UNIVERSITÄT KLAGENFURT

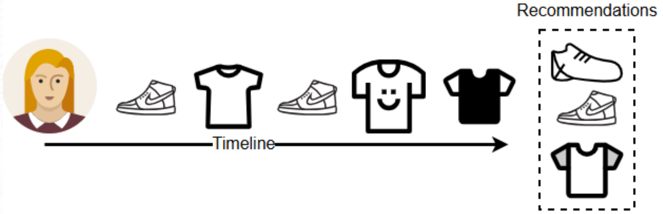
Pyramid Mixer: Multi-dimensional Multi-period Interest Modeling for Sequential Recommendation

Zhen Gong, Faisal Shehzad (Proxy presenter)
Emails: Gongzhen.666@bytedance.com

Introduction

- **Recommender system**
  - A recommender system is software that suggests items (such as videos, songs, or other content) to users based on their preferences

- **Sequential Recommender system**
  - Makes suggestions by analyzing the ***order*** in which a user interacts with items over time

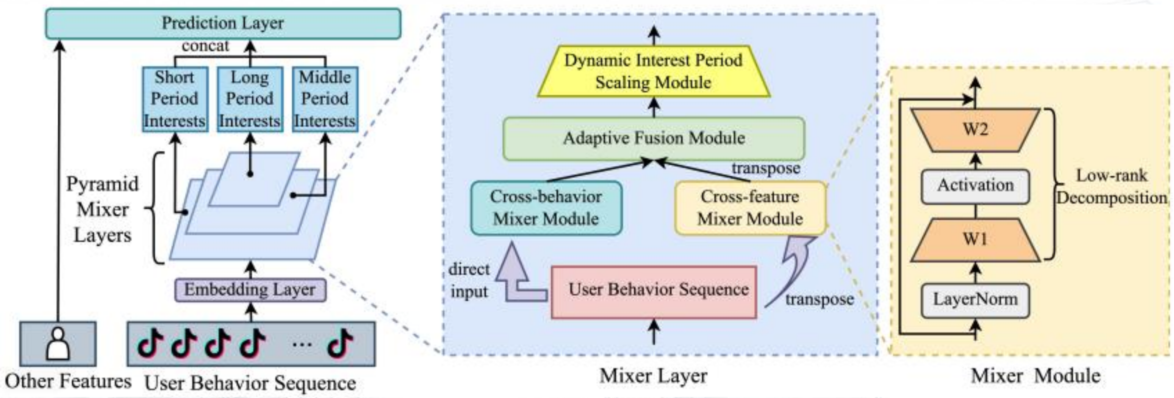


Motivation and problem statement

- Existing studies use attention mechanisms to capture hidden patterns in a user behavior [1, 2, 3], but neglect relationship among and within:
  - Behaviors
  - Features
  - Short and long interests
- Secondly, the computational complexity of the self-attention mechanism makes difficult to apply them in large-scale industrial applications
- However, MLP architectures could be a potential solution for this problem
- MLP architectures demonstrate competitive performance in both academia and industry due to their efficiency and simplicity

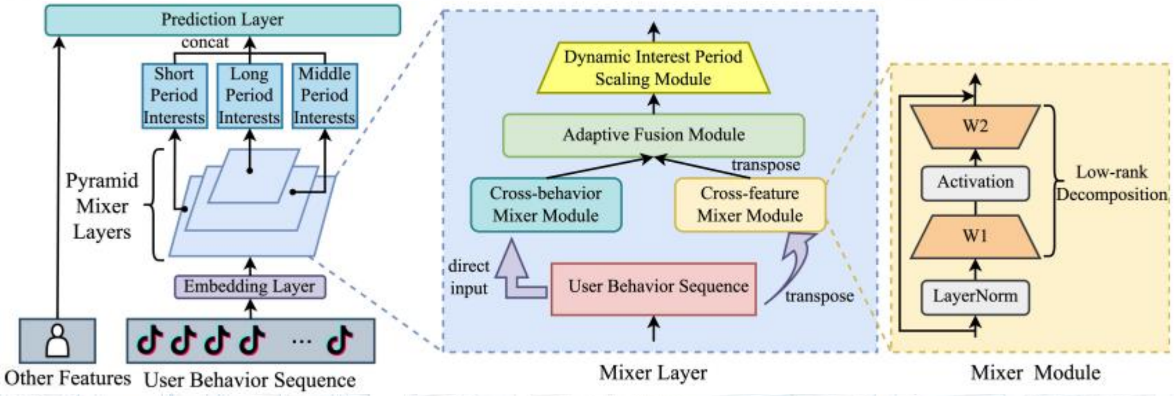
Proposed model (1 / 2)

- In this study, we propose a Pyramid Mixer model



- Contains mixer layers, which process user-embedding sequence at multiple scales
- Within each mixer layer, cross-behavior and cross feature mixer modules are employed
- Both cross-behavior and cross feature modules integrate user behaviors and item features, respectively

Proposed model (2 / 2)



- ❑ Adaptive Fusion module: used to combine output of both modules
- ❑ Pyramid network helps to retrieve short, middle and long period interest of a user
- ❑ Low rank decomposition used to overcome over-parameterization problem
- ❑ Multi-scale cross period module used to learn temporal patterns within and among different time periods

Experiment settings

- Datasets
  - MovieLens-100k
  - MovieLens-1M
  - Beauty dataset from Amazon
  - One private dataset from Douyin app
- Metrics
  - HR@10
  - MRR@20
  - NDCG@10
  - AUC and UAUC
- Baselines
  - Pop, BPR [4], FPMC [5], GRU4Rec [6], SeSRec [7], BERT4Rec [8], FMLP [9], MLP-Mixer [10] and MLP4Rec [11]

Results (1/3)

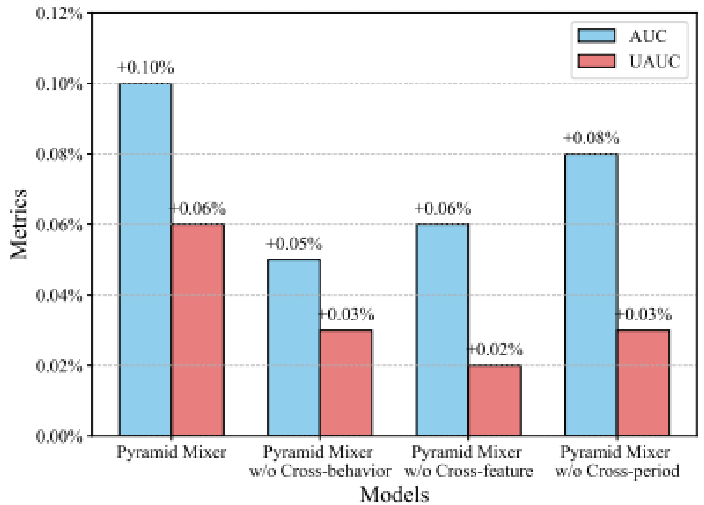
- RQ1: Can our proposed model outperform the state-of-the-art baselines in sequential recommendation tasks?

Datasets / Metrics	HR@10	NDCG@10	MRR@20
MovieLens-100k	Pyramid Mixer	Pyramid Mixer	Pyramid Mixer
MovieLens-1M	Pyramid Mixer	Pyramid Mixer	Pyramid Mixer
Beauty	Pyramid Mixer	Pyramid Mixer	Pyramid Mixer

- Findings
  - Pyramid Mixer outperforms all baseline models on accuracy measures for all three datasets
  - On larger dataset, such as Beauty, the model performs even better that demonstrates its scalability

Results (2/3)

- RQ2: What are the effects of key components of Pyramid Mixer?



- Findings
  - Measure the performance of Pyramid Mixer on a industrial dataset using AUC and UAUC measures
  - Complete Pyramid mixer shows competitive, which demonstrates the effectiveness of each module/component

Results (3/3)

- RQ3: How does the Pyramid Mixer perform in the online settings?

Online experiment core metric results				
Metric	Active days	Active hours	Stay duration	Playtime
A/B result	+0.0403%	+0.0476%	+0.0853%	+0.1106%

Online experiments interactions results				
Metric	Public	Play	Finish	Like
A/B result	+0.093%	0.0538%	+0.1801%	+0.1761%
Metric	Dislike	Comment	Share	Follow
A/B result	-0.0145%	+0.2074%	+0.2284%	+0.4876%

Conclusions / Future work

- In this study, we introduce the Pyramid Mixer model for sequential recommendation
  - In the future, we will test its performance in top-n collaborative filtering and session-based recommendation
  - We plan to conduct further experiments using additional side information such as location and other contextual features
  - Combine the MLP mixer with an attention mechanism
  - Evaluate the performance of the MLP-mixer using embeddings from LLMs models

References

1. Fan, Xinyan, et al. "Lighter and Better: Low-rank Decomposed Self-attention Networks for Next-item Recommendation." *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2021.
2. Kang, Wang-Cheng, and Julian McAuley. "Self-attentive Sequential Recommendation." 2018 IEEE International Conference on Data Mining (ICDM). IEEE, 2018.
3. Wu, Chuhan, et al. "Neural News Recommendation with Multi-head Self-attention." Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2019.
4. Rendle, Steffen, et al. "BPR: Bayesian Personalized Ranking from Implicit Feedback." arXiv preprint arXiv:1205.2618 (2012).
5. Rendle, Steffen, Christoph Freudenthaler, and Lars Schmidt-Thieme. "Factorizing Personalized Markov Chains for Next-basket Recommendation." Proceedings of the 19th International Conference on World Wide Web. 2010.
6. Hidasi, Balázs, et al. "Session-based Recommendations with Recurrent Neural Networks." arXiv preprint arXiv:1511.06939 (2015).
7. Kang, Wang-Cheng, and Julian McAuley. "Self-attentive Sequential Recommendation." 2018 IEEE International Conference on Data Mining (ICDM). IEEE, 2018.
8. Sun, Fei, et al. "BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations From Transformer." Proceedings of the 28th ACM International Conference on Information and Knowledge Management. 2019.
9. Tolstikhin, Ilya O., et al. "Mlp-mixer: An all-mlp Architecture for Vision." Advances in Neural Information Processing Systems 34 (2021): 24261-24272.
10. Li, Muyang, et al. "MLP4Rec: A Pure MLP Architecture for Sequential Recommendations." arXiv preprint arXiv:2204.11510 (2022).