World Academics Journal of Engineering Sciences
Research Paper
Vol.8, Issue.3, pp.15-20, September (2021)
E-ISSN: 2348-635X

An Analytical Study on Music Listener Emotion through Logistic Regression

Md. Anzir Hossain Rafath[1*], Fatema Tuz Zohura Mim[2], Md. Shafiur Rahman[2]

[1,2]Dept. of Computer Science and Engineering, University of South Asia, Dhaka, Bangladesh

[*]Corresponding Author: rafatanzir83@gmail.com, Tel.: +8801966186115

Available online at: www.isroset.org

Received: 30/Aug/2021, Accepted: 20/Sept/2021, Online: 30/Sept/2021

Abstract— Now-a-days music listeners face immense obstacles when trying to find suitable music for a specific context. This is what actually motivated us to find a suitable solution on it. Here we presented a music classification system that provides the listener an opportunity to browse their music by mood. Mood classification while listening to music is an emerging domain of music information retrieval. This paper investigates the challenging issue of recognizing emotions while listening to music. Specifically, we formulate the music types as a regression problem to predict the human emotional state. No categorical taxonomy is used that's why the regression approach is free of the ambiguity inherent to conventional categorical approaches. This will help to build an intuitive and contextually aware playlist generation tool. Talking of play list tool we have come to the point that most people are not sure that if their mood has been changed or not while listening to their favorite music, people tend to listen to emotional music when they deal with sorrow or disgust. In the time fear and anger they don't listen any music. In the contrast they listen to music while having good feelings. We have achieved 78% of accuracy by taking the regression approach for the research.

Keywords— Data Mining, Machine learning, Regression approach, Emotion prediction, Music analysis

I. INTRODUCTION

Music has been a part and parcel of our civilization from ancient times. But now-a-days listeners can hardly find the suitable music according to their taste and mood. At the same time even though music has become very much available, it's very unlikely that listeners like to hear the same type of music all the time. We have our personal preferences but it also differs from time to time depending on our emotional state. But it is really difficult and time-consuming to choose a song relevant to our mood. So as a helping hand we need is an additional parameter – Mood which will signify the emotion of that particular music. The aim is to prepare a model which will predict the human emotion based on their current searched song or current playing song to help suggest the next songs based on their mood. Our main objective is to combine previous data and current data in order to design a good predictive model. Understanding the different attributes is also needed in order to get more accurate results. There are numerous models describing human emotions. One of them is the Hevner model [1]. It categorizes various attributes into 8 different groups where each of them represents a class of mood. That means each class consists of the same type of emotions. But, Russell J.A. had a different idea [2]. He came up with a circle model where all the mode classes are separated from each other and have polar coordinates. Then there is Thayer's model which is a two-dimensional model [3]. IT has two axes representing stress and energy. In recent times Emery Schubert provided an update on Hevner's adjective [4]. In his study, several musically experienced people gave their opinion about 91 musical adjectives. It consists of Hevner's 67 adjectives with the addition of Russell's circumplex model and Whissel's dictionary of effect [5].Out of these 91 emotional adjectives 46 were kept. In recent time Doris Baum introduces EmoMusic [6] in their paper which was a study about the usefulness of 'PANAS-X'. It categorizes music according to mood using different machine learning algorithms. In this study, they showed that emotion may be derivable in an automatic way. And we can use Na ̈ive Bayes and Random Forest classifiers to predict the emotion of music with a success rate. K.C. Dewi and Harjoko [7] came up with a music classification system based on mood parameters where they used K-Nearest Neighbor and Self Organizing Map. But for better understanding, we used music genre and popular music to classify music type.

II. RELATED WORK

In recent years there's been a huge amount of research on music emotion recognition due to the requirement of the entertainment and digital media industry. In 2008, Yang et al[8] used the Daubechies wavelets coefficient histogram (DWCH) algorithm [9], spectral contrast algorithm with PsySound [10] and Marsyas [11] extracting the features as previous work. Then he trained the data with features by multiple linear regression (MLR) [12], support vector regression (SVR) [13], and AdaBoost [14].

Later, he used the MEVD (music emotion variation detection) to do the same ground true data and features for fear comparison and at last he got his result that the regression approach has more promising prediction operation than normal arousal-valence (AV) computation algorithms in doing MER [15], [16], [17], [18]. And also the regression approach can be applied in MEVD. In 2009, Han et al [19] used Juslin's theory [20] along with Thayer's emotion model to analysis the 11 kind of emotions (angry, bored, calm, excited, happy, nervous, peaceful, pleased, relaxed, sad and sleepy) and their cause: 7 music characteristics (pitch, tempo, loudness, tonality, key, rhythm and harmonics). Then he compared his result using support vector regression (SVR) as a classifier to train two regression functions for predicting arousal and valence values based on the low-level features, with the GMM [21] and SVM [22]. The final results showed that the SVR can increase the accuracy from 63.03% to 94.55%, but the GMM can only grow 1.2% (91.52% to 92.73%).In 2010, Kim et al [23] did comprehensive work in music emotion recognition review. He started with the psychology research on emotion with the Valence-Arousal space and the perceptual considerations. In the feature part he described the lyrics feature selection and the acoustic features. Lyrics features selection was based on the plea-sure (valence), arousal and dominance (PAD) [24], Affective Norms for English Words and Affective Norms for Chinese Words [25] to select the affective features from the signals. For the acoustic features, he gave 5 types (Dynamics Timbre, Harmony, Register, Rhythm and Articulation) 17 features (RMS energy, Mel-Frequency cepstral coefficients [26], spectral shape, and spectral contrast. Roughness, harmonic change, key clarity, major ness Chromagram, chroma centroid and deviation Rhythm strength, regularity, tempo, beat histograms Event density, attack slope, attack time). But he focused the MFCCs (Mel-Frequency cepstral coefficients) to comment. And in themachine learning part, he mentioned the SVM, Logistic Regression, Random Forest, GMM, K-NN, Decision Trees and Naive Bayes Multinomial classifiers. In the end of his paper, he gave examples of some combinations of the emotion types. For example, Yang and Lee [27] combined the both audio and lyrics for emotion classification with 12 low level MPEG-7 descriptors and increased 2.1% (82.8% vs 80.7%with only audio) of the results. And other review was for Bischoff [28] combining audio and tags [29] by Bischoff with 240 dimensional features and 178 mood tags. Then Bischoff prove that the combination is better than the single one. In 2013, "The MediaEval 2013 Brave New Task: Emotion in Music" attracted a large number of challengers to take part in the task. In this task, they split the 1000 songs as 700 for development and 300 for tests. And the first task was for the continuously time to determine the emotional dimensions, arousal and valence, and the automatically detection for the arousal and valence of the songs were the second task in static emotion characterization. Later on, this challenge was held in 2014 and 2015. In 2014, Gao et al.[30] used the 6552 standard large emotion feature which extract from OpenSMILE to compare with his multi-feature (GMM super vector– GSV [31], positive constrain in matching pursuit PCMP[32], multiple candidates and optimal path) with the APM dataset. Then he did the machine learning by SVM classifiers with RBF kernel and got the confusion matrix of the results. For the results, his working proved that his multiple features are more effective than only use the openSMILE one by 5.25% (74.9% to 80.15%) .

III. METHODOLOGY

Prediction modeling using Logistic Regression Method

Logistic Regression is used to classify data into two classes. One is binary and another is multi-class logistic regression. Here the binary class has two classes but the multi-class has more than two classes. It predicts a probability within 0 and 1 using linear combination of variables.

Logistic Function

Using logistic function we can squash numbers into a reasonable range like between 0 and 1. In the case of Logistic Regression it is the Sigmoid function.

$S(Z) = 1/1 + e^{-z}$…………………….……..……………… (1)

Here Z define our regression equation which helps us to map our function,

$Z = \beta_0+\beta_1\chi_1+...+\beta_n\chi_n$…………………………………….… (2)

The resulting output from this equation now looks something like this. It's a s-shaped curve. Values can only be between 0 and 1.

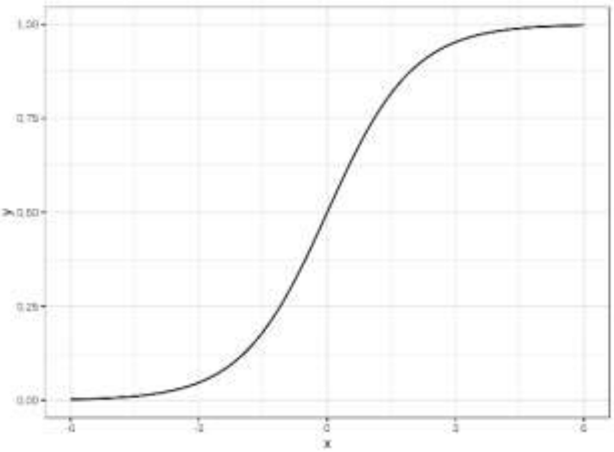

Fig 1: Logistic Regression Function [33]

Log-Odds

We have used the odds and probability to find the coefficient of our equation.

odds = p/(1-p)……………………………………....….(3)

p = odds/odds + 1……………………………….…... (4)

In regression function 'Z' is a linear combination of variables and ‘S(Z)’ is probability. So, we can say and if we isolate Z we get,

p = 1=1 + e^{-z}……………………………….…………..... (5)

This is the equation of “log of the odds”. We can take the inverse to get the coefficients in terms of odds.

Z = log (p/1 - p)………..……………………….……. (6)

After modeling and pulling some coefficients from our data if we calculate log-odds using the regression equation we get the following chart.

Fig 3: Log - odds vs Sphericity

Log-odds gives us a linear model. We need the inverse of the log-odd function to get our target variable.

Information Gain

After collecting the Firstly we transform our data to log-odds. But log-odds of 0 and 1 are positive and negative infinity. So, we plot a random line to fit our data. Since all of our points are at infinity we can't work out how well the line fits the data. So in order to assess the quality of our line, we project each point onto it so that we can read off the log-odds for each point. Since we have the log-odds of every point we can transform odds back to probabilities using the rearranged logistic regression equation. Our variable points are on a probability axis and it indicates a sigmoid line. At this point we can read off the probability from the graph and use it to calculate the total likelihood for the line. It is generally calculated in log form. After calculation if we look closely it produces a steeper sigmoid curve. Although the _rst line is random the maximum likelihood estimation finds the optimum line very quickly which makes this process highly efficient.

Data Training

After collecting the data we removed outliers and possibly unclassified instances from the training data. But logistic regression is a linear algorithm with a nonlinear transform in output. It does assume a linear relationship between the input variables with the output. So we have tried to use box-cox method to transform input data for better expose of this linear relationship. After that we converted those data into an attributed relation file format (.arff) and then we have used Weka for classification. After classification using some algorithm we got some result and later we have analyzed those result. Here is the simple work ow chart given.

IV. Results and Discussion

Study Description

The term emotion is changeable through hearing a piece of music. In our study, we tried to recognized emotion, based on music by using Logistic regression (LR). The Study mainly consists of three steps: (a) feature extraction: eighth distinct music types are extracted and analyzed. (b)Mapping: Studied Thayer's model [2], Russell's model [24], Hevener's model [5], and after getting some idea from Schubert's model we mapped musical types into eight categories. (c) Training: This study used extracted musical type as input to train the LR.

Dataset

After studying on human emotion and music genera we classified them into eight categories. In the case of emotion we decided to use sad, happy, disgust, anger, fear, surprise, relax and bored. And for music genre we stick with Country Music, Romantic Music, Emotional/Sad Music, Classical Music, Heavy Metal, Rock and Pop Music, Electronic Dance Music, and Rap Music. We did a survey and asked people some questions online to collect the data. We got 712 unique data to work with. We filtered these raw data and converted into attributed relational _file format (.arff).

Music Types

In our study, we consider various types of music from Thayer's model, Russell's model, Hevener's model including Schubert's model. Those types are Rap, Heavy metal, Pop and Rock, Romantic Music, Emotional/Sad Music, Country Music, Electronic Dance Music(EDM). We used them as an input of LR In the emotion recognition system.

Data Analysis

To analyze the data, moods were categories were re-categorized into eight categories. They are sad, happy, disgust, anger, fear, surprise, relax, and bored. A total result was calculated for each category. We obtained the general result of the survey, 85% of participants have said that their mood changes from listening to music, 8 % of participants said that music does not affect their mood and 7% participants not sure about their mood changes.

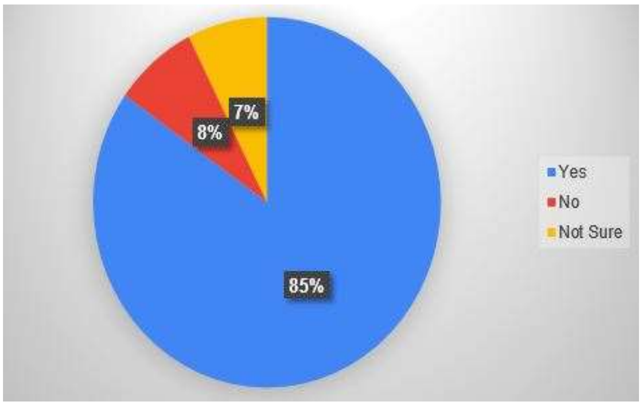

Fig 4: Mood changed from listening music

We established the relationship among variables to test the hypotheses. We achieved that when they are sad they try to hear Emotional or sad Music most which is 38%.

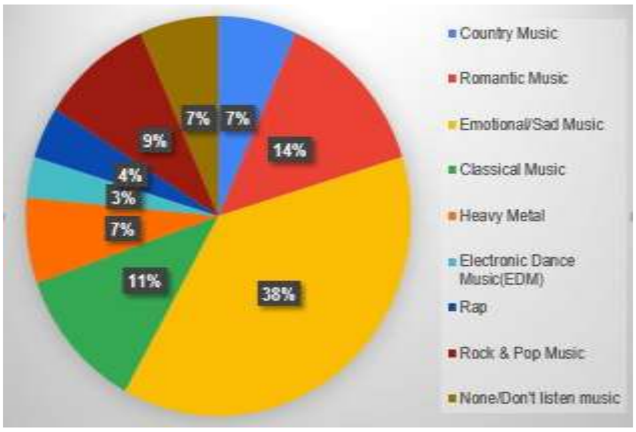

Fig 5: Types of music listen when feeling sad

We also tried find when people usually feel happy which type of song they listen most. After analysis the data we found that general participants tend to hear Romantic Music when they are happy which is 29%.

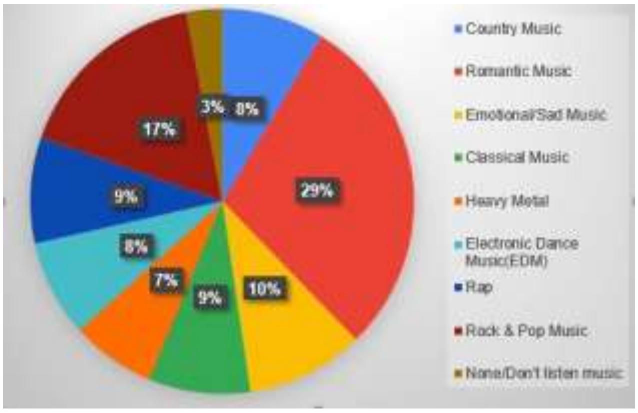

Fig 6: Types of music listen when feeling happy

Anger is another stage, it this situation some people don't know how to control their anger. From the data most participants didn't want to hear any music when they were feeling anger.

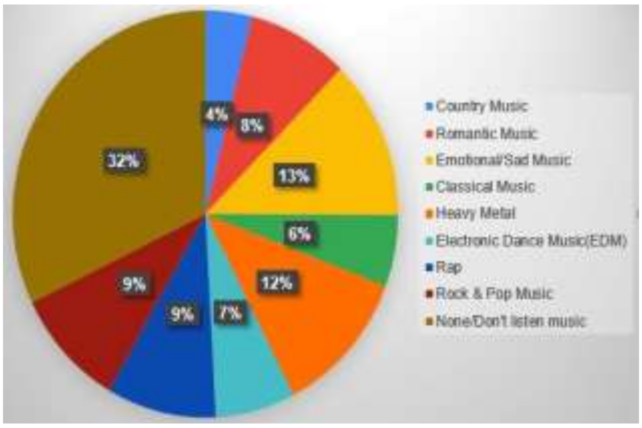

Fig 7: Types of music listen when feeling anger

It varied person to person what types of music people listened when they bored. From our collected data we find out that all types of music participants were listened when they were bored but most of them listened emotional/sad music which percentage were 16%.

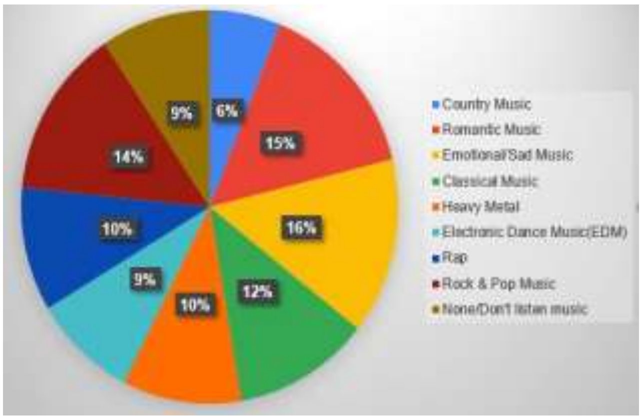

Fig 8: Types of music listen when feeling bored

Most 30% participants didn't want to hear any music when they were feeling disgusted.

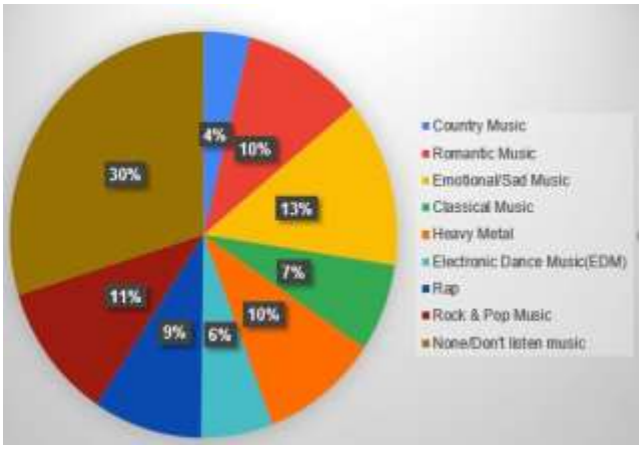

Fig 9: Types of music listen when feeling disgusted

Most 28% participants didn't want to hear any music when they were feeling fear.

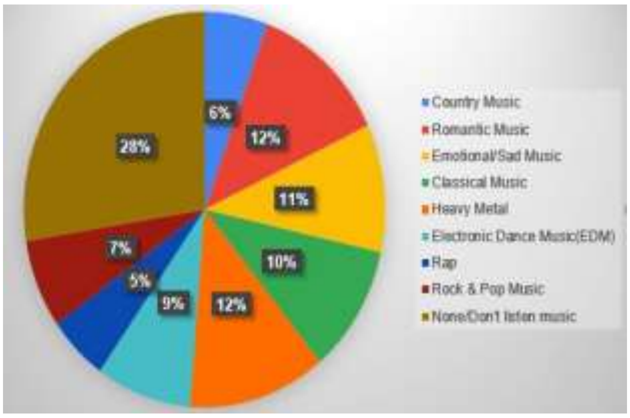

Fig 10: Types of music listen when feeling fear

Most 30% participants want to hear romantic music when they were feeling relax.

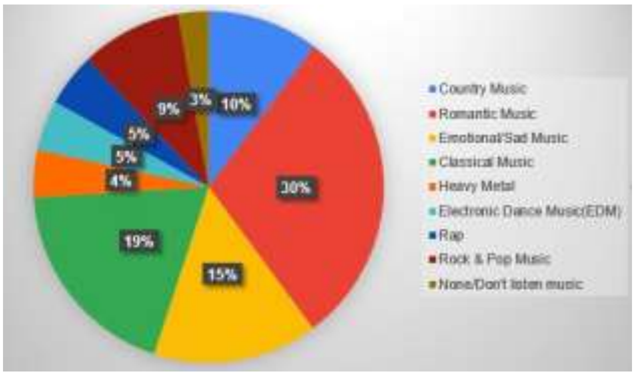

Fig 11: Types of music listen when feeling relax

With the advent of the ubiquitous era, mood-based music recommendation has become one of the rapidly emerging applications. There are some other context-based or lyrics-based music recommendations but all of these fall on the same domain. Here we have tried to add some extra dimension to that domain which will help in future research and make music access easier for us. This study proposes a music recommendation system that extracts mood information based on music type which will later recommend us music based on our mental state. Here we have used logistic regression to predict the mood swing with an accuracy of 78%.

V. Conclusion and Future Scope

In this paper, we proposed an effective method to predict human emotion while hearing a music. Considering the high subjectivity of emotion recognition the paper's precision is pretty high. Due to the low number of classes, the performance of logistic regression is better than Random forest algorithm and neural network. After analysis our survey data, we found that 85% of participants were agree to say that their mood changes from listening to music and also understand that what kinds of music usually people listen based on their emotional states. Since the experiment was performed on a small dataset and limited number of features, it still can be improved by adding more features like weather, current environment, current health state. Even we can also use some features like heart beat rate, music frequency etc. with proper equipment. More attributes will improvise decision making and predicting mood. Each person has its own preferences about what kind of song is to be played for corresponding mood. Collecting data from more user may help us to build better user specific radio application or a personal play list. Implementing our data in current music applications can provide better music experience in the future. In future, we also intend to develop a lyrics based as well as multi-modal music mood classification system and want to detect and improve our flaws in the near future to improve our classification accuracy.

References

[1] K. Hevner, "Experimental studies of the elements of expression in music," The American Journal of Psychology, **vol. 48, no. 2, pp. 246{268, 1936.**

[2] J. A. Russell, "A circumplex model of affect.," Journal of personality and social psychology, vol. 39, no. 6, p. 1161, 1980.

[3] R. E. Thayer, The biopsychology of mood and arousal. Oxford University Press, **1990.**

[4] E. Schubert, "Update of the hevner adjective checklist," Perceptual and motor skills, vol. 96, no. 3 suppl, **pp. 1117-1122, 2003.**

[5] C. M. Whissell, "The dictionary of affect in language," in The measurement of emotions, Elsevier, 1989, **pp. 113{131.**

[6] D. Baum, "Emomusic-classifying music according to emotion," in Proceedings of the 7th Workshop on Data Analysis (WDA2006), Citeseer, **2006.**

[7] K. C. Dewi and A. Harjoko, "Kid's song classification based on mood parameters using k-nearest neighbor classification method and self organizing map," in 2010 International Conference on Distributed Frameworks for Multimedia Applications, IEEE, **pp. 1-5, 2010.**

[8] Y.H. Yang, Y.C. Lin, Y.F. Su, and H. H. Chen, "A regression approach to music emotion recognition," IEEE Transactions on audio, speech, and language processing, **vol. 16, no. 2, pp. 448-457, 2008.**

[9] T. Li and M. Ogihara, "Content-based music similarity search and emotion detection," in 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, IEEE, **vol. 5, pp. V-705, 2004.**

[10] D. Cabrera et al., "Psysound: A computer program for psychoacoustical analysis," in Proceedings of the Australian Acoustical Society Conference, **vol. 24, pp. 47-54, 1999.**

[11] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," IEEE Transactions on speech and audio processing, **vol. 10, no. 5, pp. 293-302, 2002.**

[12] A. Sen and M. Srivastava, Regression analysis: theory, methods, and applications. Springer Science & Business Media, 2012.

[13] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," Statistics and computing, **vol. 14, no. 3, pp. 199-222, 2004.**

[14] D. P. Solomatine and D. L. Shrestha, "Adaboost. rt: A boosting algorithm for regression problems," in 2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541), IEEE, **vol. 2, pp. 1163-1168, 2004.**

[15] E. Schubert, Measurement and time series analysis of emotion in music. University of New South Wales Sydney, vol. 1, 1999.

[16] A. Hanjalic and L.-Q. Xu, "Affective video content representation and modeling," IEEE transactions on multimedia, **vol. 7, no. 1, pp. 143-154, 2005.**

[17] Y.H. Yang, C.C. Liu, and H. H. Chen, "Music emotion classification: A fuzzy approach," in Proceedings of the 14th ACM international conference on Multimedia, **pp. 81-84, 2006.**

[18] M. D. Korhonen, D. A. Clausi, and M. E. Jernigan, “Modeling emotional content of music using system identification,” IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), **vol. 36, no. 3, pp. 588-599, 2006.**

[19] B.j. Han, S. Rho, R. B. Dannenberg, and E. Hwang, “Smers: Music emotion recognition using support vector regression.,” in ISMIR, Citeseer, **pp. 651-656, 2009.**

[20] P. N. Juslin and J. A. Sloboda, Music and emotion: Theory and research. Oxford University Press, **2001.**

[21] R. Sridharan, “Gaussian mixture models and the em algorithm,” Available in: http://people. csail. mit. edu/rameshvs/content/gmm-em. pdf, **2014.**

[22] C.C. Chang and C.J. Lin, “Libsvm: A library for support vector machines,” ACM transactions on intelligent systems and technology (TIST), **vol. 2, no. 3, pp. 1-27, 2011.**

[23] Y. E. Kim, E. M. Schmidt, R. Migneco, B. G. Morton, P. Richardson, J. Scott, J. A. Speck, and D. Turnbull, “Music emotion recognition: A state of the art review,” in Proc. ismir, **vol. 86, pp. 937-952, 2010.**

[24] A. Mehrabian and J. A. Russell, An approach to environmental psychology. the MIT Press, **1974.**

[25] M. M. Bradley and P. J. Lang, “Affective norms for English words (anew): Instruction manual and affective ratings,” Technical report C-1, the center for research in psychophysiology., Tech. Rep., **1999.**

[26] Y. Hu, X. Chen, and D. Yang, “Lyric-based song emotion detection with affective lexicon and fuzzy clustering method.,” in ISMIR, **pp. 123-128, 2009.**

[27] L. Mion and G. De Poli, “Score-independent audio features for description of music expression,” IEEE Transactions on Audio, Speech, and Language Processing, **vol. 16, no. 2, pp. 458-466, 2008.**

[28] D. Yang and W.-S. Lee, “Disambiguating music emotion using software agents.,” in ISMIR, **vol. 4, pp. 218-223, 2004.**

[29] K. Bischo, C. S. Firan, R. Paiu, W. Nejdl, C. Laurier, and M. Sordo, “Music mood and theme classi_cation-a hybrid approach.,” in ISMIR, **pp. 657-662, 2009.**

[30] M. Soleymani, M. N. Caro, E. M. Schmidt, and Y.-H. Yang, “The mediaeval 2013 brave new task: Emotion in music.,” in MediaEval, Citeseer, **2013.**

[31] F. Zhang, H. Meng, and M. Li, “Emotion extraction and recognition from music,” in 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), IEEE, **pp. 1728-1733. 2016.**

[32] A. Aljanaki, F. Wiering, and R. C. Veltkamp, “Mirutrecht participation in mediaeval 2013: Emotion in music task.,” in MediaEval, Citeseer, **2013.**

[33] J. Brownlee, Logistic Regression for Machine Learning, en-US, Mar.2016.[Online].Available:https://machinelearningmastery.com/logistic-regression-for-machine-learning

AUTHORS PROFILE

Mr. Md. Anzir Hossain Rafath completed his Master program from Jahangirnagar University in Information Technology. He earned his B.Sc. degree from American International University-Bangladesh (AIUB) in Computer Science and Engineering. He is now working as a Lecturer in University of South Asia, Bangladeeh Since 2019.

Fatema Tuz Zohura Mim earned her degree Bachelor of Science in Computer Science and Engineering from University of South Asia, Bangladeeh. She is passionate and dedicated. She believes that constant effort will turn to something great. Her goal for her career is to work abroad and contribute to the majestic infrastructures to the world.

Md. Shafiur Rahman Diganta completed his degree Bachelor of Science in Computer Science and Engineering from University of South Asia, Bangladeeh. He is passionate and dedicated. He believes that constant effort will turn to something great. His goal for his career is to work abroad and contribute to the majestic infrastructures to the world.

