Optimizing Session-Aware Recommenders

Subjects: Computer Science, Artificial Intelligence
Contributor: Ming-Yen Lin , Ping-Chun Wu , Sue-Chen Hsueh

Recommendation mechanisms have emerged as vital tools for the filtering of information in various aspects of life. They are widely used in commercial platforms, including e-commerce sites like Amazon. Session-based or session-aware recommendation is more attractive due to the recommendation accuracy.

recommender system | session-aware recommendation | latent-context information

1. Introduction

Our preferences and purchases change over time. To ensure the recommended results align more closely with actual needs, the sequential recommendation system (SRS) has gained prominence [1]. SRS emphasizes continuous interaction records with time-series characteristics, based on the assumption that dependencies exist between interactions. Therefore, all user interaction records are essential for a comprehensive understanding. Traditionally, research in this area often used the recurrent neural network (RNN) as the network architecture, yielding positive results [2][3][4].

The sequential recommendation method utilizes a series of interaction records as its reference basis. To avoid learning incorrect information from irrelevant interaction records, the concept of a session has been introduced. Interaction records within a defined period are considered part of the same session, with interactions processed separately based on the session. This led to the development of both session-based and session-aware recommendations.

The session-based recommender system (SBRS) [2][5][6] aims to reflect users' actual thinking and behavior patterns. It considers only the interaction records within a short period, meaning recommendations are based solely on one session's data. This approach's limitation is its reliance on a narrow data range, leading to less personalized recommendations. It is, however, beneficial for users seeking recommendation system convenience without needing to register or log in.

To facilitate personalized recommendations, the session-aware recommender system (SARS) was developed [7]. SARS involves recommendations comprising multiple sessions [8][9][10][11]. Personalized recommendations typically rely on the user's current interaction to infer their short-term behavioral intentions. While short-term behaviors significantly influence future interests and are thus considered short-term preferences, sessions too distant from the short-term period are classified as long-term preferences. Users have both long-term and short-term preferences integrated into their overall intention preference for recommendations [1][11]. This method's limitations include overlooking context information and a rigid definition of short-term preferences based on the last session, which can limit recommendation adaptability.

The correlation between a product and the current product represents potential information that should be considered in understanding the real intentions of consumers. Additionally, previous studies have often limited the definition of short-term preferences to the last session [11]. This approach can lead to recommendations that are too rigid and inflexible. This rigidity arises because a session's definition is based on time intervals, as mentioned earlier. If a user's intentions span a period that extends beyond the confines of a single session, this behavior pattern may not accurately represent the user's true intentions.

Addressing the relevance of auxiliary contextual information to sequential interaction records, the hidden state in the gated recurrent unit (GRU) is suggested to be used as the model's contextual information [1]. This method ensures the generated contextual information, based on past interaction records, is relevant. It also mitigates the issue of the RNN framework forgetting older information in long-sequence data.

To address the rigidity of defining long-term and short-term user preferences based on session duration, researchers propose a flexible window-based approach. This method utilizes a designated percentage of a user's interaction history to categorize preferences. For example, as illustrated in **Figure 1**, by setting the window (W) at 45%, researchers consider 45% of the interaction records closest to the most recent interaction as indicative of short-term preferences. The remaining interactions are categorized as long-term preferences. This approach more accurately reflects the user's actual preferences by dynamically adjusting the range between long-term and short-term interactions.

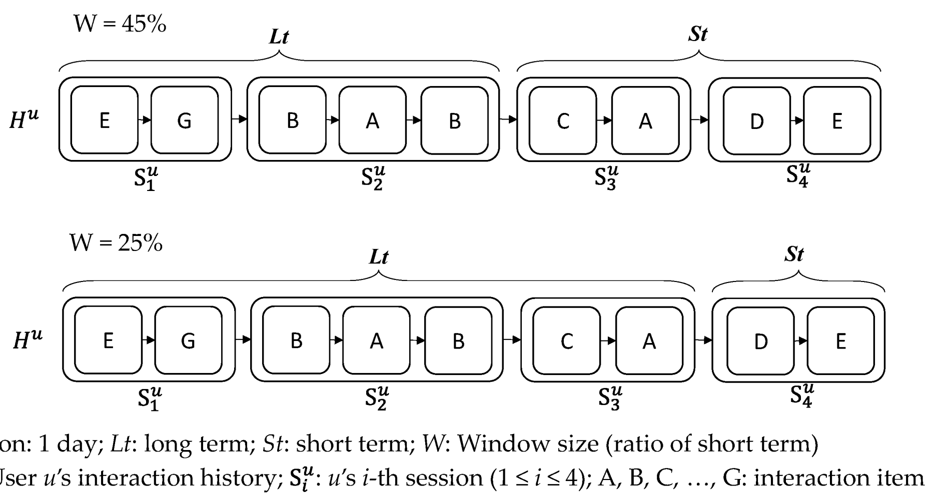

Figure 1. Long-term vs. short-term preferences by varying window scope.

2. Optimizing Session-Aware Recommenders

In traditional recommendation systems, prevalent methods include POP (Most-Popular) and item-based nearest neighbor (Item-kNN) [12]. Subsequent developments have introduced methods like collaborative filtering (CF) [13], matrix factorization (MF) [14][15], and Markov chain (MC) [8]. These methods utilize user-click data as the basis for recommendations, where clicks during browsing represent varying preferences or interests of users, to predict their next item of interest.

Sequential recommendation, a primary branch of recommendation systems, is further divided into session-based and session-aware recommendations. HCA [1], a hierarchical neural network model, focuses on enhancing short-term interests by capturing the complex correlations between adjacent data within each time frame. As RNN tends to gradually forget past information due to long-term dependencies, this method helps the system retain information.

GRU4Rec [2], is a recommendation system model employing the RNN approach. This model continually learns from past features through RNN, combined with the timing of data clicks, to construct a highly effective recommendation method at that time. Compared to traditional methods, RNN adds a sequential consideration, with experimental results underscoring its effectiveness and establishing a neural methods' status in recommendation systems.

Neural session-aware recommendation (NSAR) [9] operates under session-aware recommendation, incorporating all sessions into the model for learning and predicting the next item.

Inter-Intra RNN (II-RNN) [16] is a two-layer RNN architecture model which can effectively enhance recommendation performance and expedite feature learning. This is because the final prediction representation of the inner network is passed to the initial hidden state of the outer network. The outer network, thus, does not start learning from scratch, a method whose effectiveness is proven by this research. CAII [17] is another two-layer GRU model which utilizes session information, including item ID, image characteristics, and item price, to compare these features and achieve a balanced CAII. The CAII-P strategy emerged as the best solution in this research, suggesting that image features do not substantially enhance recommendation quality. This also indicates that session information is not directly correlated with sequential data.

References

1. Cui, Q.; Wu, S.; Huang, Y.; Wang, L. A Hierarchical Contextual Attention-Based GRU Network for Sequential Recommendation. arXiv 2017, arXiv:1711.05114. Available online: https://arxiv.org/abs/1711.05114 (accessed on 23 September 2022).
2. Hidasi, B.; Karatzoglou, A.; Baltrunas, L.; Tikk, D. Session-Based Recommendations with Recurrent Neural Networks. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015; Available online: https://arxiv.org/abs/1511.06939 (accessed on 20 October 2022).

3. Hidasi, B.; Karatzoglou, A.; Quadrana, M.; Tikk, D. Parallel Recurrent Neural Network Architectures for Feature-rich Session-Based Recommendations. In Proceedings of the 10th ACM Conference on Recommender Systems (RecSys '16), Boston, MA, USA, 15–19 September 2016; pp. 241–248.
4. Hidasi, B.; Karatzoglou, A. Recurrent Neural Networks with Top-k Gains for Session-Based Recommendations. In Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM '18), Torino, Italy, 22–26 October 2018; pp. 843–853.
5. Li, J.; Ren, P.; Chen, Z.; Ren, Z.; Lian, T.; Ma, J. Neural Attentive Session-Based Recommendation. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management (CIKM '17), Singapore, 6–10 November 2017; pp. 1419–1428.
6. Song, W.; Wang, S.; Wang, Y.; Wang, S. Next-Item Recommendations in Short Sessions. In Proceedings of the 15th ACM Conference on Recommender Systems (RecSys '21), Amsterdam, The Netherlands, 27 September–1 October 2021; pp. 282–291.
7. Guo, Y.; Zhang, D.; Ling, Y.; Chen, H. A Joint Neural Network for Session-Aware Recommendation. IEEE Access 2020, 8, 74205–74215.
8. Gu, W.; Dong, S.; Zeng, Z. Increasing Recommended Effectiveness with Markov Chains and Purchase Intervals. Neural Comput. Appl. 2014, 25, 1153–1162.
9. Phuong, T.M.; Thanh, T.C.; Bach, N.X. Neural Session-Aware Recommendation. IEEE Access 2019, 7, 86884–86896.
10. Seol, J.J.; Ko, Y.; Lee, S.G. Exploiting Session Information in BERT-Based Session-Aware Sequential Recommendation. In Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '22), Madrid, Spain, 11–15 July 2022; pp. 2639–2644.
11. Gu, G.Y.; Yanxiang, L.; Chen, H. A Neighbor-Guided Memory-Based Neural Network for Session-Aware Recommendation. IEEE Access 2020, 8, 120668–120678.
12. Sarwar, B.; Karypis, G.; Konstan, J.; Reidl, J. Item-Based Collaborative Filtering Recommendation Algorithms. In Proceedings of the 10th International Conference on World Wide Web, Hong Kong, China, 1–5 May 2001; pp. 285–295.
13. Linden, G.; Smith, B.; York, J. Amazon.com Recommendations: Item-To-Item Collaborative Filtering. IEEE Internet Comput. 2003, 7, 76–80.
14. Luo, X.; Zhou, M.; Li, S.; Shang, M. An Inherently Nonnegative Latent Factor Model for High-Dimensional and Sparse Matrices from Industrial Applications. IEEE Trans. Ind. Inform. 2018, 14, 2011–2022.

15. Luo, X.; Zhou, M.; Li, S.; Xia, Y.; You, Z.H.; Zhu, Q.; Leung, H. Incorporation of Efficient Second-Order Solvers Into Latent Factor Models for Accurate Prediction of Missing QoS Data. IEEE Trans. Cybern. 2018, 48, 1216–1228.
16. Ruocco, M.; Skrede, O.S.L.; Langseth, H. Inter-Session Modeling for Session-Based Recommendation. In Proceedings of the 2nd Workshop on Deep Learning for Recommender Systems (DLRS 2017), Como, Italy, 27 August 2017; pp. 24–31.
17. Hsueh, S.C.; Shih, M.S.; Lin, M.Y. Context Enhanced Recurrent Neural Network for Session-Aware Recommendation. In Proceedings of the 28th International Conference on Technologies and Applications of Artificial Intelligence, Yunlin, Taiwan, 1–2 December 2023.