Proceedings of the Fifth International Conference on Electronics, Communication and Aerospace Technology (ICECA 2021)
IEEE Xplore Part Number: CFP21J88-ART; ISBN: 978-1-6654-3524-6

Multi-Class Classification of Different Region Pop Songs using Spotify Database

Ankit Datta
Department of Electronics and Telecommunications
Sardar Patel Institute of Technology
Mumbai, India
ankit.datta@spit.ac.in

Mrityunjay Joshi
Department of Electronics and Telecommunications
Sardar Patel Institute of Technology
Mumbai, India
mrityunjay.joshi@spit.ac.in

Jai Dalvi
Department of Electronic and Telecommunications
Sardar Patel Institute of Telecommunication
Mumbai, India
jai.dalvi@spit.ac.in

Dr. Reena Kumbhare
Department of Electronics and Telecommunications
Sardar Patel Institute of Technology
Mumbai, India
reena_kumbhare@spit.ac.in

2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA) | 978-1-6654-3524-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICECA52323.2021.9675998

***Abstract*— Music Information Retrieval is a field that is interested in extracting information from musical compositions whether it is songs, instrumental pieces or scores of various genres. With the development of Machine Learning (ML) and Artificial Intelligence (AI) techniques, better models are used to infer songs and scores. These inferences are later used in detecting copyright breaches, categorizing songs as per user, developing music recommending applications and many more. Though our understanding of music has increased tremendously, linguistic based research has been limited. The work in this paper focuses on analyzing data-sets of songs of different Indian linguistic communities on the basis of some predefined musical parameters like speechiness, loudness, acousticness etc. The only similarity is the songs in comparison of different languages have the same mood or valence (positiveness). This work will help to understand how different linguistic music having the same mood, has different musical features. The method used for this project is Multi Class Neural Network for training and thus testing the data-set. The data-set is collected from a popular music app using the inbuilt Application Programming Interface (API) of the app.**

Keywords*— *Information Retrieval, Music analysis, Categorical encoding, Regression model.

I. INTRODUCTION

Music is a pervasive subject in our culture and nearly everyone loves listening to it and many people even love making it. Art and humanity are inextricably connected. We all have a deep drive to make something, no matter how small or big it is. It is difficult to resist engaging with music, whether making it or listening to it. Music has always played an important role in people's lives, whether for fun, emotional reaction, success, or development.

In Indian society, music occupies a special position: it is often referred to as "the food of the soul" in conventional aesthetics. It embodies the historical tides that have shaped India's modern pluralism and symbolizes India's extraordinary diversity in terms of ethnic, linguistic, and religious terms. Traces of earliest forms of art and music are found in Indus Valley civilization dated back in 2500-1500 BE [1]. Since, music has touched various communities, ethics groups and empires and has been molded in various forms from folk music, tribal music to modern day multi-linguistic western form.

For better understanding music and musical parameters scientifically, a field called Music Information Retrieval comes into picture. Music Information Retrieval or MIR is foremost interested in extracting and inferring meaningful features from music (from the audio signal, symbolic representation, or external references such as web pages), indexing music using these features, and the creation of various search and retrieval schemes (for instance, content-based search, music recommendation systems, or user interfaces for browsing large music collections) as defined by Downie [2]. Music is a highly multi modal human artifact. It can take the form of audio, a symbolic representation (score), text (lyrics), a picture (photo of a singer or album cover), a gesture (performer), or even just a mental model of a specific tune.

In certain cases, an individual's model of a musical object is made up of a blend of these representations. As Schedl et al. [3] points out, human interpretation of music, and especially music resemblance, is affected by a broad range of factors such as lyrics, beat, the perception of the artist by the user's friends, and the user's current mental state. Music perception is usually defined by one or more of the following categories of music perception: music content, music context, user properties, and user context. Computational MIR approaches typically use functionality to construct models to characterize music by one or more of the following categories of music perception: music

content, music context, user properties, and user context. Such kind of detailed analysis of Indian music across different linguistic groups is missing and needs to be given attention.

For effective analysis of music, numerous machine learning techniques are used. Proper prepossessing of data is necessary to get the audio in form useful. Some of the most reliable techniques include Hidden Markov Model (HMM) [4], KNN, Decision Tree, N-gram [5] and SVM [6] [18]. The method opted in the paper is Multi-Class Neural Network. The reason to opt for a Multi-Class Neural Network is because the work focuses on multiple musical parameters namely acousticness, speechiness, energy, liveliness, valence, tempo and loudness. All of these parameters are explained in subsequent sections. The Multi-Class Deep Neural Network method effectively handles multiple such parameters together and gives accuracy at par with other models.

The work in the paper focuses on the musical context of pitch, beat and loudness of different regional linguistic music of India. A particular emotion is common in these songs but the difference in musical parameters and users' perception highlighted in the research.

The remaining paper proceeds as follows: Section II describes the existing research work on musical analysis done till date, Section III provides a brief introduction to the problem statement. Section IV discusses the methodology used for analysis. Section V highlights the major findings and observations of the research. Section VI outlines the conclusion of the study.

II. Previous Work

Music was always an integral part of human civilization. Humans have always associated music with different aspects of life, especially emotion. S. Koelsch [7] work discussed the neurological changes observed while listening to music. The work highlighted how much biological we are stimulated by different types of music.

Woods et al [8] studied different perspectives of emotions developed in humans because of music. The poster extended paper throws light on three major perspectives: neurological, psychological and most importantly musicological. The paper emphasizes modern technologies used in music and emotion tagging.

Musicological analysis needs a heavy backing of various data analysis models as seen in S. Shukla et al [6] work. The paper studied different models for emotion analysis based on lyrics of the song. Da-chuan Wei [9] took musical keys for emotion analysis. The model opted here is Hidden Markov Model and classifies songs based on the emotion evoked.

The prominent musicological research specific to Indian music is done in classical music majorly Hindustani classical [10] and Carnatic classical. Mathura et al [11] work focuses on relating certain notes and chords of ragas to emotion they evoke. Koduri et al [12] work is on the same ground but with south Indian or Carnatic classical music. Both of their works have thrown light on the deepest, oldest and well documented form of music in India.

Ujlambkar et al [13] work focuses more on popular Indian music and emotion an individual can relate. Their work is in a much more abstract form of song unlike classical music or ragas. The model developed works by splitting audio and lyrics and then analysis and classifying songs into 5 major categories namely-Sad, Happy, Excited, Romantic and Silent. The work effectively associates lyrics with the musical parameter to the emotion it refers to.

The prominent findings from previous studies is clear that extensive work has been done in relating musical aspects to corresponding emotion, but little is the comparison of the same among different linguistic groups especially in India. The research in this paper revolves around how musical parameters vary for a particular emotion among different dialects.

III. Research Methodology

Song tracks are taken from renowned songs across various popular Indian and Western songs. The work outlines which parameters have been majorly involved in determining the popularity of the songs. The entire methodology can be divided into 3 major sections- A.Prepossessing of data-set, B.Model Development, C.Analysis of Graphs

Finally, complete content and organizational editing before formatting. Please take note of the following items when proofreading spelling and grammar:

A. Preprocessing of data-set

The data curated for the paper is taken from the popular music app named Spotify. Spotify has an inbuilt music sorting criteria [18]. The work takes inspiration from the same parameters and used them for comparing songs [14].

The parameters and their significance are as follows-

Acousticness: A scale from 0.0 to 1.0 that shows whether or not the track is acoustic. The number 1.0 denotes a high level of certainty that the track is acoustic.

Dance ability: Dance ability is a musical concept that refers to how appropriate a track is for dancing depending on a variety of factors such as speed, rhythm stability, beat power, and overall regularity. The least dance able value is 0.0, and the most dance able value is 1.0.

Energy: Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, “death metal”, for example, has a high energy level, whereas a Bach prelude has a low one. Perceptual features contributing include dynamic range, timbre, onset rate, and general entropy.

Instrumentalness: Determines whether or not a track has no vocals. In this case, the sounds “ooh” and” aah” are viewed as instrumental. Tracks like rap or spoken word are simply “vocal.” The higher the instrumental-ism score gets to 1.0, the more often the track is devoid of vocals. Instrumental tracks are represented by values above 0.5, but as the value reaches 1.0, optimism increases.

Liveliness: The participation of an audience in the video is detected. Higher liveliness values indicate a greater likelihood of the track being played live. If the value is greater than 0.8, the track is almost certainly live.

Loudness: Loudness values are summed over the whole track and can be used to compare the relative loudness of different tracks. The main psychological correlate of physical power is loudness, which is the quality of sound (amplitude). Typical values vary from -60 to 0 decibels.

Speechiness: The appearance of spoken words in a track is detected by speechiness. The attribute rating is similar to 1.0 the more purely speech-like the recording is (e.g. talk show, audio book, poetry). Tracks with a value greater than 0.66 are almost always completely made up of spoken language. Tracks with values between 0.33 and 0.66 which include both music and speech in parts or layers, such as rap music. Music and other non-speech-like songs are likely to have values below 0.33.

Valence: A scale ranging from 0.0 to 1.0 that describes how positive a track is musically. Tracks with a high valence sound more optimistic (e.g. joyful, cheerful, euphoric), whereas tracks with a low valence sound gloomier (e.g. tragic, melancholy)

Tempo: The overall estimated tempo of a track in beats per minute (BPM). The tempo is the rhythm or rhythm of a piece in musical terms, and it is taken directly from the average beat length.

After understanding the parameters of the app, two data-sets one of English songs and other for Hindi songs are prepared. These songs are a list of songs popular in India amongst the respective linguistic community. After the songs and its parameters are obtained, a new parameter namely “mood” is added. A new parameter namely “Mood” is computed. Mood depends mainly on valence but also on other previously mentioned parameters.

B. Model Development

The curated data-set of songs are divided into 2 files- popular English songs and popular Hindi songs. Librosa libraries are used for refining the data values and a new data-set is created from the same. Eventually, Exploratory Data Analysis techniques with Linear regression is used to determine the parameters that majorly factor the popularity of a song and factors that don’t. The Linear Regression model helps in finding the one to one corresponds of parameters with the popularity of the song. The sklearn packages are used for developing this model. The flow of the algorithm is as flows: -

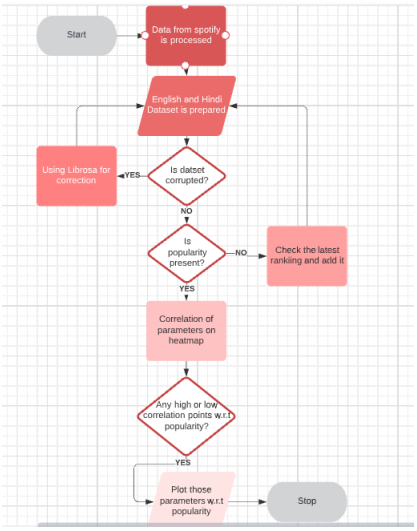

Fig. 1. Flow-Diagram of the model

The training data-set model is then used to develop a correlation matrix or heat-map to find how all parameters vary with each other. The test is later used the validate the same. The following graphs is the result of correlation of parameters: -

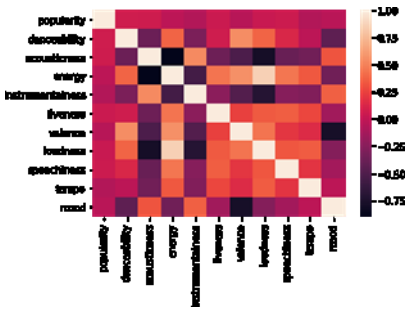

Fig. 2. Heat-map of parameters of English song

C. Analysis of Graphs

The last part of methodology focuses on analysis. After the train and testing completes, the model clearly shows maximum songs with less positive tone. Making it imperative that both the data-set has low positive or sad songs as the general trend.

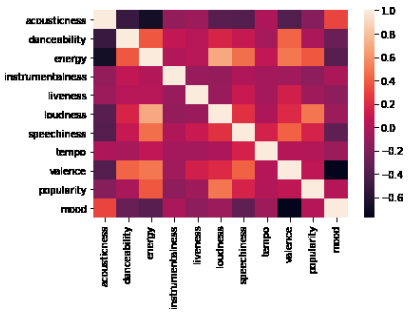

Fig. 3. Heat-map of parameters of Hindi songs

Further analysis is on the musical parameters namely acousticness, liveliness, dance ability, tempo, instrumentalness, loudness, speechiness and energy. The following parameters are plotted with respect to popularity for both the data-set to check how they fare against each other. In the following graphs, parameters are plotted alongside X-axis while the popularity on Y-axis.

The scatter plots of both the sets show some striking differences in fields of loudness, instrumentalness and acousticness. The graphs of each of these parameters are as follows: -

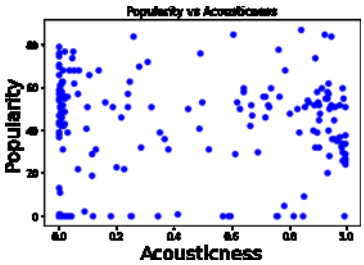

Fig. 4. Popularity vs Acousticness of English songs

Fig.3 describes the relation of English song's acousticness with respect to that of its popularity. The plot shows English songs are predominantly of 2 categories-high acoustic and low acoustic songs. Popularity doesn't hold a significant relation with increase or decrease of popularity.

Fig.4 shows clearly there is no apparent linear relation with acousticness of Hindi songs. Moreover, there is a more uniform spread of acousticness of the songs meaning a wider range of acousticness is used in Hindi songs.

Fig.5 and Fig.6 show the relation of loudness with popularity. The pattern for both English and Hindi songs are similar, popularity increases with increase in the loudness of the song. The English songs however show a gradual rise in popularity with loudness. Whereas, Hindi songs show much steeper rise in popularity with increase in loudness. Another striking difference between English and Hindi songs shows how English songs vary across the entire loudness scale while Hindi songs vary mostly on the higher end.

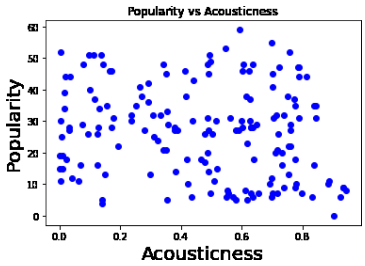

Fig. 5. Popularity vs Acousticness of Hindi Songs songs

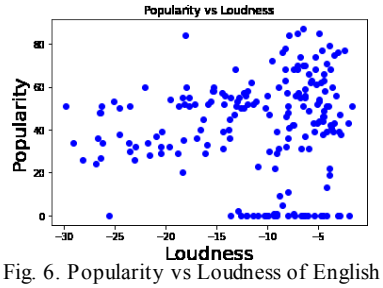

Fig. 6. Popularity vs Loudness of English songs

Thus, Hindi songs are generally louder and higher on music than English songs. Fig.7 shows the relation between instrumentalness and popularity. Instrumentalness defines the variety and frequency of instruments used in the songs. English songs show a very low number of songs that have high instrumentalness. Most of the songs have low instrumentalness Moreover, none of the highly popular songs have high instrumentalness.

Fig. 7. Popularity vs Loudness of Hindi Songs songs

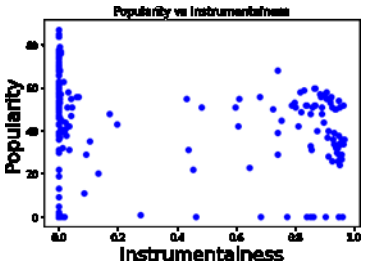

Fig. 8. Popularity vs Instrumentalness of English Songs

Fig. 8. shows the relation between the popularity and instrumentalness of Hindi songs. The plot shows Hindi songs have low instrumentalness parameter. Moreover, popularity apparently have very little correlation with instrumentalness of Hindi songs, even songs with slightly higher instrumentalness are not highly popular.

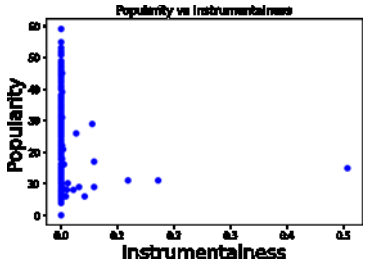

Fig. 9. Popularity vs Instrumentalness of Hindi Songs

IV. CONCLUSION

The work covered focuses on understanding the current trend of songs of a particular language. These songs are current top charters hence, there is little to no doubt that knowing the musical parameters will give an idea to artists on what all parameters must be high and low if the song has to become popular.

We can conclude from the study that English songs in particular need low instrumentalness and acousticness and must focus more on dance ability. So the song in spite being less instrumental, the rhythm and uniform use of those instruments make the song peaceful to hear. For sad songs though, there are significant sections of songs purely instrumental keeping the dance ability parameter comparatively lower. Speechiness in both cases are generally low with some songs being a little more speech.

Hindi songs are low with instrumentality and speechiness. Dance ability and acousticness are parameters not much focused on. The following analysis gives ideas to artists on making songs less instrumental and focus on getting catchy acoustics. It is common that Hindi songs are highly dance-able because they are created for Bollywood movies which undoubtedly have dance sequences. Hence, the rhythm and stability of new Hindi songs must preferably be high.

The novel study effectively understands the trends and gives better ideas on how to target the audience of a particular language. Otherwise, creating a song for example Hindi happy song with very low dance ability is quite likely to be less popular. Similarly, an English sad song with very high speechiness will not be received well by an audience who understand English songs. In comparison to the existing models, the following model takes into consideration all possible features of a song and compare them rather than focusing only on selective parameters like lyrics or pitch as given in [11] and [13]. The use of information for open-source and reliable database of Spotify helps give accurate result and sufficient volume of data samples for developing the model.

The major drawback of this study is the criteria of non-musical. Songs depend on a lot of non-musical parameters like the popularity of an artist, marketing of the song, album or movie that uses it as background score and many more. These nonmusical parameters can make it difficult to get a high percentage of accuracy accurate analysis of a song, maybe in future a separate model or parameter can help to sum these non-musical parameters well.

REFERENCES

[1] G. Eason, B. Noble, and I.N. Sneddon, "On certain integrals of James Nissen, 'India: Food of Soul,' *World of Music article*, 2017.[online]
Available:https://www.guidetotheworldofmusic.com/peopleandplaces/the - music-of-india-the-food-of-the-soul

[2] J. Stephen Downie, 'Music Information Retrieval,' Annual *Review of Information Science and Technology*, vol. 37, pp. 295-340, Jan. 2003.

[3] M. Schedl, E. Gómez and J. Urbano, 'Music Information Retrieval: Recent Developments and Applications,' *Foundations and TrendsR in Information Retrieval*, vol. 8, no. 2-3, pp. 127–261, Sep. 2014.

[4] Da-chuan Wei, 'Recognition algorithm of musical chord based on keynote-dependent HMM,' *Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE)*, pp. 2504-2507, Dec. 2011.

[5] Heng-Tze Cheng, Yi-Hsuan Yang, Yu-Ching Lin, I-Bin Liao and H. H. Chen, 'Automatic chord recognition for music classification and retrieval,' *2008 IEEE International Conference on Multimedia and Expo*, pp. 1505-1508, Jun-Apr. 2008.

[6] S. Shukla, P. Khanna and K. K. Agrawal, 'Review on sentiment analysis on music,' *2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS)*, pp. 777-780, Dec. 2017.

[7] S. Koelsch, 'Towards a neural basis of music-evoked emotions,' *Trends in cognitive sciences*, vol. 14, no. 3, pp. 131–137, Mar. 2010.

[8] Pamela A. Wood, Sudhanshu K. Semwal, 'An algorithmic approach to music retrieval by emotion based on feature data,' *Future Technologies Conference (FTC)*, pp. 140–144, Dec. 2016.

[9] Da-chuan Wei, 'Recognition algorithm of musical chord based on keynote-dependent HMM,' *Proceedings 2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE)*, pp. 2504-2507, Dec. 2011.

[10] Mathur A, Vijayakumar SH, Chakrabarti B and Singh NC, 'Emotional responses to Hindustani raga music: the role of musical structure,' *Front. Psychol*, Apr. 2015.

[11] R. Malheiro, R. Panda, P. Gomes and R. P. Paiva, 'Emotionally-Relevant Features for Classification and Regression of Music Lyrics,' *IEEE Transactions on Affective Computing*, vol. 9, pp. 240-254, Apr-Jun. 2018.

[12] Koduri, Gopala Indurkhya, Bipin, 'A behavioral study of emotions in south indian classical music and its implications in music recommendation systems,' *Trends in cognitive sciences*, pp. 55–60, Oct. 2010.

[13] Aniruddha M. Ujlambkar and Vahida Z. Attar, 'Mood Classification of Indian popular music,' '*CUBE '12: Proceedings of the CUBE International Information Technology Conference*, pp. 278–283, Sept. 2012.

[14] Developers Spotify,' *Web API reference*[online] Available:https://developer.spotify.com/documentation/web-api/reference

[15] A. S. Bhat, V. S. Amith, N. S. Prasad and D. M. Mohan, 'An Efficient Classification Algorithm for Music Mood Detection in Western and Hindi Music Using Audio Feature Extraction,' *2014 Fifth International Conference on Signal and Image Processing*, pp. 359-364, Jan. 2014.

[16] H. Sharma, S. Gupta, Y. Sharma and A. Purwar, '"A New Model for Emotion Prediction in Music,' *2020 6th International Conference on Signal Processing and Communication (ICSC)*, pp. 156-161, Mar. 2020.

[17] Y. An, S. Sun and S. Wang, 'Naive Bayes classifiers for music emotion classification based on lyrics,' *017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS)*, pp. 635-638, May. 2017.

[18] D. R. Ignatius Moses Setiadi et al., "Effect of Feature Selection on The Accuracy of Music Genre Classification using SVM Classifier," *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)*, 2020, pp. 7-11.

[19] G. A. Sandag and A. M. Manueke, "Predictive Models for Popularity of Solo and Group Singers in Spotify Using Decision Tree," *2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS)*, 2020, pp. 1-5.

[20] Vijayakumar, T. ”Comparative study of capsule neural network invarious applications.”*Journal of Artificial Intelligence 1*, no. 01 (2019):19-27.