Music and Movie Recommendation System

Article *in* International Journal of Engineering Trends and Technology · July 2018

DOI: 10.14445/22315381/IJETT-V61P229

CITATION
1

READS
1,995

2 authors, including:

Geetanjali Kale
Pune Institute Of Computer Technology
27 PUBLICATIONS **95** CITATIONS

SEE PROFILE

Music and Movie Recommendation System

Puja Deshmukh[#1], Geetanjali Kale[*2]
[#1] P.G. Student, [#2] Professor, *Department of Computer Engineering, Pune Institute of Computer Technology, Maharashtra, India*

Abstract

The advancement in technology has led to massive use of internet. People use social networking sites almost every day. These sites help people to express themselves to the online society. People use various posts to express themselves, these posts are nothing but short informal texts having positive, negative or neutral emotions. Music is an important aspect of human life. People prefer to listen to music more often than any other activity. With the internet technology, a huge amount of music content containing music of various genres has become easily available to millions of users around the world. Music collection since decades and comprising of various genres of music is available. The major difficulty that the users face is to select appropriate music from such huge collection. Similarly, a huge collection of movies comprising of various genres is also available. Music and movie recommender will recommend music and movies to the user based on their mood along with an emoticon symbolizing their mood. The mood of the user will be derived by performing sentiment analysis of the user posts. In order to provide better recommendations, sentiment analysis will be performed on the lyrics of the songs using NLP. Random forest algorithm will be used for classifying the lyrics into various categories (happy, sad, joyful).

Keywords : *Data Mining, Recommendation System, Social Network, Sentiment Analysis, Movie Recommendation.*

I. INTRODUCTION

Music is an integral part of human life. Music varies according to personal tastes or opinions. It not only can convey emotions but also can also affect listener's mood. Music is easy to listen, but hard to find. What makes it more difficult is that the tastes in music vary from one person to another and also their mood. Also, there is a huge collection of various genres of music composed by many singers. With so many options it is often difficult to know what song to hear next.

In the last few decades the access to multimedia content was limited because of less availability of resources, but now due to the huge advancement in the network, internet plays a very important source in retrieving multimedia information as music, video, books etc. Music is a very important aspect of human life. Research has also shown that people prefer to listen to music more often than any other activity [1] (i.e. watching television, reading books and watching movies). With the internet technology, a huge amount of music content has become available to millions of users around the world.

In recent years, personal music consumption behavior has also changed dramatically. Due to the vast availability of music, user's personal music collection contains plenty of songs. Efficient searching and labeling techniques are required to manage this huge collection [2]. There is a huge collection of various genres of music since decades. With millions of artists and songs in the market, it is becoming increasingly difficult for people to search for music. The problem with such huge amount of music is to organize and manage millions of music titles produced by various musicians over the decades. To overcome these problems music recommender systems are used.

Music recommender system helps users to filter and select songs as per their music preference and mood. A recommendation system would be considered good if it automatically determines the user mood and recommend music accordingly. Music recommendation systems are decision support tools that reduce the information overload by retrieving only those items that are estimated as relevant to users based on the user profile that is generated by the recommendation system [3]. Similar is the case with movies, a huge collection of movies of various genres is available. It becomes difficult for the user to decide which movie to watch.

The proposed recommendation model will determine the mood of the user from their posts and then recommend them music and movies.In order to provide a better recommendation, the songs will be analyzed based on their lyrics and then categorized accordingly. Along with the recommendation, the system would display an emoticon representing the mood of the user. The music category(happy/sad) is determined by performing sentiment analysis on the lyrics of the music.

II. RELATED WORK

An ideal recommendation system must try to automatically determine the user mood and then recommend them accordingly [4][5]. The following

section surveys various techniques that are used by the recommendation systems.

A. *Metadata-Based Information Retrieval*

This is the most basic method and it is the simplest way to search and retrieve music [6]. Metadata information mostly consists of textual metadata, such as the song title, name of the artists, and lyrics of the songs to find the appropriate songs.

B. *Collaborative Filtering*

These systems recommends songs based on the interests of similar users [7]. Collaborative filtering is one of the most successful and commonly used approach in recommendation systems. Collaborative filtering assumes that if user A and B rate k items similarly or have similar behaviour, they will rate other items also in that similar fashion [8]. In this technique the past ratings are considered and the strongest co-relation among them is found instead of computing the nearest neighbour between the users. Hence, we can predict the scores of the unseen items. Memory-based, Model-based, and Hybrid collaborative filtering are the three main subcategories of collaborative filtering [9][10].

C. *Content-Based Music Retrieval*

In content-based filtering the recommendations are provided by carefully analysing the music tracks[11][12]. After the analysis, various features are extracted from the music tracks and these features are used to compute the distance between the music tracks. K-means clustering, Expectation-Maximization with Monte Carlo Sampling, Average Feature Vectors with Euclidean Distance are the three main measures that are used to compute the distance.

D. *Emotion-Based Model*

Music is self-expressive. Music is rich in content and is self-sufficient to express itself. The emotion model is used by a web service named Musicovery. Using this service, the users can locate their emotion is 2D space as valence (how positive or negative) and arousal (how energetic or soothing) [13] [14]. Emotion model uses various acoustic features. Energy, rhythm, temporal, spectral, and harmony are the various acoustic features that have been widely used in emotion recognition. [19] proposes an android application 'EmoPlayer', which captures the facial expressions of the user through camera and determines their mood from the expressions and recommends music accordingly.

E. *Context-Based Model*

Context-based model considers the public opinions to recommend music [15]. Public opinions can be obtained through social networking sites such as Facebook, Twitter, YouTube etc. in the form of comments, reviews, posts etc. In context-based model the information is obtained by using various document mining techniques. Research has suggested that this model performs well due to the collection of social information [16] [17].

F. *Hybrid Model*

Hybrid model combines two or more of the above-stated models to improve the performance of the music recommendation system. Broke [18] proposed that various methods such as weighted, switching, mixed, feature combination and cascade can be used to develop a hybrid recommendation model.

The existing music recommendation system recommends music based on the pre-classified music collection. The songs are already classified into different genres. In the proposed recommendation system the songs are classified based on the lyrics. Analysis of the lyrics is performed to determine their sentiment value and then they are classified into different genres accordingly. Music and movies are then recommended according to the mood of the user.

III.SYSTEM ARCHITECTURE

The proposed model identifies the category of music based on the lyrics of the songs. The lyrics are manually collected and provided to the user interface. The collected lyrics are then processed to remove the stop words. After the removal of stop words, sentiment analysis is performed on the lyrics and then the lyrics are classified into categories (happy, sad, energetic) using random forest algorithm. After the completion of the analysis of lyrics, they are stored in the music database according to their category.

The sentiment value of the entire sentence is obtained by an arithmetic sum of the sentiment value obtained for each word in the sentence. Sumsenti(F) describes the total sentiment value calculated for the given sentence. The variable Wi describes each word in the sentence and Dictionary. Value gives the sentiment value for each word in the dictionary.

$$Sumsenti(S) = \sum_{n=1}^{N} Dictionary.value(Wi)$$

The meta-data of movies is manually collected which include movie title and genre. The movies will be categorized as romantic, thriller and comedy based on the meta-data and then stored in the database accordingly.

The user would be requested to either enter a post in the provided text box or to select a post from the provided list according to their current mood. These posts will then be processed to remove the stop words. After the removal of stop words, they will be analysed to determine the sentiments from the posts.

The recommendation system will then consider the sentiments derived from the user posts and then recommend the user with appropriate music and movies that suits their mood. The system will also display an emoticon (happy face, sad face, pensive face) to represent the mood of the user.

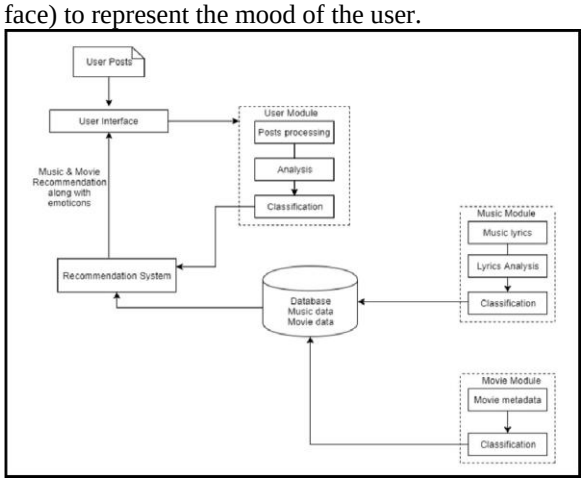

Figure1: System Architecture

A. Dataset

The lyrics of 200 Hollywood songs are collected manually. The songs are in the English language. The user posts will be collected dynamically through the user interface provided by the recommendation system. For recommending movies, the metadata of 100 Hollywood movies is collected. The movies are in the English language.

B. Software Used

The proposed system has been implemented using Java. The software as Java Development Kit and Java Runtime Environment are required. The data regarding music, user posts and movies is stored using MySQL Server 5.0. NetBeans have been used for the integrated development environment. The system is tested and successfully executes on windows and Linux platforms.

C. Pre-processing

Pre-processing activity will be performed on user posts and lyrics of songs. The first pre-processing activity will be the removal of stop words. After the stop word removal, Porter Stemming algorithm will be used to perform stemming. The pre-processing activity will provide the keywords that would be used for further analysis.

D. Analysis

The keywords obtained from the pre-processing activity will be used for sentiment analysis and to determine the sentiments of the lyrics and user posts. NLP algorithm will be used to perform sentiment analysis.

E. Classification

The results of sentiment analysis will be used to perform classification of the music lyrics. Random forest algorithm is used for classification. Random forest algorithm will generate multiple decision trees, and the trees having maximum votes will be consider for classification. The songs will be classified as happy, sad and energetic. After the classification, the songs will be stored in the database according to their category. This data will be used by the proposed system to provide recommendations.

IV. PSEUDO CODE

Step 1: Begin
Step 2: Collect the user posts and lyrics. Posts and lyrics are nothing but simple English sentences.
Step 3: Remove stop words from the posts/lyrics.
Step 4: Determine the sentiment of each word using SentiWordNet.
Step 5: Determine the sentiment of the entire sentence as described below:

$$Sumsenti(S) = \sum_{n=1}^{N} Dictionary.value(Wi)$$

Step 6: Classification of posts and lyrics is done using Random Forest algorithm.
Step 7: Recommend music and movies based on the user posts.
Step 8: End.

V. RESULTS

The following section describes the results obtained from the proposed music and movie recommendation system. Table I describes the classification of songs based on their lyrics into different categories. Table II describes the classification of user posts into different categories based on their derived sentiments. Table III briefly describes the recommendations of songs and movies that are provided to the user. The post used in Table III have been classified as happy post in Table II, hence movies and songs that have been classified as happy are recommended.

Lyrics	Category
When you are happy and you know it clap your hands.	Joyful
Sometimes everything is wrong.	Sad
So happy together.	Joyful
Everybody hurts sometimes.	Sad
Don't worry be happy.	Joyful
The happy organ.	Joyful
You've made me so very happy.	Joyful
Wish you were here.	Sad
Love can make you happy.	Joyful

Table 1: Classification of Songs

User Posts	Category
Today we have promised each other to be bound in love. & Happy	Happy
Sometimes everything is wrong.	Sad
Whenever i am alone, i think, i remember and become sad & Sad	Sad

I am roaming on the silent roads. silence makes me happy & Happy	Happy
Proud prouder proudest. Hats off to the whole team. & Joyful	Joyful
Results are out.....Supper happy!! I passed my exams. & Joyful	Joyful

Table 2: Classification of user Posts

User posts	Music	Movie
Results are out. Supper happy!	The happy organ.	Roman holiday
	Love can make you happy.	Midnight in Paris
	So happy together	As good as it gets
	Don't worry be happy.	Home alone

Table 3: Music and Movie Recommendation

VI. CONCLUSION AND FUTURE WORK

Music and movie recommendation system will reduce human efforts by searching the huge media collection containing many songs and movies of various genres. Recommendations to the user will be provided according to their mood. The mood of the user will be determined by their posts. Hence this system will provide higher user satisfaction in less time and efforts as they will be automatically provided a recommendation for music and movies based on their mood.

The proposed recommendation system works on songs and movies in English language. This system can be further extended to recommend songs and movies in Hindi language or other native languages since the users can better express themselves using their native languages. The emotion category can be further improved to consider more complex emotions as hatred, anxiety, jealousy, excitement, etc.

REFERENCES

[1] Rentfrow, Peter J., and Samuel D. Gosling. "The do re mi's of everyday life: the structure and personality correlates of music preferences." Journal of personality and social psychology 84.6 (2003): 1236.

[2] Celma Herrada, Òscar. Music recommendation and discovery in the long tail. Universitat Pompeu Fabra, 2009.

[3] Ricci, Francesco, Lior Rokach, and Bracha Shapira. "Introduction to recommender systems handbook." Recommender systems handbook. springer US, 2011. 1-35.

[4] Lamere, Paul. "Social tagging and music information retrieval." Journal of new music research 37.2 (2008): 101-114.

[5] Pachet, François, and Daniel Cazaly. "A taxonomy of musical genres." Content-Based Multimedia Information Access-Volume 2.

[6] Downie, J. Stephen. "Music information retrieval." Annual review of information science and technology 37.1 (2003): 295-340.

[7] Sarwar, Badrul, et al. "Item-based collaborative filtering recommendation algorithms." Proceedings of the 10th international conference on World Wide Web. ACM, 2001.

[8] Su, Xiaoyuan, and Taghi M. Khoshgoftaar. "A survey of collaborative filtering techniques." Advances in artificial intelligence 2009 (2009): 4.

[9] Resnick, Paul, et al. "GroupLens: an open architecture for collaborative filtering of netnews." Proceedings of the 1994 ACM conference on Computer supported cooperative work. ACM, 1994.

[10] Adomavicius, Gediminas, and Alexander Tuzhilin. "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions." IEEE transactions on knowledge and data engineering 17.6 (2005): 734-749.

[11] Li, Qing, Byeong Man Kim, and Dong Hai Guan. "A music recommender based on audio features." Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval. ACM, 2004.

[12] Kuo, Fang-Fei, et al. "Emotion-based music recommendation by association discovery from film music." Proceedings of the 13th annual ACM international conference on Multimedia. ACM, 2005.

[13] Mann, Mark, Trevor J. Cox, and Francis F. Li. "Music Mood Classification of Television Theme Tunes." ISMIR. 2011.

[14] Song, Yading, Simon Dixon, and Marcus Pearce. "A survey of music recommendation systems and future perspectives." 9th International Symposium on Computer Music Modeling and Retrieval. Vol. 4. 2012.

[15] Song, Yading, Simon Dixon, and Marcus Pearce. "A survey of music recommendation systems and future perspectives." 9th International Symposium on Computer Music Modeling and Retrieval. Vol. 4. 2012.

[16] Wang, Dingding, Tao Li, and Mitsunori Ogihara. "Are tags better than audio features? The effect of joint use of tags and audio content features for artistic style clustering." 11th International Society on Music Information Retrieval Conference, number ISMIR. 2010.

[17] Burke, Robin. "Hybrid recommender systems: Survey and experiments." User modeling and user-adapted interaction 12.4 (2002): 331-370.

[18] Balabanović, Marko, and Yoav Shoham. "Fab: content-based, collaborative recommendation." Communications of the ACM 40.3 (1997): 66-72.

[19] Iyer, Aurobind V., et al. "Emotion based mood enhancing music recommendation." Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2017 2nd IEEE International Conference on. IEEE, 2017.