AIRec: Attentive Intersection Model for Tag-Aware Recommendation

Bo Chen[1], Dong Wang[1], Yue Ding[1], and Xin Xin[2]

[1] Shanghai Jiao Tong University {chenbo.31, wangdong, dingyue}@sjtu.edu.cn
[2] University of Glasgow x.xin.1@research.gla.ac.uk

Abstract. Tag-aware recommender systems (TRS) utilize rich tagging information to better depict user portraits and item features. Existing methods fail to capture multi-aspect user preferences and lack of exploration of tags intersection. In this work, we propose attentive intersection model (AIRec) to address these issues. User representations are constructed via a hierarchical attention network, where the item-level attention differentiates the contributions of interacted items and the preference-level attention discriminates the saliencies between explicit and implicit preferences. Besides, the tags intersection is exploited to enhance the learning of conjunct features. Finally, we combine factorization machines (FM) with BPR for score prediction. Experiments on two real-world datasets demonstrate significant improvements of AIRec over state-of-the-art methods for tag-aware top-n recommendation.

1 Introduction

Social tagging systems, also known as folksonomies, are widely used in various websites, where users can freely annotate online resources (e.g., movies, artists) with arbitrary tags. These tags are composed by laconic words or phrases, which can not only indicate user preferences, but also summarize features of items. Consequently, user-defined tags can be introduced into recommender systems for alleviating the cold-start problem and improving recommendation quality.

To solve the problem of sparsity, ambiguity and redundancy in tag space, some neural networks-based methods are proposed by converting the tag space into dense latent space, such as CFA [5], DSPR-NS [3] and TRSDL [1]. Although these models have made some progress, there are some weaknesses that hinder their performance. They construct user representations by either explicit tagging behaviors (e.g., DSPR-NS) or implicit interacted items (e.g., TRSDL), which is inadequate to capture multi-aspect user preferences. The intersection of user and item tags reflects the diverse focuses of different users, which is the key incentive of user-item transactions. Unfortunately, seldom research has explored this field.

Dong Wang is the corresponding author.

In this paper, we focus on developing solution to address the drawbacks mentioned above and propose an Attentive Intersection Recommendation model (AIRec) for TRS. Compared to the previous models, our method not only takes both explicit and implicit preferences into consideration for capturing more accurate user portrait via hierarchical attention network, but also makes full use of the tags intersection to improve performance.

2 The AIRec Model

In this section, we will present the architecture of our proposed AIRec model and explain the training procedure. Figure 1 illustrates the structure of our model.

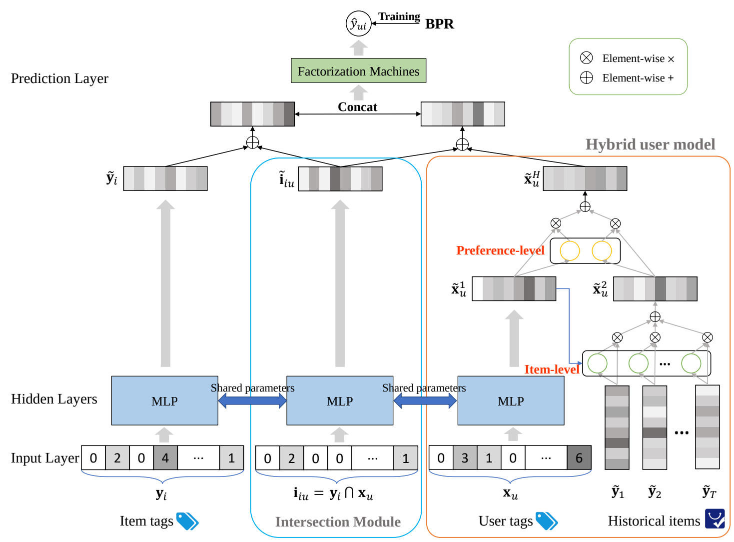

Fig. 1. The structure of AIRec model.

Input Layer and Hidden Layers The user feature vector is constructed as $\mathbf{x}_u = (p_1^u, p_2^u, ..., p_V^u)$, where V is the size of tag set and $p_j^u = |\{(u, i, t_j) \in A | i \in I\}|$ is the number of times that user u annotates items with tag t_j. Similarly, the item feature vector can be represented as $\mathbf{y}_i = (q_1^i, q_2^i, ..., q_V^i)$.

To solve the problem of sparsity and high-dimension, $\mathbf{x}_u$ and $\mathbf{y}_i$ are fed into the multi-layer perceptrons (MLPs) with shared parameters. Sharing parameters can not only obtain better generalization capability and less computational overhead, but force networks to use the same feature space to describe user and item. The latent representations of user and item are $\tilde{\mathbf{x}}_u^1 = \mathbf{h}(\mathbf{x}_u)$ and $\tilde{\mathbf{y}}_i = \mathbf{h}(\mathbf{y}_i)$.

Hybrid User Model To capture multi-aspect user preferences, we should consider not only the explicit preferences $\tilde{\mathbf{x}}_u^1$ reflected by user's own tagging behaviors, but also the implicit preferences $\tilde{\mathbf{x}}_u^2$ conveyed by the historical interacted items. In this part, we elaborate a hybrid user model with hierarchical attention network. The item-level attention aims to depict user implicit preferences $\tilde{\mathbf{x}}_u^2$ by differentiating contributions of historical items, while the preference-level attention dynamically discriminates the saliencies between explicit tagging behaviors and implicit preferences for obtaining hybrid user representation $\tilde{\mathbf{x}}_u^H$.

In the item-level attention, we leverage an additive attention network to differentiate contributions of items by investigating the similarities between item representations and explicit preferences $\tilde{\mathbf{x}}_u^1$. Suppose the historical items set of user u is $\mathcal{I}_u$, the representation of kth item $i_k \in \mathcal{I}_u$ is $\tilde{\mathbf{y}}_k$. The attention weight $\alpha(u,k)$ can be interpreted as the contribution of the kth item to the implicit preferences, which is shown as:

$$\alpha(u,k) = softmax(\mathbf{v}_1^T ReLU(\mathbf{W}^0 \tilde{\mathbf{x}}_u^1 + \mathbf{W}^1 \tilde{\mathbf{y}}_k + \mathbf{b}^1)), \tag{1}$$

where matrices $\mathbf{W}^0$, $\mathbf{W}^1$ and vectors $\mathbf{b}^1$, $\mathbf{v}_1$ are the trainable parameters. Finally, the implicit preferences $\tilde{\mathbf{x}}_u^2$ can be represented as $\tilde{\mathbf{x}}_u^2 = \sum_{i_k \in \mathcal{I}_u} \alpha(u,k) \tilde{\mathbf{y}}_k$.

The hybrid user representation can be obtained by fusing $\tilde{\mathbf{x}}_u^1$ with $\tilde{\mathbf{x}}_u^2$. Different from manually setting a hyper-parameter β for all users to determine the trade-off, we design a self-attentive fusion mechanism for complying with individual diversity. Similarly, the attention weight $\beta(u,k)$ of the kth part is:

$$\beta(u,k) = softmax(\mathbf{v}_2^T ReLU(\mathbf{W}^2 \tilde{\mathbf{x}}_u^k + \mathbf{b}^2)). \tag{2}$$

And the hybrid user representation is formulated as $\tilde{\mathbf{x}}_u^H = \beta(u,1)\tilde{\mathbf{x}}_u^1 + \beta(u,2)\tilde{\mathbf{x}}_u^2$.

Intersection Module Item features are multi-dimensional and have diverse attractions for different users. The intersection of user and item tags reveals the deep reason why the user focuses on the item and which are the vital dimensions when modeling this transaction. Motivated by this observation, we elaborate an intersection module to extract the intersection for further enhancing the recommendation performance. Firstly, we calculate the tags intersection by $\mathbf{i}_{iu} = \mathbf{y}_i \cap \mathbf{x}_u = (r_1^{iu}, r_2^{iu}, ..., r_V^{iu})$, where $r_j^{iu} = min(q_j^i, p_j^u)$ means the minimum occurrences of tag t_j. Then $\mathbf{i}_{iu}$ is fed into a MLP that shares parameters with the previous MLPs for further training the networks. At last, the latent representation $\tilde{\mathbf{i}}_{iu}$ is added to the user/item representations, that is, $\tilde{\mathbf{y}}_i = \tilde{\mathbf{y}}_i \oplus \tilde{\mathbf{i}}_{iu}$ and $\tilde{\mathbf{x}}_u^H = \tilde{\mathbf{x}}_u^H \oplus \tilde{\mathbf{i}}_{iu}$, where operation $\oplus$ means element-wise addition.

Due to the shared parameters, intersection module can constrain MLPs to focus on the conjunct features, obtaining more concrete user/item representations under a certain user-item transaction scenario.

Training Details At the prediction stage, feature vectors $\tilde{\mathbf{y}}_i$ and $\tilde{\mathbf{x}}_u^H$ are concatenated into a single vector $\mathbf{z} = \left[\tilde{\mathbf{y}}_i, \tilde{\mathbf{x}}_u^H\right]$, and passed through a prediction layer consisting of a factorization machine [2], which captures the second-order interactions in a fine-grained manner, i.e, $\hat{y}_{ui} = FM(\mathbf{z})$.

We optimize the model with the BPR framework and the loss function is $L_\Theta = \sum_{\langle u,i_+,i_- \rangle} -\ln \sigma(\hat{y}_{ui_+} - \hat{y}_{ui_-})$, where i_+ and i_- are the positive and negative items of user u respectively. The negative items are randomly sampled from a uniform distribution. Besides, dropout is also used to prevent overfitting.

3 Experiments

We conduct experiments on two public datasets: Last.Fm and Delicious and adopt the same preprocessing as [5, 3, 4] to remove infrequent tags. For each dataset, we randomly select 80% of the assignments as training set and 20% as test set. The training set is used to construct tag-based user and item profiles.

We compare the performance of AIRec with FM[2], CFA[5], DSPR-NS[3] and HDLPR[4]. Precision (P), Recall (R), F1-score (F) and Mean Reciprocal Rank (MRR) are used to evaluate the results. Table 1 illustrates the top-n recommendation performances. It's obvious that AIRec achieves the best performance in all metrics, which demonstrates the effectiveness of our model.

Table 1. Comparison between different models.

Last.Fm	**P@10**	**P@20**	**R@10**	**R@20**	**F@10**	**F@20**	**MRR**
FM	0.1470	0.1237	0.0945	0.1410	0.1151	0.1318	0.0306
CFA	0.1389	0.1055	0.0970	0.1349	0.1142	0.1184	0.0287
DSPR-NS	0.1693	0.1340	0.1667	0.2234	0.1680	0.1675	0.0362
HDLPR	0.1641	0.1328	0.1483	0.1984	0.1558	0.1591	0.0357
AIRec	**0.3074**	**0.2417**	**0.2670**	**0.3437**	**0.2857**	**0.2838**	**0.0651**
Delicious	**P@10**	**P@20**	**R@10**	**R@20**	**F@10**	**F@20**	**MRR**
FM	0.0369	0.0352	0.0103	0.0172	0.0161	0.0231	0.0058
CFA	0.0168	0.0110	0.0098	0.0125	0.0124	0.0117	0.0031
DSPR-NS	0.3656	0.3196	0.0897	0.1437	0.1441	0.1982	0.0423
HDLPR	0.2546	0.2148	0.0554	0.0885	0.0910	0.1254	0.0301
AIRec	**0.4052**	**0.3505**	**0.1165**	**0.1838**	**0.1810**	**0.2417**	**0.0480**

4 Conclusion

In this work, we propose a novel tag-aware top-n recommendation model AIRec. We design a hybrid user model with a hierarchical attention network for better user modeling and leverage the tags intersection for constraining neural networks to focus on the conjunct features. Extensive experiments shows that AIRec significantly outperforms the state-of-the-art baselines.

References

1. Liang, N., Zheng, H.T., Chen, J.Y., Sangaiah, A.K., Zhao, C.Z.: Trsdl: Tag-aware recommender system based on deep learning–intelligent computing systems. Applied Sciences **8**(5), 799 (2018)
2. Rendle, S.: Factorization machines. In: ICDM. pp. 995–1000. IEEE (2010)
3. Xu, Z., Chen, C., Lukasiewicz, T., Miao, Y., Meng, X.: Tag-aware personalized recommendation using a deep-semantic similarity model with negative sampling. In: CIKM. pp. 1921–1924. ACM (2016)
4. Xu, Z., Lukasiewicz, T., Chen, C., Miao, Y., Meng, X.: Tag-aware personalized recommendation using a hybrid deep model. In: IJCAI. pp. 3196–3202 (2017)
5. Zuo, Y., Zeng, J., Gong, M., Jiao, L.: Tag-aware recommender systems based on deep neural networks. Neurocomputing **204**, 51–60 (2016)