

Session-aware Information Embedding for E-commerce Product Recommendation

Chen Wu¹, Ming Yan¹, Luo Si¹

¹Alibaba Group, 969 West Wenyi Road, Hangzhou 311121, China
{wuchen.wc, ym119608, luo.si}@alibaba-inc.com

ABSTRACT

Most of the existing recommender systems assume that user's visiting history can be **constantly recorded**. However, in recent online services, the user identification may be usually unknown and only **limited online user behaviors can be used**. It is of great importance to **model the temporal online user behaviors and conduct recommendation for the anonymous users**. In this paper, we propose a list-wise deep neural network based architecture to **model the limited user behaviors within each session**. To train the model efficiently, **we first design a session embedding method to pre-train a session representation, which incorporates different kinds of user search behaviors such as clicks and views**. Based on the learnt session representation, we further propose a list-wise ranking model to generate the recommendation result for each anonymous user session. We conduct quantitative experiments on a recently published dataset from an e-commerce company. The evaluation results validate the effectiveness of the proposed method, which can outperform the state-of-the-art.

CCS CONCEPTS

•Computer systems organization → Embedded systems; Redundancy; Robotics; •Networks → Network reliability;

KEYWORDS

session-aware; e-commerce recommendation; **product embedding**

1 INTRODUCTION

Nowadays, hundreds of millions of people use the e-commerce website for satisfying their daily demands. Due to some privacy reason, more and more users now prefer to surf the e-commerce website anonymously without login in. For example, in the released dataset from real search engine of DIGNETIA¹, about 62.3% of all user sessions are non-logged users. For these users, there exist no historical behavior records for them and it is difficult for the recommender system to provide accurate recommendation due to a lack of adequate user behaviors.

However, to find the right product, users in e-commerce website usually leave valuable online footprints when searching for it. For

example, they may trigger multiple queries within each session or conversation to find a better result and click different products of the same category back and forth to compare the prices. Although limited, these instant search behaviors provide valuable information to understand the user's search intention, which can also help the system to better adapt to the user's subsequent information needs. Therefore, the aim of this paper is to generate recommendations for the anonymous users in e-commerce website, by leveraging the user's temporal behavioral information within the current session, such as clicks, views and purchases.

Traditional recommender systems mainly rely on the long-term user behavior history for recommendation, which follow the idea of collaborative filtering [8]. However, the instant online user-item interactions are not well considered, especially in the e-commerce search scenario. The challenge is two-fold: (1) the available user online information is sparse and limited; (2) different kinds of user behaviors exist in the e-commerce website, such as clicks, views and purchases, how to combine all the different user behaviors for a better user understanding.

To address the above challenges, we **propose a list-wise deep learning model for solution**. First, we present a session embedding method to represent each session with all the available user online behaviors. Different kinds of user behaviors are separately embedded and combined together with multi-layer deep neural networks. Based on the learnt session representation, a dependent list-wise ranking model is then used to calculate the relevance between the session and the candidate products. To accelerate the training process, the pre-train mechanism is introduced in list-wise ranking.

The main contributions of this paper can be summarized as: 1) we propose to address the cold-start recommendation problem for the anonymous users in e-commerce scenario by leveraging the diverse online search behaviors within each session; 2) a list-wise deep neural network based framework with pre-training mechanism is designed to model the user's online session behaviors for recommendation; 3) quantitative experiments on a real-world dataset demonstrate the effectiveness of the proposed method over the state-of-the-art recommendation methods.

2 RELATED WORK

Research on recommender systems has been studied for years, in which collaborative filtering is one of the most popular techniques. It relies on the users' collaborative behaviors on items for recommendation, among which KNN [7] and latent factor models [5] are the traditional state-of-the-art methods. To further base the recommendation on the implicit feedback, a generic learning algorithm BPR [6] is proposed and applied to the KNN and MF methods. Recently, deep learning has also been applied to the recommender

¹<https://competition.codalab.org/competitions/11161>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions.acm.org.

CIKM'17, Singapore, Singapore

© 2017 ACM. 978-1-4503-4918-5/17/11...\$15.00

DOI: 10.1145/3132847.3133163

systems. For example, [2] presents a wide and deep learning method to jointly combine the id features and continuous features.

Traditional collaborative filtering-based methods usually face the problem of data sparsity and will fail for the cold-start users. The proposed solution aims to leverage the instant online search behaviors within session for recommendation, which helps to address the new user problem. Some recent research also attempts to model the user’s sequential session behaviors. In [3], a RNN-based method is proposed to model the sequential session behaviors for more accurate recommendation. [9] provides a dynamic recurrent model for next basket recommendation. These methods mainly generate recommendations based on a sequence of explicit behavior records, but not distinguish among different types of behaviors. Our proposed method formulate this task as a learning to rank problem, and different types of session information are combined in a list-wise ranking framework.

3 SESSION-AWARE RECOMMENDATION

DEFINITION 1 (SESSION): We define *user session* as an *uninterrupted sequence of user activities (e.g. clicks, purchases) in the e-commerce system, which is common in the web search field.*

Given a collection of anonymous user session logs in e-commerce website (i.e., user clicks, views and purchases), and a collection of presented products, the goal of the session-aware recommendation is to re-rank the presented products according to the user’s previous online behaviors within the same session.

3.1 Deep ListNet Ranking Framework

The overall solution framework mainly consists of two parts: *Session Information Embedding (S-IE)* and *List-wise Ranking with S-IE*. In the first part, different types of session behaviors are combined to obtain a compact representation for each session. The main reason behind this part lies in that: 1) to obtain an accurate session embedding in a supervised way, which can help give a better initialization for the subsequent ranking model; 2) to provide a session embedding and item embeddings, which serves as a pre-training mechanism to accelerate the ranking procedure in the next part. To improve the recommendation performance, the second part is designed to further re-rank the presented products under a multi-layer list-wise ranking framework.

3.2 Session Information Embedding (S-IE)

This section introduces *how we combine the different types of behaviors within each session to obtain a compact session representation.* The overview of this model is illustrated in Figure 1, which mainly consists of two parts: *session feature embedding* and *supervised session representation learning.*

In session feature embedding, given a target user u and a presented item i in session s , it is reasonable to assume that user u ’s click and view behaviors before this item presentation within the session are very significant. Therefore, we design an embedding method to take advantage of all the available session behaviors with behavior vectors. Specifically, we aggregate the user’s previous clicks and views separately, and use a pooling strategy (max pooling or average pooling) to generate a session click vector and view vector. Then concatenated with the user id embedding and item id embedding, the final embedding vector can well capture the user intention in this session.

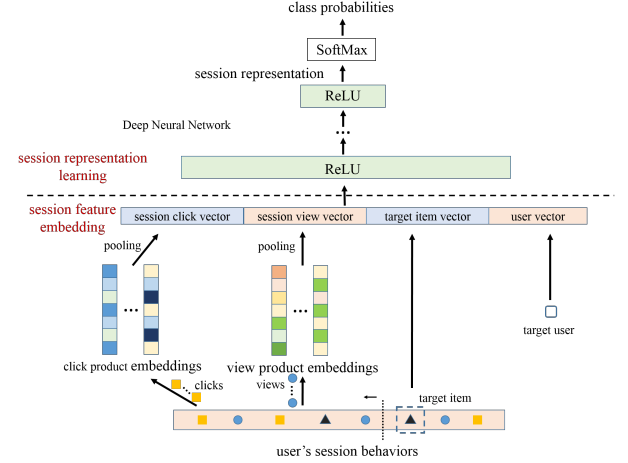


Figure 1: session information embedding model.

To better combine these different types of user behaviors for a comprehensive session understanding, a deep neural network method with full-connected relu layers is further used for learning combination weights on the derived behavior vectors. We supervise the learning of session representation as a two-class classification problem and the real user’s click and purchase behaviors are used as supervision. Specifically, we use the total click behaviors and re-sample the purchase behaviors three times as the positive samples². The presentation behaviors with no clicks are used as negative samples and negative sampling is also used for efficiency consideration. We use the cross-entropy loss as our objective function and the network parameters are updated accordingly.

Finally, we extract the last relu layer representation as the session representation. For each session s , after the session representation learning, we obtain its completed session representation \mathbf{s} . Actually, the class probabilities from the softmax layer can already give a coarse ranking for the candidate products in each session.

3.3 List-wise Ranking with S-IE

In session information embedding stage, the obtained coarse ranking does not consider the relative item order within each user session. Therefore, to further refine the recommendation result, we propose a list-wise dependent ranking method to calculate the relevance score between the obtained session representation and the candidate items. To make our list-wise ranking method more computationally efficient and take the most advantage of the obtained session representation, we depend our ranking model on the session representation from the S-IE part and fine-tune the top item embeddings in a list-wise ranking framework. The overview of the ranking framework is presented in Figure 2.

The main goal of this part is actually to match the obtained session representation with representations of the candidate items. The idea is analogous to that of item retrieval, only that there is no query information in our recommendation scenario and we rely on the derived session representation for matching. The proposed list-wise learning method is based on the typical ListNet method [1],

²Resampling the purchase behaviors shows good performance in practice.

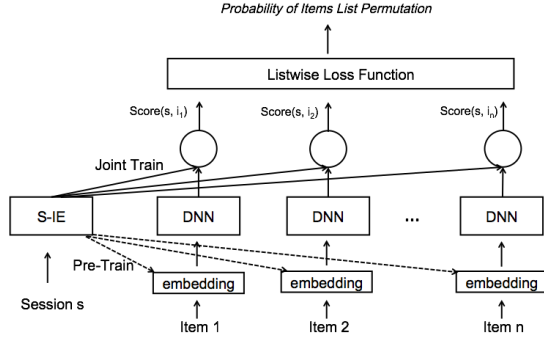


Figure 2: Deep ListNet with SIE

and a full-connected DNN network is introduced to fine-tune the item representations with a list-wise ranking loss.

Specifically, given a user session s and a list of presented items $I = \{i_1, \dots, i_n\}$, we first pass the items I to an embedding layer, so as to obtain a fixed-length vector representation of the items. To better couple with the session representation from the first part, we also pre-train the item embeddings with the SIE model and use the embeddings as the initial item representations. Then, to make it possible to match the session representation s with the item representation i in the same latent space, a full-connected DNN layer is further introduced to project the item embeddings to the session space, which share the same dimension size with the session representation. The score function for a typical session s and item i can be given directly as: $score(s, i) = s \cdot i$.

For the list-wise ranking, we denote the ranking function based on the deep neural network model w as f_w . Given a training sample x_j^i for item j of session i , $f_w(x_j^i) = score(i, j)$ assigns a score for each sample. For the total item list of session i , the ranking function f_w can generate a score list $z^i(f_w) = (f_w(x_1^i), f_w(x_2^i), \dots, f_w(x_n^i))$.

In e-commerce recommendation, people usually just search for the first several pages for their needs and the top-ranked results are what actually matters much. Therefore, we use the typical top-k probability model to calculate the list-wise loss function, which only considers the list order of the top-ranked results. Ranked by the scores, the top k probability of items is calculated as

$$P_{z^i(f_w)}(\mathbb{G}_k(j_1, j_2, \dots, j_k)) = \prod_{t=1}^k \frac{\exp(f_w(x_{j_t}^i))}{\sum_{l=t}^n \exp(f_w(x_{j_l}^i))} \quad (1)$$

where $\mathbb{G}_k(j_1, j_2, \dots, j_k)$ is the top k sub-group and n^i is the total sample number of session i .

With Cross Entropy as metric, the loss for each session list becomes:

$$L(y^i, z^i(f_w)) = - \sum_{\forall g \in \mathbb{G}_k} P_{y^i(i)}(g) \log(P_{z^i(i)}(g)) \quad (2)$$

where y^i is the actual behavior label, and $y_j^i \in \{0, 1, 2\}$ represent the presentation, click and purchase behavior, respectively.

The gradient of the loss function with respect to parameter ω is

$$\Delta\omega = \frac{\partial L(y^i, z^i(f_w))}{\partial \omega} = - \sum_{\forall g \in \mathbb{G}_k} \frac{\partial P_{z^i(i)}(g)(f_w)}{\partial \omega} \frac{P_{y^i(i)}(g)}{P_{z^i(i)}(g)} \quad (3)$$

Algorithm 1 Learning Algorithm of ListNet-SIE

Input:

Training data: $\{(x^1, y^1), (x^2, y^2), \dots, (x^n, y^n)\}$

Parameters: number of iterations T and learning rate η

Initialize parameter ω

1: **for** $t = 1$ to T **do**

2: **for** $i = 1$ to n **do**

3: Input x^i joint S-IE s^i to Neural Network and compute score list $z^i(f_w)$ with current ω

4: Compute gradient $\Delta\omega$ using Eq. (4)

5: Update $\omega = \omega - \eta \times \Delta\omega$

6: **end for**

7: **end for**

Output: Neural Network model ω

Each session with the contained item behaviors is treated as a list-wise training sample and the Gradient Descent method can be then used to update the model parameters ω accordingly. Algorithm 1 shows the learning algorithm of ListNet-SIE. Finally, with the learnt DNN weight W_{DNN}^T and the item embedding i_e , the ranking score for the target item i in session s is:

$$Score(s, i) = s \cdot W_{DNN}^T i_e \quad (4)$$

The presented results are finally re-ranked according to $score(s, i)$.

4 EXPERIMENTS

4.1 Dataset

For the task of e-commerce product recommendation, we perform experiments on a publicly released dataset of the track 2 of CIKM Cup 2016³. This dataset contains anonymized users' six-month click, view and purchase records on an e-commerce search engine, and the user session is extracted using a 1-hour of inactivity heuristic. To focus on our session-aware recommendation scenario, we selected all the "query-less" sessions in this dataset and filtered out the sessions with less than two queries. The last query in each user session is reserved as the test set, the others are used for training. The statistics of the dataset is summarized in Table 1.

4.2 Experimental Settings

In session information embedding, the embedding size of each type of behavior is 200. Three relu layers with hidden unit size of 800, 200, 100 are used to get good performance. In supervised learning, negative sampling is adopted and the ratio of positive v.s. negative samples is 1:5, which can already give an optimal result. In list-wise ranking, we use two full-connected neural network layers with sigmoid activation function, where the hidden unit sizes are both 100. For the top k probability model, we choose k as 10 for the page presentation situation and efficiency consideration. The learning rate η is set to 0.001 to ensure the convergence of the algorithm.

To evaluate the effectiveness of the proposed two-stage solution, we implemented five typical baselines and two different settings of our solution. The seven examined methods are listed as below:

- **Popularity:** the simple personalized re-ranking baseline provided by the official site;

³<https://competitions.codalab.org/competitions/11161>

Table 1: Basic statistics of the dataset.

Statistics	Value
#users	21,930
#sessions	27,644
#query-less queries	82,368
#presented products	112,685
#click logs	160,521
#view logs	181,180
#purchase records	43,666
#avg.(show items) per query	146.5

Table 2: The NDCG@all results with different behavior vector representations in SIE part.

	no click & view	only click	only view	both
SIE	0.359	0.385	0.376	0.393
ListRank	0.368	0.393	0.384	0.400

- *KNN*: the typical item-based collaborative filtering recommendation algorithm [4];
- *LFM*: state-of-the-art Latent Factor Model [5], which is mainly designed to address the sparsity problem;
- *BPR*: the typical bayesian personalized ranking method with the implicit feedback [6];
- *S-RNN*: a typical session-based recommendation method with recurrent neural networks [3];
- *SIE*: the proposed solution that directly use the class probabilities from session information embedding part for ranking;
- *ListRank*: the proposed solution considers both the session information embedding and deep ListNet ranking.

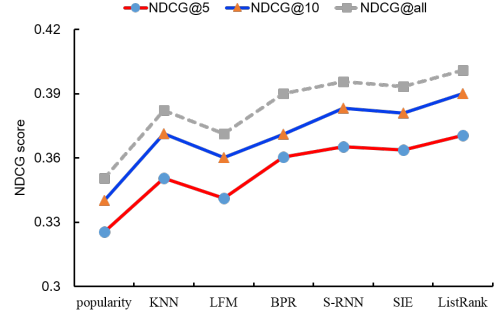
We adopt the NDCG metrics, which also serves as the final competition goal. The NDCG score is defined on each test query, and the final result is averaged over all the test queries.

4.3 Experimental Results and Analysis

The NDCG results of all the examined methods are shown in Fig. 3. Thus it can be seen: (1) compared with other typical methods, LFM and KNN get the worst performance, which may be due to that the user behaviors are not adequate in our session-based scenario; (2) by considering the online session information, *SIE* and our proposed methods can outperform other baselines, while the proposed *ListRank* method is slightly better for the consideration of combination of different types of session behaviors; (3) *ListRank* further improves the performance over the *SIE* method, which validates the effectiveness of the list-wise ranking procedure.

Influence of different behavior embeddings. To further investigate what session information contributes most to the final performance, we also examine with different session vector representations in the SIE part: (1) with no previous session information; (2) with only previous click vector representation; (3) with only previous view vector representation; and (4) with both representations. The comparative results are shown in Table 2.

From the results, we can see that: (1) with either click or view information embedded in session representation, the recommendation performance gets improved, which shows the importance of the previous session behaviors. (2) The click behavior is slightly more important than the view behavior, which may be due to that only click and purchase behaviors indicate the positive labels to

**Figure 3: The NDCG results of all the examined methods.**

calculate the NDCG formula. (3) Combining both the click and view behaviors for session representation learning leads to better recommendation performance in both proposed methods, which also shows the effectiveness of our method for combining different types of online session behaviors.

5 CONCLUSIONS

In this paper, we have proposed a novel session-aware recommendation method to address the recommendation problem of anonymous users surfing in e-commerce website. The overall framework mainly consists of two parts: session information embedding and list-wise ranking. Firstly we embed each session with multiple user session behaviors and different types of online session behaviors are combined for a comprehensive session understanding. Then a deep list-wise ranking model is used to re-rank the presented results for better recommendation. The experiment results validate the effectiveness of the proposed method in modeling different types of online session information. In the future, we are also considering designing dynamic neural network models such as LSTM to better understand the user's sequential online behaviors.

REFERENCES

- [1] Z. Cao, T. Qin, T.-Y. Liu, M.-F. Tsai, and H. Li. Learning to rank: from pairwise approach to listwise approach. In *Proceedings of the 24th international conference on Machine learning*, pages 129–136. ACM, 2007.
- [2] H.-T. Cheng, L. Koc, J. Harmsen, T. Shaked, T. Chandra, H. Aradhye, G. Anderson, G. Corrado, W. Chai, M. Ispir, et al. Wide & deep learning for recommender systems. In *Proceedings of the 1st Workshop on Deep Learning for Recommender Systems*, pages 7–10. ACM, 2016.
- [3] B. Hidasi, A. Karatzoglou, L. Baltrunas, and D. Tikik. Session-based recommendations with recurrent neural networks. *arXiv preprint arXiv:1511.06939*, 2015.
- [4] G. Karypis. Evaluation of item-based top-n recommendation algorithms. In *Proceedings of the tenth international conference on Information and knowledge management*, pages 247–254. ACM, 2001.
- [5] Y. Koren. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *ACM SIGKDD 2008*, pages 426–434. ACM, 2008.
- [6] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme. Bpr: Bayesian personalized ranking from implicit feedback. In *Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence*, pages 452–461, 2009.
- [7] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl. Item-based collaborative filtering recommendation algorithms. In *WWW*, pages 285–295. ACM, 2001.
- [8] X. Su and T. M. Khoshgoftaar. A survey of collaborative filtering techniques. *Advances in artificial intelligence*, 2009:4, 2009.
- [9] F. Yu, Q. Liu, S. Wu, L. Wang, and T. Tan. A dynamic recurrent model for next basket recommendation. In *SIGIR 2016*, pages 729–732, 2016.