

Improving Music Recommendation Using Distributed Representation

Dongjing Wang, Shuiguang Deng*
 College of Computer Science
 and Technology, Zhejiang University
 Hangzhou, Zhejiang, China
 {tokyo1, dengsg}@zju.edu.cn

Songguo Liu
 Hangzhou National E-commerce
 Product Quality Monitoring and
 Management Center, China
 Sgliu@zju.edu.cn

Guandong Xu
 Advanced Analytics Institute,
 University of Technology Sydney
 Sydney, Australia
 Guandong.xu@uts.edu.au

In this paper, a music recommendation approach based on distributed representation is presented. The proposed approach firstly learns the distributed representations of music pieces and acquires users' preferences from listening records. Then, it recommends appropriate music pieces whose distributed representations are in accordance with target users' preferences. Experiments on a real world dataset demonstrate that the proposed approach outperforms the state-of-the-art methods.

music recommendation; distributed representation

Nowadays, digital music market is growing rapidly due to the prevalence of mobile devices and advance in the Internet technology. It is more important than ever to help people find the interested music pieces from massive music contents available on the Internet. How to extract the feature of music and incorporate them into music recommendation is still a challenging task. To address this problem, we present a music recommendation approach based on distributed representation. Firstly, the proposed approach learns the distributed representations (vectors in real-valued, low-dimensional space) of music pieces from all users' historical listening records. Then, it infers users' music preferences from their listening records with these distributed representations. Finally, our approach recommends appropriate music pieces according to target users' preferences to satisfy their requirements.

Music recommendation is to find music pieces that the target user would probably enjoy. Formally, let $U = \{u_1, u_2, \dots, u_n\}$ be the user set and $T = \{t_1, t_2, \dots, t_m\}$ be the music set. For each user u_i , his/her historical listening record is denoted as $R_i = \{t_{i1}, t_{i2}, \dots, t_{i|R_i|}\}$, where $t_p \in T$. Music in R_i are sorted according to the corresponding playing time. Then, our

*The corresponding author is Shuiguang Deng and this work was supported in part by the National Key Technology Research and Development Program of China under Grant 2014BAD10B02.

task specifies to be seeking for music that user u_i may enjoy given his/her listening record R_i . To address this task, we propose a music recommendation approach, which consists of three steps: distributed representation learning, users' preferences acquisition, and recommendation.

Firstly, we propose the $t_{j-1:j} | t_p$ model to learn the distributed representations of all music pieces. Specifically, the $t_{j-1:j} | t_p$ model adopt a skip-gram model [1], which is much more efficient as well as memory-saving than other approaches, to learn the distributed representation by maximizing the objective function over music sequences in all users' listening records. The underlying idea of $t_{j-1:j} | t_p$ is that similar music pieces should have similar contexts. Formally, the objective function is defined as follows:

$$\mathcal{L} = \sum_{t_p \in T} \sum_{-j \leq j \leq j, j \neq 0} \log (t_{j-1:j} | t_p) \quad (1)$$

where j is the length of the context window. $(t_{j-1:j} | t_p)$ represents the probability of observing a neighbor music piece $t_{j-1:j}$ given the current music item t_p in u_i 's record R_i , which is formally defined using the soft-max function as follows:

$$(t_{j-1:j} | t_p) = \exp(\mathbf{t}_{j-1:j} \cdot \mathbf{t}'_{t_p}) / \sum_{t \in T} \exp(\mathbf{t}_{j-1:j} \cdot \mathbf{t}'_t) \quad (2)$$

where $\mathbf{t}_{j-1:j}$ and \mathbf{t}'_t are the input and output distributed representations of music t , respectively. In the learning phase, we need to maximize the objective function defined in Equation 1 over all users' historical listening records. However, the complexity of computing corresponding soft-max function defined in Equation 2 is proportional to the music set size, which can reach millions easily. In this paper, we adopt negative sampling [1] to increase computation efficiency by generating a few noise samples for each input music to estimate the target music. Therefore, the training time yields linear scale to the number of noise samples. Finally, the distributed representations of all music pieces can be obtained.

Then, music preference of the target user is inferred from his/her historical listening record using the following formula:

$$= \frac{1}{|R_i|} \sum_{t_p \in R_i} t_p \quad (3)$$

where t_p is the learned distributed representation of music t_p using $t_{j-1:j} | t_p$ model.

Finally, we propose a music recommendation method which can recommend music pieces appropriate music pieces whose distributed representations are in accordance with target users'

preferences. Specifically, the predicted interest p of the target user in music piece t is the cosine similarity between t 's music preference and t 's distributed representation t_p , which is defined as follows:

$$\hat{r}(t_p | t) = \cos(t_p, t) \quad (4)$$

Therefore, the ranking of music pieces t_p in our approach is defined as

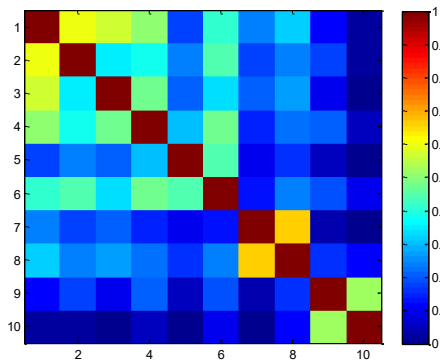
$$t_p > t'_p \Leftrightarrow \hat{r}(t_p | t) > \hat{r}(t'_p | t) \quad (5)$$

We then can recommend the music pieces with high ranking scores (similar to user's musical preference) to the target user.

The experiments consist of two parts: evaluation of t_p and the comparison of the proposed approach with baselines.

Firstly, we illustrate the effect of t_p model by visualizing similarity among music pieces given in Table 1. As shown in Figure 1, music pieces with similar styles, such as singers, tags, and genres, have similar distributed representations. For example, "Summer" and "Moonlit Sea of Clouds", which have the same genre and player, do lie nearby in the real-valued distributed representation space. Besides, neither of these two music pieces has similar distributed representations with the other music pieces in Table 1. Therefore, the learned distributed representations with t_p capture useful features effectively and depict music pieces well.

1	Hero-Mariah Carey	pop, female vocalists, 90s, ballad
2	Without You-Mariah Carey	pop, female vocalists, soul, love
3	Drowning-Backstreet Boys	pop, boy bands, ballad
4	My Love-Westlife	pop, boy bands, Irish
5	Don't Cry-Guns N' Roses	classic rock, hard rock, ballad
6	Hotel California-Eagles	classic rock, rock, 70s
7	Fall Again-Kenny G	smooth jazz, R&B, Soul
8	Heart and Soul-Kenny G	smooth jazz, Rhythm and blues
9	Summer-Joe Hisaishi	sound track, Japanese, anime, instrumental, classical
10	Moonlit Sea of Clouds-Joe Hisaishi	sound track, Japanese, anime, instrumental, classical



Then, we compare the proposed method with three state-of-the-art recommendation algorithms, including Bayesian Personalized Ranking (BPR) [2], FISMauc (FISM) [3], and user based collaborative filtering method (UserKNN) [4] on a real

world dataset collected from Xiami Music (<http://www.xiami.com/>). From the comparison results in Table 2, we can see that our approach outperforms baselines in terms of F1 score and hitrate. Taking the F1 score as an example, when compared with BPR, FISM, and UserKNN with the recommending number being 10, the relative performance improvement achieved by the proposed approach is around 33.3%, 20.7%, and 42.2%, respectively. The improvements indicate that our approach is more effective than baselines in acquiring users' preference and assisting music recommendation.

In conclusion, the proposed approach can effectively recommend music pieces appropriate for target users and satisfy their preferences well.

Our approach				
BPR	5.91	7.13	26.96	30.21
FISM	6.53	7.50	29.68	33.52
UserKNN	5.54	6.34	22.72	25.61

We present a music recommendation approach, which learns the distributed representations of music pieces from users' historical listening records, and utilizes these distributed representations to acquire users' music preferences and recommend appropriate music pieces. Experimental results show the effectiveness of the proposed approach. There are two possible future directions. Firstly, we plan to combine the distributed representation with more advanced recommendation techniques [5, 6], to further improve the performance. Secondly, we will try to evaluate our approach by online experiments.

- [1] T. Mikolov, I. Sutskever, K. Chen, G.S. Corrado, and J. Dean, Distributed representations of words and phrases and their compositionality. In *Hk hu jl p Ul yhsRmvt h pu W y w j l p n l t*, 3111-3119, 2013.
- [2] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, BPR: Bayesian personalized ranking from implicit feedback. In *W y w j l l k p n m o l 9 o v u n t y l u j l v u u j l y h p p Hy p p j p h s R u l s s p l u j l*, 452-461, 2009.
- [3] S. Kabbur, X. Ning, and G. Karypis, Fism: factored item similarity models for top-n recommender systems. In *W y w j l l k p n m o l o H T p l y u h p u h s j v u n t y l u j l v u R u v s k n l k p j v l y h u k k h h t p p n*, 659-667, 2013.
- [4] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, GroupLens: an open architecture for collaborative filtering of netnews. In *W y w j l l k p n m o l H T j v u n t y l u j l v u v t l y v y l k j v v l y h p l v y r*, 175-186, 1994.
- [5] S. Deng, L. Huang, and G. Xu, Social network-based service recommendation with trust enhancement. *l y l t p o H s p h p u*, 41, 18 (12/15/), 8075-8084, 2014.
- [6] A. Karatzoglou, X. Amatriain, L. Baltrunas, and N. Oliver, Multiverse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering. In *W y w j l l k p n m o l m y o H T j v u n t y l u j l v u l j v t l u k l y l t*, 79-86, 2010.