Fig. 3. Motivated by the news encoder in (Wu et al., 2019c), we first use a word embedding layer to convert the sequence of words in a news title into a sequence of semantic vectors, and then use a Transformer (Vaswani et al., 2017) to capture the contexts of words and build a unified $\mathbf{r}$ representation of news texts. However, the news representations directly learned by the Transformer are usually not sentiment-bearing. In fact, the sentiment information of news is very important for understanding

[2] The candidate news set is usually recalled from the entire news pool.

[3] We do not strictly require the recommendation results in an impression to be diverse in sentiment. We expect the sentiment of recommended news in a long term (e.g., multiple impressions in months) is diverse.

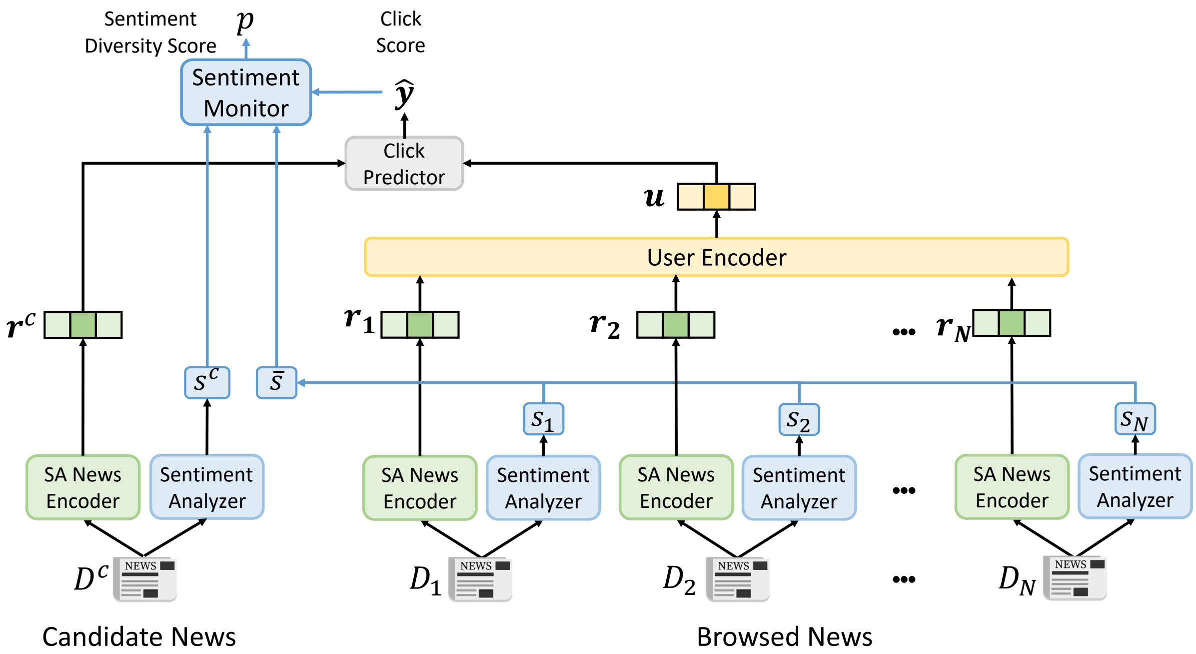


Figure 2: The framework of our *SentiRec* approach.

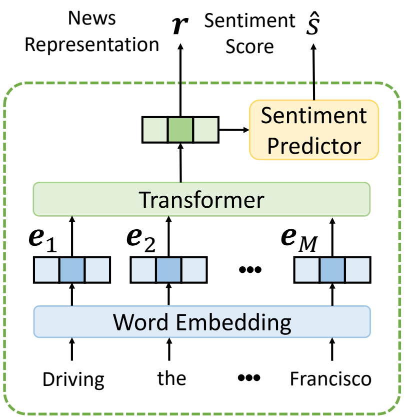


Figure 3: The architecture of the sentiment-aware news encoder.

the content of news. For example, in Fig. 1, although the news "Early morning..." and "Snack Man..." are both related to the homeless, they have opposite sentiment polarity, and modeling the sentiment of them can help understand their content better. In addition, the sentiment of news can also provide useful clues for user modeling and news ranking. For example, if a user frequently clicks negative news as the second user in Fig. 1, it may be more appropriate to recommend several positive news to this user rather than continuously recommending similar negative news. Thus, modeling news sentiment has the potential to enhance news recommendation. However, since the sentiment scores of news are numerical variables, simply regarding them as model input may be not optimal. Thus, we propose an auxiliary sentiment prediction task, and we jointly train the news encoder with this task to encourage it to learn sentiment-aware news representations. The real-valued sentiment score $\hat{s}$ is predicted as follows:

$$\hat{s} = \mathbf{V}_s \times \mathbf{r} + \mathbf{v}_s, \tag{1}$$

where $\mathbf{V}_s$ and $\mathbf{v}_s$ are parameters. The loss function of sentiment prediction we use is the mean absolute error (MAE), which is formulated as follows:

$$\mathcal{L}_{senti} = \frac{1}{S}\sum_{i=1}^{S} |\hat{s}_i - s_i|, \tag{2}$$

where $\hat{s}_i$ and s_i respectively stand for the predicted sentiment score and sentiment label of the i-th news, and S denotes the number of news. The sentiment labels are obtained by the sentiment analyzer modules in Fig. 2, which can be implemented by many sentiment analysis methods.

3.4 Sentiment Diversity Regularization

To further improve the sentiment diversity of news recommendation, we propose a sentiment diversity regularization method to penalize the recommendation model according to the overall sentiment score of browsed news, the sentiment score of candidate news, and its predicted click score. As shown in Fig. 2, we first use the sentiment analyzer to obtain the sentiment scores of the candidate news (denoted as s^c) and browsed news (denoted as $[s_1, s_2, ..., s_N]$). We then compute an overall sentiment score[4] of browsed news to indicate the historical sentiment preference of a user as follows:

$$\bar{s} = \frac{1}{N}\sum_{i=1}^{N} s_i. \tag{3}$$

A positive $\bar{s}$ indicates that the user has read news with more positive sentiment and a negative $\bar{s}$ indicates the negative sentiment is dominant in the browsed news. If the news recommender intensively recommends news with the same sentiment polarity with the overall sentiment s of a user's browsed news, it is difficult for this user to receive diverse news information. Thus, it is important to recommend news with diverse sentiment to users. To solve this problem, we propose a sentiment diversity regularization method. We first propose to compute a sentiment diversity score p with a sentiment monitor, which is formulated as follows:

$$p = max(0, \bar{s}s^c\hat{y}), \tag{4}$$

where a larger score of p indicates less sentiment diversity. In this formula, for a candidate news that shares the same sentiment polarity with s, the score p is larger if the model assigns it a higher click score or its sentiment and the overall browsed news sentiment are more intense, which indicate that the recommendation is less diverse in sentiment. Then, we propose a sentiment diversity loss function to regularize our model as follows:

$$\mathcal{L}_{div} = \frac{1}{|\mathcal{S}|}\sum_{i\in\mathcal{S}} p_i, \tag{5}$$

where $\mathcal{S}$ is the data set for model training, and p_i denotes the sentiment diversity score of the i-th sample in $\mathcal{S}$.

3.5 Model Training

In this section, we introduce how to train the models in our *SentiRec* approach. Following (Huang et al., 2013; Wu et al., 2019c), we use negative sampling techniques to construct labeled data for the news recommendation task from the user impression logs. More specifically, for each news clicked by a user, we randomly sample K news displayed in the same impression which are not clicked by this user. We denote the click scores of the i-th clicked news as $\hat{y}_i^+$ and the associated K non-click news as $[\hat{y}_{i,1}^-, \hat{y}_{i,2}^-, ..., \hat{y}_{i,K}^-]$. we use the click predictor to jointly predict these scores, and normalize these scores via the softmax function to compute the click probability scores. The news recommendation loss we used is the negative log-likelihood of the clicked news samples, which is computed as:

$$\mathcal{L}_{rec} = \sum_{i\in\mathcal{S}} \log(\frac{\exp(\hat{y}_i^+)}{\exp(\hat{y}_i^+) + \sum_{j=1}^{K}\exp(\hat{y}_{i,j}^-)}), \tag{6}$$

where $\mathcal{S}$ is the data set for model training. We jointly train the news recommendation model with the auxiliary sentiment prediction task and meanwhile regularize it using the sentiment diversity loss. The final unified loss function of our approach is a weighted summation of the three loss functions, which is formulated as follows:

$$\mathcal{L} = \mathcal{L}_{rec} + \lambda\mathcal{L}_{senti} + \mu\mathcal{L}_{div}, \tag{7}$$

where λ and μ are coefficients to control the relative importance of the sentiment prediction loss and sentiment diversity regularization loss.

4 Experiments

4.1 Datasets and Experimental Settings

Our experiments were conducted on a real-world news recommendation dataset provided by (Wu et al., 2019b), which is constructed from MSN News[5] logs from Oct. 31, 2018 to Jan. 29, 2019. We use the logs in the last week as the test set and the rest are used for training and validation, where the split ratio is 9:1.[6] To obtain the sentiment labels of the news in this dataset, we use the VADER algorithm (Hutto and Gilbert, 2014) as the sentiment analyzer in our approach.[7] It is a famous sentiment analysis method based on a set of sentiment lexicons such as LIWC (Pennebaker et al., 2001), ANEW (Nielsen, 2011) and GI (Stone et al., 1966). We use VADER to compute an overall sentiment orientation score of each news as the gold label, and these scores are ranged in [-1, 1]. The detailed statistics of the news recommendation dataset are shown in Table 1. We also plot the distribution of news sentiment scores and the overall sentiment orientation of users' browsed news

[4]We do not incorporate the numbers of positive and negative news because they cannot take the sentiment intensity into consideration.

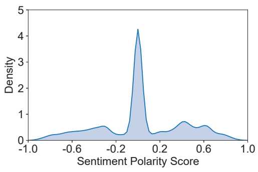
[5]https://www.msn.com/en-us/news

[6]The numbers of constructed samples for training and validation are 277,811 and 30,868, respectively. The number of samples for test is 1,707,588.

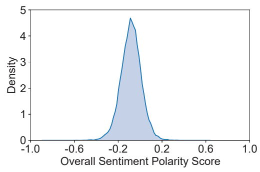
[7]We choose this algorithm because it can compute the real-valued sentiment scores rather than polarity only.

# users	10,000	**avg. # words per title**	11.29
# news	42,255	**# click samples**	489,644
# impressions	445,230	**# non-click samples**	6,651,940
# samples	7,141,584	**avg. sentiment score**	0.0314

Table 1: Statistics of the dataset.



(a) News sentiment polarity scores.



(b) Overall sentiment orientation of users' browsed news.

Figure 4: Distributions of the news sentiment polarity scores and the overall sentiment orientation of users' browsed news in our dataset.

in Figs. 4(a) and 4(b), respectively. We find that although positive and negative news are almost balanced, the overall sentiment orientation of users' browsed news is negative. In addition, we show the average click-through rate (CTR) of news with different ranges of sentiment scores in Fig. 5. As the saying goes, *"evil news rides fast, while good news baits later"*. We find it is interesting that more negative news have higher CTRs, which indicates that negative news has stronger ability in attracting news clicks. Thus, it is important to recommend news with diverse sentiment to avoid overwhelming users with too much negative news information.

Following Wu et al. (2019b), in our experiments the word embeddings were initialized by the 300-dimensional Glove embeddings (Pennington et al., 2014). The negative sampling ratio K was set to 4. Adam (Kingma and Ba, 2014) was chosen as the optimizer and the size of a minibatch was 30. In addition, the loss weights λ and μ were respectively set to 0.4 and 10. These hyperparameters were tuned on the validation set. To evaluate the performance of news recommendation, we use metrics including AUC, MRR, nDCG@5 and nDCG@10[8].

Since there is no off-the-shelf metric to evaluate the sentiment diversity of news recommendation, motivated by the MRR and hit ratio metrics, we propose three metrics named $Senti_{MRR}$, $Senti@5$ and $Senti@10$ to quantitatively measure sentiment

[8]The relevance grade is binary, i.e., 0 for non-clicked news and 1 for clicked news.

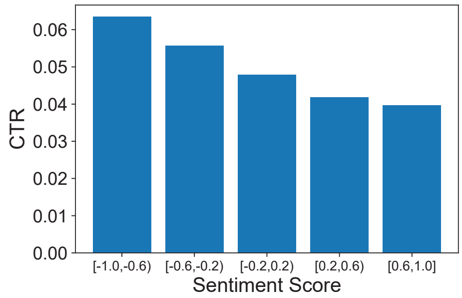


Figure 5: Click-through rates of news with different sentiment polarity scores.

diversity. They are computed as follows:

$$\begin{aligned} Senti_{MRR} &= max(0, \bar{s}\sum_{i=1}^{C}\frac{s_i^c}{i}), \\ Senti@5 &= max(0, \bar{s}\sum_{i=1}^{5} s_i^c), \\ Senti@10 &= max(0, \bar{s}\sum_{i=1}^{10} s_i^c), \end{aligned} \tag{8}$$

where C is the number of candidate news in an impression, s_i^c denotes the sentiment score of the candidate news with the i-th highest click score. In these metrics, higher scores indicate that the recommendation results are less diverse from the browsed news in their sentiment.[9] We repeated each experiment 10 times and reported the average results over all impressions in terms of the recommendation performance and sentiment diversity.

4.2 Performance Evaluation

We evaluate the recommendation performance and sentiment diversity of our approach by comparing it with several baseline methods, including: (1) *LibFM* (Rendle, 2012), a feature-based recommendation method based on factorization machine. TF-IDF features are used to represent the textual content of news. (2) *EBNR* (Okura et al., 2017), an embedding-based neural news recommendation method. It uses denoising autoencoders to learn news representations and a GRU network to encode user representations. (3) *DKN* (Wang et al., 2018), a knowledge-aware news recommendation method, which learns news representations via knowledge-aware CNN networks and learns user representations with a candidate-aware attention network.

[9]The scores are positive if the top ranked news have the same sentiment orientation with the overall sentiment, and are higher if these sentiments are more intensive.

Methods	AUC	MRR	nDCG@5	nDCG@10
LibFM	0.5661	0.2414	0.2689	0.3552
EBNR	0.6102	0.2811	0.3035	0.3952
DKN	0.6032	0.2744	0.2967	0.3873
Conv3D	0.6051	0.2765	0.2987	0.3904
DAN	0.6154	0.2860	0.3093	0.3996
NPA	0.6240	0.2952	0.3185	0.4094
NAML	0.6205	0.2902	0.3144	0.4060
NRMS	0.6275	0.2985	0.3217	0.4139
SentiRec	0.6294	0.3013	0.3237	0.4165
SentiRec-same	0.6299	0.3017	0.3240	0.4171

Table 2: Results of recommendation performance. Higher scores indicate better performance.

Methods	$Senti_{MRR}$	$Senti@5$	$Senti@10$
Random	0.0262	0.0442	0.0687
LibFM	0.0843	0.1192	0.2579
EBNR	0.0989	0.1476	0.2868
DKN	0.0954	0.1389	0.2810
Conv3D	0.0973	0.1431	0.2830
DAN	0.1005	0.1520	0.2897
NPA	0.1044	0.1583	0.3015
NAML	0.1030	0.1569	0.2967
NRMS	0.1066	0.1592	0.3034
SentiRec	**0.0046**	**0.0083**	**0.0115**
SentiRec-same	0.3271	0.4963	0.9373

Table 3: Results of sentiment diversity. Lower scores indicate better sentiment diversity.

(4) *Conv3D* (Khattar et al., 2018), a neural news recommendation method which learns news representations using 2-D CNN models and learns user representations using a 3-D CNN model. (5) *DAN* (Zhu et al., 2019), a neural news recommendation method which learns news representations from title and entities with two independent CNN models and learns user representations using attentive LSTM network. (6) *NPA* (Wu et al., 2019b), a neural news recommendation method which learns news and user representations via personalized attention mechanism. (7) *NAML* (Wu et al., 2019a), a neural news recommendation method which learns news representations with CNN models and learns user representations using attention networks. (8) *NRMS* (Wu et al., 2019c), a neural news recommendation method which learns news and representations using multi-head self-attention and additive attention networks. For fair comparison, in all methods we used news titles to learn news representations. In addition, we compare the sentiment diversity of random news ranking, which aims to show the benchmark sentiment diversity without news and user modeling. Besides, we also compare a variant of our method (denoted as *SentiRec-same*) which only recommends the news with the same sentiment polarity with the browsed news (filter the candidate news with different sentiment polarity), which aims to show the scores of an extreme case with minimal sentiment diversity.

The results of recommendation performance and sentiment diversity are summarized in Tables 2 and 3. From these results, we find that neural news recommendation approaches achieve better recommendation performance than *LibFM*. This is probably because neural networks can learn more informative news and user representations than traditional matrix factorization methods. However, compared with random ranking, we find that the diversity scores of all baseline methods are much larger, especially those based on neural networks. This is probably because the compared baseline methods mainly recommend news based on the relevance between candidate news and browsed news, and will tend to recommend news with similar sentiment orientation with browsed news, which is harmful for users to receive diverse news information. Different from baseline methods, our *SentiRec* approach can achieve much better sentiment diversity even than random ranking. These results show that our approach can actively recommend news with diverse sentiment from browsed news. In addition, our approach can also achieve better recommendation performance than baseline methods. These results validate that our approach can achieve the goal of improving sentiment diversity in news recommendation without hurting the recommendation performance. Besides, by comparing *SentiRec* and its variant *SentiRec-same*, although the recommendation performance of *SentiRec-same* is slightly better, the sentiment of its recommendation results are minimally diverse from browsed news, which may amplify the problem of filter bubble and hurt user experience.

4.3 Ablation Study

In this section, we conduct ablation studies to verify the influence of the auxiliary sentiment prediction task in the sentiment-aware news encoder and the sentiment diversity regularization method on the recommendation performance and sentiment diversity. The results are shown in Fig. 6. From Fig. 6, we find that the sentiment prediction task can improve both sentiment diversity and recommendation performance. This may be because this auxiliary task can encourage the news encoder to

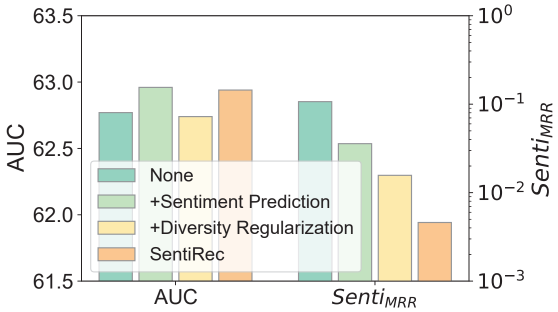


Figure 6: Influence of the sentiment prediction task and sentiment diversity regularization method.

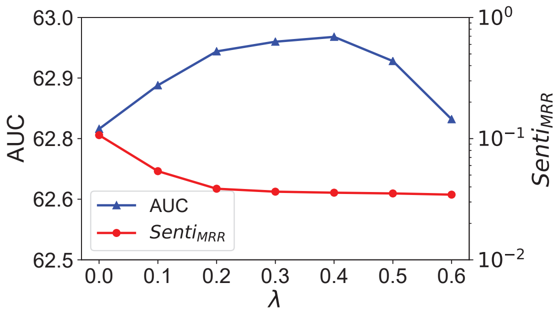


Figure 7: Influence of the hyperparameter λ.

exploit the sentiment information in news texts to encode sentiment-aware news representations, which is beneficial for predicting news clicks more accurately and further improving sentiment diversity by modeling users' dynamic preferences on news sentiment. In addition, the sentiment diversity regularization can also effectively improve the sentiment diversity of news recommendation and meanwhile keep the recommendation performance. This is because this regularization method can enforce the model to recommend news with different sentiment orientations with the browsed news. Moreover, combining both techniques can further improve sentiment diversity, which verifies the effectiveness of our *SentiRec* method.

4.4 Influence of Hyperparameters

In this section, we will explore the influence of two important hyperparameters on our approach, i.e., the loss coefficients λ and μ in Eq. (7) on the performance and sentiment diversity of our approach. Since there are two hyperparameters, we first vary the value of λ to find the optimal one to learn sentiment-aware news representations in our approach. The results are illustrated in Fig. 7. According to Fig.7, we find both sentiment diversity and recommendation performance of our approach

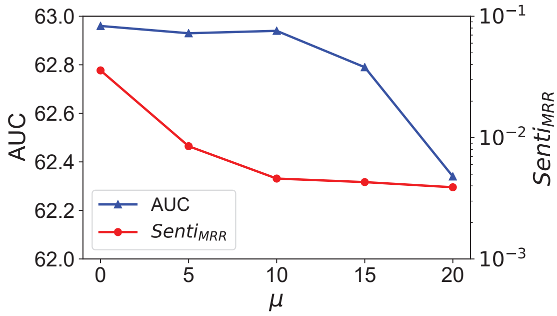


Figure 8: Influence of μ under $\lambda = 0.4$.

improves when λ increases from 0. This is probably because when λ is too small, the useful sentiment information in news cannot be fully exploited. However, the performance of our approach starts to decline when λ is too large. This may be because when λ is too large, the auxiliary sentiment prediction task is over-emphasized and the news recommendation task is not fully respected. Thus, a moderate λ (e.g., 0.4) is more appropriate for our approach to make a tradeoff between recommendation performance and sentiment diversity.

Then, we vary the value of μ under $\lambda = 0.4$ to evaluate the recommendation performance and sentiment diversity of our approach.[10] The results are illustrated in Fig. 8. According to the results, we find the sentiment diversity can be consistently improved when μ increases. This is probably because when μ is larger, the model is regularized more intensively and may tend to recommend more news with diverse sentiment from browsed news. However, when μ goes too large, the performance in terms of AUC declines significantly, which may hurt user experience. Thus, a moderate selection on μ (e.g., 10) is appropriate to achieve the goal of recommending news with diverse sentiment and meanwhile keep good recommendation performance.

4.5 Case Study

In this section, we present several case studies to better demonstrate the effectiveness of our approach in improving sentiment diversity of news recommendation. The clicked news of a randomly selected user as well as the top ranked candidate news recommended by a state-of-the-art method *NRMS* and our *SentiRec* approach are shown in Table 4. We can see that the historical browsed news of this user are mainly about negative topics such

[10] We find that the scale of the regularization loss is relatively small and the magnitude of μ needs to be larger.

Browsed News	Top Ranked Candidate News	
	NRMS	SentiRec
Woman Arrested for Alleged California Wildfire Scam	Sheriff: California officer's killer is in the US illegally	Eight 2018 Fashion Trends We're Ready to Move On From
Guns are the second leading killer of kids, after cars	Professional golfer and his caddie arrested for poaching at a tiger reserve	Josh Duhamel Wants to Date Someone Young Enough to Have Kids
From international fashion model to suspect in racist attack on Kansas toddler	Trump threatens years-long shutdown for his wall as GOP support begins to fracture	58 Amazing After-Christmas Deals Happening Right Now

Table 4: The browsed news of a user and top ranked candidate news provided by different methods.

as crime, which usually convey negative sentiment. However, the *NRMS* method still intensively recommends news with negative sentiment such as "Sheriff: California officer's killer...". It indicates that *NRMS* tends to recommend news with similar sentiment to the browsed news, which is not suitable for users to acquire diverse news information. Different from *NRMS*, our approach can effectively recommend news with diverse sentiment from browsed news, and the recommended news also has some inherent relatedness with browsed news in their content (e.g., both the first candidate news and the third browsed news mention "fashion"). It shows that our approach can improve the sentiment diversity of news recommendation and meanwhile keep recommendation accuracy.

5 Conclusion and Future Work

In this paper, we propose a sentiment diversity-aware neural news recommendation approach which can effectively recommend news with diverse sentiment from browsed news. We propose a sentiment-aware news encoder to learn sentiment-aware news representations by jointly training it with an auxiliary sentiment prediction task. We learn user representations from representations of browsed news, and compute click scores based on user and candidate news representations. In addition, we propose a sentiment diversity regularization method to regularize the model according to the overall sentiment orientation of browsed news as well as the click scores and sentiment scores of candidate news. Extensive experiments on real-world benchmark dataset validate that our approach can effectively enhance the sentiment diversity of news recommendation without hurting the recommendation performance.

In our future work, we plan to analyze the sentiment on the entities in news and explore to improve the entity-level sentiment diversity of news recommendation. In addition, we plan to extend sentiment polarities to more kinds of emotions, such as angry, happiness, sad and surprise, to enhance the emotion diversity of news recommendation.

Acknowledgments

This work was supported by the National Key Research and Development Program of China under Grant number 2018YFB2101501, and the National Natural Science Foundation of China under Grant numbers U1936208 and U1936216.

References

Mingxiao An, Fangzhao Wu, Chuhan Wu, Kun Zhang, Zheng Liu, and Xing Xie. 2019. Neural news recommendation with long-and short-term user representations. In *ACL*, pages 336–345.

Trapit Bansal, Mrinal Das, and Chiranjib Bhattacharyya. 2015. Content driven user profiling for comment-worthy recommendations of news and blog articles. In *RecSys.*, pages 195–202. ACM.

Michel Capelle, Flavius Frasincar, Marnix Moerland, and Frederik Hogenboom. 2012. Semantics-based news recommendation. In *WIMS*, page 27. ACM.

Cheng Chen, Xiangwu Meng, Zhenghua Xu, and Thomas Lukasiewicz. 2017. Location-aware personalized news recommendation with deep semantic analysis. *IEEE Access*, 5:1624–1638.

Abhinandan S Das, Mayur Datar, Ashutosh Garg, and Shyam Rajaram. 2007. Google news personalization: scalable online collaborative filtering. In *WWW*, pages 271–280. ACM.

Florent Garcin, Christos Dimitrakakis, and Boi Faltings. 2013. Personalized news recommendation with context trees. In *Proceedings of the 7th ACM conference on Recommender systems*, pages 105–112. ACM.

Suyu Ge, Chuhan Wu, Fangzhao Wu, Tao Qi, and Yongfeng Huang. 2020. Graph enhanced representation learning for news recommendation. In *WWW*, pages 2863–2869.

Po-Sen Huang, Xiaodong He, Jianfeng Gao, Li Deng, Alex Acero, and Larry Heck. 2013. Learning deep structured semantic models for web search using clickthrough data. In *CIKM*, pages 2333–2338. ACM.

Clayton J Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *ICWSM*.

Margarita Karkali, Dimitris Pontikis, and Michalis Vazirgiannis. 2013. Match the news: A firefox extension for real-time news recommendation. In *SIGIR*, pages 1117–1118. ACM.

Dhruv Khattar, Vaibhav Kumar, Vasudeva Varma, and Manish Gupta. 2018. Weave& rec: A word embedding based 3-d convolutional network for news recommendation. In *CIKM*, pages 1855–1858. ACM.

Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

Jianxun Lian, Fuzheng Zhang, Xing Xie, and Guangzhong Sun. 2018. Towards better representation learning for personalized news recommendation: a multi-channel deep fusion approach. In *IJCAI*, pages 3805–3811.

Chen Lin, Runquan Xie, Xinjun Guan, Lei Li, and Tao Li. 2014. Personalized news recommendation via implicit social experts. *Information Sciences*, pages 1–18.

Jiahui Liu, Peter Dolan, and Elin Rønby Pedersen. 2010. Personalized news recommendation based on click behavior. In *IUI*, pages 31–40. ACM.

Finn Årup Nielsen. 2011. A new anew: Evaluation of a word list for sentiment analysis in microblogs. In *Workshop on'Making Sense of Microposts: Big things come in small packages*, pages 93–98.

Shumpei Okura, Yukihiro Tagami, Shingo Ono, and Akira Tajima. 2017. Embedding-based news recommendation for millions of users. In *KDD*, pages 1933–1942. ACM.

James W Pennebaker, Martha E Francis, and Roger J Booth. 2001. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001.

Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global vectors for word representation. In *EMNLP*, pages 1532–1543.

Rui Ren, Lingling Zhang, Limeng Cui, Bo Deng, and Yong Shi. 2015. Personalized financial news recommendation algorithm based on ontology. *Procedia Computer Science*, 55:843–851.

Steffen Rendle. 2012. Factorization machines with libfm. *TIST*, 3(3):57.

Jeong-Woo Son, A Kim, Seong-Bae Park, et al. 2013. A location-based news article recommendation with explicit localized semantic analysis. In *SIGIR*, pages 293–302. ACM.

Philip J Stone, Dexter C Dunphy, and Marshall S Smith. 1966. The general inquirer: A computer approach to content analysis.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *NIPS*, pages 5998–6008.

Hongwei Wang, Fuzheng Zhang, Xing Xie, and Minyi Guo. 2018. Dkn: Deep knowledge-aware network for news recommendation. In *WWW*, pages 1835–1844.

Chuhan Wu, Fangzhao Wu, Mingxiao An, Jianqiang Huang, Yongfeng Huang, and Xing Xie. 2019a. Neural news recommendation with attentive multi-view learning. In *IJCAI*, pages 3863–3869. AAAI Press.

Chuhan Wu, Fangzhao Wu, Mingxiao An, Jianqiang Huang, Yongfeng Huang, and Xing Xie. 2019b. Npa: Neural news recommendation with personalized attention. In *KDD*, pages 2576–2584. ACM.

Chuhan Wu, Fangzhao Wu, Suyu Ge, Tao Qi, Yongfeng Huang, and Xing Xie. 2019c. Neural news recommendation with multi-head self-attention. In *EMNLP-IJCNLP*, pages 6390–6395.

Guanjie Zheng, Fuzheng Zhang, Zihan Zheng, Yang Xiang, Nicholas Jing Yuan, Xing Xie, and Zhenhui Li. 2018. Drn: A deep reinforcement learning framework for news recommendation. In *WWW*, pages 167–176.

Qiannan Zhu, Xiaofei Zhou, Zeliang Song, Jianlong Tan, and Li Guo. 2019. Dan: Deep attention neural network for news recommendation. In *AAAI*, volume 33, pages 5973–5980.

Morteza Zihayat, Anteneh Ayanso, Xing Zhao, Heidar Davoudi, and Aijun An. 2019. A utility-based news recommendation system. *Decision Support Systems*, 117:14–27.