Priyadarshini College of Engineering Nagpur

Department of Computer Technology Session

2021-22

Project Defining

Seminar Synopsis

Report

On

—Prediction of Song Mood Through Lyrics‖

Project Guide

Prof. Ms. Priyanka Padmane

Projectees

7th Sem C.T Department

273 – Krishi Agrahari

259 – Rajsi Kesharwani

203 – Kirti Mohitkar

266 – Nikhil Kamale

234 – Shazia Khan

VISION OF INSTITUTE

To become one of the India’s leading engineering institutes in both education & research. We are committed to provide quality & state of the art technical education to our student, so that they become technologically superior & in turn contribute for creating a great society.

MISSION OF INSTITUTE

We commit ourselves to the pursuit of excellence in technical education & promise to uphold the spirit the spirit of professionalism to serve humanity.

VISION OF DEPARTMENT

To excel in creating outstanding academicians and technocrats for valuable contribution to the society and computer engineering field worldwide as per the ever changing needs.

MISSION OF DEPARTMENT

To ensure technical proficiency to meet industrial needs.

To impart ethical and value based education for social cause.

To encourage the students for carrying out research activities and lifelong learning.

PRIYADARSHINI COLLGE OF ENGINEERING

Department of Computer Technology

NAGPUR-440019 **CERTIFICATE**

This is to certify that project report entitled - "**Project Title**" is a bonafide work done by the student **PROJECTEES-1, PROJECTEES-2, PROJECTEES-3, PROJECTEES-4 and PROJECTEES-5.** The project report is submitted to **Rashtrasant Tukadoji Maharaj Nagpur University, Nagpur** in partial fulfillment of the requirements for the degree of **Bachelor of Engineering in Computer Technology.**

Session 2021-2022

GUIDED BY- Guide Name

Dr. N.M. Thakare
H.O.D, Computer Technology, P.C.E.

Dr. S.A. Dhale
Principal, P.C.E.

PRIYADARSHINI COLLGE OF ENGINEERING,

NAGPUR-440019 (INDIA)

Department of Computer Technology

DECLARATION

We, the undersigned, declare that the project entitled "**Project Title**", being submitted in partial fulfillment for the award of Degree in Computer Technology, affiliated to RASHTRASANT TUKDOJI MAHARAJ NAGPUR UNIVERSITY, is the work carried out by us.

Place: Nagpur **PROJECTEES: -**

Date: **PROJECTEES-1**

PROJECTEES-2

PROJECTEES-3

PROJECTEES-4

ACKNOWLEDGEMENT

It is our pleasure to acknowledge our sincere thanks with a deep sense of gratitude towards our project guide "**Guide Name**", Assistant Professor/ Associate Professor, Computer Technology Department for her continuous knowledge and support in conducting this dissertation work. She has a whole heartedly helped us in this endeavor at all stages of this work.

We are thankful to **Dr. S. A. Dhale**, Principal, Priyadarshini College of Engineering, Nagpur, for providing the facilities at the institute.

We thank **Dr. N.M. Thakare**, Professor and Head of Computer Technology Department of Priyadarshini College of Engineering, Nagpur.

We herewith express our immense thanks to "Guide Name", Assistant Professor/ Associate Professor, Assistance Professor, Computer Technology Department, Priyadarshini College of Engineering, Nagpur for giving us suggestions and co-ordination with us from time to time as the project in charge.

We also take the opportunity to thank all, who have directly or indirectly extended help and encouragement in executing this project.

PROJECTEE:

PROJECTEES-1

PROJECTEES-2

PROJECTEES-3

PROJECTEES-4

PROJECTEES-5

INDEX

Contents

1. Abstract: 8
2. Literature survey 9
3. Motivation 13
4. Aim of problem statement 14
5. Objectives 15
6. Project Methodology 16
   6.1 Training dataset preparation 16
   6.2 Test Dataset Preparation 17
   6.3 Data Preprocessing 17
   6.4 Feature Engineering 18
   6.5 Model Selection 18
7. Expected Result: 21
8. Conclusion 23
9. Future Work 24

Date: 13/11/2021

Abstract:

Human Beings are vexed up with a variety of problems such as anxiety, stress, work tensions, and emotional outbreaks. This may be caused due to Job, family problems,responsibilities, and pressure from friends. During the battle with an emotional crisis, a human desperately searches for means to riddance the problem. Music, being one of the most popular means of entertainment, can help during such situations. It provides a wayto express our feelings and to enhance our state of mind. The core part of music is the mood. Every situation we go through has a spirit associatedwith it. There are many songs that are written on emotions. Many public places, such as restaurants, tourist places, and cultural events, have a theme song in the background. This enhances the mood of the customers. In this regard, we are performing a mood classification of songs using lyrics alone. We are implementing Decision Tree and Random Forest models for the problem. The exploratory outcomes through training and testing the model show that music related to —happy‖ and —sad‖ states of mind can be anticipated with sensible accuracy dependent on features extracted from tune verses. Anxiety, worry, job conflicts, and emotional outbursts are just a few of the issues that plague humans. This might be due to a job, family issues, obligations, or peer pressure. During a fight with an emotional crisis, a person urgently seeks solutions to the situation. Music, being one of the most popular forms of entertainment, may be beneficial in such situations. It allows us to express our emotions and improve our mental condition. The mood is the most important aspect of music. Every scenario we face is accompanied by a spirit. Many songs are composed with emotions in mind. Many public venues, such as restaurants, tourist attractions, and cultural events, have a backdrop theme tune. This improves the clients' mood. In this regard, we are performing a mood categorization of songs based solely on words. For the problem, we are using Decision Tree and Random Forest models. The exploratory results from training and testing the model suggest that music associated with joyful and sad states of mind may be predicted with reasonable accuracy based on characteristics derived from song verses.

Literature survey

An overview of the literature survey of the relevant work done by the researchers. Many existing techniques have studied by the researchers on lyric mood prediction problems, few of them are discussed below. In Mood prediction of music from lyrics can have a good range of applications in modern society and daily lives. For example, selecting music for public corporations like hospitals or restaurants to improve the mental well-being of people, patients, and customers, respectively. The music recommendation system built upon on Machine Learning algorithms such as Decision Tree and Random Forest classifiers is trained to predict the mood of songs based on the sentiment of the lyrics alone.

The experimental results obtained by training and testing the model show that music corresponding to happy and sad moods can be predicted with reasonable accuracy based on features extracted from song lyrics. This paper lists out the performances of music genre and mood classification using only lyric features. In this research study, the Part-of-Speech (POS) feature is utilized for the classification of a set of 600 songs. Ten music genre and mood categories were selected respectively supported by an overview from the literature. Experiments show that accuracies for mood categories outperform genres. In this research, the music suggestion framework based upon algorithms such as Decision Tree and Random Forest classifiers trained to foresee the state of mind of melodies dependent on the lyrics alone. We have implemented the algorithms listed bythis work to get the desired results. This study also suggests that using ML, predictions, and recommendations can be made accurately. This study also suggests that using ML, predictions, and recommendations can be made accurately,
A replacement method is proposed to create an outsized ground truth set of 5,585 songs and 18 mood categories supported social tags so on reflect a practical, user-centered perspective.

A complete set of lyric features and representation models were investigated. The most straightforward performing lyric feature set was also compared to the number one audio-based system. By observing statistics of various mood dimensions, we examine to what extent the linguistic a neighborhood of music reveals adequate information for assigning a mood category and which aspects of mood are often classified best. The word-oriented metrics provide a valuable source of data for automatic mood classification of music, supported lyrics only. Properties like term frequencies and

TF- IDF values are used in various mood classes. These metrics are incorporated during a machine learning classifier setup. Predictions on the valence, tension, and combinationsof aspects cause similar performance. Extensive online music databases have recently been created by vendors, but they typically lack content-based retrieval methods. Human experts say there are several thousand songs categorized into 183 moods. During this paper, machine learning techniques are used rather than human experts to extract emotions in Music. The classification is predicated on a psychological model of emotion that's extended to 23 specific emotion categories.

Mining lyrics focused during this paper is one aspect of research that mixes different classifiers of musical emotion like acoustics and lyrical text. This paper suggests the research work done on music automation. Employing an assortment of verses and comparing client labeled states of mind, we manufacture classifiers that order verses of tunes into dispositions. By comparing different mood methods and techniques, various characteristics of emotions are determined. This has uncovered what part of the data is to be used for further classification. The outcomes of this study denote that word- arrangements and their measurements play an important role in classifying the problem. Feature extraction of the lyric is an essential aspect of this study. The full verse is divided into several tokens, and a feature extraction method is performed. According to the accuracies of different feature extraction methods, one among many methods is selected to input the data. There are various programmes that give facilities and services for music playlist generation or playing a certain song, and all manual effort is included in this process. There are now a variety of strategies and approaches that have been proposed and developed to characterise human emotional states of behaviour. The proposed methodologies, such as Viola and Jones', have only addressed a subset of the basic emotions. Several scientific publications that provide a summary of the concept are:

[1]According to the authors of this research, music plays a vital function in human existence and inside current technological technology. Typically, the user must actively go through the playlist of music to select one. In this paper, we propose an efficient and accurate approach for generating a playlist based on the user's current mood and behaviour. Existing approaches for automating the playlist building process are computationally sluggish, less precise, and may necessitate the use of extra gear such as EEG or sensors. Speech is that the most ancient and natural way of expressing feelings, emotions and mood and its and its processing requires high computational, time, and cost. This system supported

real-time extraction of facial expressions also as extracting audio features from songs to classify into a selected emotion which will generate a playlist automatically such the computation cost is comparatively low.

[2] This study presents an intelligent agent that organises a music collection based on the emotions communicated by each song and then recommends a suitable playlist to the user based on his or her current mood. The user's local music collection is first grouped based on the emotion conveyed by the song, i.e. the mood of the song. This is frequently assessed by taking into account the song's words as well as the music. When the user wants to acquire a mood-based playlist, the user snaps a picture of themselves at the time. This photograph is subjected to face detection and emotion identification methods, which recognise the user's emotion. The music that best suits this feeling is then offered as a playlist to the user..

[3] According to the authors of this article, people are becoming increasingly stressed as a result of the terrible economy, excessive living expenditures, and so on. Taking note of music may be a significant action that aids in stress reduction. However, it will be ineffective if the music does not match the listener's current emotional state. Furthermore, there is no music player that can select songs based on the user's emotions. To address this issue, this study presents an emotion-based music player that may recommend songs depending on the user's emotions: sad, joyful, neutral, and furious. The device gets the user's pulse or a face picture via a sensitive band or mobile camera. It then uses the classification method to spot the user's emotion. This paper presents 2 sorts of the classification method; the guts rate-based and therefore the facial image-based methods. Then, the appliance returns songs which have an equivalent mood because the user's emotion. The experimental results show that the proposed approach is in a position to exactly classify the happy emotion because the guts rate range of this emotion is wide.

[4] According to the authors, digital audio is simple to record, play, process, and maintain. Because of its pervasiveness, gadgets for handling it are inexpensive, allowing more individuals to record and play music and voice. Furthermore, the web has made it easier to access recorded audio. As a result, the amount of recorded music that people own has rapidly expanded. The majority of today's audio players compress audio files and store them in internal memory. Because storage prices have constantly reduced, the amount of music that will be stored has expanded significantly. If each song is saved in compressed format and contains 5 Mbytes, a player with 16 Gbytes of memory may carry around 3,200 songs. Effectively organizing such large volumes of music is difficult. People often listen repeatedly to alittle number of favorite songs, while others remain un justifiably neglected. We've developed Affection, an efficient system for managing music collections. Affection groups pieces of music that

convey similar emotions and labels each group with a corresponding icon. These icons let listeners easily select music consistent with its emotional Content. Experiments have demonstrated Affection' effectiveness.

Motivation

Predicting a small set of emotions in music from the audio content is technically feasible. However the difficulty comes when introducing more subjectivity together with more complex semantic descriptions. For example, what are the differences between sad and melancholic. What is the overlap between both concepts? Would we all agree on this? We will probably not. Moreover how can our system be aware of the personal history of users, their social or cultural contexts, and their current status? We should also take care about the drawbacks of such systems. The marketing and social control issues can be frightening. However all the promising applications in everyday life, and especially in art and therapy are definitely strong arguments to continue these investigations. Detecting automatically emotion in music is at its early stage but we can expect many improvements and exciting applications in the future.

Aim of problem statement

With the fast development of computerized music libraries just as progressions in innovation, music characterization and proposal has expanded prevalence in the music business and among the audience members. There are many applications which utilize the AI techniques in their models. They are used to group the music based on Artist, Genre, Instruments used, Title, and Year of release, Artist similitude, and type. Modern studies suggest that human beings use music as a means of riddance of their tensions and stress. The Web platform is an ocean of musical content, so there is a difficulty to the people to group according to their need. Hence by using MLalgorithms, this can be automated and can be done quickly. As the web is a vast platform, the successful preparation of a model using ML algorithms to classify the mood of a song using verses can be done. Digital music has been expanded beyond our imagination. As it increases, it becomes hard to collect the songs we desire. So we use cutting edge technologies to provide an option for clients to get what they need . For example, clients might want to have the opportunity to look for tunes dependent on different properties like the title of the track, its artist, genre of the song, and the year it was released. Sometimes, the client may even want to search through recommendation systems present on the web. Few individuals might wish to distinguish their playlist according to the emotion of the lyrics. Here we explore the chance of allocating such data without client collaboration. Previously, It was a manual process which involved listener in shuffling the playlist which consumed a lot of time manually. Our paper addresses not only the speed but also the efficiency and involvement of the user. Thus, massive trials have been led in the field ofthe music investigation dependent on its verses and emotion. While Data mining showed a few promising results, yet it was not consistent. So we ML provided a way to achieve this. With the rapid expansion of digital music libraries, as well as advancements in innovation, music characterisation and suggestion has grown in popularity in the music industry and among audience members. Many applications include AI approaches into their models. They are used to categorise music based on the following criteria: artist, genre, instruments utilised, title, and year of release, artist similitude, and type. According to recent research, humans utilise music to relieve their worries and stress. Because the Web platform is a sea of musical information, it is difficult for users to categorise it based on their needs. As a result, utilising MLalgorithms, this may be automated and completed swiftly.

Objectives

Music has long been an effective way to communicate to the masses, and lyrics have played a massive role in delivering this communication. Yet the opportunity for research on the role lyrics play in well-being is vastly underutilized. This paper is an exploration of the relationship between lyrics and positive psychology. I will discuss a brief origin of lyrics, examine the body of literature on lyrics as well as its gaps, and finally suggest potential application of lyrics to increasing various aspects of well-being. We are only beginning to have the language to discuss the positive and negative effects of lyrics. The results of this exploration indicate that lyrics have the potential to increase two of the five elements of well-being in the PERMA model, positive emotions and meaning. It is suggested that you can increase well-being by mindfully listening to meaning- filled lyrics bolstered by music's ability to influence emotion.

Project Methodology

6.1 Training dataset preparation

The dataset that we start with is a 10000 Song subset of the Million SongDataset. Now we dothe following :

- Store the dataset into a Pandas Dataframe with features File Name,Artist Name
- Using the Artist Name and Song Title, We are fetching lyrics for allthe songs using PyLyrics

package, which uses LyricWikia.com API to get lyrics for songs

- We are creating our model using english lyrics. So all the song lyrics containing any otherlanguage
- Now we will use Last.FM API to extract Tags for the remain-ing 3000 songs in our

dataset. Tags can be based on Genre, Mood, Artist Type etc. For getting the song tags we request to "http://ws.audioscrobbler.com/2.0/" as endpoint with the parametersas follows :

– method = track.getTopTags

– api.getKeys = 0f6916aff634cb3e768baa9d5ee89341

– artist = artists fetched from our csv file

– tracks = tracks

● In the paper Lyric Text Mining in Music Mood Classification, Hu et.al, Last.FM tags are being grouped into 18 categories according to different human moods. We have taken 10 groups from it and distributed theminto our Mood Categories - Happy, Sad, Angry, Relax.

• Happy Tags : cheerful, cheer up, festive, jolly, jovial, merry, cheer, cheering, cheery, get happy, rejoice, songs that are cheerful, sunny,happy, happiness, happy songs, happy music, glad, mood: happy, upbeat, gleeful, high spirits, zest, enthusiastic, buoyancy, elation, mood: upbeat,excitement, exciting, exhilarating, thrill, ardor, stim-ulating, thrilling, titillating

• Sad Tags : sad, sadness, unhappy, melancholic, melancholy, feel- ing sad, mood: sad - slightly, sad song, depressed, blue, dark, de- pressive, dreary, gloom, darkness, depress, depression, depressing, gloomy, anger, angry, choleric, fury, outraged, rage, angry music,grief, heartbreak, mournful, sorrow, sorry, doleful, heartache, heartbreaking, heartsick, lachrymose, mourning, plaintive, regret, sorrowful

- Angry Tags : anger, angry, choleric, fury, outraged, rage, angry music,aggression, aggressive, angst, anxiety, anxious, jumpy, nervous, angsty,pessimism, cynical, pessimistic, weltschmerz, cynical/sarcastic

- Relaxed Tags : calm, comfort, quiet, serene, mellow, chill out, calm down, calming, chillout, comforting, content, cool down, mellow mu- sic, mellow rock, peace of mind, quietness, relaxation, serenity, so- lace, soothe, soothing, still, tranquil, tranquility, tranquility,brooding, contemplative, meditative, reflective, broody, pensive, pondering, wist-ful,desire, hope, hopeful

By correlating the Tags that we found from Last.FM and the tag groups generated by us, we are creating the Class Labels for Moods in our Dataset. From the paper Multimodal Music Mood Classification by Fusion of Audio and Lyrics, Hao et.al , we get to know about 777 other songs, al- ready categorized into Happy, Sad, Angry, Relaxed. We append this dataset with our previous dataset for our final training dataset.

6.2 Test Dataset Preparation

- We manually collected over 250 Hindi song lyrics and stored into dataframe
- Using Google Translate API, we auto-translated them into Englishlyrics and manually labelled them for performance checking.
- As our test dataset is labelled manually and training dataset is auto- labelled from the tags of Last.FM and Hao’s paper, we add a fractionof translated-to-english lyrics to our training dataset to reduce the bias of the testing dataset.

6.3 Data Preprocessing

We have lyrics in our data. To Preprocess the lyrics column we did following preprocessing steps on both training and test datasets.

- Tokenization : Taking a text or set of text and breaking it up into its individual words
- Stop-word Removal : Stop words such as *the, a, an, in* are removed from lyrics
- Punctuation Removal : Removes punctuation from lyrics
- Stemming : Reducing inflected (or sometimes derived) words to theirword stem, base or root form
- Lemmatization : Process of grouping together the inflected forms of a word so they can be analyzed as a single item, identified by the word'slemma, or dictionary form

6.4 Feature Engineering

We created training, validation and testing datasets for these 3 models usingnatural language processing.

- CountVectorizer
- TfIdfVectorizer
- NGram Vector Model
- Word2vec Embedding Vector

6.5 Model Selection

The following classifiers are used for model preparation and testing :

- Random Forest
- Multinomial Naive Baye
- Logistic Regression
- Ensemble - Bagging
- Ensemble - Boosting
- Support Vector Machines
- Convolutional Neural Network

Final Training Dataset : It contains 1482 english song lyrics prepro- cessed and feature-engineered using the previous steps

Final Validation Dataset : It contains fraction of training dataset containing english song lyrics preprocessed and feature-engineered using the previous steps

Final Testing Dataset : It contains 243 Hindi Bollywood translated- to-english song lyrics preprocessed and feature- engineered using the same previous steps.

Music has historically been an efficient means of communicating with the public, and lyrics have played a significant role in this communication. However, the possibility for study on the impact of lyrics in happiness is largely neglected. This research investigates the connection between lyrics and positive psychology. I will provide a brief history of lyrics, evaluate the corpus of research on lyrics and its deficiencies, and lastly suggest prospective applications of lyrics to improve various elements of well-being. We are only now developing the vocabulary to discuss the good and negative consequences of songs. According to the findings of this study, lyrics have the capacity to boost two of the five dimensions of well-being in the PERMA model: good emotions and meaning. It is stated that you can improve your well-being by deliberately listening to lyrics with meaning, which is aided by music's power to alter emotion.

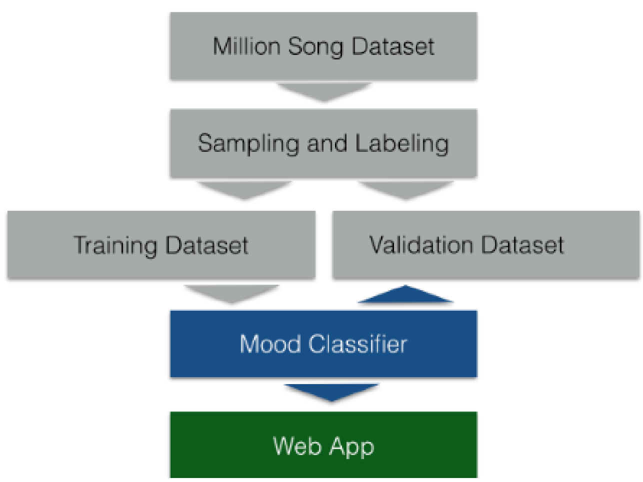


Fig 1 Flow Chart of system

We already have the data in the proper format because we are utilising a csv file. Feature engineering, also known as feature selection, assists the model in achieving the required performance. When developing a machine learning model, it is critical to pick a strong collection of characteristics that will assist us in more accurately predicting the proper conclusion. Filter methods, wrapper methods, embedding methods, and hybrid methods are the four types of feature selection strategies used in machine learning. We employed Random Forest Importance, which is a strategy in the embedded approach. We found the top 20 characteristics with the greatest influence on the output label using random forest significance. We also trained another set of classifiers with all of the data save the name and length of the audio and compared the results. We next divide the dataset in an 80:20 ratio, with 80 percent going to the training set and the remaining 20 percent going to the test set. The data is then scaled. We scale the test and training data independently because we want our test data to be entirely new to the model and free of bias. Scaling the training data yields scaling parameters like as mean and variance, which are then utilised to scale the test data. The fundamental purpose for scaling the data is to avoid biassing the model to a certain aspect of the dataset. This data is now being used to train our classifiers.

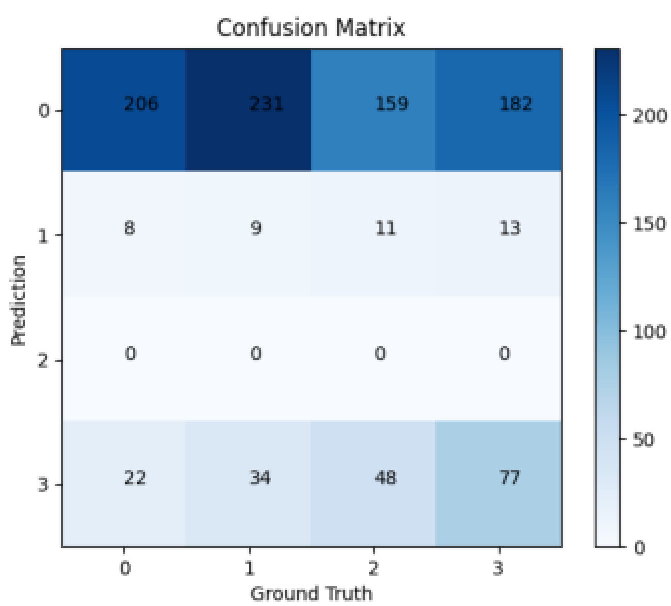
Expected Result:

The programming language used to implement our paper is ―Python 3.7. Various packages like Scikit-learn, Pandas, Matplotlib, and other necessary packages have been imported for our code. Dataset considered for our paper is the lyric dataset. After training and testing our ―Mood Prediction‖ model using Decision Tree and Random Forest Algorithms, the following results have been obtained.in Table 1. The findings mostly suggest that even a for all kinds of reasons naive Bayes model that used mood classifier - based lyrics may accurately predict the pretty positive class (glad), which may be beneficial for filtering a large music collection for cheerful music while reducing the kind false positives, which is fairly significant. A music collection that has been specifically filtered in this manner would be used as input to genre categorization, allowing music to be specifically filtered according to various tastes, or so they believed. Future work will entail integrating more ROC curves of multiple lyrics classification methods to the moods identification web service, which will also be verified using 10-fold cross-validation here on lyrics training dataset, which is thought to consist of 1,000 random songs. The real absolutely positive rate was calculated using songs that were correctly labelled as cheerful, whereas the essentially false sort of positive rate was calculated using songs that were wrongly classed as joyous. In a considerable way, our study was conducted predominantly with Python 3.7 as the programming language. demonstrating how the findings kind of imply that a naive Bayes model applied to mood categorisation based lyrics may effectively forecast the very positive class (happy), that might primarily be beneficial for filtering a large music selection for cheerful lyrics while reducing false positives, or so they thought. Several packages, including Scikit-learn, Pandas, Matplotlib, and others, have been heavily used in our code. For our research, we focused solely on the lyric dataset, which is unquestionably substantial. We obtained the following outcomes in a huge manner after utilising Decision Tree and certainly Random Forests Algorithms to train an algorithm our Mood Prediction model. The count vectorization approach as well as the TF-IDF vectorization model were both used extensively. This machine learning technique was used to categorise lyric mood as "sad" or "glad." The findings suggest that, for the most part, the Random Forest algorithm is capable of accurately predicting lyric mood. The TF-IDF encoding and beautiful Random forest, in example, have a general accuracy of 72.68 percent, or so they believed.

We used the vectorization model of count and TF-IDF. We used this machine learning algorithm and predicted lyric mood as binary class value "sad" or "happy". The result shows that lyric mood prediction is efficient using the Random Forest algorithm. In specific, the TF-IDF vectorization and Random forest achieve 72.68% accuracy.

Vectorization	Algorithm	Accuracy (%)
Count	Decision Tree	68.04
	Random Forest	69.07
TF-IDF	Decision Tree	67.52
	Random Forest	72.68

Table 1: Experimental Results of proposed system



Conclusion

In recent days music plays an important role in human entertainment, by listening music the listener get some relaxation and refreshment in their busy life. So many music applications are also available online to recommend a song based on the user's mood. In this paper, the major research area was focused on song lyrics to predict the mood of a song, then that song is recommended to the listener. We handled the challenge of identifying musical mood in this study by evaluating lyrics and acoustic data using supervised learning methods and reasoning. Deep learning system implementations have been proposed in combination with several data representations based on the natural language processing and digitizer processing techniques. The procedure was finished with the training and assessment of the three suggested systems—lyrics only, audio only, and multi-modal. The experimental approach validates the basic hypothesis: multi-modal systems outperform uni-modal systems. In the case of recognising the mood created by music, that both lyrics and indeed the audio include relevant information for building deep learning models. The results demonstrate some emotion uniformity in playlists, implying that emotion identification generates useful information for developing Recommender Systems. The short dataset size limits the model's training. With more datasets, more complicated and successful categorization algorithms would be conceivable. This method literally is commonly used to filter a huge music collection for happy music with a low very false generally positive rate, so this method really is commonly used to filter a huge music collection for happy music with a very low very false positive rate in a subtle way. In the future, our classifier will really be developed on a web platform to for all intents and purposes include a wider range of Music Information databases. The TF-IDF feature extraction along with the Random Forest algorithm has shown a better accuracy of classifying the mood based on the lyrics. By using our research study and paper work, a wide range of music libraries can be grouped or clustered according to the predicted mood. As per our researchand analysis of the results, „Happy" and „sad" moods can be differentiated accurately. This work is often useful to filter an outsized music library for happy music with a coffee false positive rate. The future work is to expand our classifier on a web platform to cover a wide area of Music Information databases.

Future Work

Future work can indeed be accomplished in a variety of ways. The first, and most importantly, direction concerns the amount of the dataset required for the effective creation of a strong deep learning system. Because the data we utilised in this study was restricted, future work would entail designing a system that executes unsupervised learn from unlabeled data, which is abundant. A better alternative for improvements would be to combine vast volumes of unlabeled and modest amounts of labelled data to create systems that include semi-supervised and/or ego learning techniques. In terms of data, using data with lyrics and music aligned will bring significant value and robustness. r. Future study might also rely on a database containing labels reflecting the degree of ambiguity of a track's mood, since we know that listener variability can be large in some circumstances. Such databases would be especially useful for delving deeper into musical emotion. Temporarily localised labels in sufficient quantities may also be of importance. Unsupervised pretraining of deep learning models might be used in future studies, as unlabeled data is easier to locate in large quantities.

References

1. Tsunoo, E. &Tzanetakis, George & Ono, Nobutaka&Sagayama, Shigeki. (2009). ―Audio genre classification using percussive pattern clustering combined with timbral features‖. 382 - 385. 10.1109/ICME.2009.5202514.
2. Chen, J. H., & Asch, S. M. (2017).―Machine learning and prediction in medicine—beyond the peak of inflated expectations‖. The New England journal of medicine, 376(26), 2507.
3. Damaschk, M., Dönicke, T., & Lux, F. (2019). ―Multiclass Text Classification on Unbalanced, Sparse and Noisy Data‖. In Proceedings of the First NLPL Workshop on Deep Learning for Natural Language Processing (pp. 58-65).Raschka, Sebastian. "MusicMood: Predicting the mood of music from song lyrics using machine learning." arXiv preprint arXiv:1611.00138 (2016).
4. Teh Chao Ying, S. Doraisamy and Lili Nurliyana Abdullah, "Genre and mood classification using lyric features," 2012 International Conference on Information Retrieval & Knowledge Management, Kuala Lumpur, (2012), pp. 260-263.
5. Amiripalli, S. S., &Bobba, V. (2019).―Impact of trimet graph optimization topology on scalable networks‖. Journal of Intelligent & Fuzzy Systems, 36(3), 2431-2442.
6. S. S. Amiripalli, V. Bobba, ―A Fibonacci based TGO methodology for survivability in ZigBee topologies‖. INTERNATIONAL JOURNAL OF SCIENTIFIC &TECHNOLOGY RESEARCH, 9(2), pp. 878-881. (2020).
7. Amiripalli, S. S., Kumar, A. K., &Tulasi, B. (2016).―Introduction to TRIMET along with its properties and scope‖. In AIP Conference Proceedings (Vol. 1705, No. 1, p. 020032). AIP Publishing LLC.
8. Amiripalli, S. S., Kollu, V. V. R., Jaidhan, B. J., SrinivasaChakravarthi, L., & Raju, V. A. (2020).―Performance improvement model for airlines connectivity system using network science‖. International Journal of Advanced Trends in Computer Science and Engineering, 9(1), 789-792. doi:10.30534/ijatcse/2020/113912020.
9. Erion C¸ ano and Maurizio Morisio. Music mood dataset creation based on last.fm tags. In International Conference on Artificial Intelligence and Applications (AIAP), 2017.
10. Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In 22nd ACM International Conference on Knowledge Discovery and Data Mining (KDD), 2016.
11. Michael Fell and Caroline Sporleder. Lyrics-based analysis and classification of music. In 25th International Conference on Computational Linguistics (COLING), 2014.
12. Youngmoo E. Kim et al. Music emotion recognition: A state of the art review. In 11th International Society for Music Information Retrieval Conference (ISMIR), 2010.
13. Owen Craigie Meyers. A mood-based music classification and exploration system. Master's

thesis, Massachusetts Institute of Technology, 2007.

14. Jeffrey Pennington, Richard Socher, and Christopher D. Manning. Glove: Global vectors for word representation. In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014.
15. Ashequl Qadir and Ellen Riloff. Learning emotion indicators from tweets: Hashtags, hashtag patterns, and phrases. In Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014.
16. Hassan Walaa, Medhat Ahmed and Korashy Hoda. Sentiment analysis algorithms and applications: A survey. Ain Shams Engineering Journal, 1994.
17.] Miller, John W, “A Python Client for the Genius.com API.” LyricsGenius, 2020, https://lyricsgenius. readthedocs.io/en/master
18. Richardson, Leonard. “Beautiful Soup Documentation.” Beautiful Soup Documentation - Beautiful Soup 4.9.0 Documentation, 2004, https://www.crummy.com/software/BeautifulSoup/bs4/doc/#
19. Google Colab, Google, https://colab.research.google.com/?utm_source=scs-index

Date: 13/11/2021

Name & Sign of Projectees

273 – Krishi Agrahari

Name & sign of Guide

259 – Rajsi Kesharwani

203 – Kirti Mohitkar

266 – Nikhil Kamale

Name & Sign Of HOD

234 – Shazia Khan