2021 International Electrical Engineering Congress (iEECON2021)
March 10-12, 2021, Pattaya, THAILAND

2021 9th International Electrical Engineering Congress (iEECON) | 978-1-7281-9584-1/20/$31.00 ©2021 IEEE | DOI: 10.1109/iEECON51072.2021.9440333

Mood classification from Song Lyric using Machine Learning

Kornkanya Siriket
Department of Computer Science
Faculty of Science
Srinakharinwirot University
Bangkok, Thailand
e-mail: kornkanya.sir@g.swu.ac.th

Vera Sa-ing
Department of Computer Science
Faculty of Science
Srinakharinwirot University
Bangkok, Thailandl
e-mail: vera@g.swu.ac.th

Subhorn Khonthapagdee
Department of Computer Science
Faculty of Science
Srinakharinwirot University
Bangkok, Thailandl
e-mail: subhorn@g.swu.ac.th

***Abstract*— Nowadays, many people change the way they listen to music by listening to the mood of the songs in the tracks. This research is interested in analyzing song extraction using natural language processing to acquire mood information. Lyrics are valuable for categorizing music. First, removing special characters and using Term-frequency/inverse-document frequency (TFIDF) and then Latent Dirichlet Allocation (LDA) are used to connect words to mood classes. We perform a lyric- based mood classification on local machine learning classifiers such as Random forest, Decision tree, Naïve Bayes, Logistic Regression, AdaBoost and XGBoost. Using grid search for tuning the best parameter yield the results XGBoost shows the highest accuracy. It can prove that boosting algorithms have better performance than local machine learning in this research.**

Keywords*— Mood, Song Lyric, Machine Learning. Boosting algorithm, natural language processing.

I. Introduction

Music has long been involved in human daily life. Nowadays, music streaming applications such as Spotify, TIDAL, Apple Music etc. have become so popular. An important feature of such applications is their ability to sort large collections of digital music by genre, artists, or albums. However, sorting music based on genre, artists, or albums may not satisfy diverse listeners. Music serves both social and psychological functions. Style, mood, and similarity information should be considered when indexing music according to [1]. Artists often express the meaning and the mood of music through lyrics. Therefore, words or repeated words that are contained in lyrics are the best way to identify the mood of the music.

In this study, we use natural language processing techniques and machine learning algorithms to analyze lyrics and classify the mood of the music. The automatic mood classification would be helpful for sorting large collections of digital music for music streaming applications that we mentioned above.

II. Related Work

A lyric-based mood classification has been studied widely in recent years. An et al. [2] presented a method for classifying Chinese music according to the mood of the song by compiling Chinese songs from Baidu website, and emotion tags from anonymous listeners. The study used Russel's emotional model [3] and divided the mood of the song into 3 categories: depression, contentment, and exuberance. Naïve Bayes algorithm was used to classify text data and yield 68% accuracy. In this study, we also use Russel's emotional model, but we classify mood of music into 4 categories: Happy, Sad, Angry and Relaxed.

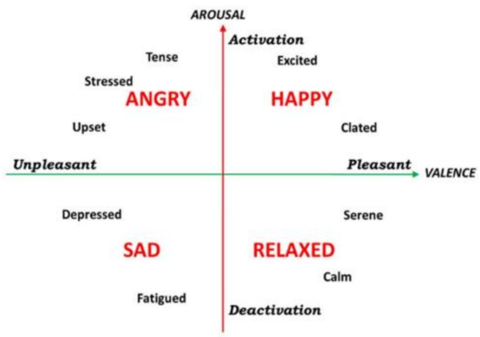

Fig. 1. Russel's emotion model. Reprinted from [8]

Akella and Moh [4] proposed mood classification models for English songs collected from Million Song website and combined with emotional tags from the Lastfm website. Various techniques for transforming from text to numerical vectors were used in this study. Additionally, the study compared the classification results from traditional machine learning techniques, such as Support vector machine (SVM), Random forest, Linear regression, and Naive Bayes, with deep learning techniques, such as Convolutional Neural Network (CNN). It was shown that CNN yielded the highest accuracy classification of 71%. In this study, instead of using deep learning, we increased the accuracy of classification results by using model ensemble techniques such as boosting methods. Recently, boosting algorithms and its variants have gained popularity and have been shown to yield good results on many classification tasks [5-7].

The remainder of the paper is organized as follows. In the methodology section, we discuss the mood model that we use to represent mood from each song. Data gathering, text preprocessing and feature extraction are also discussed in this section. Next, the details of classifiers such as machine learning algorithm and boosting algorithm are described in the model classification section. Finally, we present the model evaluation in the results section and follow by the conclusion.

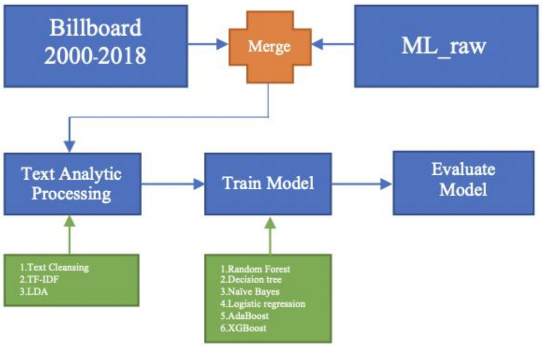

Fig. 2. The overall process of the proposed research

III. Methodology

A. Mood model

Russel's emotion model is 2 -dimensional valence-arousal model as shown in Fig. 1. It is a most widely used emotion model. Valence which can be positive or negative, is a measure of the intensity of an emotion. In contrast, Arousal is a measure of how strongly or rhythmically an emotion is felt.

B. Data gathering

We use Billboard dataset [8] which consist of Lyrics of English songs from Spotify between 2000-2018. We then combine it with mood labels from MoodyLyrics corpus [9] which is a dataset of lyrics classified into 4 categories corresponding to Russel's model. The workflow of this study can be seen in Fig. 2. Note that ML_raw in the figure is referred to MoodyLyrics corpus. The Mood categories in this study are happy, relaxed, sad and angry. There are 265 happy songs, 113 angry songs, 75 sad songs, and 183 relaxed songs after matching mood and song title. Fig. 3 shows the distribution of each mood between 2000 to 2017.

C. Text preprocessing

Lyric contains a much smaller set of words than news or books. However, these words contain emotion that can be used to classify mood of the song. We started by removing special characters for example "\n". It is often that we see the use of common contractions in lyrics such as "gonna", "wanna". We then remove all those contractions and stop words such as "a", "an", "the".

D. Feature selection

Feature selection is a crucial part in the classification process. In this study, lyrics are converted into feature space using TF-IDF technique. TF-IDF or Term Frequency - Inverse Document Frequency can be viewed as a measure of importance of a term in each document in a corpus. The benefit of using TF-IDF technique is that it gives a larger weight to terms which are less common in the corpus. The importance of very frequent terms will then be lowered, which helps to balance off those words which appear more frequently in general. To reduce the dimension of the feature set, we employ Latent Dirichlet Allocation (LDA) technique. LDA is an unsupervised model that does not take the categorization of the documents into account when extracting latent topics. The number of topics (K) is a hyperparameter that needs to be determined.

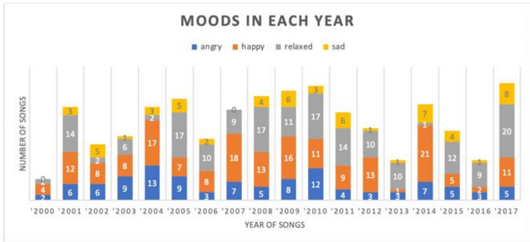

Fig. 3. The distribution of each Mood between 2000 to 2017

IV. Model Classification

The problem that we focus on can be viewed as multi- class classification. Multi- class classification describes the task of classifying data into classes, whereby there are a lot of classes and each data point can be assigned to only one class

A. Traditional machine learning

Well known algorithms such as Decision tree, Random forest, Linear regression, and Naive Bayes have been widely used for text classification. In this study, we compared the performance of the above algorithms. A grid search was performed to find the optimal hyperparameters.

To increase the accuracy of model classification, we use ensemble methods. Model combinations are usually identified as model ensembles. They are the most effective approaches in machine learning, generally achieving better performance than the single models [10]. Two types of ensemble methods are bagging and boosting. Random forest is an example of a bagging ensemble model.

B. Adaptive boosting algorithm

The idea behind adaptive boosting algorithms or AdaBoost can be found in many literature [11,12]. We also briefly explained here. In a boosting algorithm, weak learning algorithms or rules are combined into a single strong rule. This algorithm fits weak rules with different subsets of training data repeatedly. Each time, the distribution of weights is applied to each weak rule by focusing on the most misclassified by the preceding one. Finally, these weak rules are simply combined by using majority votes of their predictions. AdaBoost is fast and simple. In this study, we use AdaBoost that's implemented in Scikit-learn [13] with the decision tree as a base learner.

C. Gradient boosting machine

Instead of increasing the weight of misclassified prediction, in Gradient tree boosting or Gradient boosting machine (GBM), the gradient is calculated from the error that happened in each iteration. Therefore, the goal is to minimize the loss function from the previous learner. Recently, gradient tree boosting has gained popularity and has been shown to

Word of sad	Number of words
im	400
lonely	217
like	177
don't	160
got	139
girl	135
na	127
baby	114
know	106
hold	104

Word of relaxed	Number of words
home	1060
im	860
got	447
youre	411
take	392
know	381
go	361
baby	352
like	346
dont	306

Word of Happy	Number of words
love	1767
im	931
know	593
dont	489
like	390
got	370
na	367
oh	356
life	337
yeah	317

Word of angry	Number of words
burn	563
im	461
na	443
got	400
dont	348
know	342
fire	330
let	302
like	293
oh	257

Fig. 4. The top 10 of the most common words in each mood

yield good results on many classification tasks. In 2016, Chen and Guestrin [14] introduced a scalable tree boosting system called Extra gradient boosting or XGBoost, which proved to be highly efficient. Grids search best parameters for XGBoost which automatically use learning rate = 0.1, estimator = 80 and max depth = 8.

V. RESULT

After text preprocessing, we found that there are on average 70 words in each song. Fig. 4 shows the most common words in each mood. In Relaxed mood and Happy mood, the words "home" and "love" appear the most, respectively.

Words that express sadness such as "lonely" or "hold" are frequently found in sad mood. Lastly, the word "burn", or "fire" are common words found in Angry mood. Next, we split 80% of the data into training set and 20% into test set. We all trained algorithms that we mentioned above and evaluate by using the calculation of percentage accuracy formula by divided the correctly classified music by all music. We found that XGBoost has outperformed all other models with 89% accuracy. The result was shown in table 1.

TABLE I. PERFORMANCE OF EACH MODEL

Classifier	Accuracy
XGBoost	89.19%
Random forest	81.01%
AdaBoost	63.68%
Decision tree	60.13%
Naïve Bayes	58.02%
Logistic regression	48.41%

VI. CONCLUSION

Expectation from this paper for general people who are interested in text analytic or loving music can study word trends in 2002 to 2007 for using text lyrics in the future. In this study, we combine the Bilboard dataset, which contains lyrics from Spotify between 2000-2018, with the mood label from the MoodyLyrics corpus. We perform text preprocessing and the extract features from those lyric data. Finally, the multi- class classification for lyric corpus is used to train and evaluate various classifiers. The experiment shows the XGBoost gives the best result with 89% accuracy.

ACKNOWLEDGMENT

We would like to thank the instructors from the department of computer science at Srinakharinwirot University, Thailand. In addition, this research was funded by the Graduate School of Srinakharinwirot University to support the publication.

REFERENCES

[1] D. Huron, "Perceptual and cognitive applications in music information retrieval," ISMIR, Plymouth, Massachusetts, USA, October 2000.

[2] Y. An, S. Sun, and S. Wang, "Naive Bayes classifiers for music emotion classification based on lyrics," in Computer and Information Science (ICIS), 2017 IEEE/ACIS 16th International Conference on. IEEE, 2017, pp. 635–638.

[3] J. A. Russell, "A circumplex model of affect," Journal of Personality and Social Psychology, vol. 39, 1980, pp. 1161–1178.

[4] R. Akella, and T. Moh, "Mood Classification with Lyrics and ConvNets,". In Proceedings of 18th IEEE International Conference On Machine Learning And Applications (ICMLA), 2019, pp. 511-514.

[5] B. Al-Salemi, S.A.M. Noah and M.J. Ab Aziz, "RFBoost: An improved multi-label boosting algorithm and its application to text categorisation," Knowledge-Based Systems, vol. 103, 2016, pp. 104-117.

[6] N. S. Elmitwally, "Building A Multi-class XGBoost Model for Arabic Figurative Language," 2nd International Conference on Computer and Information Sciences (ICCIS), Sakaka, Saudi Arabia, 2020, pp. 1-4.

[7] J. Salminen, M. Hopf, S.A. Chowdhury, S.-G. Jung, H. Almerekhi, B.J. and Jansen, " Developing an online hate classifier for multiple social media platforms," Human-centric Computing and Information Sciences, vol 10, 2020, pp. 1-34.

[8] M. Tauberg, Billboard Hot-100 Songs 2000-2018 Spotify Data + Lyrics. Retrieved from https://data.world/typhon /billboard-hot-100-songs-2000-2018-w-spotify-data-lyrics

[9] E. Çano and M. Morisio, "MoodyLyrics: A Sentiment Annotated Lyrics Dataset," In Proceedings of the 2017 International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence (ISMSI '17), 2017, pp. 118–124.

[10] A. Subasi, "Practical Machine Learning for Data Analysis Using Python," Elsevier Science & Technology, 2020.

[11] R. Schapire, "The boosting approach to machine learning – an overview," MSRI Workshop on Nonlinear Estimation and Classification, 2002.

[12] Y. Freund, and R. Schapire, "A Short Introduction to Boosting," Journal of Japanese Society for Artificial Intelligence, Vol 4, 1999, pp. 771-780.

[13] Pedregosa et al, Scikit-learn: Machine Learning in Python, JMLR 12, 2011, pp. 2825-2830.

[14] T. Chen, and C. Guestrin, " XGBoost: A Scalable Tree Boosting System," ArXiv e-prints. 2016.

[15] Laoh, Enrico, Isti Surjandari, and Limisgy Ramadhina Febirautami. "Indonesians' Song Lyrics Topic Modelling Using Latent Dirichlet Allocation." Paper presented at the 2018 5th International Conference on Information Science and Control Engineering (ICISCE), 2018, pp. 270-274.