

mechanism to items and attributes respectively, which can effectively model users' varying appetite towards items. To empirically evaluate the effectiveness of the attention mechanism and attribute information respectively, we compare our ANAM with DREAM - an RNN model without attention and attribute information, and a degenerated Neural Attentive Model (NAM) - our attentive RNN model without attribute information.

Results and analysis. We present the results of F1-score@5 and NDCG@5 on the next basket recommendation performance in Table 2.

Table 2: Performance comparisons of different methods on the next basket recommendation task.

Datasets	Ta-Feng		JingDong	
	F1-score@5	NDCG@5	F1-score@5	NDCG@5
TOP	0.051	0.084	0.0066	0.0094
NMF	0.052	0.072	0.0069	0.0097
FPMC	0.059	0.087	0.0078	0.0099
HRM	0.062	0.089	0.0095	0.0174
DREAM	0.065	0.084	0.0122	0.0123
DREAM*	0.133	0.173	0.1046	0.1542
NAM	0.142*	0.187*	0.1283*	0.1826*
ANAM	0.146*	0.190*	0.1313*	0.1842*

Note: “*” indicates the statistically significant improvements (i.e., two-tailed t -test with $p < 0.01$) over the best baseline (i.e., DREAM*).

In our experiments, the embedding dimension in HRM and DREAM is set to 50. For JingDong dataset, we set the learning rate to 0.001 in DREAM, 0.0003 in HRM and 0.001 in FPMC which yield the best results. For Ta-Feng, we use the same training and testing data as HRM, so we report the baseline results as in [5]. We re-implement the DREAM, and modify the objective function as Eq. 8. This change leads to a huge improvement of the original results in [6]. The parameters of our modified DREAM (denoted as DREAM*) are the same as in our ANAM model. We report the results of DREAM* in Table 2.

We can make the following observations:

- TOP is the weakest baseline, since it is a non-personalized method. NMF performs better than TOP, but it does not consider any sequential information of users. FPMC outperforms slightly NMF by taking into account adjacent baskets.
- HRM further improves the effectiveness by using neural network. This shows the ability of neural network to model complex interactions between user's general taste and their sequential behavior. Compared with HRM, DREAM achieves better effectiveness on F1-score@5 and NDCG@5 due to use of whole sequential information.
- DREAM*, which uses a modified loss function, leads to a large boost in effectiveness. This indicates the great importance to weigh the training examples in the training process.
- Our degenerated model NAM consistently and significantly outperforms all baseline methods, showing the effectiveness of our attentive mechanism on item to capture user's evolving appetite for items.

4 CONCLUSION

This paper presented a novel attribute-aware neural attentive model for next basket recommendation, which utilizes a hierarchical attentive architecture to integrate the attribute information of items. ANAM effectively captures the user's evolving appetite for the item by using a joint learning function combining the attentive information of items and attributes. Experimental results on two public datasets (i.e., Ta-Feng and JingDong) demonstrated the effectiveness of our ANAM model for next basket recommendation. This work shows the necessity to model the baskets in the purchase history of a user, and to incorporate the attribute information about items. As future work, we will investigate more attributes of item (e.g., price) and explore the effects of multiple factors on the user's purchase decision. This series of experiments confirms the previous results that the sequential information about baskets provides some useful information for next basket prediction. The superior performance of our model shows that attribute information can further boost the effectiveness.

ACKNOWLEDGMENTS

This work was partially supported by the National Natural Science Foundation of China under Grant No. 61502502, the National Basic Research 973 Program of China under Grant No. 2014CB340403 and the Beijing Natural Science Foundation under Grant No. 4162032. Ting Bai was supported by the Outstanding Innovative Talents Cultivation Funded Programs 2016 of Renmin University of China, and partly by an NSERC discovery grant.

REFERENCES

- [1] Ting Bai, Ji-Rong Wen, Jun Zhang, and Wayne Xin Zhao. 2017. A Neural Collaborative Filtering Model with Interaction-based Neighborhood. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*. ACM, 1979–1982.
- [2] Riccardo Guidotti, Giulio Rossetti, Luca Pappalardo, Fosca Giannotti, and Dino Pedreschi. 2017. Next Basket Prediction using Recurring Sequential Patterns. *arXiv preprint arXiv:1702.07158* (2017).
- [3] Daniel D Lee and H Sebastian Seung. 2001. Algorithms for non-negative matrix factorization. In *Advances in neural information processing systems*. 556–562.
- [4] Steffen Rendle, Christoph Freudenthaler, and Lars Schmidt-Thieme. 2010. Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of the 19th international conference on World wide web*. ACM, 811–820.
- [5] Pengfei Wang, Jiafeng Guo, Yanyan Lan, Jun Xu, Shengxian Wan, and Xueqi Cheng. 2015. Learning hierarchical representation model for nextbasket recommendation. In *Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval*. ACM, 403–412.
- [6] Feng Yu, Qiang Liu, Shu Wu, Liang Wang, and Tieniu Tan. 2016. A dynamic recurrent model for next basket recommendation. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*. ACM, 729–732.
- [7] Xin Wayne Zhao, Yanwei Guo, Yulan He, Han Jiang, Yuexin Wu, and Xiaoming Li. 2014. We know what you want to buy: a demographic-based system for product recommendation on microblogs. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 1935–1944.