Sentiment-guided Sequential Recommendation

Lin Zheng, Naicheng Guo, Weihao Chen, Jin Yu, Dazhi Jiang
{lzheng,19ncguo,19whchen,jinyu,dzjiang}@stu.edu.cn
Department of Computer Science, Shantou University
Shantou, Guangdong, China

ABSTRACT

The existing sequential recommendation methods focus on modeling the temporal relationships of user behaviors and are good at using additional item information to improve performance. However, these methods rarely consider the influences of users' sequential subjective sentiments on their behaviors—and sometimes the temporal changes in human sentiment patterns plays a decisive role in users' final preferences. To investigate the influence of temporal sentiments on user preferences, we propose generating preferences by guiding user behavior through sequential sentiments. Specifically, we design a dual-channel fusion mechanism. The main channel consists of sentiment-guided attention to match and guide sequential user behavior, and the secondary channel consists of sparse sentiment attention to assist in preference generation. In the experiments, we demonstrate the effectiveness of these two sentiment modeling mechanisms through ablation studies. Our approach outperforms current state-of-the-art sequential recommendation methods that incorporate sentiment factors.

CCS CONCEPTS

• **Information systems → Recommender systems**; *Personalization*; *Sentiment analysis*; Collaborative filtering.

KEYWORDS

sequential behavior; sequential recommendation; human temporal sentiment; sentiment-guided attention; sparse sentiment attention

ACM Reference Format:
Lin Zheng, Naicheng Guo, Weihao Chen, Jin Yu, Dazhi Jiang. 2020. Sentiment-guided Sequential Recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20), July 25–30, 2020, Virtual Event, China.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3397271.3401330

1 INTRODUCTION

In modern recommender systems such as Amazon and YouTube, users generate large amounts of sequential behavior and related information. Therefore, processing a large number of sequential behaviors is not only required to meet the actual scenario but also presents a new challenge to the traditional recommendation techniques. Fortunately, research on sequential recommendations that specifically model sequential behavior has yielded many fruitful results [9].

As an initial attempt at applying deep learning to sequential recommendation, GRU4Rec [2] concatenates the behavior sequences of different users in a reasonable manner, thereby solving the difficult problem of uniformly modeling user behavior sequences of different lengths. This method illustrates the important role of recurrent neural networks (RNNs) [7] in sequential recommendation. Taking this research as a starting point, subsequent advanced approaches, such as memory-based RNNs [3, 12] and self-attention-based RNNs [5, 11], have taken sequential recommendation technology to a new level. These advanced technologies not only complete the modeling of pure behavior (item) sequences but also fully characterize additional information such as knowledge [3] and features [11] related to items to improve the sequence prediction performance. Some recent models based on the self-attention mechanism [8] perform particularly well in sequential recommendation. For example, SASRec [5] effectively captures users' long-term preferences from both sparse and dense data sets, and in FDSA [11], two separated self-attention blocks (one for item sequences and one for features) were designed that achieved remarkable predictive effects.

Despite their enhanced performances, *most existing methods consider only the objective information of sequential items; they rarely consider the subjective influence of human sentiments on sequential recommendation.* Specifically, there is a potential correlation between the temporal changes in a user's sentiments and that user's sequential behavior. However, most of the existing models based on sentiment factors [4, 10] focus only on non-sequential recommendations. In other words, *they ignore the influence of the temporal sentiment change pattern on the sequence of user behavior; however, this temporal influence sometimes plays a decisive role in generating the final preference.* For example, under the influence of continuous positive sentiments, a user may review a series of purchased items positively; in contrast, under the influence of continuous negative sentiments, the same user is more likely to consume some items that can usher in a good mood.

Different from the existing approaches, in this work, we focus on how to capitalize on sequential sentiments to improve the sequential recommendation task. To this end, we set up a new preference prediction mechanism from the perspective of users' subjective sentiments and propose a sequential recommendation method that employs temporal sentiments to facilitate generating users' future preferences. Our work makes the following contributions:

Novelty: We design two types of fused sequential sentiment modeling channels: one uses temporal sentiment to guide user behavior; the other generates pure sentiment preferences based on sparse attention. To the best of our knowledge, this is the first work on human sentiment-guided sequential recommendation.

SIGIR '20, July 25–30, 2020, Virtual Event, China

ACM ISBN 978-1-4503-8016-4/20/07.
https://doi.org/10.1145/3397271.3401330

Effectiveness: The dual-channel modeling structure can be easily transplanted into existing RNNs. We report the results of an ablation study to prove experimentally that either or both of the two sentiment channels can effectively improve the performance.

Performance: We report the results of a method comparison to demonstrate that our model outperforms other competing state-of-the-art sequential methods.

2 SENTIMENT-GUIDED SEQUENTIAL RECOMMENDATION

In this section, we formalize the sequential recommendation task based on the sentiment state, introduce the network architecture of the Sentiment-Guided Sequential (SGS) recommender, and then describe the model implementation details.

2.1 Problem Formalization

In the traditional sequential recommendation setting, the goal is to recommend the next item given a user's historical sequence of behaviors $\mathcal{I}^u = (i_1^u, i_2^u, \cdots, i_{|\mathcal{I}^u|}^u)$, where $\mathcal{I}^u$ represents a sequence of items in chronological order that user u has previously interacted with. When considering sentiment factors in sequential recommendation, however, the situation becomes somewhat more complicated. Specifically, each item in the sequence uniquely corresponds to a sentiment state $\mathcal{S}_t^u = (s_t^1, s_t^2, \cdots, s_t^M)^u$, where $\mathcal{S}_t^u$ corresponds to the item i_t^u at time t and M denotes the number of sentiment factors. In a simple case, a sentiment state consists of three factors: positive, neutral, and negative; in this case, $M = 3$, and each factor appears with a certain probability. For instance, at time t, the positive sentiment may account for a probability of 0.6 (represented as $s_t^1 = 0.6$), and the neutral (represented by s_t^2) and negative (represented as s_t^3) sentiments may account for 0.3 and 0.1, respectively. Therefore, the sequential recommendation task based on the sentiment state can be formalized as follows: Given a sequence of (*item, sentiment_state*) tuples $\big((i_1^u, \mathcal{S}_1^u), (i_2^u, \mathcal{S}_2^u), \cdots, (i_{|\mathcal{I}^u|}^u, \mathcal{S}_{|\mathcal{I}^u|}^u)\big)$, the task is to recommend the next item that user u is most likely to interact with. In the following sections, we omit the identity of user u and use a fixed-length sequence $\big((i_1, \mathcal{S}_1), (i_2, \mathcal{S}_2), \cdots, (i_n, \mathcal{S}_n)\big)$ suitable for training to describe our approach.

2.2 The Network Architecture

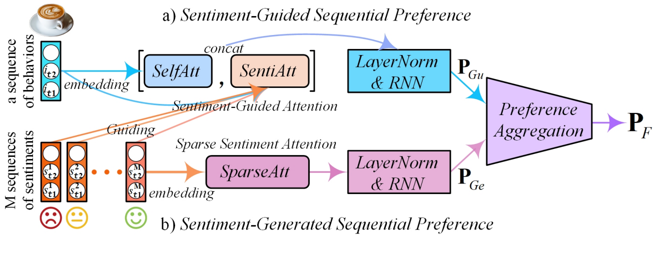

Figure 1: The Architecture of the SGS recommender

The network architecture of the sentiment-guided sequential recommender is shown in Figure 1. We use subscripts $t1, t2, \cdots$ to represent a sequence in chronological order. SGS models users' behavior-sentiment sequences in two channels. In the first channel, to reflect the guiding role of sentiment on the behavior sequence, we design a special attention mechanism based on sentiment queries and then perform sequential preference mining on both self-attention and sentiment-guided attention to obtain the preference of the first channel, $\mathbf{P}_{Gu}$. The second channel emphasizes capturing the sequential preference, $\mathbf{P}_{Ge}$, which is independently generated by sentiments. To this end, we incorporate a variant sparse sentiment attention [1] to model the sentiment factors in the sequence separately because a certain behavior in the sequence corresponds to M sentiment factors, and sparse attention can improve the calculation efficiency by reducing the number of attention interactions in the sequence. Similar to $\mathbf{P}_{Gu}$, we employ a layer-normalization RNN to complete sequential preference modeling to obtain $\mathbf{P}_{Ge}$. Finally, the two preferences are aggregated into a fully connected layer based on layer normalization to generate the final preference, $\mathbf{P}_F$.

2.3 The Implementation Details

2.3.1 Embedding Layer. We divide the fixed-length training sequence $\big((i_1, \mathcal{S}_1), (i_2, \mathcal{S}_2), \cdots, (i_n, \mathcal{S}_n)\big)$ into one item sequence and M sentiment sequences to construct embedding matrices. Because the elements of these $M + 1$ sequences are aligned individually in chronological order, we add zero padding to the left of these sequences to ensure that they have a fixed length of n. Then, we created an item embedding matrix $\mathbf{E} \in \mathbb{R}^{|\mathcal{I}| \times d}$ and M sentiment embedding matrices $\{\mathrm{S}_1, \mathrm{S}_2, \cdots, \mathrm{S}_M\}$, where $\mathrm{S}_1, \mathrm{S}_2, \cdots, \mathrm{S}_M \in \mathbb{R}^{|\mathcal{S}| \times d}$ and d is the number of latent dimensions.

2.3.2 Two types of sentiment attention. The self-attention mechanism has been shown to be highly effective at sequential recommendation [5, 11]. It is based on the scaled dot-product attention [8], which is defined as follows:

$$Att\,(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \mathrm{softmax}(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}})\mathbf{V}\,, \tag{1}$$

where $\mathbf{Q}$, $\mathbf{K}$, and $\mathbf{V}$ represent queries, keys, and values, respectively, and $\sqrt{d}$ is a scale factor that changes with the input dimension. Given the current item embedding $\mathbf{E}$ as input, the self-attention operation transforms $\mathbf{E}$ into queries, keys, and values through linear projection and feeds them to the attention layer:

$$\mathbf{A}_E = SelfAtt\,(\mathbf{E}) = Att\,(\mathbf{E}\mathbf{W}^Q, \mathbf{E}\mathbf{W}^K, \mathbf{E}\mathbf{W}^V)\,, \tag{2}$$

where $\mathbf{W}^Q, \mathbf{W}^K, \mathbf{W}^V \in \mathbb{R}^{d \times d}$ are projection matrices. The self-attention mechanism considers only items as queries; it ignores the role of sentiment factors. To emphasize the guiding role of sentiment factors, we define the sentiment-guided attention of the m-th sentiment factor to items as follows:

$$\mathbf{A}_S^m = SentiAtt\,(\mathrm{S}_m, \mathbf{E}) = Att\,(\mathrm{S}_m\mathbf{W}^Q, \mathbf{E}\mathbf{W}^K, \mathbf{E}\mathbf{W}^V)\,. \tag{3}$$

Because there are M sentiment factors, we perform a linear projection transformation on item self-attention $\mathbf{A}_E$ and sentiment-guided attentions $\{\mathbf{A}_S^1, \mathbf{A}_S^2, \cdots, \mathbf{A}_S^M\}$ and then concatenate them according to the last dimension as follows:

$$\mathbf{G}_{ES}^A = ConCat\,([\mathbf{A}_E\mathbf{W}^E, \mathbf{A}_S^1\mathbf{W}^{S^1}, \mathbf{A}_S^2\mathbf{W}^{S^2}, \cdots, \mathbf{A}_S^M\mathbf{W}^{S^M}])\,, \tag{4}$$

where $\mathbf{W}^E, \mathbf{W}^{S^1}, \cdots, \mathbf{W}^{S^M} \in \mathbb{R}^{o \times o}$ and o is the size of output dimension of the attention layer. $\mathbf{G}_{ES}^A$ is the aggregation of attention to items and sentiments. To model the sentiment-generated preference, we applied a simplified sparse attention mechanism [1] to model pure sentiments and specifically designed its inputs to facilitate interactions between sentiment factors. In particular, we concatenate the M sentiment embedding matrices to obtain $\mathbf{S}_C$ as the input for sparse sentiment attention:

$$\mathbf{A}_P^C = \mathit{SparseAtt}\left(\mathbf{S}_C, r\right) = \mathit{Att}\left(\mathbf{S}_C\mathbf{W}^Q, \mathbf{S}_C\mathbf{W}^K, \mathbf{S}_C\mathbf{W}^V\right) \mid\mid r \ , \qquad (5)$$

where the || operator indicates that only two behaviors with a distance less than or equal to r are used to calculate the attention. We term r as the "sparsity rate" and adjust its value according to the different datasets in the experiments.

2.3.3 Sequential Aggregation to Obtain a Final Preference. We first apply layer normalization to $\mathbf{G}_{ES}^A$ and then use an RNN to model the attention on behavior sequence to obtain the user preferences guided by sentiment as follows.

$$\mathbf{P}_{Gu} = \mathit{RNN}\left(\mathit{LayerNorm}\left(\mathbf{G}_{ES}^A\right)\right) \ . \qquad (6)$$

Here, using a simple RNN for sequential modeling to obtain preference $\mathbf{P}_{Gu}$ involves excluding other factors to directly study the guiding effect of sentiment on behavior sequences. We capitalize on LSTM to instantiate RNN throughout the experiments.

Similar to sentiment-guided user preferences, the preference generated by pure sentiment is calculated as follows:

$$\mathbf{P}_{Ge} = \mathit{RNN}\left(\mathit{LayerNorm}\left(\mathbf{A}_P^C\mathbf{W}^P\right)\right) \ . \qquad (7)$$

Here, $\mathbf{W}^P \in \mathbb{R}^{o \times o}$ and o is the size of output dimension of the attention layer. The two types of preferences generated thus far by the two sequential channels are aggregated into the final preference as follows:

$$\mathbf{P}_F = \mathbf{W}^F \mathit{LayerNorm}\left(\mathit{ConCat}\left(\left[\mathbf{P}_{Gu}, \mathbf{P}_{Ge}\right]\right)\right) + \mathbf{B}^F \ , \qquad (8)$$

Thus, $\mathbf{P}_{Gu}$ and $\mathbf{P}_{Ge}$ are concatenated and then layer normalized as the input to the fully connected layer, which takes the weight $\mathbf{W}^F$ and the bias $\mathbf{B}^F$ as parameters. Finally, we employ binary cross entropy loss and the Adam optimizer to train the SGS network.

3 EXPERIMENTS

In this section, we describe experiments intended to answer the following research questions:

RQ1: Do temporal sentiments have a positive guiding effect on sequential recommendation?

RQ2: Is pure sequential sentiment modeling helpful in generating user preferences?

RQ3: Does SGS outperform current state-of-the-art models, including RNN-based methods?

3.1 Datasets

We conducted experiments on four publicly available datasets. The first two datasets, *Restaurant* and *Bars*, are from the *Yelp Challenge Dataset*[1]. The other two datasets, *Baby* and *Videogames*, were selected from the *Amazon dataset*[2]. We first extracted temporal sentiments from reviews corresponding to sequential behaviors by using the Stanford NLP tools [6] and obtained five categories of sentiment factors: *Very Positive*, *Positive*, *Neutral*, *Negative*, and *Very Negative*, which constitute a sentiment state. Then, to facilitate training, we adopted the data augmentation technique reported in [7] to generate subsequences as input data. The maximum length of the subsequences was set to 50 for all four datasets. Finally, the input data were randomly split into a training set, a validation set, and a test set (8 : 1 : 1) for five-fold cross validation. The dataset statistics after these preprocessing operations are listed in Table 1, where Avg.B/user and Avg.B/item denote "average behavior per user" and "average behavior per item", respectively.

Table 1: The statistics of the four datasets.

Dataset	Users	Items	Avg.B/user	Avg.B/item	Behaviors
Restaurant	1, 264	37, 352	94.9	3.2	0.12M
Bars	1, 263	10, 426	31.7	3.9	0.04M
Baby	470, 361	64, 421	1.7	12.3	0.79M
Videogames	613, 076	41, 361	1.6	23.7	0.97M

3.2 Baselines and Metrics

In these experiments, we compare SGS with the following relevant RNN-based sequential recommendation models:

- **RNN** - an LSTM-based recurrent neural network for sequential recommendation,
- **GRU4Rec**[2] - a representative GRU-based sequential recommender,
- **GRU4Senti**[2] - an improved GRU4Rec model by concatenating user sentiment embeddings with item embeddings,
- **SASRec**[5] - a state-of-the-art self-attention-based sequential recommender,
- **CFSA**[11] - a state-of-the-art self-attention-based sequential recommender that considers sentiments as its features,
- **SGS** - our Sentiment-Guided Sequential recommender.

We fixed $\{\mathit{batch\_size}, \mathit{embedding\_dimension}\}$ of all models to $\{128, 20\}$ on the two smaller datasets (*Restaurant*, *Bars*) and $\{256, 30\}$ on the two larger datasets (*Baby*,*Videogames*), respectively. The initial learning rate was uniformly set to 0.001 and cooperated with the Adam optimizer to complete the training. For the multihead self-attention-based methods (**SASRec**, **CFSA**, **SGS**), we adopted a grid search to find the optimal values for the attention-based parameters. Subsequently, the $\{\mathit{head\_number}, \mathit{size\_per\_head}\}$ of attention were set to $\{5, 4\}$ and $\{6, 5\}$ for the small and large datasets, respectively. A hyperparameter specific to **SGS** is the sparsity rate r, whose optimal values are 3 and 5 for the small and large datasets, respectively. As evaluation metrics, we adopted HitRate (Hit for short) and NDCG to evaluate the performances of all the approaches, and we choose $K = \{10, 20\}$ to report the different results for $\mathit{Hit@K}$ and $\mathit{NDCG@K}$.

3.3 Sentiment Ablation Study

The goal of the sentiment ablation study is to answer both **RQ1** and **RQ2**. We designed three variants of SGS and analyzed their performances:

- **1)** *Remove Sentiment Attention*: a version of SGS without sentiment-guided attention.
- **2)** *Remove Sparse Attention*: a version of SGS without sparse sentiment attention.
- **3)** *Remove All Sentiments*: a version of SGS without any sentiment factors.

[1] https://www.yelp.com/dataset/challenge

[2] http://jmcauley.ucsd.edu/data/amazon/

Figure 2 illustrates that the two SGS versions with sentiment factors outperformed the SGS version without sentiments, which indicates the effectiveness of temporal sentiments in sequential recommendations. On the one hand, adding a sparse attention of pure sentiments to the behavior sequence modeling channel is helpful for improving the preference prediction performance. On the other hand, the performance of the sentiment-guided attention mechanism is much better than that of the self-attention of only behaviors on several metrics, which means that it has a positive influence on the self-attention mechanism. Fusing two types of sentiment attention further improves the performance of SGS and provides positive answers to **RQ1** and **RQ2**.

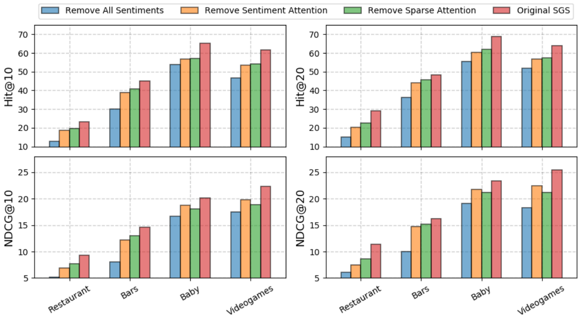

Figure 2: The Performance of sentiment ablation study

3.4 Method Comparisons

Table 2: The performance of method comparisons. The best performances of each dataset are boldfaced, whereas the second best performances are underlined.

Dataset	Methods	Hit@10	NDCG@10	Hit@20	NDCG@20
Restaurant	**RNN**	6.923	2.627	6.491	2.397
	GRU4Rec	7.628	2.804	7.820	3.036
	GRU4Senti	10.085	3.656	11.112	3.917
	SASRec	14.394	5.702	16.548	6.809
	CFSA	18.473	7.288	22.943	8.592
	SGS	**23.216**	**9.325**	**29.017**	**11.436**
	Improvement ↑	25.68%	27.95%	26.47%	33.10%
Bars	**RNN**	15.439	6.065	16.689	7.360
	GRU4Rec	15.932	6.390	16.147	7.009
	GRU4Senti	17.020	7.817	18.025	8.620
	SASRec	32.427	10.836	39.790	12.826
	CFSA	34.500	11.475	39.579	12.785
	SGS	**45.138**	**14.647**	**48.354**	**16.181**
	Improvement ↑	30.83%	27.64%	21.52%	26.16%
Baby	**RNN**	18.945	9.197	20.232	9.948
	GRU4Rec	23.982	10.164	24.200	10.907
	GRU4Senti	24.533	11.378	25.035	11.021
	SASRec	56.862	17.459	59.977	20.093
	CFSA	59.810	19.633	60.761	21.005
	SGS	**65.049**	**22.176**	**68.915**	**23.327**
	Improvement ↑	8.76%	12.95%	13.42%	11.05%
Videogames	**RNN**	23.707	10.099	29.534	11.862
	GRU4Rec	26.435	11.037	34.660	13.058
	GRU4Senti	30.935	12.125	36.640	14.843
	SASRec	54.734	17.781	57.577	19.974
	CFSA	56.904	19.361	58.284	21.086
	SGS	**61.598**	**22.317**	**63.975**	**25.439**
	Improvement ↑	8.25%	15.27%	9.76%	10.19%

Table 2 shows that SGS outperforms all baselines on the four datasets; gaining average improvements of {18.38%, 20.95%, 17.79%, and 20.125%} on {*Hit*@10, *NDCG*@10, *Hit*@20, and *NDCG*@20}, respectively, compared to the strongest baseline. In particular, SGS performs relatively better on the datasets *Restaurant* and *Bars* which include more average user behaviors. Because each user behavior corresponds to 5 sentiment factors, more behaviors means that the model can learn more sentiment information. To answer **RQ3**, we first observed two types of comparative models: {GRU4Rec, GRU4Senti} and {SASRec, CFSA}. As a representative model of sequential recommendation, GRU4Rec becomes GRU4Senti after concatenating the user sentiment embeddings with the item embeddings, and the latter achieves better results. Similarly, the self-attention network based on sentiment features (CFSA) performs relatively better than that without sentiments (SASRec). SGS is superior to both types of approaches because it does not simply organize sentiment factors but instead promotes sentiment factors as preference guides modeled in two channels. Thus, the resulting preferences are more human-centric and more meaningful to users.

4 CONCLUSION AND FUTURE WORK

The intent of this work is to model the subjective preferences of users in sequential recommendation. In future work, we will further consider additional subjective factors (such as emotions) to improve sequential preference mining based on human study.

ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China under Grant No. 61902231, the Guangdong Basic and Applied Basic Research Foundation under Grant No. 2020A1515010531, and the Research Startup Fund Project of Shantou University under Grant No. NTF18017.

REFERENCES

[1] Rewon Child, Scott Gray, Alec Radford, and Ilya Sutskever. 2019. Generating Long Sequences with Sparse Transformers. arXiv:1904.10509 [cs.LG]

[2] Balázs Hidasi, Alexandros Karatzoglou, Linas Baltrunas, and Domonkos Tikk. 2016. Session-based Recommendations with Recurrent Neural Networks. In *Proceedings of ICLR 2016*. arXiv:1511.06939 [cs.LG]

[3] Jin Huang, Wayne Xin Zhao, Hongjian Dou, Ji-Rong Wen, and Edward Y. Chang. 2018. Improving Sequential Recommendation with Knowledge-Enhanced Memory Networks. In *The 41st ACM SIGIR Conference (SIGIR '18)*. 505–514.

[4] Dongmin Hyun, Chanyoung Park, Min-Chul Yang, Ilhyeon Song, Jung-Tae Lee, and Hwanjo Yu. 2018. Review Sentiment-Guided Scalable Deep Recommender System. In *The 41st ACM SIGIR Conference (SIGIR '18)*. 965–968.

[5] W. Kang and J. McAuley. 2018. Self-Attentive Sequential Recommendation. In *2018 IEEE International Conference on Data Mining (ICDM '18)*. 197–206. https://doi.org/10.1109/ICDM.2018.00035

[6] Christopher Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard, and David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit. In *Proceedings of the 52nd ACL*. 55–60.

[7] Yong Kiam Tan, Xinxing Xu, and Yong Liu. 2016. Improved recurrent neural networks for session-based recommendations. In *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*. 17–22.

[8] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. 2017. Attention is All you Need. In *NeurIPS*. Curran Associates, Inc., 5998–6008.

[9] Shoujin Wang, Liang Hu, Yan Wang, Longbing Cao, Quan Z. Sheng, and Mehmet Orgun. 2019. Sequential Recommender Systems: Challenges, Progress and Prospects. In *Proceedings of IJCAI '19*. 6332–6338.

[10] Yuanbo Xu, Yongjian Yang, Jiayu Han, En Wang, Fuzhen Zhuang, Jingyuan Yang, and Hui Xiong. 2019. NeuO: Exploiting the sentimental bias between ratings and reviews with neural networks. *Neural Networks* 111 (2019), 77 – 88.

[11] Tingting Zhang, Pengpeng Zhao, Yanchi Liu, Victor S. Sheng, Jiajie Xu, Deqing Wang, Guanfeng Liu, and Xiaofang Zhou. 2019. Feature-level Deeper Self-Attention Network for Sequential Recommendation. In *Proceedings of IJCAI '19*. 4320–4326.

[12] Lin Zheng, Naicheng Guo, Jin Yu, and Dazhi Jiang. 2020. Memory Reorganization: A Symmetric Memory Network for Reorganizing Neighbors and Topics to Complete Rating Prediction. *IEEE Access* 8 (2020), 81876–81886.