BanglaMusicMooD: A Music Mood Classifier from Bangla Music Lyrics

Nazia Parvin Urmi, Nasir Uddin Ahmed, Md. Habibur Rahman Sifat, Salekul Islam, Abu Shafin Mohammad Mahdee Jameel

Abstract This research work aimed at using word-level features for mood classification. The goal of this project is to build a classifier model to detect the mood of the song by analyzing the lyrics of Bangla Song. We have used the most famous approach "Bagging" (standing for "bootstrap aggregating") that targets at constructing an ensemble model that is more robust than the individual models composing it. We used this learning to predict the probability of different classes based on various attributes. Our work proposed Support Vector Machine (SVM) ensembles with bagging that outperforms better performance in terms of classification accuracy greatly. We have worked on two types of Bangla moods such as happy and sad, corresponding to the Bangla meaning. At the end of the experiment, our proposed method performed well to classify the mood.

Keywords: Ensemble Learning (EN), Support Vector Machine (SVM), Term Frequency-Inverse Document Frequency (TF-IDF), Bootstrap Aggregating (Bag-

Nazia Parvin Urmi
United International University, Dhaka, Bangladesh, e-mail: nurmi161201@bscse.uiu.ac.bd

Nasir Uddin Ahmed
United International University, Dhaka, Bangladesh, e-mail: nahmed151086@bscse.uiu.ac.bd

Md.Habibur Rahman Sifat
United International University, Dhaka, Bangladesh, e-mail: habibsifatuiu@gmail.com

Salekul Islam
United International University, Dhaka, Bangladesh, e-mail: salekul@cse.uiu.ac.bd

Abu Shafin Mohammad Mahdee Jameel
United International University, Dhaka, Bangladesh, e-mail: mahdee@cse.uiu.ac.bd

ging).

1 Introduction

In our leisure and recreation time, we usually do some activities such as reading books, listening to music, playing. Among these activities, listening to music is one of the popular ones that have an amplifying effect on social consistency, the emotional state of mind and mood of the listeners [1]. Among music lovers, music classification and recommendation model has obtained large popularity.

The power of music to influence mood, create scenes, routines and occasions are widely recognized and this is reflected in a strand of social theory that portrays music as an influence on character, social structure, and action [2]. ***Music*** is an important art that is very soothing and relaxing which can make our bad days better. The more we will be able to examine what we want in music the more music will be appealing to us. Around 68 percent of people who have aged between 18 and 34 years old announce that they listen and enjoy music every day.

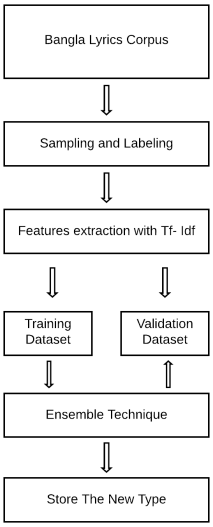

Fig. 1 Flowchart of the Proposed Method

In our approach, we created our own Bangla Song dataset for filtering and music mood prediction of listeners that can be associated with happy or sad emotions. Our main focus is to create a system that can be able to predict song mood that will be applied to the song database to select music by its lyrics analysis.

The **ensemble technique** can be said as a hybrid approach. Multiple classifiers can be used. The ensemble technique is divided into two types. One is called "Bagging" and the other one is called "Boosting". **Bagging** is helpful in the way to reduce variance. In bagging, samples are drawn with replacement. Advantage of bagging is that we don't suffer from random errors that classifier is bound to make. In bagging each trees are independent of each other because they consider different subset of features and samples. As bagging reduces variance it helps to avoid overfitting. We have used bagging with **Support Vector Machine (SVM)**. SVM gives high accuracy. It also works well in case the dataset is not linearly separable. Another key point is SVM can help to avoid overfitting by regularization of parameter.

Term Frequency-Inverse Document Frequency (TF-IDF) was used for feature extraction. TF counts the occurrence of a term within a document. IDF finds out the importance of a term. TF-IDF becomes large when there is a large count of a term in a document and the low number of documents with a term in them. TF-IDF becomes smaller if the term appears in many documents in the corpus.

In 2, we provide our project-related work information. In 3 methods that we are using in our project. The process and results we obtained in our project are discussed in 4. In 5, we summarize our whole works and also provided future works of our project.

2 Literature Review

In this section, we will discuss some related works on music classification, also some recent literature review about mood prediction in text [3], very little has been done so far to address the classification of lyrics according to their mood.

2.1 Music Classification

Widowati et al. [4] classified the music into moods like happy, peaceful, angry and sad. They performed this classification with 200 track database for training and 50 for testing. Kashyap et al. [5], in their paper, worked on text mining. They used naïve Bayes for strong independent assumptions to divide the model. They also focused on parts of speech, so that word can be easily tokenized. In previous, the majority of work-related with music mood have done by categorical (Happy, Sad) models of emotions [6]. In [7] [8], they classified music mood through audio lyrics.

2.2 Mood Prediction

Widowati et al. [4] used some music features like pitch, pulse clarity, tempo, key and scale. To classify, Convolutional Neural Network (CNN) is used in this paper. For this, the accuracy they have got is 82 percent. Kashyap et al. [5] took motivation from mood detection challenges. The features they focused on Mood perception, Mood cataloging, Acoustic cues. Acoustic cues [9] made their work significant. They followed the divide and conquer approach to break a model into multiple ones. By gathering the information they trained the model. They divided their dataset based on models in the categories like stop words, sad words, angry words, aggressive words, average words, happy words, etc. It is like a bunch of trees. So feature mapping becomes easier [10]. In this paper [6], they predict the mood of a song based on song lyrics. Happy or Sad mood can be detected by high precision based on the text lyrics features [11].

2.3 Bangla Mood Processing

Bangla's mood processing depends on word embedding [12]. A single Bangla word can occur in a sad song or a happy song. The divide and conquer approach can be followed. At the same time, the probabilistic approach applied to find the best match. Though not much information available for Bangla mood processing.

3 Proposed Methodology

3.1 Method Overview

Figure 1 shows a complete overview of our project. At first, we created a corpus that contains Bangla song lyrics. Corpus [13] mainly appears in Natural Language Processing (NLP) area. In the next step, we sampled and labeled the corpus. Which means we have preprocessed the data. The next staff is feature extraction. Feature extraction is mainly the transformation of raw data into features. Which can be used as the input of a learning algorithm. In this project, we used the Term Frequency-Inverse Document Frequency (TF-IDF) for feature extraction. Then divided the data into training and validation set. In the next stage, we have applied an ensemble learning technique [14]. More clearly we have applied a bagging algorithm. Ensemble technique is the process of combing different classification techniques to build a powerful composite model from the data. Bagging combines different classifiers into a single prediction model. We have used a bagging classifier with a Support Vector Machine (SVM) model.

3.2 Our Dataset

For classifying lyrics mood we did not find a suitable dataset. So, We have built our dataset. In our dataset we have two attributes they are lyrics and mood, following that we have 40 objects. In the train set, we have 32 objects and in the test set, we have 8 objects.

3.3 Feature Used

TF-IDF: Term Frequency-Inverse Document Frequency (TF-IDF) is used to know how important a word is to a corpus. It is mainly used for feature extraction. Term-Frequency (TF) finds how often a word appears within a document. Inverse Document Frequency (IDF)down scales words that appear a lot across documents. From the lyrics, tokenization has to be performed. For tokenization, the lyrics were converted into vectors. But it's not enough. Because it could have different n-gram sequences. A Porter-Stemmer-Tokenizer hybrid to splits sentences into words (tokens) and applies the Porter Stemming algorithm to each of the obtained tokens [15]. Tokens that are only consisting of punctuation characters are removed as well. Only tokens that consist of more than one letter are being kept. Binarization is used to transform both the discrete attributes and continuous attributes into binary attributes. For counting the number of times each word occurs in each lyrics term frequency (tf) has been used.

$$tf - idf(t,d) = tf(t,d) * idf(t) \tag{1}$$

The equation was used to calculate term frequency depending on normalized term frequency. Term document frequency used as a form of inverse document frequency.

$$idf(t) = log(\frac{(1+nd)}{1+df(t,d)}) + 1 \tag{2}$$

"nd" is the total number of lyrics. In the lyrics corpus, the columns correspond to terms and rows correspond to documents in the collection. The value of "nd" increases proportionally with the number of times along with the word appears as a document.

Tokenization: From the lyrics corpus it is important to extract desired content. For breaking lyrics into tokens we have used porter tokenizer. After tokenizing the lyrics the texts were kept into a vocabulary. Stopwords from the corpus were removed.

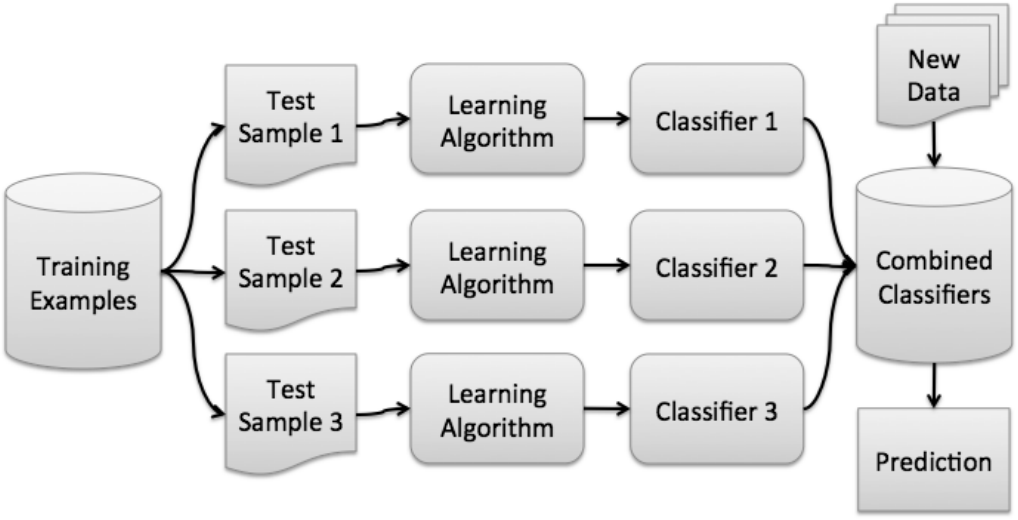

Fig. 2 Workflow of the Bagging Technique

3.4 Procedure

Fig 2 shows the workflow of our work. The training examples are indicating the lyrics' corpus. The corpus is divided into small bags. Which contains Bangla lyrics randomly from the lyrics corpus. The bags are being used as the input of the learning algorithm. Here we have applied Support Vector Machine (SVM) as the learning algorithm. In the bagging procedure by Support Vector Machine (SVM) we need to fix the "n_estimators". The term "n_estimators" indicates the number of trees we want to build before taking the maximum voting or averages of predictions. We have used the value of "n_estimators" equal to 10. Which means it would build 10 classifier model. Then the algorithm would take combined voting. And thus it will give the prediction.

3.5 Model Used

Support Vector Machine (SVM): In the project to build classifier we have used Support Vector Machine (SVM) as the model Fig 3. It is a supervised learning technique. The support vector machine tries to find the optimal separating hyperplane, as a result, the margin of the training data can be maximized.

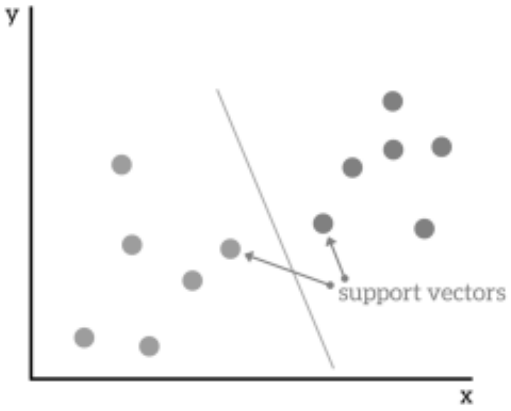

Fig. 3 Support Vector Machine Diagram

4 Result and Experiment

By using precision, recall, and f1 score we have evaluated the performance of our model. The formulas of the following terms as follows:

$$F1 = 2 * \frac{precision * recall}{precision + recall} \tag{3}$$

where

$$precision = \frac{TP}{TP + FP} \tag{4}$$

and

$$recall = \frac{TP}{TP + FN} \tag{5}$$

TP = number of true positives, FP = number of false negatives, and FN = number of false negatives.

Table 1 Results of training and test accuracy :

Training Accuracy	0.9375
Test Accuracy	0.875

Table 2 Model Evaluation :

	Precision	Recall	F1-score	Support
Happy	.50	.33	.40	3
Sad	.67	.80	.73	5
Accuracy	-	-	.62	8
Macro avg	.58	.57	.56	8
Weighted avg	.60	.62	.60	8

Table 1 shows the test and training accuracy of the model. Table 2 shows the evaluation result. Macro average (macro avg) take the average of the precision and recall of the system on different sets, calculating metrics for each label, and find their unweighted mean. This does not take label imbalance into account. The weighted average (weighted avg) calculates metrics for each label and find their average weighted by the number of true instances for each label. Fig 4 shows the word cloud for happy class and Fig 5 shows the word cloud for sad class.

Fig. 4 Happy Word Cloud Visualizations of the Most Frequent Words.

Fig. 5 Sad Word Cloud Visualizations of the Most Frequent Words.

5 Conclusion

We have tried to develop a classifier to detect the mood of the Bangla song. Due to the scarcity of the Bangla labeled lyrics dataset, we have created a dataset for the

project. We have used Bagging with Support Vector Machine (SVM). We tried to show the fruitfulness of Bangla song mood prediction by using lyrics. For feature extraction, we have used Term Frequency-Inverse Document Frequency (TF-IDF). Stop word removal was also in the act. However, we have created low dimension dataset containing lyrics and moods as features. In the future, we want to enlarge the dataset and we want to apply the neural network to increase the performance.

Acknowledgments this research work was supervised by Prof Dr. Salekul Islam (Professor Head, Department of Computer Science and Engineering, United International University) and Mr. Abu Shafin Mohammad Mahdee Jameel (Assistant Professor, Department of Computer Science and Engineering, United International University).

References

1. Västfjäll, Daniel: Emotion induction through music: A review of the musical mood induction procedure.In :Musicae Scientiae volume=5.pp: 173–211.SAGE Publications Sage UK: London, England(2010).
2. DeNora, Tia : Music in everyday life.In :Cambridge University Press (2000).
3. Alm, Cecilia Ovesdotter and Roth, Dan and Sproat, Richard :Emotions from text: machine learning for text-based emotion prediction.In :Proceedings of the conference on human language technology and empirical methods in natural language processing.pp:579–586.Association for Computational Linguistics (2005).
4. Widowati, Fika Ulfa and Nugroho, Fitriyanto Santi and Shidik, Guruh Fajar: Classification of Music Moods Based on CNN. In: 2018 International Seminar on Application for Technology of Information and Communication.pp318–321.IEEE (2018).
5. Kashyap, Nirbhay and Choudhury, Tanupriya and Chaudhary, Dev Kumar and Lal, Roshan : Mood based classification of music by analyzing lyrical data using text mining.In: 2016 International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE). pp.287–292.IEEE (2016).
6. Raschka, Sebastian : MusicMood: Predicting the mood of music from song lyrics using machine learning. In :arXiv preprint arXiv:1611.00138 (2016).
7. Gong, Zheng and Yu, Tian :Chinese web text classification system model based on Naive Bayes. In: 2010 International Conference on E-Product E-Service and E-Entertainment.pp:1–4. IEEE (2010).
8. Laurier, Cyril and Grivolla, Jens and Herrera, Perfecto: Multimodal music mood classification using audio and lyrics. In: 2008 Seventh International Conference on Machine Learning and Applications. pp:688–693. IEEE (2008).
9. Acoustic cues to linguistic profiling? Machine learning of phonetic features of African American English, Arjmandi, Meisam K and Dilley, Laura and Wagner, Suzanne, The Journal of the Acoustical Society of America, ASA (2018).
10. Classification of music into moods using musical features, Kartikay, Avijeet and Ganesan, Harish and Ladwani, Vandana M, 2016 International Conference on Inventive Computation Technologies (ICICT), IEEE (2016).
11. Searching for musical features using natural language queries: the C@ merata evaluations at MediaEval, Sutcliffe, Richard and Hovy, Eduard and Collins, Tom and Wan, Stephen and Crawford, Tim and Root, Deane L, Language Resources and Evaluation, Springer (2019).
12. From word embeddings to document distances, Matt and Sun, Yu and Kolkin, Nicholas and Weinberger, Kilian, International conference on machine learning (2015).

13. The JHU parallel corpus filtering systems for WMT 2018, Khayrallah, Huda and Xu, Hainan and Koehn, Philipp, Proceedings of the Third Conference on Machine Translation: Shared Task Papers (2018).
14. Ensemble learning for data stream analysis: A survey, Krawczyk, Bartosz and Minku, Leandro L and Gama, João and Stefanowski, Jerzy and Woźniak, Michał, Information Fusion, Elsevier (2017).
15. Text mining: open source tokenization tools-an analysis, Vijayarani, S and Janani, R and others, Advanced Computational Intelligence: An International Journal (ACII) (2016).