The 33rd Annual Conference of the Japanese Society for Artificial Intelligence, 2019

Proposal of Context-aware Music Recommender System Using Negative Sampling

Jin-cheng ZHANG[*1] Yasufumi TAKAMA[*1]

[*1] Graduate School of System Design, Tokyo Metropolitan University

This paper proposes a method for recommending music items considering listeners' context information. Recently, users can enjoy music easily regardless of time and a place due to evolution of online music services such as Spotify. However, it is difficult for us to find appropriate music items from enormous resources. On the other hand, because of listening style and characteristic of music items, music items do not usually have explicit rating. Therefore, implicit feedback such as playing count has been popular to construct recommender systems. As additional information, this paper considers listeners' context. The proposed method employs FMs (Factorization Machines), in which the context information is treated as factors. Negative sampling is applied to reduce the number of negative samples (music items a user has yet to be listened). The effectiveness of the proposed method and the effect of negative sampling are shown with an offline experiment.

1. Introduction

This paper proposes a method for recommending music items considering listeners' context information. Recently, with the development of communication technology and portable electronic devices, online music services such as Spotify, AWA, Amazon Music has grown rapidly and become popular in a short period of time. The number of users who utilize online music services has exceeded 17 millions on 2017 in Japan.

To help users find their favorite music items, music recommendation systems have been studied. One of the challenges of music recommendation is that explicit feedback such as rating is not usually available due to the way of listening to music and characteristic of music items. Therefore, existing music recommender systems often use implicit feedback such as playing count[Koren 08]. Another challenge is that the influence of users' context while listening to music on their preference is higher than other domains. By extending the conventional matrix factorization approach, a tensor factorization method can consider context axis in addition to user and item axes [Adomavicius 11]. However, it is difficult to optimize the use of contexts due to high time complexity. Therefore, this paper employs Factorization Machines (FMs), in which listeners' contexts are used as factors. It is shown that FMs can flexibly consider the interaction of users, items, and context with relatively low time complexity [Rendle 12].

As FMs needs both positive and negative samples, LEs (Listening events) are used as positive samples, and music items a user has yet to be listened are used as negative samples. To reduce the number of negative samples, this paper employs negative sampling. Effectiveness of the proposed method is shown by comparing it with wALS (weighted Alternating Least Squares). The effect of negative sampling is also investigated.

Contact: Yasufumi TAKAMA, Graduate School of System Design, Tokyo Metropolitan University, ytakama@tmu.ac.jp

2. Related work

2.1 Implicit collaborative filtering

Koren et al. proposed a collaborative filtering method called wALS for implicit feedback datasets[Koren 08]. As an implicit feedback such as play count is not as reliable as explicit rating, it is difficult to make recommendation by solving the same loss function as methods based on Matrix Factorization. Eq. (1) shows the loss function of wALS,

$$\mathrm{L}(x_u, y_i) = \sum_{u,i} c_{u,i}(p_{u,i} - x_u^T y_i)^2 + \lambda(\sum_u ||x_u||^2 + \sum_i ||y_i||^2), \tag{1}$$

where x_u, y_i is the latent factors of users and items, $p_{u,i}$ is the element of user-item binary matrix representing whether or not a user u played a music item i. $c_{u,i}$ is the element of confidence matrix: $c_{u,i} = 1 + \alpha r_{u,i}$, where α is a hyperparameter and $r_{u,i}$ is the play count of i by u.

2.2 Factorization Machines

FMs[Rendle 12] can not only learn the latent factors of user and items like Matrix Factorization methods, but also flexibly consider the interaction of users, items, and any other features with relatively low time complexity. The model equation for a FMs of degree $d = 2$ is defined as:

$$\hat{r}(\boldsymbol{x}) := \omega_0 + \sum_{i=1}^{n} \omega_i x_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} (\sum_{f=1}^{k} v_{i,f} \cdot v_{j,f}) x_i x_j, \tag{2}$$

where $x_i \in \boldsymbol{x}$ is rating, which corresponds to a LE in this paper, ω_0 is the global bias, ω_i is the strength of x_i. $\sum_{f=1}^{k} v_{i,f} \cdot v_{j,f}$ shows the interaction of x_i and x_j with k factors. The $\hat{r}(\boldsymbol{x})$ are the prediction values for all LEs.

3. Proposed Method

This paper considers season, time zone, time stamp and mother language when a user listened music item as context information. Those are used as additional factors $v_{i,f}$.

We set $r(x_i) = 1$ for positive samples, and $r(x_i) = 0$ for negative samples.

Regarding all music items a user has yet to be listened as negative is a naive method. Although it could be effective in Matrix Factorization method such as wALS, the number of negative samples becomes huge, and cause a problem of time complexity when training a model by FMs.

To solve the problem, this paper employs negative sampling, which selects a part of music items that a user has yet to be listened as negative samples. In particular, this paper proposes to select negative samples according to the hypothesis that **if a music item is not played by a user in some context, s/he would not be interested in it in that context.** Because a negative sample is not real LE, it is not used in evaluation process for ensuring the fairness.

4. Experiment

Dataset We use nowplaying-RS dataset[*1] that is generated from all tweets tagged by hashtag #nowplaying of 2014. This dataset also includes above-mentioned users' listening context. For avoiding the influence of tweet bot and extremely not active users, we removed the users who listened less than 10 and larger than 5000 music items from the dataset. After preprocessing, the number of users, music items, and LEs are 18,946, 22,023, and 1,835,993, respectively.

Evaluation metric This paper employs MPR (Mean Percentage Ranking) (Eq. (3)) to evaluate the performance of each method. $rank_{u,i}$ represents the percentile-ranking of a music item in the recommend list, and $r_{u,i}$ corresponds to $r(x_i)$ in Eq. (2): $r_{u,1} = 1$ if a user u actually listened a music item i, otherwise 0. Smaller MPR indicates better performance.

$$MPR = \frac{\sum_{u,i} r_{u,i} \times rank_{u,i}}{\sum_{u,i} r_{u,i}} \tag{3}$$

Experiments are done with 5-fold cross-validation based on time series. Experiments are repeated 5 times, and use t-test to confirm whether or not significant difference to wALS is observed. Used parameters are: $\alpha = 250$, $\#latent\,factor = 15$, $\lambda = 0.05$ for wALS by hyperparameter tunning, and $k = 3$ for the proposed method.

Table 1 shows MPR with the p-value among wALS, UO (user-oriented negative sample)+FMs and UO+FMs+Contexts. Table 1 shows that UO+FMs is more effective than wALS, and its performance is improved by introducing context information.

Fig. 1 shows the MPR against the ratio of negative samples to positive samples (denoted by k). It is observed that setting k larger than 2 can obtain more effective result than wALS. Although MPR decreases as k increases, the effect is gradually reduced. The computation time when $k = 2$ and 6 are 688(s) and 2142(s) per fold in cross-validation, respectively. This mean that k should be set by considering the balance between time complexity and accuracy of recommendation.

Table 1: Experimental result

p-value	wALS	UO+FMs	UO+FMs +Contexts
wALS	-	2.363e-05	5.368e-07
UO+FMs	2.363e-05	-	0.0003
UO+FMs +Contexts	5.368e-07	0.0003	-
MPR	0.1092	**0.1041**	**0.1017**

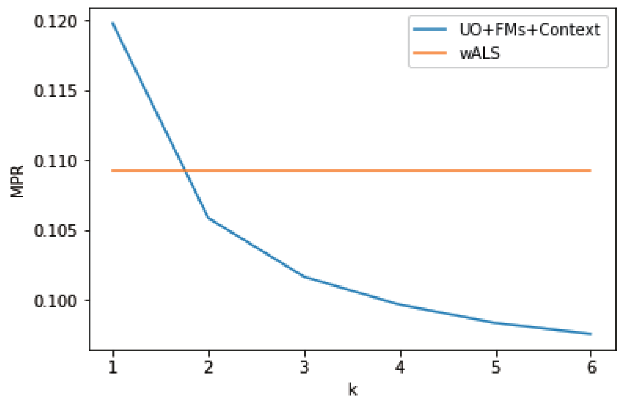

Figure 1: Effect k on MPR

5. Conclusion and Future work

This paper proposes a method for recommending music items considering listeners' context information. An experimental result shows the proposed method is more effective than wALS. In future work, we will investigate other negative sampling methods to improve performance, and evaluate the proposed method with other metrics such as diversity and novelty.

References

[Koren 08] Hu Yifan, Yehuda Koren, and Chris Volinsky, "Collaborative filtering for implicit feedback datasets," ICDM'08, pp.263-272, 2008.

[Rendle 12] Rendle Steffen, "Factorization machines with libfm," ACM Transactions on Intelligent Systems and Technology, Vol.3, No.3, Article No.57, 2012.

[Adomavicius 11] Adomavicius Gediminas and Alexander Tuzhilin, "Context-aware recommender systems," Recommender systems handbook, pp.217-253, 2011.

*1 http://dbis-nowplaying.uibk.ac.at/#nowplayingrs