International Journal of Multimedia and Ubiquitous Engineering
Vol.11, No.8 (2016), pp.163-170
http://dx.doi.org/10.14257/ijmue.2016.11.8.18

User Musical Taste Prediction Technique Using Music Metadata and Features

Minseo Gong[1], Jae-Yoon Cheon[2], Young-Suk Park[3], Jeawon Park[4] and Jaehyun Choi[5*]

[1,4,5]Graduate School of Software, Soongsil University, Seoul, Korea
[2,3]Graduate School of IT Policy & Management, Soongsil University, Seoul, Korea
gutssoul1@gmail.com[1], jycheon2@naver.com[2], psk1@lgnsys.com[3], jwpark@ssu.ac.kr[4], jaehyun@ssu.ac.kr[5]

Abstract

The digital music market has been growing significantly in the past years. In music streaming services, music recommendation plays a key role, but Korean users' recognition about their music service is not high because the service's recommendation accuracy is not good. Therefore, this paper suggests technique to predict the user's musical taste. This technique proceeds through a four-step process; data collection, data pre-processing, feature extraction, and machine learning. Collection of data was taken from TOP 100 chart in 'Melon', the number one music service provider in Korea from December 2013 to March 2015. Then, collected MP3 file format is converted into WAV file format. In the stage of feature extraction, we classify the genre from the music's metadata and extract factors that can be taken using STFT's ZCR, Spectral Rolloff, Spectral Flux. In the stage of machine learning, we produce a prediction model in a variety of classification techniques. To measure the performance of the created prediction model, 456 data were used for training dataset and 130 data were used for validation dataset. Since the results of experiment show an average of 78% of accuracy, the proposed technique seems to be effective.

Keywords*: musical taste, prediction technique, machine learning*

1. Introduction

The development of Korea's IT technologies and fast Internet environment, and the popularization of smart phones have brought a big change in how to consume music content [1]. In particular, consumer's media consumption was accelerated by the popularization of LTE (Long-Term Evolution). As the growth of digital music sales in the world's music market led to the growth of the entire music market [2], the digital music market has since had significant growth. KOCCA (Korea Creative Content Agency) has forecast that the domestic market i on the scale streaming music about 5 trillion. Therefore, there is a high growth potential in digital music market [3]. In music streaming services, recommendation and discovery function play key role [4]. However, users' recognition about Korea's music recommendation services is not good since, as shown in a statistical survey on usefulness and credibility of Korea music recommendation service, users experience each about 47.8% and 44.1% [5]. In this analysis, it seems that accuracy of recommendation service cannot meet user's expectations.

In general, the typical technique used in recommendation systems use two techniques, which are collaborative filtering and content-based method. Collaborative filtering is a

* Corresponding author

ISSN: 1975-0080 IJMUE

technique that recommends each other in groups that have similar preference after dividing the group using users' evaluation information about an item. Content-based method is a technique that recommend items, similar to the user profile after creating user profile analyzing contents regarding that user's preferences in the past [6]. Collaborative filtering takes time in collecting the data to compare it to the user; it is impossible to know the preferences of the individual. In addition, there is a limit to ignore the information of the other user by like based on the user information with a portion of the same tendency. On the other hand, the content-based method has the advantage that it can be recommended only if sufficient data and the user profile for the item does not have this problem. Therefore, in this paper, we propose a technique to predict the user's musical taste using content- based method to provide more personalized music recommendation service.

2. Related Research

2.1 Related Researches about Music Recommendation

Until now, a number of studies have been conducted on music recommendation. One of the studies about music recommendation, [7], was implemented through vector similarity after analyzing the sound waves of one hundred of domestic and international music and using the user's music download lists. Based on this, [8] variable cluster analysis was performed after extracting features of one hundred of domestic and international music, and recommended music similar to the user's musical taste using time weighted method. Verification of the proposed method has shown that the proposed method is effective throughout the experimental verification. In [9], the study provided personalized music recommendation using AHP, or analytic hierarchy tool, in approximately 200 songs and support user feedback based on Bayesian network. In [10], after creating genre vector using time weights with the music dataset of Last.FM, recommend music through preference prediction using genre vector was presented. However, the above studies were verified from a small number of users and music data. In this paper, we utilize 586 music data sources to the targets for 26 Korean people.

2.2 Extracting Features of Music

Music consists of waveform. Kinds of waveforms feature are pitch, loudness, duration, timbre and so on. More specific features are ZCR (Zero Crossing-Rate), Spectral centroid, Spectral Rolloff, Spectral Flux based on STFT (Short Time Fourier Transform). Also, MFCC (Mel-Frequency Cepstral Coefficients) express perceptual features, LPC (Linear Prediction refection Coefficients) [11]. In this paper, we use a combination of ZCR, Spectral Rolloff, Spectral flux. This combination shows high accuracy in algorithm to find the music within a certain period [11].

2.2.1. Zero Crossing Rate (ZCR)

The ZCR measures the number of times the audio signal goes through the zero base line in a given interval. In other words, consecutive sampling values in discrete signal generate when these values are different from each other, which is very useful to the voice in division, analysis, recognition [11]. The formula is shown in Figure 1 represents the ZCR.

2.2.2. Spectral Roll-Off

The spectral rolloff is a feature that distinguishes between the speech interval and non-speech interval, and therefore is another method of measuring the spectral shape with the Spectral Centroid. The spectral rolloff is defined as the frequency Rt below which 85% of the magnitude distribution is concentrated. In other words, this shows how much of the signal's energy is concentrated in the lower signals [11]. The formula shown in Figure 2 represents the Spectral Roll-off.

$$Z_i = \sum_{n=1}^{N} \frac{|sgn|s_i(n)| - sgn|s_i(n-1)||}{2}$$

$$sgn|s_i(n)| = \begin{cases} 1 & s_i > 0(n) \\ -1 & s_i > 0 \end{cases}$$

s = noisiness of signal, n = sample, i = frame

Figure 1, Formula of ZCR

$$\sum_{n=1}^{R_t} M_t[n] = 0.85 \times \sum_{n=1}^{N} M_t[n]$$

M_t [n] is the size of Fourier Transform on the frequency n and frame t

Figure 2. Formula of Spectral Roll-Off

2.2.3. Spectral Flux

$$F_t = \sum_{n=1}^{N} (N_t[n] - N_{t-1}[n])^2$$

N_t[n], N_t[n-1] are the normalized magnitude of the Fourier Transform at the previous frame t-1, and the current frame t respectively.

Figure 3. Formula of Spectral Flux

3. Suggested Technique

In this chapter, we explain the user's musical taste prediction technique for music recommendation. The overall process of the suggested technique consists of four stages; first stage is Data Collection, second is Data Preprocessing, third is Music feature extraction, last stage is Machine learning, as shown in Figure 4.

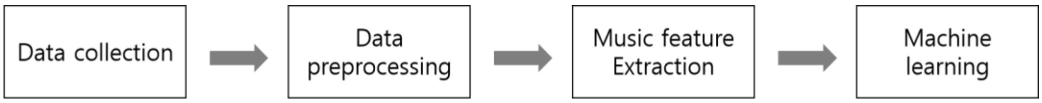


Figure 4. Overall Process of Suggested Technique

3.1. Data Collection

To predict the user's musical taste, there are two requirements. First, it is necessary to create Source list that reflects the public. Second, the preferences of a larger part of the public should be reflected in the Source list. Since there are a very large number of music sources in the world, this source list with the number of preference reflecting the public can choose the source of the much preferred. In addition, accuracy of the user profile is also going up. Therefore, in order to satisfy these requirements, this paper collects music sources corresponding to the list of the TOP 100 chart in 'MELON', the number one music service provider in Korea, which has a membership of approximately 26 million people. We collect music sources of the TOP 100 chart in 'MELON' from December 2013 to March 2015. Figure 5 shows a portion of the TOP 100 chart of 'MELON'.

Figure 5. TOP 100 Chart of MELON

3.2. Data Preprocessing

In order to extract features from the music source, it is necessary to go through preprocessing, which converts the compressed sound file to an uncompressed sound file called WAV. In this paper, we used the "media.io" [12] web site, which provides a service for converting a number of compressed music. The properties of a WAV file completing conversion are shown in Table 1.

Table 1. Properties of WAV File Completing Conversion in Media.Io

Column	Value
Average bit rate	1411 Kbps
Bit-depth	16bit
Channel	2
Sampling rate	44,100 Hz

3.3. Music Feature Extraction

This paper used jAudio [13], a Java-based music feature extraction system, for extracting music features from a converted WAV file. JAudio supports both GUI and command-line interface, and WAV files are put in jAudio as input value for calculating music feature and produces result lists in the ACE, XML, ARFF file format. This paper extracted three kinds of features mentioned in Chapter 2 above, which are ZCR (Zero-crossing rate), Spectral Roll-off, and Spectral Flux in STFT.

3.4. Machine Learning

Machine learning is the method wherein a computer program improves performance gradually like humans. In machine learning, the program improves learning capabilities of the analysis model through learning using new incoming data [14]. The input data for machine learning is training data consisting of records including class label. Records are made up of property values, and property values can be divided into categorical type and numerical type. Classification algorithm in machine learning generates a classification model that describes class label from the property value point using training data as input value [15]. This paper generated a prediction model using classification algorithms based on machine learning theory, and used Bayesian Network, Random Forest, Support-Vector Machine (SVM). Bayesian Classifier is a classifier that predicts the probability that a given set of attributes belong to a certain class based on Bayesian theory, while the Bayesian Network can express a subordinate relationship where the subset of the set of properties combining the graphic theory with Bayesian classifier [16]. Random Forest algorithm was first proposed by Ho, *et al*. [17]. It demonstrated operation of the algorithm by "Law of large numbers". Support Vector Machine (SVM) is one of the statistical learning theories and have been recognized as one of the effective classification methods as compared to supervised machine learning algorithms [18]. Also, it can be introduced into solving pattern recognition problems with small samples and learning problems such as function estimation [19].

4. Experimental Results and Analysis

To verify the usefulness of the proposed technique, an experiment was performed. Prior to experiment, 586 MP3 files that were removed from overlapping sources in lists of TOP 100 chart of 'MELON' from Dec. 2013 to Mar. 2015 were converted WAV files. Each converted WAV file consisted of 16-bit, 2 channels, 44100Hz form. Three types of music features, which are ZCR, Spectral Roll-off, Spectral Flux, were extracted from each of the WAV files. And the genre tag was extracted from metadata each sources. Prior to making a prediction model by machine learning, each of the users' profiles were obtained through the questionnaire. The form of the questionnaire is shown in Figure 6. Among user preference data were about 586 songs, user preference data about 456 songs from Dec. 2013 to Dec. 2014 were used as training dataset and user preference data about 130 songs from Jan. 2015 to mar. 2015 were used as validation dataset to validate the prediction model. This paper used open source software (the 'WEKA' developed by the University of Waikato in New Zealand) to measure the performance of prediction models through machine learning process. Classification algorithms used were Bayesian Network, Random Forest, and SVM.

singer	title	Streaming Address	Bad : 0 , Good : 1
2AM	Just stay	https://www.youtube.com/wat	0
2AM	regret	https://soundcloud.com/melor	0
Apink	NoNoNo	https://soundcloud.com/wenin	1
M&N (Miryo&Narsha)	Tonight	https://soundcloud.com/enjoy-	0
San E	Story of Someone	https://soundcloud.com/oliun/	0
San E	Where Did you Sleep	https://soundcloud.com/hanga	0
San E	To my Ex-Girlfriend	https://soundcloud.com/refrida	0
T	Touch Love	https://soundcloud.com/kimro	1
YB & Lena Park	December	https://soundcloud.com/melor	0
Ylvis	The Fox	https://soundcloud.com/afterli	1
Kim Bo kyung	Want to go back in time	https://soundcloud.com/mark-	0
Lim Kim(Togeworl)	Goodbye 20	https://soundcloud.com/k9808	0

Figure 6. Form of Questionnaire about Music Preference

4.1. Result Analysis and Conclusion

Experimental results using training dataset and validation dataset are shown in Table 2. As a result of comparing the arithmetic mean of precision and F1 score of prediction model, using Support Vector Machine showed the best performance with accuracy of 80.8% and F1 score of 0.812. Results of this paper can be seen as an advanced result when comparing with results of previous research, using cluster analysis [8][9], using AHP method [10], or using genre vector [11]. In this paper, we suggested one way that can be applied to the recommendation function, which has a key role in the digital music market with high growth potential. Moreover, it showed that the technique of extracting music features and genre in metadata can be used in the process of refining data necessary for the prediction method that is based on machine learning theory. On the other hand, it is expected to require more music features DB and user preference analysis method in order to increase the accuracy of the presented prediction model in this paper. In future research, there is a need to take action to complement the limitations of the prediction model with a progressing study of music features affecting user preference and a study of reliable analysis techniques of user preference.

Table 2. Performance Comparison of Classification Algorithms

	Precision	F1 score
BayesNet	0.7994	0.804
Random Forest	0.767	0.773
SVM	**0.808**	**0.812**

References

[1] H. Kwon, "A research on impact of free music streaming service on music industry", KOCCA15-14 in Korea Creative Content Agency(KOCCA), **(2015)**, pp. 12.
[2] J. G. Kang, "Recent trends in world music industry sector", KOCCA Statistics Briefing 13-16 in KOCCA, **(2013)**, pp.1-2.
[3] K. Lee, "Streaming changed music market landscape", ETNEWS, **(2014)**.
[4] http://www.etnews.com/20141015000154
[5] "Major music streaming service and technical analysis of changing landscape of digital music market", CT- issue analysis in KOCCA, **(2014)**.
[6] M. Kim, "A research on impact analysis and recommendation system architecture analysis of online digital music retailers", Ministry of Culture, Sports and Tourism (MCST), **(2013)**.
[7] J. Basicilico and T. Hofmann, "Unifying Collaborative and Content-based Filtering", Proceedings of International Conference on Machine Learning (ICML), **(2004)**.
[8] D. Kim and J. Lee, "User adapted music recommendation system using sound wave", Proceedings of Korean Institute of Information Scientists and Engineers(KIISE), vol. 33, no. 2, **(2006)**.
[9] J. Kim, T. Yoon, D. Kim and J. Lee, "A personalized Music Recommendation System with a Time-weighted Clustering", Journal of Korean Institute of Intelligent Systems(KIIS), vol. 19, no. 4, **(2009)**, pp. 504-510.
[10] S. Y. Lee and N. K. Kim, "A Music Recommendation System based Context Awareness using Bayesian Network", Proceedings of Korean Institute of Intelligent Systems(KIIS) Fall Conference, vol. 22, no. 2, **(2012)**.
[11] T. Hwang, C. Park, S. Kim and S. Kim, "Music recommendation System using Genre Vector", Proceedings of KIISE, **(2013)**, pp. 1491-1493.
[12] G. Tzanetakis, "Manipulation, Analysis and Retrieval systems for audio signals", **(2002)**, pp32-33, 42-43
[13] Music file Conversion Web Site, www.media.io
[14] M. Corym F. Ichiro and D. Philippe, "jAudio: A feature extraction library", in: Proceedings of the International Conference on Music Information Retrieval, **(2005)**, pp. 600-3.
[15] H. Kim, "Introduction to Data Science", Hongrung publishing company, **(2014)**, pp. 170.
[16] K. H. Kim, "Study on Classification Algorithm based on Weight of Support and Confidence Degree", Journal of Information and Communication Convergence Engineering, vol. 13, no. 4, **(2009)**, pp. 700-713.
[17] D. Y. Kim, "A study on Stock Price Prediction Model Using Sentiment Analysis and Machine Learning Based on SNS and News Articles", Thesis of Soongsil University, **(2014)**.
[18] T. K. Ho, "Random Decision Forests", Proceedings of the IEEE Document Analysis and Recognition, **(1995)**.

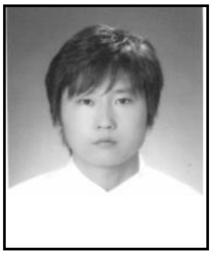
[19] S. Y. Park, J. Chang and T. Kihl, “Document Classification Model Using Web Documents for Balancing Training Corpus Size per Category”, Journal of Information and Communication Convergence Engineering, vol. 11, no. 4, **(2013)**, pp. 268-273.
[20] F. G. Wang, M. C. Kim, S. K. Park and G. P. Kwak, “Adaptive States Feedback Control of Unknown Dynamics Systems Using Support Vector Machines”, Journal of Information and Communication Convergence Engineering, vol. 6, no. 3, **(2008)**, pp. 310-314.

Authors

Minseo Gong received his BFA degree in year 2011 in Ceramic-design from Sangmyung University, Korea. Currently, he is a graduate student in the Graduate School of Software, Soongsil University, Korea. His research area includes Knowledge Discovery in Databases, Artificial Intelligence, web-based applications and data mining.

Jae-Yoon Cheon received his bachelor's degree of Computer Science in Korea National Open University, Seoul (2000) and master's degree of Computer Science in Yonsei University(2003). Now, he is studying his doctoral degree IT Policy & Management in Graduate Soongsil University. From 1988 to present is currently in LG Nsys and working as a Network consultant and Project Manager. His research interests focus on Open Source Software and Software Defined Networking.

Young-Suk Park received his bachelor's degree of Mathematics Education in Sungkyunkwan University, Seoul (1985) and master’s drgree of Engineering in Graduate Yonsei University, Seoul (2012). He is studying his doctoral degree IT Policy & Management in Graduate Soongsil University. From 1990 to present is currently in LG Nsys and working as a system engineer, the main areas of interest include Information Technology Project Manager, Cloud, IOT, *etc.*



Jeawon Park received his Ph.D. degree in Computer Science from Soongsil University in Korea, 2011. He is a profressor at the Graduate School of Software, Soongsil University. His research interests are in areas of Software Testing, Software Process, Web Services, and Project Management.

Jaehyun Choi received his Ph.D. degree in Computer Science from Soongsil University in Korea, 2011. He is a profressor at Graduate School of Software, Soongsil University. His research interests are in areas of Data Processing, Service Engineering, Software Engineering, and Text Mining.

