2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW) | 978-1-6654-3328-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICCE-TW52618.2021.9603205

# Music Mood Classification System for Streaming Platform Analysis via Deep Learning Based Feature Extraction

Yu-Chia Chen[1], Zih-Ching Chen[2], and Chih-Hsien Hsia[3,*], *Member, IEEE*
[1]Department of Economics
National Taiwan University, Taipei, Taiwan
[2]Graduate Institute of Communication Engineering
National Taiwan University, Taipei, Taiwan
[3]Department of Computer Science and Information Engineering
National Ilan University, Yilan, Taiwan
Email: chhsia625@gmail.com

*Abstract*— **The proliferation of smart phones and the Internet has led to the growth of streaming music services. Streaming music services that provide vast libraries of songs have changed the way music lovers listen and leave a digital footprint when they are listening to the music. Music can arouse the emotional feelings of the listeners, thus providing emotional rewards for the listeners and achieving physiological adjustment effect [1]. In general, the study of music mood is based on the melody, rhythm, timbre and other feature of music. However, the lyrics of a song also provides information that helps the listener to have a higher level of understanding of the emotion expressed by the music. In this work, we focus on Spotify, which is the world's biggest music streaming platform. Using the information retrieved from web crawler, we obtain the data which contains the audio features of the song from Spotify playlist, while analyzing the audio feature and the lyrics through semantic analysis.**

## I. Introduction

Music emotion is directly related to the content of the music. In an era of the Internet, the music industry has gone online and streaming music services have become mainstream. Streaming Music platforms provide a huge library of songs in the cloud, users leave a digital footprint as they listen to the music using the service. These behavior and interactions in the digital space can be utilized by the platform providers to understand the user's likes and dislikes, characteristics, and habits to explore business opportunities.

There are many studies exploring the relevance of music preferences and personality between [2-4], which show the preference of peoples who are good at communication tend to be vibrant and lively music. However, different music lovers may have distinct definitions of genres of music. For this reason, [5] suggest that people's musical preferences can indeed be used to predict their personality.

The analysis of song mood includes two aspects: audio and lyric. Audio features such as Mel spectrum coefficient can be used to analyze music mood [6]. The most used lyric feature is TF-IDF [7]. For deep learning (DL) approach, Liu et al. achieved good results by using convolutional neural network (CNN) to process music sonograms in 2017 [8]. However, lyrics can also be deeply related to the mood of the music. Our study is to focus on predicting the mood of a song through its both audio and lyrics features provided by the streaming platform. By integrating these two features, we can analysis the mood of a song with the crawled information online.

## II. Algorithms

In this work, we proposed a deep neural network architecture which not only considers the audio features of a song provided by Spotify API, but also extracted the lyrics of the song crawled from Genius. Firstly, we extract the lyrics feature by BERT encoder. After adopting the pre-trained BERT [9] for emotion classification, we concatenate both audio features and lyrics for music mood classification.

### A. Spotify API Data

As the largest music streaming service in the world, Spotify provides API for users to access. Using Spotify Web API, we can retrieve JSON relays of artists, albums, and tracks from Spotify's playlists with the help of Web crawlers, as shown in Fig. 1. For this study, we collected 1200 songs that are mostly mainstream popular music from 12 playlists on Spotify.

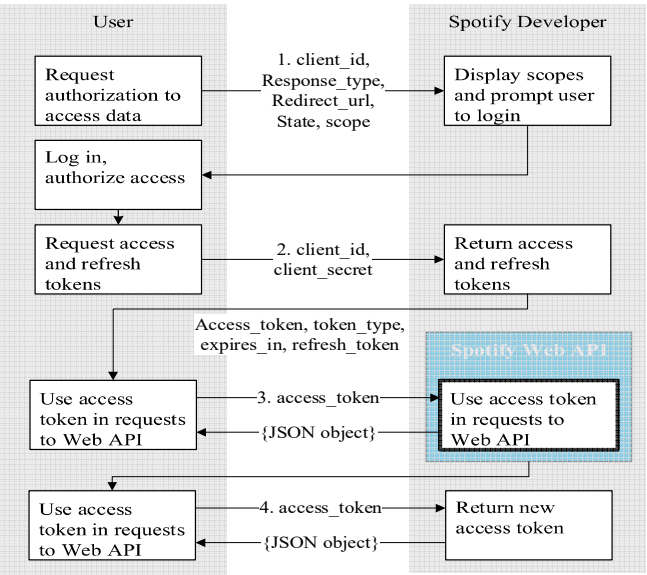


Fig. 1. The workflow of Spotify API

### B. Deep Learning Based Music Mood Classification

Both audio and lyrics are helpful in the task of estimating the mood of a song. In this study, we combine these two features to create a music mood classification system that analysis the emotion of the song from its audio and lyrics information. The model architecture is shown in Fig. 2.

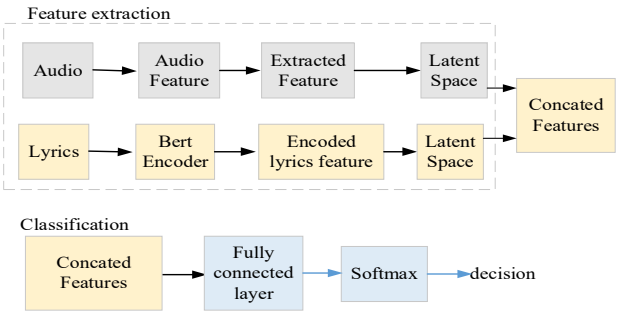


Fig. 2. The architecture of our DL-based music mood classification system

To analysis the emotions from the lyrics, we utilize BERT, a pre-trained language presentation model for natural language processing task, to transfer knowledge to our model. We first encode each of the song lyrics into vectors that represent the information of emotions. We then integrate the extracted lyrics features with the audio features and predict the result with a fully connected classifier. The architecture of Bert for sequence classification is shown in Fig. 3.

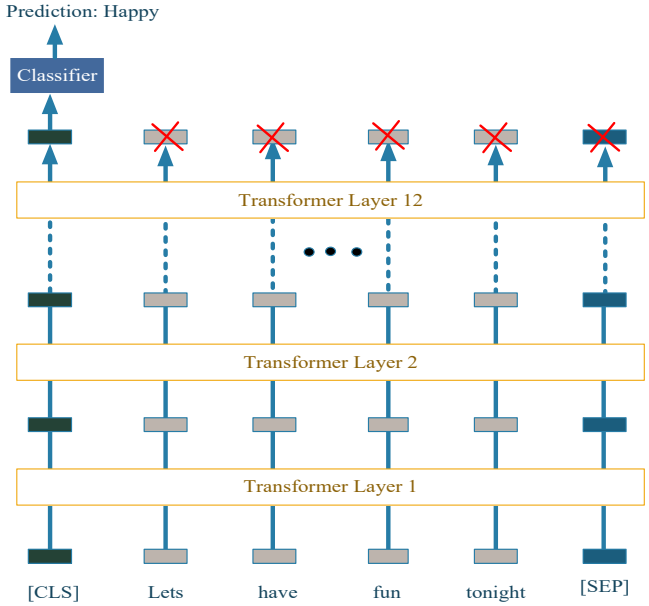


Fig. 3. The architecture of Bert for sequence classification.

## III. Experiment and Result

In this study, we use DL development framework PyTorch for the training and testing stage of our model. The dataset is split into three parts, there are 874 songs in the training set, 99 songs in the validation set, and 193 songs for testing. For the hyperparameters, we set optimizer Adam's learning rate to 0.01, Batch Size to 32, and train for 100 epoch. Our model achieves 0.81 AUC score, as shown in Fig. 4.

TABLE I
Accuracy of Machine Learning Approaches

| Classifier | Accuracy | AUC score |
|---|---|---|
| Logistic regression | 0.7395 | 0.6217 |
| XGB classifier | 0.8235 | 0.7715 |
| Decision tree classifier | 0.7227 | 0.6666 |
| Random forest classifier | 0.8403 | 0.7959 |
| Bert for sequence classification | 0.6182 | 0.5874 |
| Proposed methods | 0.8452 | 0.8125 |

AUC = Area Under Curve

Comparing with the results using audio feature with algorithms including logistic regression, XGB classifier, Decision tree classifier, and random forest classifier, and lyrics feature, which is BERT with a fully conntected layer, our method achieved 84.52 accuracy and 0.8125 AUC score. The result of the above approaches are presented in Tabel I.

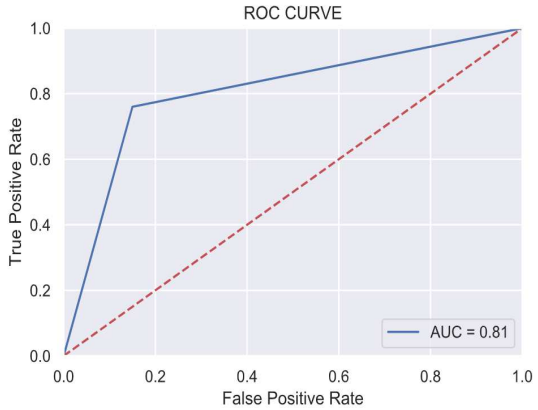


Fig. 4. The ROC curve of the proposed method.

## IV. Conclusion

Music is considered as the best form of expression of emotions. The music that people listen to is governed by what mood they are in. In this study, we proposed a DL model architecture which extracts the features from both audio and lyrics from a song for mood classification based on the data obtained from streaming platform. Through the audio features provided by Spotify API and the lyrics crawled from Genius, we achieve 0.81 AUC score on 193 testing data. With the help of music classification system, the music information retrieval for the streaming platform can be achieved, the platform can also understand the users' current emotional state, thus helping to have a better understanding of the users' needs.

## V. Reference

[1] M. Zentner, D. Grandjean, and K. R. Scherer, "Emotions evoked by the sound of music: Characterization, classification, and measurement," *Emotion*, Vol. 8, No. 4, pp. 494–521, 2008.

[2] D. M. Greenberg, S. Baron-Cohen, D. J. Stillwell, M. Kosinski, and P. J. Rentfrow. "Musical preferences are linked to cognitive styles." *PLOS One*, Vol. 10, No. 7, 2015.

[3] D. M. Greenberg, M. Kosinski, D. J. Stillwell, B. L. Monteiro, D. J. Levitin, and P. J Rentfrow, "The song is you: Preferences for musical attribute dimensions reflect personality." *Social Psychological & Personality Science,* Vol. 7, pp. 597–605, 2016.

[4] T. Schäfer and C. Mehlhorn, "Can personality traits predict musical style preferences? A meta-analysis." *Personality and Individual Differences*, Vol. 116, pp. 265–273, 2017.

[5] G. Nave, J. Minxha, D.M. Greenberg, M. Kosinski, D. Stillwell, and J. Rentfrow, "Musical Preferences Predict Personality: Evidence from active listening and Facebook likes," *Psychological Science*, Vol. 29, No. 7, pp. 1145-1158, 2018.

[6] S. Hizlisoy, S. Yildirim, and Z. Tufekci, "Music emotion recognition using convolutional long short term memory deep neural networks," *Engineering Science and Technology, an International Journal*, Vol. 24, pp. 760-767, 2021.

[7] M. Zaanen and P. Kanters, "Automatic Mood Classification Using TF-IDF Based on Lyrics," *11th International Society for Music Information and Retrieval Conference*, August 2010.

[8] X. Liu, Q. Chen, X. Wu, Y. Liu, and Y. Liu, "CNN based music emotion classification," *arXiv*:1704.05665v1, 2017.

[9] J. Devlin, M.W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv*: 1810.04805v2, 2019.