SIFN: A Sentiment-aware Interactive Fusion Network for Review-based Item Recommendation

Kai Zhang[1], Hao Qian[3], Qi Liu[1,2], Zhiqiang Zhang[3], Jun Zhou[3], Jianhui Ma[1], Enhong Chen[1,2]

[1] Anhui Province Key Lab of Big Data Analysis and Application, School of Data Science

[2] School of Computer Science and Technology, University of Science and Technology of China

[3] Ant Group, Hangzhou, China

Review-based Item Recommendation？

- E-commerce platforms (e.g., Amazon, Alibaba) allow users to post their reviews towards products. The reviews may contain the opinions of users and the features of the items.

- Value of the reviews:
    - User-generated content
    - Carrying explanation
    - Rich textual information

Outline

1. **Background**
    a. Formal definition & previous methods
    b. Some existing problems
2. **Our method: Sentiment-aware Interactive Fusion Network**
3. **Experiments**
4. **Conclusion**

Outline

1. **Background**
   a. Formal definition & previous methods
   b. Some existing problems
2. **Our method: Sentiment-aware Interactive Fusion Network**
3. **Experiments**
4. **Conclusion**

Background – formal definition

To begin with, we first define the review-based item recommendation. It estimates the probability that a user rating at the candidate items based on the input feature representation.

predicts the rating $R_{u,i}$ of item i by user u. General schema:

$$R_{u,i} = f(\mathrm{user}\_reviews, item\_reviews)$$

CNN / RNN

Capsule network

Attention methods

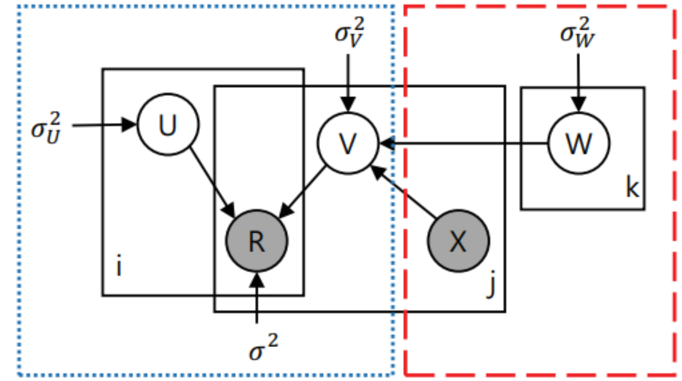
Local & Global Info

Aspect info

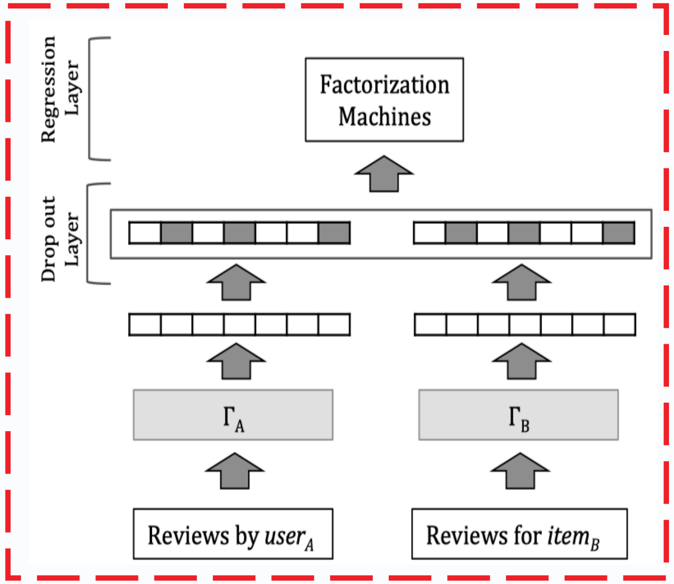
…

Background – previous methods

CNN, RNN based-methods



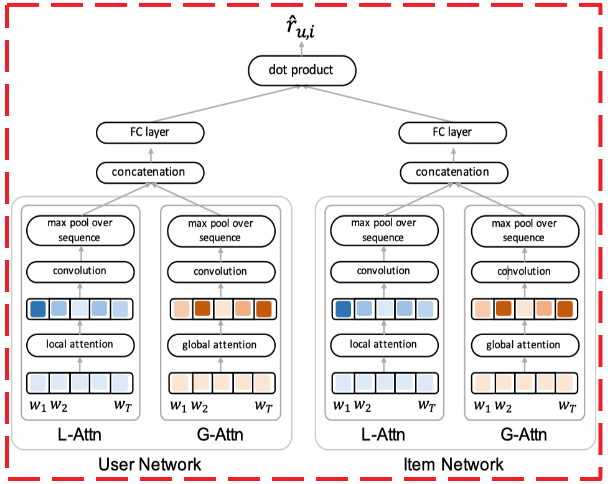
- ✓ ConvMF [D.H. Kim et al. Recsys'16]
  - Convolutional Matrix Factorization
  - Utilizing CNN to extract item latent features from item reviews
- ✓ Parallel CNN model: DeepCoNN [L. Zhang et al. WSDM'17]
  - Deep Cooperative Neural Networks
  - Jointly modeling user & item reviews; Dual CNN
- ✓ TransNet [R. Catherine et al. Recsys'17]



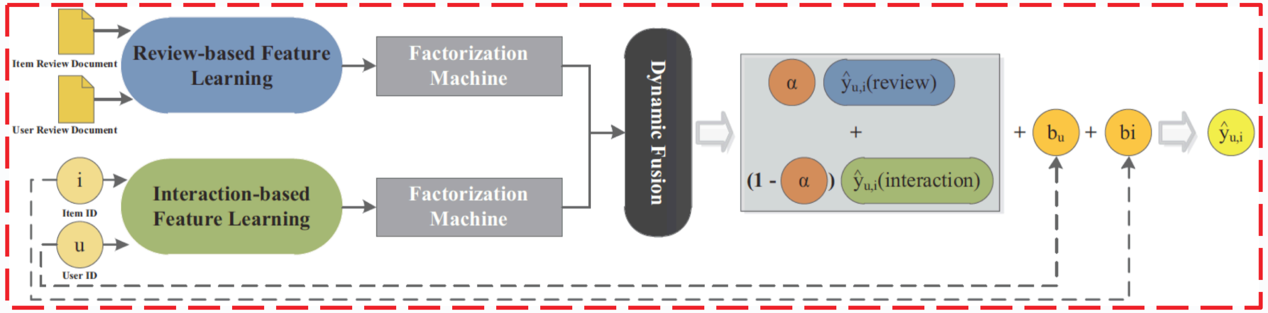
Background – previous methods

Deep Models with Attention for review-based recommendation

- ✓ D-Attn [S. Seo et al. Recsys’17
  - Dual Local and Global attention for selective features
  - Combining local and global attention on review text



- ✓ NARRE [C.Chen et al. WWW’18]
  - Predicting the rating and learning the usefulness of reviews simulateneously
- ✓ MPCN [Tay et al. KDD’18], CARL [Wu et al. TOIS’19], DAML [Liu et al. KDD’19]



Background – shortcomings

Although these works achieved significant performance improvement, they still suffer from **two intrinsic issues**:

1. largely ignoring the explicit sentiment polarity of reviews
2. neglect the personalized interaction of reviews with user/item

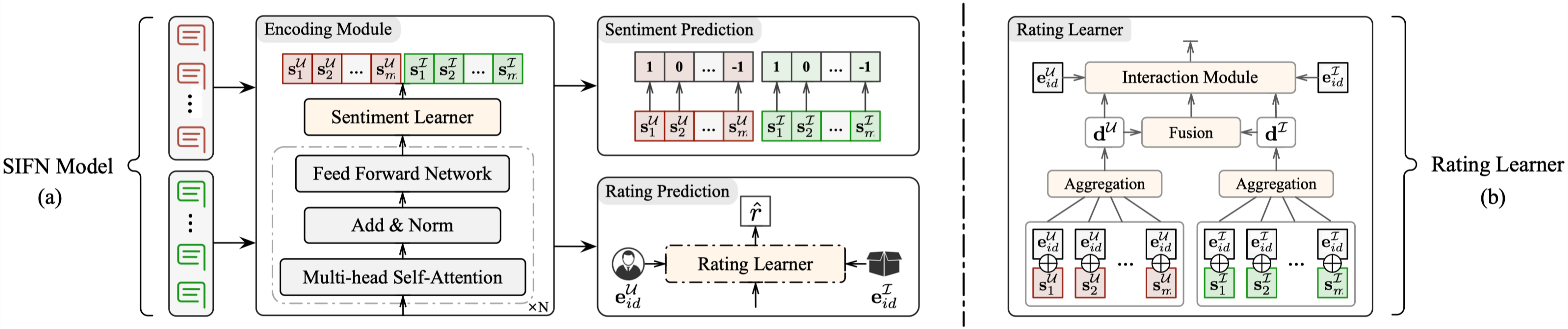
↓

implicitly mining the semantic information and interactions of reviews may lead to sub-optimal prediction because the reviews' sentiment label has not been applied to the training process

Outline

1. Background
    a. Formal definition & previous methods
    b. Some existing problems
2. **Our method: Sentiment-aware Interactive Fusion Network**
3. Experiments
4. Conclusion

Our solution – Sentiment-aware Interactive Fusion Network



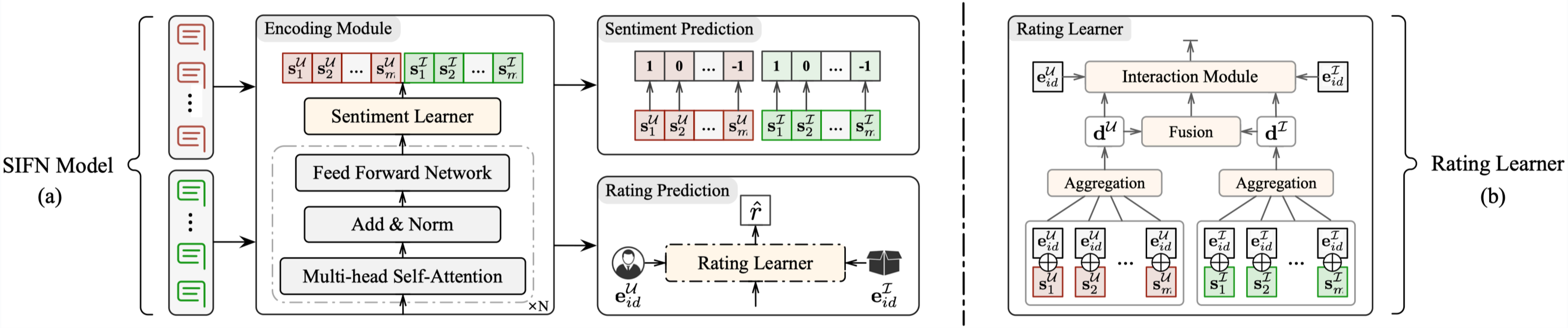
Exploring the explicit sentiment and personalized interaction :

e.g., sentiment for each review

1. There exists a large amount of sentiment labels which carry the user attitudes and preferences (i.e., which kinds of item user may like or dislike)
2. it is necessary to model the interactions between each review and user/item features.

e.g., weigh and fuse

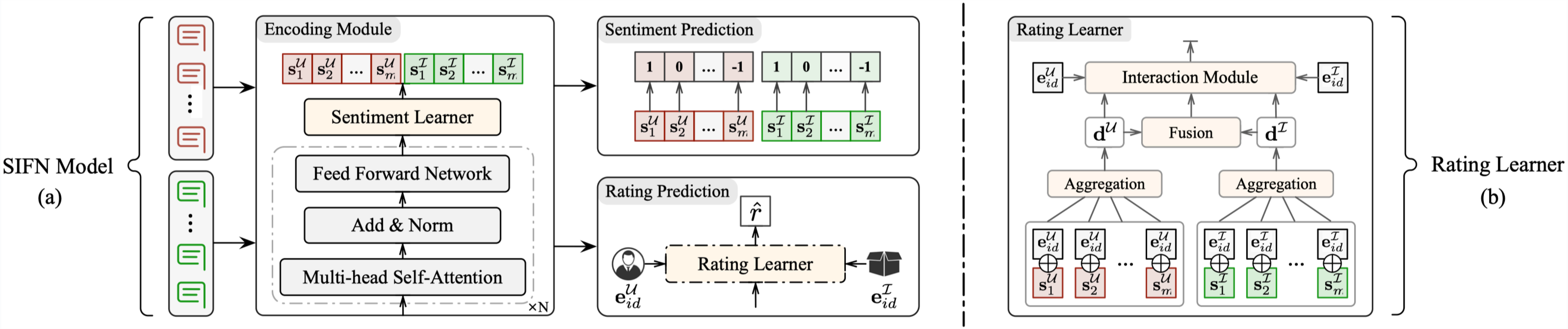
Our solution – Sentiment-aware Interactive Fusion Network



Overall architecture of SIFN:

1. An Encoding Module (BERT embedding + Sentiment Learner)
2. Sentiment Prediction (MLP + softmax)
3. Rating Prediction (Rating Learner : aggregation + fusion)

Our solution – Sentiment-aware Interactive Fusion Network



Overall architecture of SIFN:

BERT$(w_1^{\mathcal{U}}, w_2^{\mathcal{U}}, \ldots, w_l^{\mathcal{U}})$,

$$\alpha_i = \frac{\exp(\tanh(\mathbf{W}_a \mathbf{e}_i^{\mathcal{U}} + \mathbf{b}_a))}{\sum_{i'=1}^{l} \exp(\tanh(\mathbf{W}_a \mathbf{e}_{i'}^{\mathcal{U}} + \mathbf{b}_a))}.$$

1. An Encoding Module (BERT embedding + Sentiment Learner)

$$\mathbf{s}^{\mathcal{U}} = \sum_{i=1}^{l} \alpha_i \mathbf{e}_i^{\mathcal{U}},$$

2. Sentiment Prediction (MLP + softmax)
3. Rating Prediction (Rating Learner : aggregation + fusion)

$$\mathbf{f} = \mathbf{d}^{\mathcal{U}} \mathbf{W}_f \mathbf{d}^{\mathcal{I}},$$

$$\mathbf{p} = \left(\mathbf{d}^{\mathcal{U}} + \mathbf{e}_{id}^{\mathcal{U}}\right) \odot \left(\mathbf{d}^{\mathcal{I}} + \mathbf{e}_{id}^{\mathcal{I}}\right) + \mathbf{W}\mathbf{f} + \mathbf{b},$$

Outline

1. Background
   a. Formal definition & previous methods
   b. Some existing problems
2. Our method: Sentiment-aware Interactive Fusion Network
3. **Experiments**
4. Conclusion

Experiments

Dataset

✓ Benchmark Amazon dataset:

- Music Instruments, Office Products, Digital Music, Tools, Video Games；

Dataset	# Users	# Items	# Ratings	Density
Music Instruments	1,429	900	10,261	0.798%
Office Products	4,905	2,420	53,228	0.448%
Digital Music	5,540	3,568	64,664	0.327%
Tools	16,638	10,217	134,345	0.079%
Video Games	24,303	10,672	213,577	0.089%

Baseline methods

① MF-based methods: PMF; ConvMF+

② Neural-based methods: DeepCoNN; D-Attn; NARRE; CARP

Mean Squared Error (MSE) as the evaluation metric

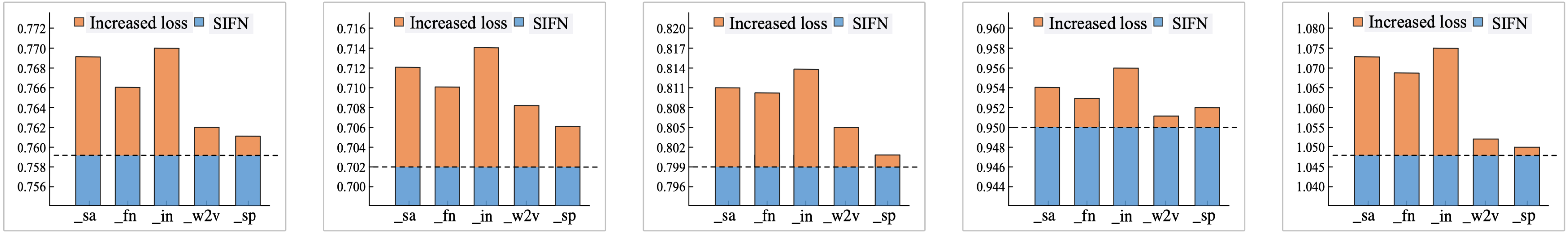
Experiments

Methods	Music Instruments	Office Products	Digital Music	Tools	Video Games	Average
PMF [8]	1.398*(+45.7%)*	1.092*(+35.7%)*	1.206*(+33.7%)*	1.566*(+39.3%)*	1.672*(+37.4%)*	1.386*(+38.6%)*
ConvMF+ [6]	0.991*(+23.4%)*	0.960*(+26.9%)*	1.084*(+26.3%)*	1.240*(+23.4%)*	1.449*+27.7%)*	1.145*(+25.7%)*
DeepCoNN [18]	0.814*(+6.76%)*	0.860*(+18.4%)*	1.058*(+24.5%)*	1.061*(+10.5%)*	1.145*(+8.56%)*	0.988*(+13.9%)*
D-Attn [12]	0.982*(+22.7%)*	0.825*(+14.9%)*	0.911*(+12.3%)*	1.043*(+8.92%)*	1.144*(+8.48%)*	0.981*(+13.3%)*
NARRE [2]	0.803*(+5.48%)*	0.848*(+17.2%)*	0.898*(+11.0%)*	1.029*(+7.68%)*	1.129*(+7.26%)*	0.941*(+9.6%)*
CARP [7]	0.773*(+1.81%)*	0.719*(+2.36%)*	0.820*(+2.56%)*	0.960*(+1.04%)*	1.084*(+3.41%)*	0.872*(+2.4%)*
SIFN	**0.759**	**0.702**	**0.799**	**0.950**	**1.047**	**0.851**

Overall performance

1. MF-based methods (e.g., PMF) consistently fall behind other methods
2. Neural-based methods (e.g., D-Attn) outperform MF-based ones by a large margin
3. our proposed SIFN model still outperforms CARP by 1.81%~3.41%, which shows the superiority of the well designed interactive& Fusion module

Experiments

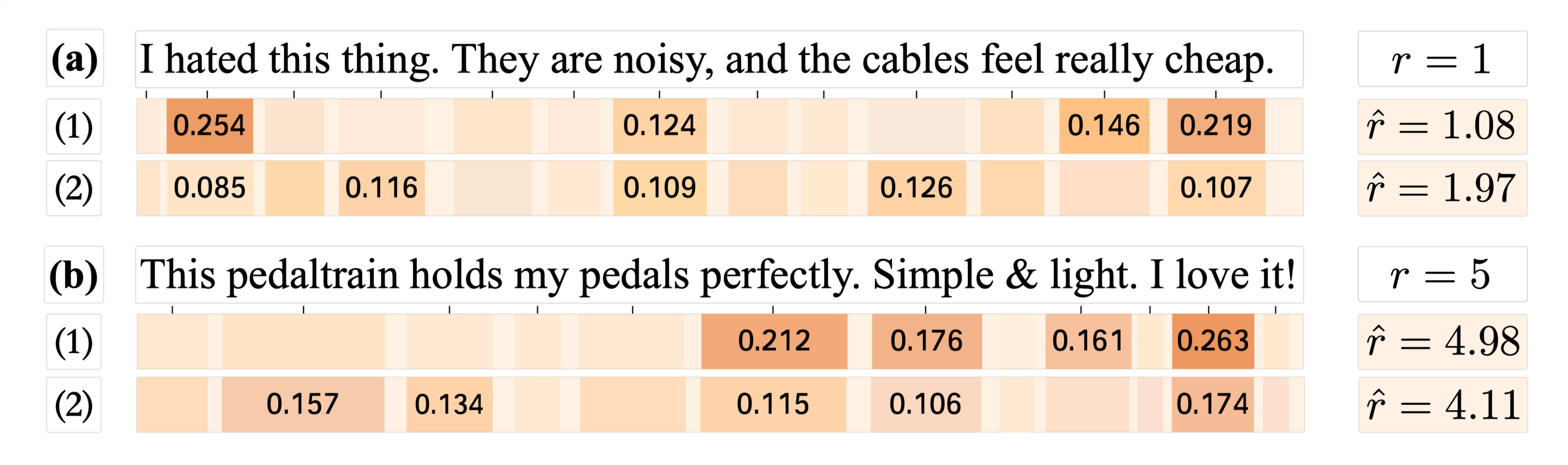


Music Instruments | Office Products | Digital Music | Tools | Video Games

- **SIFN_sa**: replaces the sentence attention with a simple average sum pooling over all the reviews.
- **SIFN_fn**: removes the fusion network so that user and item features are disentangled without explicit interactions.
- **SIFN_in**: replaces the interactive network with commonly used Factorization Machine (FM) [11] to estimate the ratings.
- **SIFN_w2v**: replaces the BERT encoding of text reviews with commonly used pre-trained word embedding GloVe [9].
- **SIFN_sp**: removes the sentiment prediction task so that the model focuses on user-item rating prediction.

- the performance of **SIFN_sa** drops as it just assumes every review contributes equally for user-item rating.
- the performance of **SIFN_fn** also declines since the interactions of user and item reviews are not fully exploited.
- **SIFN_in** also declines because it is inefficient to perform second-order operations in the same space with FM.
- Not surprisingly, without BERT encoding, **SIFN_w2v** is incapable of representing deep semantic of reviews.
- Without sentiment prediction in **SIFN_sp**, there is no supervision towards attending sentiment-aware words in reviews, which are vital for the rating prediction.

Experiments



Case study

1. In review (a), SIFN aligns the sentiment words, e.g., "hated" and "cheap", with a rating of 1.08, which is consistent with actual value
2. In review (b), SIFN accurately predicts a rating of 4.98 by extracting the sentiment words, e.g., "love" and "perfectly", while SIFN_sp is incapable of achieving this

Outline

1. Background
   a. Formal definition & previous methods
   b. Some existing problems
2. Our method: Sentiment-aware Interactive Fusion Network
3. Experiments
4. Conclusion

Conclusion

1. We highlight the explicit sentiment polarity in each review, and focus on modeling the multiple feature interactions between each review and user/item.
2. We propose a novel Sentiment-aware Interactive Fusion Network (SIFN) model with two main components, Sentiment Leaner and Rating Learner.
3. We conduct extensive experiments on five datasets so that the results demonstrate the effectiveness of our proposed method.

References

[1] Donghyun Kim, Chanyoung Park, Jinoh Oh, Sungyoung Lee, Hwanjo Yu, Convolutional Matrix Factorization for Document Context-Aware Recommendation, Proceedings of RecSys, 2016

[2] Zheng L, Noroozi V, Yu P S. Joint deep modeling of users and items using reviews for recommendation, Proceedings of the Tenth ACM International Conference on Web Search and Data Mining (WSDM), 2017: 425-434.

[3] Catherine, Rose, and William Cohen. "Transnets: Learning to transform for recommendation." Proceedings of the Eleventh ACM Conference on Recommender Systems (RecSys). ACM, 2017.

[4] Seo S, Huang J, Yang H, et al. Interpretable convolutional neural networks with dual local and global attention for review rating prediction, Proceedings of the Eleventh ACM Conference on Recommender Systems (RecSys), 2017: 297-305

[5] Chen C, Zhang M, Liu Y, et al. Neural attentional rating regression with review-level explanations, Proceedings of the 2018 Conference on World Wide Web (WWW), 2018: 1583-1592.

[6] Libing Wu, Cong Quan, Chenliang Li, Qian Wang, Bolong Zheng, Xiangyang Luo, A Context-Aware User-Item Representation Learning for Item Recommendation, ACM Transactions on Information Sysems (TOIS), 2019, 37

[7] Donghua Liu, Jing Li, Bo Du, Jun Chang, Rong Gao. DAML: Dual Attention Mutual Learning between Ratings and Reviews for Item Recommendation, Proceedings of KDD 2019.

Thanks !