Hybrid personalized music recommendation method based on feature increment

Guimei Liu[1] and Wenjun Jiang[2]

[1] Hunan University, Changsha, China
lgm@hnu.edu.cn
[2] Hunan University, Changsha, China
jiangwenjun@hnu.edu.cn
https://sites.google.com/site/happywenjunjiang2012/

Abstract. Sentiment analysis is a key factor in music recommendation systems. Existing work usually focuses on sentiment analysis based on music audios,which faces three issues, namely, high cost and the emotions in the song lyrics and the comment text, followed by not considering the user's social relationship, which may lead to songs that are difficult to discover the user's potential interest. In this paper,we study the listening intent by exploring listening behavior, which may lead to the user's listening intention. We make three progresses. (1) We analyzed a real music platform data from multiple perspectives. This data is concentrated on the user's listening record and the user's followers. Taking the user's song listening as the research object, the relationship between the user's listening behavior and the emotion of the song, and the relationship with followed is studied. (2) We analyze the correlation between the user's preference value of the song and the emotional category of the user listening to the song. (3) We construct the classification model based on the KMeans and adjusting different parameters to predict whether a user will listen a song. The results of this experiment show that it is better to consider the relationship of followed. We distinguish the importance of each feature. Based on the importance, we find that user's listening behavior has relationship with followed and the user's preference for the song has relationship with the emotion of the song.

Keywords: Social relationship · Emotion analysis · Music recommendation.

1 INTRODUCTION

With the continuous development and enrichment of music content in recent years,it is difficult for users to find music suitable for themselves in the huge ocean of music. The purpose of personalized music recommendation is to provide users with tailor-made music services. It is a research topic that benefits users and music platforms. With the development of natural language processing [15] in recent years,the computer is able to automatically analyze and understand music content, which makes lyrics-based music recommendations become possible. Emotional description [7, 8] is very useful and effective in describing music taxonomy. Since songs generally have a certain emotion, we can get the emotions of songs by emotionally analyzing song lyrics and comment texts. As a pioneering effort to describe human emotions,Russel [9] proposed a rotation model,and each of these emotions is shown in two bipolar dimensions Degree. These two dimensions represent unpleasant to pleasure, calmness to excitement. Therefore, each emotional word can be defined as some combination of these two dimensions. Later, Thayer [19] adapted Russel's model based on music. The two main dimensions of the Thayer model are"arousal" and "valence". In this model, emotional terms are described as calm to excitement in the arousal dimension and described as negative to positive in the valence dimension. The two-dimensional emotion plane of the Thayer model can be divided into four quadrants, composed of 11 emotions, such as Figure 1 shows.

There are many types of music attributes, such as song singers, song lyrics, song audios, etc. More and more research based on music attributes to explore the intention of user listening. Sanchez-Moreno [20] et al use the KNN algorithm to find similar users based on the singer as a property to achieve music recommendation. Bu [2] et al use the song audio as an attribute to perform sentiment analysis and combine the user's social relationship to achieve music recommendation. These works use music attributes to build a music recommendation model. Sentiment analysis based on song audio, usually from the melody, beat, rhythm, and range of music et. It requires a certain amount of music expertise and faces issue of high cost. Meanwhile, using the song singer as a music attribute to achieve music recommendation ignores the user's social relationship,which makes it difficult to mine users who are potentially interested but not too familiar. In order to recommend music more effectively, we need music attribute information, user behavior information and user's social relationship information. In this paper, we combine emotion analysis with the user's social relationships [11, 12] to solve these problems in the field of music recommendation. To be specific, we aim to (1) find the user's musical interest by analyzing the correlation between the user's preference for the song and the emotional category of the user listening to the song and (2) analyze the correlation between the listening behavior of the user and the user's follower to find the song that the user has potential interest but is not familiar with.

Our contributions are threefold.
1.We analyze the correlation between user listening behavior and song emotion. We find that the listening behavior of different users is related to the emotion of the song. At the same time, the user's preference for the song is also related to the emotion of the song. These analyses can give us an idea of how to identify the user's listening intentions.
2.We analyze the listening behavior of both the user and the user's followers. We find that is difference in the degree of relationship between the listening behavior of user and the different followers, which indicates that the degree of relevance between the listening interest of the user and the different followers is diversity. Through the above analysis, we find that the user's listening intention can be

reflected from the user's followers listening behavior.
3.We construct the hybrid personalized music recommendation model. We construct the classification model based on the KMeans and adjusting different features (the relationship between the emotional category of a song and user preferences, and the degree of relationship between users and followers) to predict whether a user will listen a song. The experiment results show that the importance of the features are different,for different features have different contributions to users' listen intentions.
The method of this paper applies to all music sites because the data types of the datasets used in this paper is common in music websites.

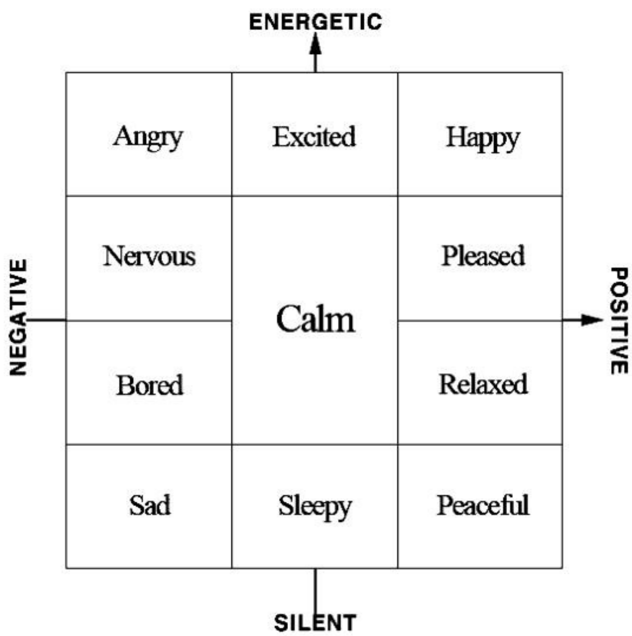

Fig. 1. Thayer emotion model.

2 RELATED WORK

Since the advent of the music platform, various music recommendation studies have been proposed. Content-based recommendation [17],for example, Sanchez-Moreno [20] et al published an paper in 2016,since each singer has his own style,by using the singer as a song attribute, combined with the KNN algorithm to find similar users and the songs that the similar user often listen to but the target user does not listen to are recommended. Model-based recommendation [5,13,16],for example, Pacula [18] published an paper in 2009, using matrix decomposition to implement music recommendation based on user's implicit feedback. User-based recommendation method [10, 14],for example, Deng [6] et al published an article in 2015,finding similar users by based on the characteristics of listening to the same song under the same emotions condition and finding the recommended items from similar users' listening records based on the user's emotions from the blog text. There are recommended methods based on hybrid models [3],for example, Bu [2] et al published an paper in 2010,by using content-based recommendations combine with collaborative filtering to build a hybrid recommendation model that uses music audio information and user's social relationship.

At present, some of the music recommendation systems use the user's social relationship and music audio information for mixed recommendation,but the method does not take into account another important attribute of music,namely the lyrics text information. It only takes into account the genre of music, but does not take into account the emotions of music. There are also sentiment analysis based on lyrics texts [1, 4, 21],but does not consider the feelings of the user after listening,that is, the emotion in the comment text. The emotions in the comment text can help to dig out the hidden emotions in the song,which can affects the accuracy of the emotional sentiment of the song. Secondly, the user's social relationship is not considered, that is, the user who is interested in the user's social relationship is ignored, which makes it difficult to find the user's potential interest song.

In conclusion, our work differs from others in mainly two aspects. (1) We use emotional analysis based on lyrics and comment texts to get attributes of songs. (2) We consider the social relationship of the target user to find songs that are not familiar to the target user but are potentially interesting.

3 Correlation Analysis of User Listening Behavior in Music Platform

In this section,we focus on the user's listening intention as the research object and the analysis of various feature attributes related to it. We aim to identify the emotions of the song and the user's social relationship how to affects the user's listening behavior. The result obtained from these analyses can help businesses better understand their users and find their personalized listening intention. Our method is suitable for most music platform websites because we only need the information of the basic user behavior information, song attribute information, and user's follower information.

3.1 Raw Data

First, we used the scrapy crawler framework to crawl data on the NetEase cloud music platform. This dataset contains the user's song record, the song's lyrics and the comment text and the user's follower as the original data. This dataset contains a total of 93,319 songs and 1103 users. Among them, 30,856 songs are

Table 1. Some records of user behavior logs.

userId	songId	score
58459	27902876	22
58459	115569	21
183103	417596830	100
183103	472109598	36
1713095	554367706	86
1713095	493663235	85
2881455	417596830	12
2881455	31152310	4
3028129	432506809	85

Table 2. Some followers information of the user.

userId	FollowerId
58459	319910944
58459	264745245
183103	48168
183103	48340
1404462	601070196
1404462	94451159
2547334	46309613

the number of listening songs for 1,103 users, and the remaining 62,463 songs are the number of listening followers of 1,103 users. The user's listening record is partially shown in Table 1. Where the score field value has a range of 0 to 100, this value indicates how much the user likes the song. The larger the value, the more the user likes the song.

In Table 2, the user's follower information is displayed. The first field represents the target user's id, and the second field indicates the id of the user's followers.In Figure 2, the id of the song, the lyrics of the song, and the comment text information are displayed. The songs here are limited to Chinese songs.

3.2 Data preprocessing

Firstly, the Chinese songs are filtered on the original data, and 16482 Chinese songs are obtained, of which 7144 songs are 1103 user listening songs, and the remaining 9338 songs are the user's followers listening songs. The songId and the userId are renumbered, which is beneficial to the subsequent operations. Secondly, the original userId and songId are sorted from small to large, and then replaced by 1 to 1103 and 1 to 16482. What's more, the Chinese and English punctuation marks and other non-Chinese symbols are removed from the lyrics and comment texts, and texts containing only Chinese characters are obtained. The THULAC Chinese word segmentation algorithm is used to process Chinese text information, and then use the stop word table to remove the stop words

song_id	lyric	comment
27902876	烽火乱世 狼烟不止...	隆中一对天下知...
115569	这晚在街中偶遇心中的她..	追的好故事追不好叫事故..
417596830	热夏 你归来 听蝉 再游于..	吹过的风有没有告诉你...
472109598	国王不在家皇后盘起了长发	你是我安稳岁月里的节外生
554367706	在夜与雾之间浓稠的是思念	这个世界缺的不是完美的人
493663235	每当站在高楼上 看着地上..	活着最重要...
31152310	你是我梦里 陌生 熟悉...	梦里你是梦 越梦越空...

Fig. 2. Some lyrics and comments information.

from the Chinese text after the word segmentation to obtain the text represented by one word.

3.3 Word embedding

Since the computer cannot directly process Chinese text information, the Chinese text should be constructed as a word vector. In this paper, we use the emotional value of the emotional words in the text to construct the word vector of the text. Since not all words in the lyrics and commentary texts have emotions, some words only serve as a link. In order to more accurately and effectively predict the emotions of the text,we use the SenticNet sentiment dictionary. The sentiment dictionary has a more granular sentiment analysis than other sentiment dictionaries (such as NRC, DUTIR). The sentiment dictionary contains more emotional words, which not only contain the polarity of the emotional words, but also include the emotional values of the emotional words in the four emotional dimensions, making the emotional analysis more fine-grained.

Firstly, we take the first 99 comments and a lyric for each song to make up 100 texts. Secondly,by matching the lyrics and comment texts after the word segmentation with the emotional dictionary SenticNet,the emotional words in the lyrics and the comments text are extracted,and the corresponding emotional values in the four emotional dimensions are extracted. We use the emotional values on the four emotional dimensions to represent the lyrics and comment text information to get four emotions matrix Edj, where d is the four emotional dimensions and j is the song id.

Since the lyrics of the song and the number of emotional words in the comment text are not all the same, for the convenience of presentation and processing, the emotions in the lyrics or comment texts matching the emotion dictionary SenticNet are matched. The maximum number of words is taken as the standard. In this paper, it is set to 169, and if the number of emotional words is less than 169, use 0 to fill. The four emotional dimensions in the SenticNet sentiment dictionary are Pleasantness, Attention, Sensitivity, and Aptitude. We use the SVD singular value decomposition algorithm to perform matrix dimensionality reduction on the emotional matrix Edj, where the Edj matrix is 100*169. Firstly, the matrix is reshaped, and it is transformed into a square matrix of 130*130 dimensions, which is reduced into 1*130 dimensional matrix edj and combining the emotional matrix edj on the four emotional dimensions to obtain a matrix of 1*520 dimensions, which is the emotional vector of the song.

3.4 Statistical analysis

From Figure 3, we can see that the number of listening songs for 1,103 users is roughly uniform, mostly from 20 to 40 songs. Only a small number of people have more than 50 songs listening. At the same time, only a small number of people listen to less than 10 songs. From Figure 4, we find that only a few songs were listened to by most users, and the majority of the songs were listened to no more than 50 users,which indicates that the song is listening to the long tail distribution, that is, the popularity of each song is different.

In Figure 5, it shows a distribution of the emotional categories to which the songs in the entire dataset. We regard the emotional category to which the song belongs as a classification problem.We use the sklearn which is a Python library as the tool to train the model. We train a cluster classifier based on the emotion vector of the song, and classify the emotion categories of the songs in the entire dataset into 11 categories. In Figure 6 ,it shows the number of samples included in each cluster. From the sample distribution, we can conclude that the number of samples contained in each cluster has a certain difference. It can indicate that the songs of different emotional categories are not the same as the user's preferences and that is a difference in the user's listening preferences.

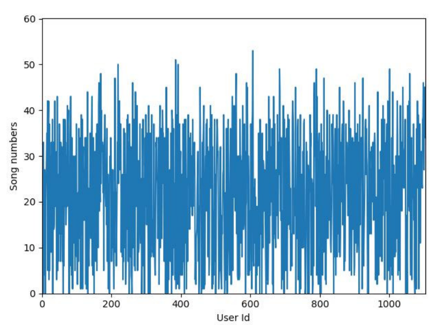

Fig. 3. Number of users listening to songs.

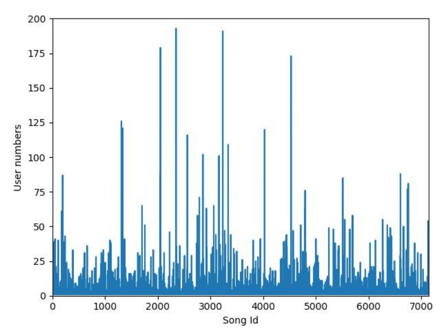

Fig. 4. The number of users whose songs are listened.

4 Correlation analysis

In this section,we analyze the correlation between the emotional category of the user's listening record and the user's preference for the song. We also analyze

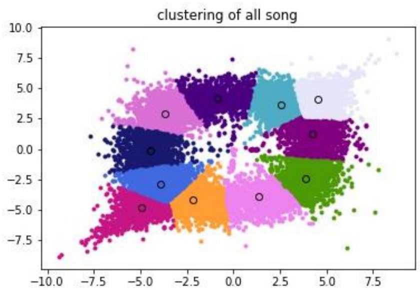

Fig. 5. Emotional clustering map of all songs.

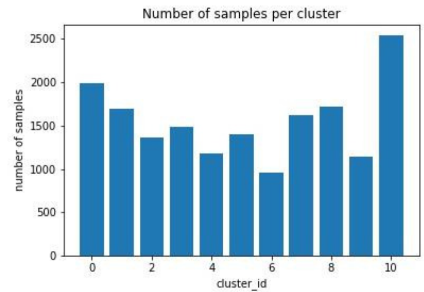

Fig. 6. Number of samples in each cluster.

the degree of relationship between users and followers. Through analysis, we try to find out the impact of these attributes on the user's listening intentions.

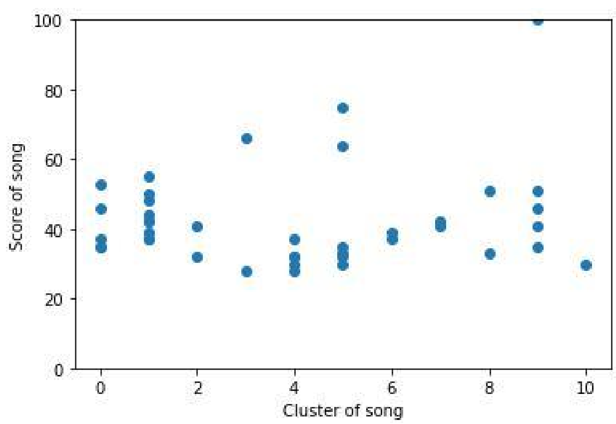

Fig. 7. The relationship between emotional categories of listening record and user preferences.

4.1 Correlation analysis between emotional categories of songs and user preferences

For each user,the emotional category of his or her like songs and the preference for the song may be different. So, when predicting the user's listening intent, we have to consider the different attributes of each user. This section analyzes the correlation between the emotional category of the user's listening record and the user's preference for listening to the song to find out the different listening preferences of different users.

Firstly, we analyze the correlation between the emotional category of each user's listening record and the user's preference for listening to the song. In the dataset, the number of listening songs per user is roughly around 30. the emotional category of the listening song is basically distributed and each user has a certain difference in the degree of preference for different songs because different users have different listening preferences.

In Figure 7,it show the correlation between the emotional category to which the user's listening record belongs and the user's preference for listening to the song. From this Figure, we can find that the user's preference for the performance of songs of different emotional categories is different. For songs of different emotional categories, the number of users who choose to listen will be different. The user usually choose to listen to more songs in the favorite emotional category, and choose less listening for songs that do not like the emotional category.

As can be seen from the above analysis, although the emotional category of the user's listening record is relatively comprehensive, the different emotional category clusters contain different listening records. Therefore, it indicates that the user's listening preferences are related to the emotional category to which the song belongs.

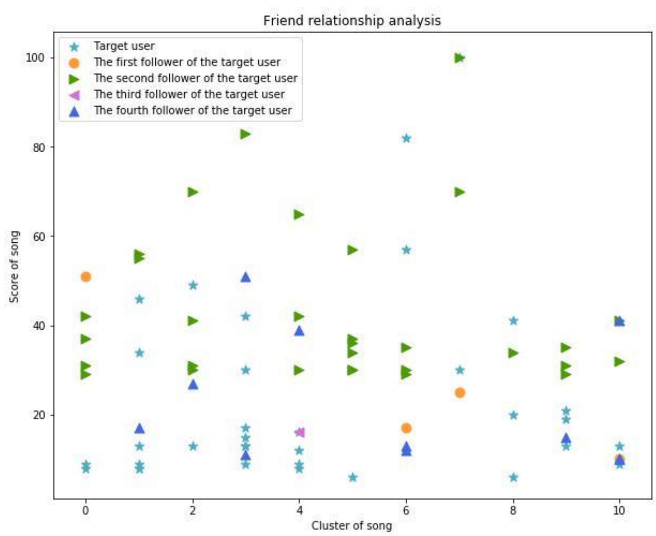

Fig. 8. the relationship between emotional categories of listening record of user and followers.

4.2 Analysis of the relationship between users and followers

For each user,the emotional category of his or her favorite songs and the preference for the song may vary. So, when looking for users who are similar to the target user's listening interests, it is also necessary to consider the similarity of the listening interests between the two. We analyze the degree of similarity between the target user and the listening record of the user's followers, and the degree of relationship between the user and the different followers.

A distribution of the emotional categories to which the user and the follower's listening record belong is shown in Figure 8. From this figure, we find that similarity of the sentiment category between the target user and the first follower is smaller than the target user and the second follower. Listening to interests of the second follower is more similar to the target audience,which indicates that is a difference in the similarity of the emotional categories to which the listening record between the user and the different followers belongs.

It can be seen from the above analysis that although different followers and the user's listening records have certain similarities, the similarity between the different followers and the user's listening records has certain differences. Therefore, it is necessary to consider the degree of relationship between followers and users.

5 Recommendation generation

In this section,we generate candidate recommendation items based on the relationship weight between the user and the follower and the preference of the follower for the listening record. We use the Gaussian function to calculate the similarity between the candidate recommended song and the category of the user's listening record. The generation of the recommended item is based on the similarity between the candidate recommended song and the target user's listening category and the weight of the target user's listening category. We use metrics such as accuracy and recall to measure recommendations.

5.1 Target users listen to song preferences categories and calculate weight values for category

First, using the trained classification model, based on the emotion vector e_j of the song in the listening record R of the target user u, the songs with similar emotional categories are grouped into clusters, that is, the target user listens to the song preference category C_{ui}. Then,we calculate a weight w_{ui} of each category C_{ui} relative to the target user u,where m_{ui} represents the number of samples in the category C_{ui}, M_u represents the total number of samples of all categories of the target user u, and i indicates the category. The weight w_{ui} is calculated by the formula (1):

$$\mathrm{w}_{ui} = m_{ui}/M_u \tag{1}$$

5.2 Selection of candidate recommended songs and calculation of similarity

Selection of candidate recommended songs. Based on the follow relationship of the target user u,finding the song that the follower v likes and the target user u has not heard as the candidate recommended song h from the listening record of the follower v. The song that the follower v likes is judged by its score_{vj} value of the song j. When the value of the score_{vj} of the user v to the song j is not less than a times the average value of the score_{vj} value of all songs that the user v has listened to, the song j is considered to be a favorite song of the user v. The discrimination formula (2) is as follows.

$$\text{score }_{\mathrm{vj}} \geq \frac{a}{\mathrm{num}} \times \sum_{j=1}^{\mathrm{num}} \text{ score }_{\mathrm{vj}} \tag{2}$$

Where num represents the number of songs that the user v has listened to, and a represents the weight of the relationship between the user v and the target user v.

Calculate the similarity between the candidate recommended song and the category. Based on the emotion vector eh of the candidate recommended song h and the classification result of the listening record R of the target user u, the Gaussian function is used to calculate the similarity between them. The similarity $\mathrm{S}_{\mathrm{hCui}}$ is calculated by the Gaussian function (3) as follows.

$$S_{\mathrm{hCui}} = \frac{1}{\sqrt{2\pi k_{\mathrm{g}} \delta_{\mathrm{Cui}}^2}} \exp\left(-\frac{(\mathrm{e}_h - \overline{p_{\mathrm{Cui}}})}{2k_{\mathrm{g}}\delta_{\mathrm{Cui}}^2}\right) \tag{3}$$

Where k_g represents a constant weight, σ_{Cui}^2 represents the variance of the category C_{ui}, $\overline{\mathrm{p}}_{\mathrm{Cui}}$ represents the mean of the category C_{ui}, and e_{h} represents the emotional vector of the candidate recommended song.

5.3 Calculate the preference value of the target user for the candidate recommended song

Based on the similarity between the candidate recommended song h and the target user's listening record category C_{ui} and the weight W_{ui} of the target user listening record category C_{ui} to calculate the preference value g_{uh} of the target user u for the candidate recommended song h,calculated by the formula (7) as follows.

$$\mathrm{g}_{\mathrm{uh}} = \frac{1}{\mathrm{c}} \sum_{\mathrm{i}=1}^{\mathrm{c}} w_{\mathrm{ui}} s_{\mathrm{hCui}} \tag{4}$$

Where c is the number of categories the target user listening record.

When preference value g_{uh} of the target user u for the candidate recommended song h exceeds the threshold t,the candidate recommended song h is added as a recommendation item to the recommendation list CL, otherwise, the candidate recommended song h is discarded.

Table 3. Experimental results under four different states.

	Precision	Recall
Random baseline	0.0468	0.1960
1	0.0822	0.5147
2	0.0562	0.2598
1+2	0.0888	0.6078

5.4 Result analysis

We use the sklearn, which is a Python library, as the tool to train the model and get the experimental result. Besides the random baseline, we present three methods to compare the effects of the two features as follows.
1. The first method (denotes 1) only uses similarity of user and follower features which is described in Section 4.2.
2. The second method (denotes 2) uses features of emotional categories of songs features which is described in Section 4.1.
3. The third method (denotes 1+2) uses all two types features.
The experimental results are shown in Table 1.

The integrated effects of different features. The improvement from random baseline to 2 and 1 to 1+2 are not obvious,but the improvement from random baseline to 1 and 2 to 1+2 are obvious. In other words, the similarity of user and follower features the improvement of the model for about 60.78

6 CONCLUSIONS

We study using sentiment analysis based on lyrics and comment texts and user social relationships to conduct music recommendations in this paper. To this end,we intensively analyze users' listening behaviors considering their changes with the emotional categories of listening to the song. Moreover,we analyze the listening behavior of users and followers, and find that the similarities between users' listening behaviors and different followers are different. We explore the features in constructing a music recommendation model,and the proposed model using two features performs the best. We also find that different features have different importance, which is in line with real life experiences. In future work,we are interested in applying our model into music recommendation.

References

1. An, Y., Sun, S., Wang, S.: Naive bayes classifiers for music emotion classification based on lyrics pp. 635–638 (2017)
2. Bu, J., Tan, S., Chen, C., Wang, C., Wu, H., Zhang, L., He, X.: Music recommendation by unified hypergraph: combining social media information and music content pp. 391–400 (2010)

3. Chen, X., Tang, T.Y.: Combining content and sentiment analysis on lyrics for a lightweight emotion-aware chinese song recommendation system pp. 85–89 (2018)
4. Choi, J., Song, J., Kim, Y.: An analysis of music lyrics by measuring the distance of emotion and sentiment pp. 176–181 (2018)
5. De Assuncao, W.G., Neris, V.P.D.A.: An algorithm for music recommendation based on the user's musical preferences and desired emotions pp. 205–213 (2018)
6. Deng, S., Wang, D., Li, X., Xu, G.: Exploring user emotion in microblogs for music recommendation. Expert Systems With Applications **42**(23), 9284–9293 (2015)
7. Feng, Y., Zhuang, Y., Pan, Y.: Music information retrieval by detecting mood via computational media aesthetics pp. 235–241 (2003)
8. Feng, Y., Zhuang, Y., Pan, Y.: Popular music retrieval by detecting mood pp. 375–376 (2003)
9. JA, R.: A circumplex model of affect. Journal of Personality Social Psychology **39**, 1161C1178 (1980)
10. Jannach, D., Kamehkhosh, I., Lerche, L.: Leveraging multi-dimensional user models for personalized next-track music recommendation pp. 1635–1642 (2017)
11. Jiang, W., Wu, J., Wang, G.: On selecting recommenders for trust evaluation in online social networks. ACM Transactions on Internet Technology **15**(4), 14 (2015)
12. Jiang, W., Wu, J., Wang, G., Zheng, H.: Forming opinions via trusted friends: Time-evolving rating prediction using fluid dynamics. IEEE Transactions on Computers **65**(4), 1211–1224 (2016)
13. Kelen, D., Berecz, D., Beres, F., Benczur, A.A.: Efficient k-nn for playlist continuation (2018)
14. Li, M., Jiang, W., Li, K.: When and what music will you listen to? fine-grained time-aware music recommendation pp. 1091–1098 (2017)
15. Mahedero, J.P.G., Martinez, A., Cano, P., Koppenberger, M., Gouyon, F.: Natural language processing of lyrics pp. 475–478 (2005)
16. Mao, K., Chen, G., Hu, Y., Zhang, L.: Music recommendation using graph based quality model. Signal Processing **120**, 806–813 (2016)
17. Nakamura, K., Fujisawa, T., Kyoudou, T.: Music recommendation system using lyric network pp. 1–2 (2017)
18. Pacula, M.: A matrix factorization algorithm for music recommendation using implicit user feedback (2009)
19. RE, T.: The Biopsychology of Mood and Arousal. Oxford Univ. Press, Oxford, U.K (1989)
20. Sanchezmoreno, D., Gonzalez, A.B.G., Vicente, M.D.M., Batista, V.F.L., Garcia, M.N.M.: A collaborative filtering method for music recommendation using playing coefficients for artists and users. Expert Systems With Applications **66**, 234–244 (2016)
21. Wang, X., Chen, X., Yang, D., Wu, Y.: Music emotion classification of chinese songs based on lyrics using tf*idf and rhyme pp. 765–770 (2011)