BanglaMusicMooD : A Music Mood Classifier from Bangla Music Lyrics

***Abstract*—The aim of this research work is to detect the music mood of the listener, by analyzing the lyrics of Bangla language based songs. We used a naive Bayes Classifier as the main feature to built the music recommendation system, trained the attitude of songs based on song lyrics to predict. We constructed our own dataset of 500 Bangla songs consisting Rabindra Sangeet, Nazrul Geeti, Folk, Modern. 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 really well to classify the mood of the song with high accuracy.**

***Index Terms*—Music recommendation system, Music classification model, naive Bayes Classifier, lyrics analysis**

I. Introduction

In our leisure and recreation time, we usually do some activities such as reading book, listening musics, playing, among them listening musics is one of the popular one that has an amplifying effect on social consistency, emotional state of mind and mood of the listeners [10]. Among music lovers, music classification and recommendation model has obtained enlarged popularity.

The power of music to influence mood, create scenes, routines and occasions is widely recognized and this is reflected in a strand of social theory from that portrays music as an influence on character, social structure and action [2]. ***Music*** makes humans human, it's an important art that is very soothing and relaxing which can make our bad days better. The rapid growth expansion of music industry we listened everyday, the more we will be able to examine what we actually want in music and also realize exactly what it means to be human. Around 68 percent of people who has aged between 18 and 34 years old announce that they listened and enjoyed music everyday.

In our approach, we created our own Bangla Song dataset for filtering and music mood prediction of listeners that can be associate 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 song database in order to select music by it's lyrics analysis.

Based on lyrics analysis, we used a **naive Bayes classification** model for mood prediction. As naive Bayes classifier performs really well when the conditional independence assumption actually hold, these classifier will converge really faster than any other models like Logistic Regression. Naive Bayes has very high learning efficiency and it can estimate all the probability just need a scan of the training data [3]

Grid Search has used for finding the optical hyper parameters of a model that outcome the most accurate predictions. Based on the hyper parameter values, the performance of the entire model is specified. Grid Search is a technique that works in an iterative way.

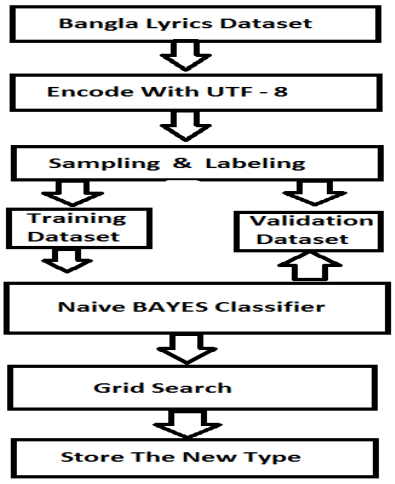

Fig. 1. Flowchart summary of our work. At first a dataset is created. As it contains bangla character so it is need to encode it in UTF-8 format. Then assigning the class label. Next dividing the dataset into test and validation. Then using naïve bayes classifier, and grid search to find out the mood. Grid search used for hyper parameter specification and evaluating the model.Then saving the newly classed lyrics.

In machine learning. the value of a **hyper parameter** is set before the learning process begins. It's just a value by which measured that how much performance can be gained by tuning it.Depend on the choice of hyper parameters, a model can required the time to train and test.

K Fold Cross validation is a numerical method used to determine that how skill a machine learning models is. In this method, the data is divided into k subsets. Then the whole method is repeated k times, each time one of the k subsets is used as the test set/ validation set and the other k-1 subsets are put to form a training set.

In II, we provides our project related work informations. III methods that we are using in our project. The process and results we obtained in our project are discussed in IV. there is a small dicussions about our work in V. In VI, we summarizes our whole works and also provided future works of our project.

II. Related Works

In this section, we discuss some related works on music classification, also some recent literature review about mood prediction in text [1], very little has been done so far to

address the classification of lyrics according to their mood.

A. Music Classification

Widowati et al. [11] 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. [6], 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 tokenize. In previous the majority of work related with music mood have done by categorical (Happy, sad) models of emotions [8]. For example this paper [8] used naive Bayes classifier for training. Some other research [9] is done the same thing but they used audio based modeling. In this paper they used a semantic computing technique which does not have genre information.

B. Mood Prediction

Widowati et al. [11] 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. [6] took motivation from mood detection challenges. The features they focused on Mood perception, Mood cataloging, Acoustic cues. Acoustic cues made their work significant. They followed divide and conquer approach to break a model into multiple ones. By gathering information they trained the model.They divided their dataset based on models in the categories like stop words, sad words, angry words, aggression words, average words, happy words, etc. It is like bunch of trees. So feature mapping becomes easier. In this paper [8], they predict the mood of song based on songs lyrics. Happy or Sad mood can be detected by high precision based on the text lyrics features.

C. Bangla Mood Processing

Bangla mood processing depends on word embedding. A single bangla word can occur in a sad song or in a happy song. Divide and conquer approach can be followed. At the same time probabilistic approach applied to find the best match. Though not much information available for bangla mood processing.

III. Methods

A. Dataset

For building the machine learning model we need dataset. But for bangla music mood dataset in .csv format was not available. So we created our own dataset. We categorized the dataset of bangla song lyrics by using lyrics, artist, mood, year, genre. We made the dataset aiming 500 songs. We divied it 60 percent for training, 40 percent for validation. Another thing is we created the dataset is .txt format because UTF-8 encoding was creating problem to fit .csv files. We labled the mood in bangla in the dataset. We also manually created the stop words list, so that the model can find map them.

B. Feature Extraction

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. Porter stemming algorithm is used for nflexional endings from words. Binarization is used to transform both the discrete attributes and the 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 dataset the columns correspond to terms and rows correspond to documents in the collection. Value of "nd" increases proportionally with number of times along with the word appears as document.

C. Tokenization

In the dataset, we have lyrics. Lyrics are like sentences. From the lyrics for it is important to trim desired content. Tokenize texts, select tokens of interests, create an NLTK text. As the work is going on Bangla music mood, so CLTK [Classical Language Toolkit] is used. By using CLTK, normalizing the words making vocabulary. TokenizeSentence named function is used to divide sentences into smaller parts of words. It works by detecting spaces between two words. Then finding most common words is the task. This thing is done by stop words. It is a kind of pre-processing. Then the words converted into vectors. Count vectorize is used to build vocabulary. In the vocabulary one word can occur many times, to remove inflation porter stream algorithm is used. We wrote a tokenization function by using CLTK to token the bangla strings.

D. Model Selection

Grid search applied to optimize the "Naïve Bayes Model". F1 score was applied on class label of the dataset. To support F1 score Recall and Precision were used. Confusion matrix was generated to see the number of true positive rate as well as negative rate.

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

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.

$$P(W_j/X_i) = \frac{P(X_i/W_j) * P(W_j)}{P(X_i)} \tag{6}$$

Bayesian theory was applied to calculate Prior and Posterior probability. It is the way to classify a test data. Binary feature vector supports the Bayesian Theory to count probabilities.

E. Software

The PyCharm [5] and scikit-learn [7] used for model training, the Google colab [4] for running the code. Notepad++ for editing the code. Matlab used to generate graph. Matplotlib along with colab was used to visualize the data in real time. The model generated a PKL file to serialize the objects .All the codes are available at Google drive.

IV. Experiment and Results

A. Experiment

Manual class labeled assigned. The data set consists of 40 percent sad , and 60 percent happy song lyrics. Cross validation was performed measuring F1 score. In figure-2, we generates happy workcloud of the frequent words as well as in figure-3, we generates sad wordcloud. After collecting the Bangla song lyrics dataset, we manually assigned class label. Then we divided the dataset into test and train. The training dataset consist of 40 percent happy and 60 percent sad song lyrics in Bangla. So it can be said that the model is a bit biased to sad lyrics, it is because of the lack of bangla lyrics resources. For model selection we applied grid search technique. For hyper-parameter specification we used grid search approach. K-fold cross validation was used. Here we used k=2, because there were some decoding problem. We used ROC curve for measuring the quality of our model. Grid search used three separate Naïve Bayes models.

For features vector mapping Multi-Variate Bernouli Bayes was used. For term frequency Multi-Variate Bernouli Bayes was used and tf-idf was used for checking the occurrence of a word.

B. Results

The word cloud showed the most frequently used Bangla words. Grouping the songs by lyrics showed the sub sample. We applied grid search by using three different naïve bayes models. Multinomial naïve Bayes classifier showed good performance. But for some technical issues, we could not generate enough number of evaluation curves.

Fig. 2. Happy wordcloud visualizations of the most frequent words.

Fig. 3. Sad wordcloud visualizations of the most frequent words.

After selecting the model the final classifier was used for training.The model is generating result. As we did not get enough number of data, so the model has some calculational porblem to find out the actual results that we were expecting. If we train the model with high amount of data then the result would become more clear.

ACC	PRE	REC	F1	ROC	AUC
100.0	100.0	100.0	100.0	100.	100.0
100.0	100.0	100.0	100.0	100.	100.0

Fig. 4. Result generated by the bangla music mood model

V. Discussion

Our model worked on our dataset. As there are limited resources on bangla music mood prediction so, comparison is little bit challanging. Our model is better because we made the dataset and class labled them manually. We used optimal grid search technique. And removal of stop words made that

more redundancy free. After deep analysis on result data, it shows that the no of sad songs increases over the years passed. It is very clear in fig . It strongly gives a knowledge that in recent years bangali people are very much comfortable with sad song rather than happy songs. Or the artists have very much interest on sad songs over happy songs. However in modern society people are having so much trouble with their emotions, that they are listening to large amount of sad songs than previous generations. It can be used as mood filter which will describe a music recommendation system. In our naive Bayes model among all three of them that we have used all perform better than before if the stop words are not used for grid search. The higher ROC AUC derives that in the used model some non-relevant words which are both describes for happy sad. As explained before the multinomial naive Bayes model worked better than Bernoulli naive Bayes model which was used to input the binary feature vectors. The mood classifier have high training and validation precision but the result for cross validation for electing the model did not completely satisfied overfitting data. However the high precision for the classifier is still the successive for reaching the goal.

VI. CONCLUSION AND FUTURE WORK

In this paper we tried to focus on Music Mood Prediction by using naïve Bayes Classifier. We worked with "Bangla Song" dataset. We focused on the strong independence assumption of Naïve Bayes. Besides Naïve Bayes need less data to train our model. Naive Bayes classifier plateaus above a certain threshold in our model. Till now the model is working with little size dataset. As we increase the instances of the dataset, it would work much better. While using Bangla Dataset encoding was a bigger issue. Tokenizing was another issue. By using Classical Language Toolkit we overcome the problem. Grid search was used to find out the combinations of "Hyper-parameters". Due to lack of Bangla resources the model still struggles to find the best solution. But in most of the cases it can identify the mood. By using this model, we can get to know how people categorize music besides, what actually a person feels. Music trends can be plotted and assumed also.

In near future, we would like to predict the genre of song and also try to improve the accuracy level of our music classification model.

REFERENCES

[1] Cecilia Ovesdotter Alm, Dan Roth, and Richard Sproat. "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*. Association for Computational Linguistics. 2005, pp. 579–586.

[2] Tia DeNora. *Music in everyday life*. Cambridge University Press, 2000.

[3] Zheng Gong and Tian Yu. "Chinese web text classification system model based on Naive Bayes". In: *2010 International Conference on E-Product E-Service and E-Entertainment*. IEEE. 2010, pp. 1–4.

[4] Gd J Hoyos-Rivera et al. "Colab: A new paradigm and tool for collaboratively browsing the web". In: *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans* 36.6 (2006), pp. 1074–1085.

[5] Qiang Hu, Lei Ma, and Jianjun Zhao. "Deepgraph: A pycharm tool for visualizing and understanding deep learning models". In: *2018 25th Asia-Pacific Software Engineering Conference (APSEC)*. IEEE. 2018, pp. 628–632.

[6] Nirbhay Kashyap et al. "Mood based classification of music by analyzing lyrical data using text mining". In: *2016 International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE)*. IEEE. 2016, pp. 287–292.

[7] Fabian Pedregosa et al. "Scikit-learn: Machine learning in Python". In: *Journal of machine learning research* 12.Oct (2011), pp. 2825–2830.

[8] Sebastian Raschka. "MusicMood: Predicting the mood of music from song lyrics using machine learning". In: *arXiv preprint arXiv:1611.00138* (2016).

[9] P. Saari et al. "Genre-Adaptive Semantic Computing and Audio-Based Modelling for Music Mood Annotation". In: *IEEE Transactions on Affective Computing* 7.2 (Apr. 2016), pp. 122–135. ISSN: 1949-3045. DOI: 10.1109/TAFFC.2015.2462841.

[10] Daniel Västfjäll. "Emotion induction through music: A review of the musical mood induction procedure". In: *Musicae Scientiae* 5.1_suppl (2001), pp. 173–211.

[11] Fika Ulfa Widowati, Fitriyanto Santi Nugroho, and Guruh Fajar Shidik. "Classification of Music Moods Based on CNN". In: *2018 International Seminar on Application for Technology of Information and Communication*. IEEE. 2018, pp. 318–321.