Applying data mining for sentiment analysis in music

Lucía Martín Gómez, María Navarro Cáceres

BISITE Research Group, University of Salamanca, Calle Espejo, S/N, 37007 Salamanca, Spain

Abstract. Listening to music can affect people emotions. They can experience simultaneous feelings, such as happiness and hope, or sadness and angry, when a song is being played. However, infering emotions that can be caused by a musical fragment is a complex task. To deduce relationships between feelings and music, we propose a sentiment analysis method based on data mining. In particular, different musical features are extracted and classified to analyze the influence of some music parameters on human emotions. In this process, data mining algorithms such as Random k-Labelsets, Multi-Label k-Nearest Neighbors or Apriori have been essential for the success of our proposal.

Keywords: Music, sentiment analysis, data mining, multi-label classification, music information retrieval

1 Introduction

Music has been considered as a mean of expressing and arousing emotions. Different musical features such as melody, rhythm, harmony or timbre causes mental human reactions, which can be feelings, emotions or moods. To demonstrate the relationships between music and sentiments, some authors have proposed different psychological analysis. [5] address the role of our emotions in the composition of music, the ways that emotions can be communicated via musical structure from different perspectives. [22] investigates if the universal appeal of music lies in the emotional rewards that music offers with four interrelated studies with different listeners. [2] carries out a deep study about the different features of music and how they influence the emotions in individuals or in the society. [8] describes principles underlying the evocation of emotion with music: evaluation, resonance, memory and social functions. [15] investigated the relationship between musical characteristics and the ability of some participants to recognize five basic emotions (happiness, sadness, tenderness, fear, and anger). All of them concludes that music is inherently related to sentiments.

With the rise of opinion mining and sentiment analysis, the fact of finding out what people think and feel has become a necessity [14]. Some techniques such as text-categorization [13] or support vector machines [11] have been widely used to solve similar problems of association in other areas. In relation to music, there are many works for predicting the genre a song belongs to [12]. Otherwise, text information (e.g., websites, tags, and lyrics) [4], and a combination of text and audio information [3, 6] have been used for music mood classification.

This work aims to classify music by the emotions produced when a person is listening to a musical fragment. Additionally, the relationships between musical features and emotions will be deeper analyzed. To do so, a dataset which contains rhythmic and timbre features for different songs is selected. Each song is described with some rhythmic and timbre features. Our final goal is to label the musical tracks with six different emotions.

The proposal will be resolved as a multi-label problem with data mining, meaning the detection of emotions in music as an approximation to sentiment analysis, which will be modelled as a multi-label task and solved with data mining techniques. As a case study, we make use of the dataset called Emotions from Mulan, which consists of 593 songs with 72 musical features. The music classification will be made using the WEKA framework and its extension MEKA, which provides support for multi-label and multi-target problems. Furthermore, some unsupervised learning techniques will be applied in order to extract some relationship patterns between the considered emotions and the music features.

The next section explains the problem solving process: preprocessing tasks and algorithms used. The results obtained with data mining techniques are shown in Section 4. Finally, last section discusses the evaluation of the sentiment analysis and future work proposed.

2 Multi-label Classification

Classification problems are widespread nowadays. Some approaches aim to classify some aspects related to biology and medicine such as tumor types [9] or to perform human recognition [1] and human detection [7] by using convolutional neural networks. All of them classify the different instances with only one class.

However, instances of some problems should be classified with more than one class at the same time; this is known as multi-label classification [17]. In case of films, such classification is easy to understand: for example, "Shrek" movie could be classified as an animated film, but also as comedy, fantastic or adventure.

Multi-label classification has been widely applied in music tasks in order to predict the genre [10] and the feelings a song produces on the listener [21]. In this work, musical features will be used to classify songs into a set of emotions.

3 Problem description

The present proposal addresses a multi-label classification problem to label instances of a musical dataset. The Tellegen-Watson-Clark model [20] is employed to classify the songs into emotions in Figure 3 for all the songs [16].

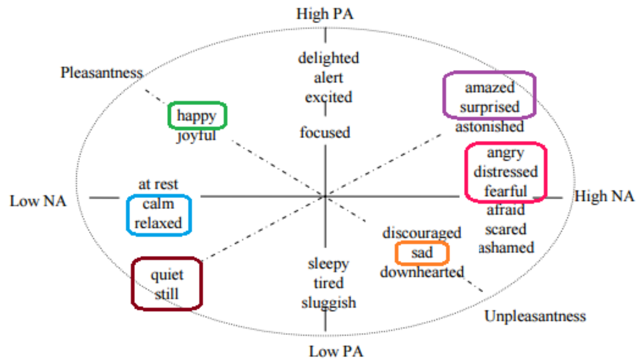

Fig. 1. The Tellegen-Watson-Clark model of mood (figure reproduced from [20])

Initially, the dataset was selected. Among the different options that we have in the Net, we finally select "Emotions from Mulan" because of the large number of instance and its richness in labels and emotions classes. Before applying any algorithm, we preprocess the data selected. To do so, those attributes that does not provide useful information for the classification task were not considered. A statistical analysis was performed to check the possible outliers, missing values and the distribution. With the data retrieved and preprocessed, different algorithms were applied to classify the data according to the classes. Two algorithms are used for the multi-label classification: Random k-Labelsets and Multi-Label k-Nearest Neighbors. Likewise, as a part of the analysis, some association rules are extracted by using the Apriori algorithm.

We will detail the data preprocessing in Sections 3.2 and 3.1. Section 3.3 will expose the application of the classification algorithms, while Section 3.4 will explain the Apriori performance to extract the association rules.

3.1 Emotions Dataset

The dataset used for this work is called Emotions from Mulan [18]. It consists on a collection of 593 songs which belong to 7 different genres: Classical, Reggae, Rock, Pop, Hip-Hop, Techno and Jazz. A period of 30 seconds after the initial 30 seconds was considered in each instance in order to extract 72 numerical type musical features. Additionally, 6 classes of emotions are considered as the sentiments that can be felt by the listener.

The musical features can be divided into rhythmic and timbre attributes. For the purposes of this work, only timbre attributes will be considered, based on a musical decision, as we considered that the sound bumps in a song (what is known as rhythm) does not have a strong influence on the emotions that it causes. An example of this could be the rhythm difference between some musical pieces such as Beethoven's Ode to Joy and Radetzky March, even though they both evoke joy and majesty. The full description of rhythmic and timbre features can be found at [16]. Regarding with timbric features, 13 coefficients of the Fast Fourier Transform are considered. They are known as MFCCs (Mel Frequency Cepstral Coefficients) in the dataset. Furthermore, 3 attributes related to timbre texture such as the spectral centroid, the spectral rolloff and the spectral flux are contemplated. For each of the 16 timbre features the mean, standard deviation (std), mean standard deviation (mean std), and standard deviation of standard deviation (std std) are calculated over all frames. In total, 64 timbre numerical type features are used in the dataset.

3.2 Data preprocessing

As a first step, a statistical analysis of the attribute values for the instances is needed. This analysis was performed with WEKA tool, and visualized in this framework. According to the results, we can conclude, there is no outliers or missing values in the dataset, and, in general, attributes have a normal distribution; thus, no filter is needed to prepare the attributes for the classification task.

However, before applying any unsupervised learning algorithm, the number of attributes for each of them should be reduced. To do so, a filter of correlation and another filter of information gain are applied on each considered class. The results will be compared to select those features that have a total higher value.

As a consequence, we reduce the list of attributes from 64 to 12 features that will be considered for Apriori algorithm. In particular, the attributes used were: Mean of spectral centroid, Mean of spectral rolloff, Mean of MFCC_0, Mean of MFCC_1, Mean of MFCC_4, Mean standard deviation of Spectral rolloff, Mean standard deviation of MFCC_11, Standard deviation of MFCC_10, Standard deviation of MFCC_11, Standard deviation of MFCC_12, Standard deviation of standard deviation of MFCC_6, Standard deviation of standard deviation of MFCC_7.

Furthermore, specifically for the rules extraction process, a discretization of the attribute values is needed, as it would not make sense looking for rules for specific numerical values of attributes in this concrete problem. Note that the value discretization could be negatively influencing the final results. Additionally, it is important not to create too large intervals to better represent reality. Therefore, we discretize the value to obtain 10 different groups.

3.3 Supervised algorithms

Once the data are deeply analyzed, we can carry out the classification procedure. Our goal is to classify the songs according to the emotions that they produces while listening. Several emotions can be felt by an individual song; therefore, this problem will be resolved as a multi-labeling classification task. Two algorithms are applied in order to compare the results.

– **Random k-Labelsets (RAkEL)**: This algorithm is based on random projections of the label space [19]. The main idea is to create an ensemble, which

will be divided in m base classifiers that are trained with a k random subset of the choosen labels. It is widely used in multi-label problems.
- **Multi-Label k-Nearest Neighbors (MLkNN)**: This algorithm is an adaptation of the k-Nearest Neighbours (kNN) to the multi-label task [23]. It is a classification algorithm that, in order to determine the classification of an instance, combines the classification of the k nearest instances.

In both cases, a cross-validation technique is used. The split percentage is 65%, which means that the analysis is performed on one subset called the training set (65%) and validated on the other subset called the validation set (35%). The parameters and the results are described in Section 4.

3.4 Unsupervised algorithms

In data mining, association rules are a type of association method in which it is sought to find out if there are facts that occur always collectively (or very repeatedly) within a set of data. The algorithm Apriori is selected to observe and to corroborate the relation and influence of the attributes in the classes (emotions). As a consequence, the multi-label feature of the dataset will not be taken into account in this task, considering the emotions as a new feature in the dataset. The final results are detailed in Section 4.

4 Results and discussion

The results can be divided according to the task that we performed. Firstly, we aim to classify the data in six different classes that represent emotions. As we explained above, we apply two different algorithms to make a comparison about the performance and final results.

On the one hand, in the RAkEL algorithm, the number of subsets in which the ensemble or set is divided in m groups and the k number of labels of each subset of the ensemble are variable. The values for this parameters are $m = 12$ and $k = 4$. The base classifier used for RAkEL is SMO (Sequential Minimal Optimization) because of its good performance in this particular case.

On the other hand, the Multi-Label k-Nearest Neighbors implementation that we selected is the Mulan Filter, which is available in MEKA tool. The base classifier will be again SMO so that the comparison between both classifiers will not be biased by the base algorithm.

The final results in both cases are shown in Table 1. The first column corresponds to quality measures obtained by RAkEL algorithm. The second column represents the values for MLkNN. The quality measures considered are the hamming score, hamming loss, one error and ranking loss. The hamming score and hamming loss captures the fraction of labels that are correctly and incorrectly predicted, respectively. One error evaluates the frequency of the top-ranked label that was not in the set of true labels, and ranking loss measures the average fraction of pairs of labels that are ordered incorrectly. The last row represents the total execution time for each algorithm.

	RAkEL	MLkNN
Hamming score	0.788	0.804
Hamming loss	0.212	0.196
One error	0.285	0.251
Ranking loss	0.193	0.154
Running time	2.385	0.138

Table 1. Comparison between RAkEL and MLkNN performance

As we can see in Table 1, both algorithms produce similar results. Hamming score and hamming loss are practically the same in both cases. The values are near 0.8, thus the accuracy is about 80%. The optima values for one-error and ranking loss is 0. In both algorithms, one error measure is less than 0.3, which means that the top-ranked predicted label is misclassified in less than 30% of cases. Ranking loss values mean that about 20% of label pairs are wrongly ordered for an instance. Finally, the most remarkable aspect is the running time; MLkNN takes close to 0.14 seconds, and RAkEL, about 2.4; thus, MLkNN is quite quicker than RAkEL.

Additionally, we performed a rules extraction process. In the final results, the minimum confidence accepted for each rule is above 90%, and their support should be greater than 10%. First column of Table 2 shows the most representative rules obtained and the second column reflects confidence values for each one. The third and fourth columns present lift and conviction values respectively.

RULE	CONF.	LIFT	CONV.
std(MFCC_12)$<$ 0.21&¬Relaxing-calm&¬Sad-lonely⇒ ¬Quiet-still	0.99	1.33	22.71
std(MFCC_12)$<$ 0.21&¬Relaxing-calm⇒ ¬Quiet-still	0.99	1.31	12.54
¬Relaxing-calm&¬Sad-lonely⇒ ¬Quiet-still	0.98	1.3	9.13
¬Amazed-surprised&Relaxing-calm⇒ ¬Angry-aggresive	0.97	1.43	10
std(MFCC_12)$<$ 0.21&¬Sad-lonely⇒ ¬Quiet-still	0.97	1.3	7.63
Angry-aggresive⇒ ¬Relaxing-calm	0.96	1.74	10.52

Table 2. Rules obtained with Apriori algorithm

As Table 2 shows, the confidence is between 96% and 99% in all the rules obtained, and lift is always greater than 1, so the rules happen with a greater probability than by chance. Finally, conviction values are not too high, thus the rules usually are unknown beforehand.

In rules 1, 2 and 5, the musical feature that appears is the same in all cases: Standard deviation of MFCC_12. It shows that this feature is very relevant and influential in the listeners feelings. The rest of rules show relationship between the emotions, which are considered as a new feature of the dataset. The obtained rules have a very strong logic base, e.g., rule 3: "if the song is not relaxing or sad, it will not be quiet" or rule 6: "if the song is aggresive, it will not be calm". This kind of associations are obtained due to the existence of opposing emotions in the data set (happy-sad, relaxing-aggresive...).

5 Conclusion

Data mining can be applied to very different problems, such as sentiment analysis in music, providing useful and interesting information that can be used as the starting point for other works. This paper has successfully solved a classification problem of musical songs and emotions. Firstly, a dataset was selected, preprocessed and analyzed using different algorithms of multi-label classification. Additionally, an analysis about unsupervised associations has been performed.

The statistical analysis and filter of the dataset make the classification task easier. Furthermore, data preprocessing is necessary to obtain consistent and optimized results. Timbre features are considered in this work, leaving apart the attributes related to rhythm.

From the application of the algorithms RAkEL and MLkNN very similar results are obtained according to different quality measures. Both of them classifies the songs succesfully according to six emotions. We finally selected the MLkNN as the best algorithm based on its running time.

A unsupervised association algorithm was applied for obtaining different relationships between emotions and musical features. This trial provides an indicator of which attributes are more influential in the emotions the song provokes.

Directions for future research include the empirical analysis of the impact of rhythmic attributes in the feelings produced by music.

Acknowledgments

This work was supported by the Spanish Ministry, Ministerio de Economía y Competitividad and FEDER funds. SURF: Intelligent System for integrated and sustainable management of urban fleets TIN2015-65515-C4-3-R, and through the FPU program FPU2013 of the Ministry of Education and Culture, number 2071.

References

1. Anbarjafari, G.: Convolutional neural network super resolution for face recognition in surveillance monitoring. In: Articulated Motion and Deformable Objects: 9th International Conference, AMDO 2016, Palma de Mallorca, Spain, July 13-15, 2016, Proceedings. vol. 9756, p. 175. Springer (2016)
2. Deutsch, D.: Psychology of music. Elsevier (2013)
3. Hu, X., Downie, J.S.: Improving mood classification in music digital libraries by combining lyrics and audio. In: Proceedings of the 10th annual joint conference on Digital libraries. pp. 159–168. ACM (2010)

4. Hu, X., Downie, J.S., Ehmann, A.F.: Lyric text mining in music mood classification. American music 183(5,049), 2–209 (2009)
5. Juslin, P.N., Sloboda, J.A.: Music and emotion: Theory and research. Oxford University Press (2001)
6. Kim, Y.E., Schmidt, E.M., Migneco, R., Morton, B.G., Richardson, P., Scott, J., Speck, J.A., Turnbull, D.: Music emotion recognition: A state of the art review. In: Proc. ISMIR. pp. 255–266. Citeseer (2010)
7. Kim, Y., Moon, T.: Human detection and activity classification based on micro-doppler signatures using deep convolutional neural networks. IEEE Geoscience and Remote Sensing Letters 13(1), 8–12 (2016)
8. Koelsch, S.: Music-evoked emotions: principles, brain correlates, and implications for therapy. Annals of the New York Academy of Sciences 1337(1), 193–201 (2015)
9. Louis, D.N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W.K., Ohgaki, H., Wiestler, O.D., Kleihues, P., Ellison, D.W.: The 2016 world health organization classification of tumors of the central nervous system: a summary. Acta neuropathologica 131(6), 803–820 (2016)
10. McKinney, M., Breebaart, J.: Features for audio and music classification (2003)
11. Mullen, T., Collier, N.: Sentiment analysis using support vector machines with diverse information sources. In: EMNLP. vol. 4, pp. 412–418 (2004)
12. Oramas, S., Espinosa-Anke, L., Lawlor, A., et al.: Exploring customer reviews for music genre classification and evolutionary studies. In: The 17th International Society for Music Information Retrieval Conference (ISMIR 2016), New York City, United States of America, 7-11 August 2016 (2016)
13. Pang, B., Lee, L.: A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42nd annual meeting on Association for Computational Linguistics. p. 271. Association for Computational Linguistics (2004)
14. Pang, B., Lee, L., et al.: Opinion mining and sentiment analysis. Foundations and Trends® in Information Retrieval 2(1–2), 1–135 (2008)
15. Taruffi, L., Allen, R., Downing, J., Heaton, P.: Individual differences in music-perceived emotions. Music Perception: An Interdisciplinary Journal 34(3), 253–266 (2017)
16. Trohidis, K., Tsoumakas, G., Kalliris, G., Vlahavas, I.P.: Multi-label classification of music into emotions. In: ISMIR. vol. 8, pp. 325–330 (2008)
17. Tsoumakas, G., Katakis, I.: Multi-label classification: An overview. International Journal of Data Warehousing and Mining 3(3) (2006)
18. Tsoumakas, G., Spyromitros-Xioufis, E., Vilcek, J., Vlahavas, I.: Mulan: A java library for multi-label learning. Journal of Machine Learning Research 12(Jul), 2411–2414 (2011)
19. Tsoumakas, G., Vlahavas, I.: Random k-labelsets: An ensemble method for multilabel classification. In: European Conference on Machine Learning. pp. 406–417. Springer (2007)
20. Yang, D., Lee, W.S.: Disambiguating music emotion using software agents. In: ISMIR. vol. 4, pp. 218–223 (2004)
21. Yang, Y.H., Liu, C.C., Chen, H.H.: Music emotion classification: A fuzzy approach. In: Proceedings of the 14th ACM international conference on Multimedia. pp. 81–84. ACM (2006)
22. Zentner, M., Grandjean, D., Scherer, K.R.: Emotions evoked by the sound of music: characterization, classification, and measurement. Emotion 8(4), 494 (2008)
23. Zhang, M.L., Zhou, Z.H.: Ml-knn: A lazy learning approach to multi-label learning. Pattern recognition 40(7), 2038–2048 (2007)