

APPLYING TEXT ANALYTICS ON MUSIC EMOTION RECOGNITION

BY

RAFAYET HOSSAIN

ID: 142-15-3994

MD. RAHMATUL KABIR RASEL SARKER

ID: 142-15-3732

AND

MEHEJABIN MIMO

ID: 142-15-3453

This Report Presented in Partial Fulfillment of the Requirements for the
Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Mr. Ahmed Al Marouf

Lecturer

Department of CSE

Daffodil International University

Co-Supervised By

Mr. Shah Md. Tanvir Siddiquee

Senior Lecturer

Department of CSE

Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

MAY 2018

APPROVAL

This Research based Project titled “**Applying Text Analytics on Music Emotion Recognition**”, submitted by Rafayet Hossain (142-15-3994), Md Rahmatul Kabir Rasel Sarker (142-15-3732) and Mehejabin Mimo (142-15-3453) and to the Department of Computer Science and Engineering, Daffodil International University, has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering (BSc) and approved as to its style and contents. The presentation has been held on 7th May 2018.

BOARD OF EXAMINERS

Dr. Syed Akhter Hossain

Professor and Head

Department of Computer Science and Engineering

Faculty of Science & Information Technology

Daffodil International University

Chairman

Dr. Sheak Rashed Haider Noori

Associate Professor and Associate Head

Department of Computer Science and Engineering

Faculty of Science & Information Technology

Daffodil International University

Internal Examiner

Md. Zahid Hasan

Assistant Professor

Department of Computer Science and Engineering

Faculty of Science & Information Technology

Daffodil International University

Internal Examiner

Dr. Mohammad Shorif Uddin

Professor

Department of Computer Science and Engineering

Jahangirnagar University

External Examiner

DECLARATION

We hereby declare that, this research project has been done by us under the supervision of **Ahmed Al Marouf, Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Ahmed Al Marouf
Lecturer
Department of CSE
Daffodil International University

Co-Supervised by:

Shah Md. Tanvir Siddiquee
Senior Lecturer
Department of CSE
Daffodil International University

Submitted by:

Rafayet Hossain
ID: 142-15-3994
Department of CSE
Daffodil International University

Md. Rahmatul Kabir Rasel Sarker
ID: 142-15-3732
Department of CSE
Daffodil International University

Mehejabin Mimo
ID: 142-15-3453
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year research project successfully.

We really grateful and wish our profound our indebtedness to **Ahmed Al Marouf, Lecturer** and **Shah Md. Tanvir Siddiquee, Senior Lecturer**, Department of CSE, Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor and co-supervisor in the field of “*Human Computer Interaction*” to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Dr. Syed Akhter Hossain, Professor & Head, Department of CSE**, for his kind help to finish our research project and also to other faculty members and the staffs of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

ABSTRACT

Music is most important part in human life. When different kind of words are prepared a new sound which is enjoyable to the human beings, it is called music. Music is not just a source of entertainment. It is something more than entertainment. Title should come after the song is finished and should reproduce a termination of the lyrical content. In this paper our proposed system have recommend a proper song title based on song lyrics. We have applied topic model algorithm Latent Dirichlet allocation (LDA) for song title recommendation. In this paper another experiment is music emotion recognition. Song feel emotionally different to listeners depending on their lyrical contents. Emotions classify is so difficult through the existing music emotion classification method .We have extracted eight features from song lyrics. We propose a method for lyrics based emotion classification using feature selection. We also proposed another experiment music personality trait. We have to generate a customize dataset based on music interest and 20 questions of big five personality model. Our proposed module would be helpful for user. Song title recommendation system produces satisfactory result. We may use this module to recommend song title from lyrics. Music emotion recognition system will help to predict the overall emotional state of a user. Music personality traits could be useful to find out the personality measurements of any user.

TABLE OF CONTENTS

CONTENTS	PAGE
Board of examiners	i
Declaration	ii
Acknowledgements	iii
Abstract	iv
Table of contents	v
List of figures	vii
List of tables	viii
 CHAPTER	
 CHAPTER 1: INTRODUCTION	 1-5
1.1 Introduction	1
1.2 Motivation	3
1.3 Rationale of the Study	3
1.4 Research Questions	4
1.5 Expected Outcome	4
1.6 Report Layout	4
 CHAPTER 2: BACKGROUND	 6-17
2.1 Introduction	6
2.2 Related Works	7
2.3 Research Summary	10
2.4 Scope of the Problem	16
2.5 Challenges	17

CHAPTER 3: RESEARCH METHODOLOGY	18-25
3.1 Introduction	18
3.2 Research Subject and Instrumentation	18
3.3 Data Collection Procedure	20
3.4 Statistical Analysis	23
3.5 Implementation Requirements	25
CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSION	26-35
4.1 Introduction	26
4.2 Experimental Results	26
4.3 Descriptive Analysis	30
4.4 Summary	35
CHAPTER 5: SUMMARY, CONCLUSION, RECOMMENDATION & IMPLICATION FOR FUTURE RESEARCH	36-38
5.1 Summary of the study	36
5.2 Conclusion	37
5.3 Recommendations	37
5.4 Implication for Further Study	38
REFERENCES	39-40
APPENDIX	41
Appendix A: Project Reflection	41
PLAGIARISM REPORT	42

LIST OF FIGURES

FIGURES	PAGE NO
Figure 1.1: Emotion Chart	2
Figure 2.1: Big Five Personality Model	6
Figure 2.2: Song Title Recommendation Procedure	10
Figure 2.3: Dataset for Title Recommendation	11
Figure 2.4: Data Filtering	11
Figure 2.5: Music Emotion Recognition Procedure	12
Figure 2.6: Personality form 1	14
Figure 2.7: Personality Form 2	15
Figure 2.8: Music Survey Website	16
Figure 3.1: Data Collection Process of STR	20
Figure 3.2: Data Collection Process of MER	21
Figure 3.3: Data Collection Process of MPT	22
Figure 3.4: Song Category Based on Emotion	23
Figure 3.5: Song Category Based on Genre	24
Figure 4.1: Result of LDA	26
Figure 4.2: User Interface of STR	27
Figure 4.3: Dataset of Personality test	29
Figure 4.4: Raw data of song	30
Figure 4.5: Result of UI of STR	31
Figure 4.6: Dataset of 8 feature	32
Figure 4.7: Music Survey	33
Figure 4.8: Survey form with ID	33
Figure 4.9: Personality form with ID	34
Figure: Plagiarism Report	42

LIST OF TABLES

TABLES

PAGE NO

Table 4.1: Features Output

28

CHAPTER 1

INTRODUCTION

1.1 Introduction

In the digital age, Music plays an important role in human's day to day life. Emotion is something very personal and maybe it is for better that computer will never completely understand a human. Music emotion recognition is a challenging problem. But we will try to get a best recognition of mode of a song. It can add a new feature of modern robotics world. We have focused on a challenging issue of recognizing the emotion content of music emotion recognition. We'll find out emotional genre of the song. It can be happy, sad, angry or excited.

Human Computer Interaction (HCI) is the study of how human being and computer interact with each other. Researchers in the arena of HCI both perceive the process in which humans collaborate with computers and scheme technologies that enable humans interact with computers in innovative way. As a field of research, human-computer interaction is located at the junction of Media Studies, Design, Behavioral Science Computer Science and different other aspects of study. In favor of get a comfortable and better life, human computer interaction systems are broadly being used in the smart home for the resident. It is important to response the joyful or stressful emotion of the resident. It's an obvious to detect the mood or emotion of the resident, as HCI here is to betterment of life. In this new era emotion have been simulated with computer. In our project we have emphasized on music emotion recognition. The emotional condition should be recognized in the situations which accord with real life by means of music and the lyrics of music.

Emotion is any relatively short deliberate practical knowledge specified by strong mental activity and a powerful degree of satisfaction or dissatisfaction. Emotion is often interweave with feeling, temper, personality, nature and motivation. In some theories, perception is an essential appearance of emotion.



Figure 1.1: Emotion Chart [13]

Emotion recognition is the process of identifying human emotion, mostly from facial expressions. This is both something that humans do automatically but computational methodologies have also been developed. But in this project, we are identifying emotion from the lyrics of the music.

At the beginning of Music Emotion Recognition (MER), maximum classification system were used audio content to emotion recognition. Now a days, researchers has begun addressing the problem of emotion detection in music lyrics (e.g., [1]). Such as, combining audio and lyrics, bimodal systems, are being researched. Different bi-modal studies have exhibited improved classification performances (e.g., [2] [3]). That's why we are using Lyrics to get a better performance here as there has some scopes to do so. Our main target is to explore the performance of music emotion recognition from lyrics.

Scopes:

Scopes of Music Emotion Recognition Using Lyrics are tremendous. Different application can be developed through this. Like:

- a. Recommendation of song
- b. Prediction System
- c. Rehabilitation
- d. Self-Assistant
- e. Music Therapy
- f. Music Player

1.2 Motivation:

The common way to enjoy with our own emotion is music. Every time we used to hear song but sometimes we cannot select our song according to our emotion. This have made us to work with music. We'll work on music emotion, based on lyrics which can add new era on modern robotics and music industry.

1.3 Rationale of study:

Text analysis for Music Emotion Recognition is a Challenging fact. From music to get lyrics or text, from text removing some unnecessary words, feature extraction, feature selection and then classification of those features to detect the emotion, that are huge procedure. To perform such things are not easy. It's really a challenging task to get a better performance. In order to get a better performance, we are working on this project.

1.4 Research Questions:

- a) Is it possible to recommend a song title from its lyrics?
- b) How Text Analysis could be applied to find out emotions from music?
- c) Can music preference a person and his/her personality traits be mapped?

1.5 Expected Output:

As we have worked on three experiment on music, so we have different expectation from different experiment. It is possible to recommend a song title from song lyrics. So, we are working on this to get a proper title from the song lyrics. Emotion is the best part in human's life. We want to make a model to find out the different emotion from the music. Because music is the best way to express anyone emotion. Human have a great relationship between music and his/her personality. We need to find out a pattern between them. It helps the human to select the proper music according to his/her personality and vise-versa.

1.6 Report Layout:

Chapter 1: Introduction

In this chapter we have discussed about music, human's emotion, human's personality and their relationship. We also discussed about introduction, motivation, rationale of study, research questions and expected outcome of our research based project. Later followed by the report layout.

Chapter 2: Background

We discuss about the background circumstances of our research based project. Here we have worked on three experiment we talked about their primary condition. We also talked about the related works, research summary, scope of the problem and challenges.

Chapter 3: Research Methodology

In this chapter we have talked about details methodology of our search based project. Here we also talked about search subject and instrumentation, data collection procedure, statistical analysis and implementation requirement.

Chapter 4: Experimental results and discussion

We have discussed about the experiment outcome and descriptive analysis of whole procedure. Here we talked in details about experimental result, descriptive analysis on our experiment and summary of the experiment.

Chapter 5: Summary, Conclusion, recommendation and implication for future research

In this chapter finally we have talked about summary of the study, conclusion, recommendation and implication of further study.

CHAPTER 2

BACKGROUND

2.1 Introduction:

Title recommendation is important to identify the category of music. Title can be defined as a summary of a song. So it is necessary to create proper title for a song. In this experiment we have focused on title recommendation, which is an obvious thing for an artist to maintain.

Emotion recognition means to recognize human emotion. In this project we have focused on music emotion recognition from the song lyrics. We have chosen lyrics because, it says a lot of about emotion in the songs. Lyrics can be an expression of emotion, which is so efficient to identify human's emotion.

Music personality traits means to find out the personality of a person according to their music taste. In this experiment we mainly focused on, to make a dataset of personality of human music interest. We have followed big five model to identify the personality. Big five model uses descriptors of common language therefore suggest five broad dimensions normally used to explain the human personality.

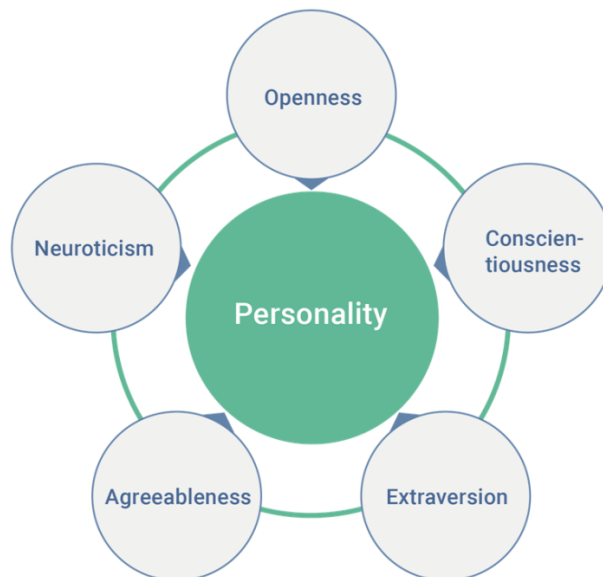


Figure 2.1: Big Five Personality Model [4]

2.2 Related Work:

Experiment 1: (Song Title Recommendation)

J. Chang et al. in [5] proposed a relational topic models for document networks. In this paper they develop the relational topic model (RTM), a model of documents and the links between them. For each pair of documents, the RTM models their link as a binary random variable that is conditioned on their contents. Their model can be used to summarize a network of documents, predict links between them, and predict words within them. They derive efficient inference and learning algorithms based on various methods and evaluate the predictive performance of the RTM for large networks of scientific abstracts and web documents.

D. Ramage et al. in [6] proposed a Topic Modeling for the Social Sciences. As textual datasets grow in size and scope, social scientists need better tools to help make sense of that data. Despite the natural applicability of topic modeling to many such problems, word counts and tag clouds are often used as the primary means of gleaning information from textual data. They characterize two barriers to adoption encountered during a collaboration between the Stanford NLP group and social scientists in the school of education: accessibility and trust. Accessibility refers to the technical barriers that make text processing and topic modeling difficult. Trust comes when practitioners can explore and validate a model being used to discover or support a hypothesis. They introduce recent work aimed at solving these challenges including the Stanford Topic Modeling Toolbox software.

D. Newman et al in [7] proposed an improving topic coherence with regularized topic models. Topic models have the potential to improve search and browsing by extracting useful semantic themes from web pages and other text documents. When learned topics are coherent and interpretable, they can be valuable for faceted browsing, results set diversity analysis, and document retrieval. However, when dealing with small collections or noisy text, learned topics can be less coherent, less interpretable, and less useful. To overcome this, they propose two methods to regularize the learning of topic models. Their regularizers work by creating a structured prior over words that reflect broad patterns in the external data. Using thirteen datasets they show that both regularizers

improve topic coherence and interpretability while learning a faithful representation of the collection of interest. Overall, this work makes topic models more useful across a broader range of text data.

Experiment 2: (Music Emotion Recognition)

Classification of music emotions first requires the choice of an emotion model. Considerations in that regard include the accuracies with which models represent emotions and emotions reflect actual music, respectively. Most of the existing research on music emotion classification has utilized.

R. Malheiro et al in [8] proposed music emotion recognition from lyrics. In this experiment, they have shown a study on music emotion recognition using lyrics. They have begun from a dataset of 764 samples and performed feature extraction using different NLP (natural language processing) techniques. Their goal was to create classifiers for the several datasets, comparing various algorithms and using feature selection. The best results (44.2% F-measure) were achieved with SVMs. They also have done a bi-modal analysis that merge the best feature sets of audio and lyrics. The compound of the best audio and lyrics features have attained better results than the best feature set from audio only (63.9% F-Measure against 62.4% F-Measure).

D.Yang et al in [9] proposed music emotion identification from lyrics. In this Experiment, machine learning techniques are used as an alternative to human experts to obtain emotions in Music. The classification is mainly based on a psychological model of emotion that is expanded to 23 distinct emotion groups. Their results for mining the lyrical text of songs for certain emotion are encouraging, result in classification models that are humanly-understandable, and create results that correspond to commonsense intuitions about particular emotions. Mining lyrics have been focused in this paper is a particular part of research which merge various classifiers of musical emotion like acoustics and lyrical text.

E. Chao-Chun Kao et al in [10] proposed a towards text-based emotion detection. In this paper they present an overview of the emerging field of emotion in text and describe the current generation of detection method that are usually divided into the following three main categories: Keyword based, learning based and hybrid recommendation approach.

They improved emotion detection capabilities in practical system which emphasize on human computer interaction. This solution include extracting keywords with semantic analysis and ontology design with emotion theory of appraisal.

M. Kim et al in [11] proposed a lyrics-based emotion classification using feature selection by partial syntactic Analysis. In this paper they proposed a method for lyrics based emotion classification using feature selection by partial systaltic analysis. Based on the existing emotion ontology, four kinds of syntactic analysis which were applied to extracted emotion feature from lyrics. They extracted emotion feature along with the NB, HMM, SVM and Machine Learning were used.

Experiment 3: Music Personality Traits

Ferwerda et al in [12] proposed a personality traits and music genre preferences. In this paper, they are focusing on identifying the relationship between personality and behavior, preferences, and needs often do not take into account deference's between age groups. With music playing an important role in our lives, deference's between age groups may be especially prevalent. In this paper they investigate whether deference's exist in music listening behavior between age groups. They analyzed a dataset with the music listening histories and personality information of 1415 users. Their results show agreements with prior work that identified personality-based music listening preferences. However, their results show that the agreements they found are in some cases divided over different age groups, whereas in other cases additional correlations were found within age groups.

T. Chamorro-Premuzic et al in [13] proposed a personality and music. In this paper reports the results of a study on the relationship between individual differences and specific uses of music, referring to why and how people use music in everyday life. Questionnaire data from 341 respondents showed that open and intellectually engaged individuals, and those with higher IQ scores, tended to use music in a rational/cognitive way, while neurotic, introverted and non-conscientious individuals were all more likely to use music for emotional regulation. Results suggest that individual differences in personality and cognitive ability may partly determine the way in which we experience music. Limitations and suggestions for future studies are discussed.

2.3 Research summary:

Experiment 1:

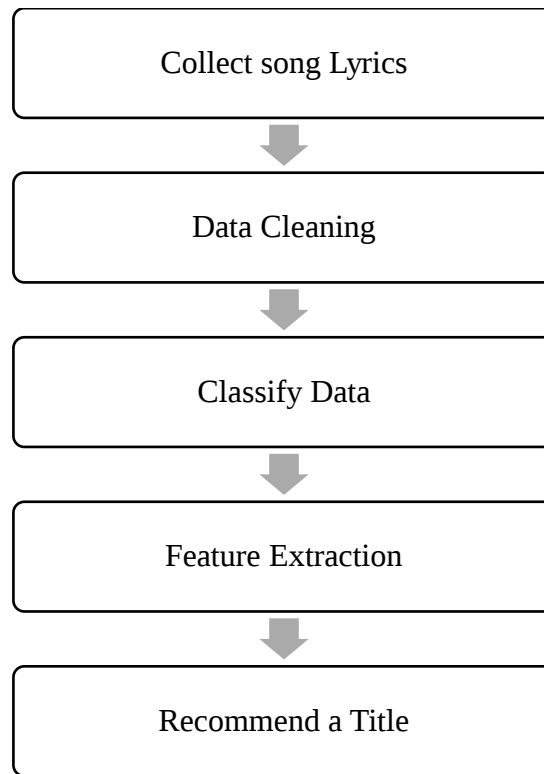


Figure 2.2: Song Title Recommendation Procedure

Here the flowchart is given, this is our total procedure. Firstly we have taken 200 random songs as our experimental data. We collected them different web site.

song1.txt	song30.txt	song59.txt	song88.txt	song117.txt	song146.txt	song175.txt
song2.txt	song31.txt	song60.txt	song89.txt	song118.txt	song147.txt	song176.txt
song3.txt	song32.txt	song61.txt	song90.txt	song119.txt	song148.txt	song177.txt
song4.txt	song33.txt	song62.txt	song91.txt	song120.txt	song149.txt	song178.txt
song5.txt	song34.txt	song63.txt	song92.txt	song121.txt	song150.txt	song179.txt
song6.txt	song35.txt	song64.txt	song93.txt	song122.txt	song151.txt	song180.txt
song7.txt	song36.txt	song65.txt	song94.txt	song123.txt	song152.txt	song181.txt
song8.txt	song37.txt	song66.txt	song95.txt	song124.txt	song153.txt	song182.txt
song9.txt	song38.txt	song67.txt	song96.txt	song125.txt	song154.txt	song183.txt
song10.txt	song39.txt	song68.txt	song97.txt	song126.txt	song155.txt	song184.txt
song11.txt	song40.txt	song69.txt	song98.txt	song127.txt	song156.txt	song185.txt
song12.txt	song41.txt	song70.txt	song99.txt	song128.txt	song157.txt	song186.txt
song13.txt	song42.txt	song71.txt	song100.txt	song129.txt	song158.txt	song187.txt
song14.txt	song43.txt	song72.txt	song101.txt	song130.txt	song159.txt	song188.txt
song15.txt	song44.txt	song73.txt	song102.txt	song131.txt	song160.txt	song189.txt
song16.txt	song45.txt	song74.txt	song103.txt	song132.txt	song161.txt	song190.txt
song17.txt	song46.txt	song75.txt	song104.txt	song133.txt	song162.txt	song191.txt
song18.txt	song47.txt	song76.txt	song105.txt	song134.txt	song163.txt	song192.txt
song19.txt	song48.txt	song77.txt	song106.txt	song135.txt	song164.txt	song193.txt
song20.txt	song49.txt	song78.txt	song107.txt	song136.txt	song165.txt	song194.txt
song21.txt	song50.txt	song79.txt	song108.txt	song137.txt	song166.txt	song195.txt
song22.txt	song51.txt	song80.txt	song109.txt	song138.txt	song167.txt	song196.txt
song23.txt	song52.txt	song81.txt	song110.txt	song139.txt	song168.txt	song197.txt
song24.txt	song53.txt	song82.txt	song111.txt	song140.txt	song169.txt	song198.txt
song25.txt	song54.txt	song83.txt	song112.txt	song141.txt	song170.txt	song199.txt
song26.txt	song55.txt	song84.txt	song113.txt	song142.txt	song171.txt	song200.txt
song27.txt	song56.txt	song85.txt	song114.txt	song143.txt	song172.txt	
song28.txt	song57.txt	song86.txt	song115.txt	song144.txt	song173.txt	
song29.txt	song58.txt	song87.txt	song116.txt	song145.txt	song174.txt	

Figure 2.3: Dataset for Title Recommendation

From the dataset we have elaborate the contraction word like “isn’t” to “is not”, “don’t” to “do not”, “can’t” to “can not” etc. We have removed special keywords, verse & chorus from the data.

life.txt - Notepad File Edit Format View Help [Verse 1] This ain't a song for the broken-hearted No silent prayer for the faith-departed I ain't gonna be just a face in the crowd You're gonna hear my voice When I shout it out loud [Chorus] It's my life It's now or never I ain't gonna live forever I just want to live while I'm alive (It's my life) My heart is like an open highway Like Frankie said, "I did it my way" I just wanna live while I'm alive It's my life	life.txt - Notepad File Edit Format View Help This is not a song for the broken-hearted No silent prayer for the faith-departed I am not gonna be just a face in the crowd You are gonna hear my voice When I shout it out loud It is my life It is now or never I am not gonna live forever I just want to live while I am alive It is my life My heart is like an open highway Like Frankie said, "I did it my way" I just wanna live while I am alive It is my life
---	--

Figure 2.4: Data Filtering

Thus we filter our experimental data. Then we classified those data using several classifier. Among the classifier which one was suitable for our experiment we choose that. Then we applied the classifier and extracted the feature which has more accuracy than other. With the help of that feature we recommend a suitable title to the song lyrics.

Experiment 2:

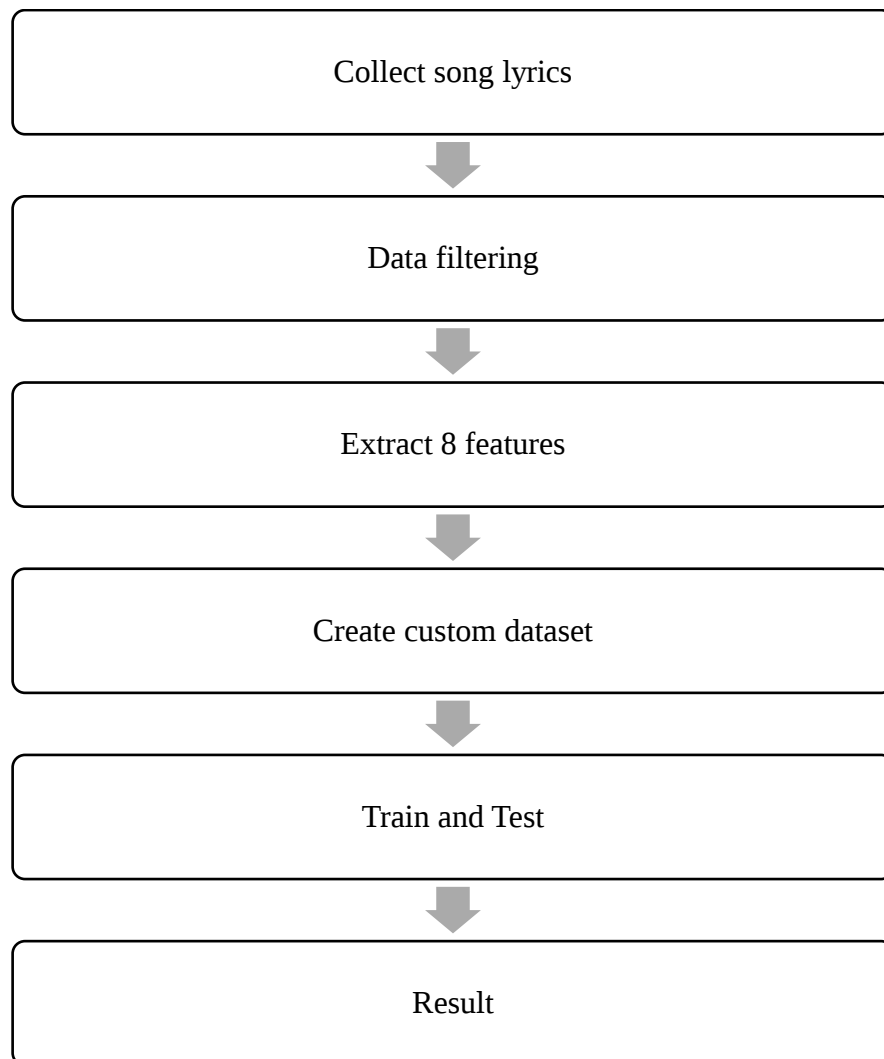


Figure 2.5: Music Emotion Recognition Procedure

For the experiment we have collected 200 random songs as we used lyrics as our input. In filter section we filtered the experimental data. We removed special keyword, auxiliary verb like “am, is, are” etc. We also deleted the digits from the lyrics. Here we extracted 8 features like,

1. First Line of the lyrics
2. Full uppercase word
3. Positive Negative word count
4. Average verb counting
5. Average parts of speech counting
6. Special Character counting
7. Average word length per sentence
8. Bag of Words counting

From this features we have created a custom dataset with all numeric and nominal data. We have taken 80% data as training data and another 20% data as test data. We performed different classifier and obtained the best accuracy among them for music emotion recognition.

Experiment 3:

We have created 20 questions to recognize the human personality according to big five model.

RealityAboutMe: A personality measuring questionnaire.

This questionnaire is based on the Mini-IPIP (International Personality Item Pool). based on this 20 interesting questions your personality could be reflected. Please try to give the answers honestly as these data will be used for solely research purpose.

Note: No personal information (name, StudentID, Address etc.) will not be stored in this dataset.

Request: Please do attend the another survey for "BanglaMusicInterest".

Footer: All the phot used in this survey are properly referenced to the concern publishers.

* Required

What is your UserID? (Ask for your UserID to the volunteers) *

Your answer

I am the life of a party *



Figure 2.6: Personality form 1

I am relaxed most of the time. *



☐ Strongly Disagree

☐ Disagree

☐ Neutral

☐ Agree

☐ Strongly Agree

SUBMIT

Figure 2.7: Personality Form 2

We have developed a survey web site to finding the music taste of a person. We have taken total 80 songs of different genre. We have cropped the song in short, from the form we get the subject age and gender.

Survey Form

Name

Age

Gender
☐ Female ☐ Male ☐ Others

SUBMIT

Figure 2.8: Music Survey Website

2.4 Scope of the problem:

Experiment 1:

Scope of title recommendation using lyrics are tremendous. Here are some application of the song title recommendation.

- Assisting the artists
- Music therapy
- To summarize the song for the listener

Experiment 2:

Scope of the music emotion recognition are remarkable, here some application are given below that could be developed using Music Emotion Recognition.

- Prediction system
- Self-Assistant
- Music therapy
- Smart music player

Experiment 3:

Some system or application may possible to construct with music personality traits. Lists are-

- Prediction system
- Smart Music player
- Self-assistance device

2.5 Challenges:

Experiment 1:

There were some challenging task to implement this such as,

- To choose the appropriate song
- Preprocessing part was a bit difficult
- Scoring and picking the accurate value

Experiment 2:

Some challenges in this aspect we faced are-

- Negative Positive Dataset are rare
- Bag of words (BOW) implementation was complicated
- Data maintenance
- Developing custom dataset

Experiment 3:

- To create 20 logical questions that actually represent personality
- Domain and hosting are costly
- To make user friendly web site
- Data Maintenance

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

A research method is an organized plan for regulating research. It composes the theoretical analysis of the figure of methods and principles related with a sector of knowledge. For three experiment we have performed different methodology. For experiment one we used some methodology of topic modeling. For experiment two the methodology was different. We extract some feature and make a database for classification. For experiment three we followed big five model for personality detection and also developed a survey web site to detect musical interest. These three experiment showed three different outcome in Human computer interaction (HCI), Natural language processing (NLP) and as well as music industry.

3.2 Research Subject and Instrumentation:

Experiment 1:

Subject: Song Title Recommendation using Lyrics

Instrumentation:

- Song Lyrics
- Tokenizer
- Stop_Words
- LDA model
- Gensim

Experiment 2:

Subject: Music Emotion Recognition using Lexical Analyzer based on lyrics

Instrumentation:

- Song Lyrics
- 8 features
- Emotion Dataset
- Positive Negative Dataset
- Bag of Words Dataset

Experiment 3:

Subject: Music Personality Traits

Instrumentation:

- Google Survey form
- web site

3.3 Data Collection Procedure:

Experiment 1:

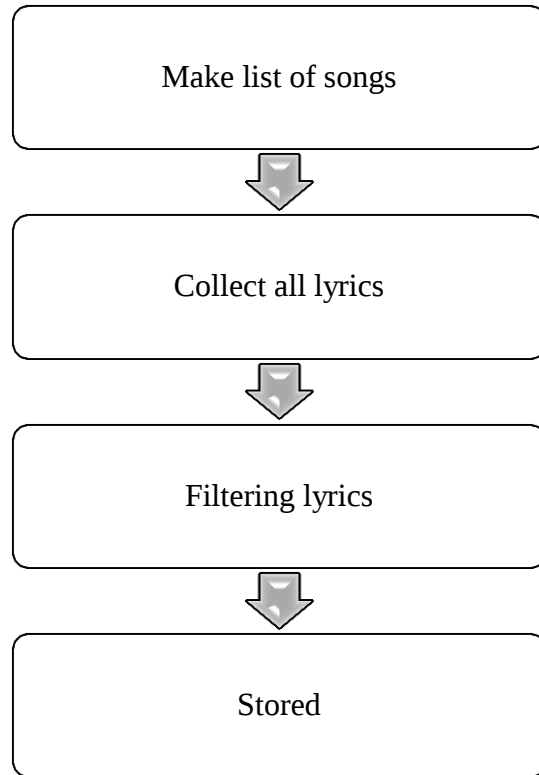


Figure 3.1: Data Collection Process of STR

Data collection is a procedure of collecting and computing information focused variable in an established organized manner. Here, firstly we make a list of some selective untitled song. Then form the list we collected all the lyrics. Then we filtered those data for our experiment, because the data was not appropriate for our experiment. Here we also removed the word verse and chorus and digit as well. We also elaborated the contraction word. Finally we stored the preprocessed data as a dataset.

Experiment 2:

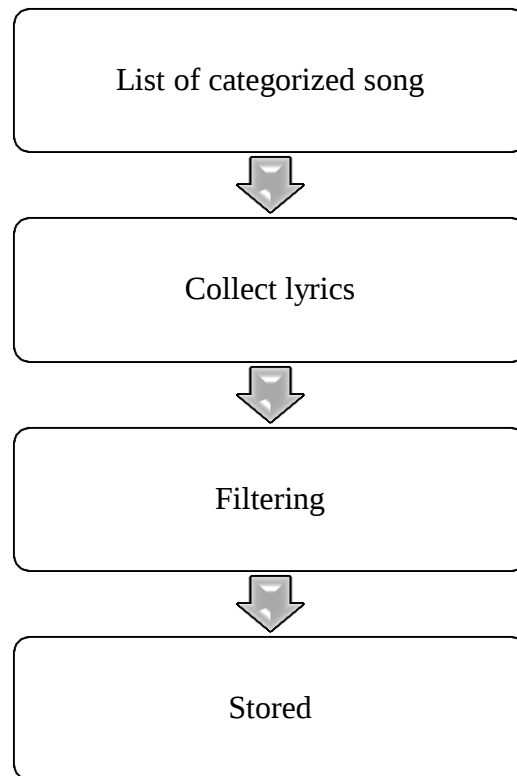


Figure 3.2: Data Collection Process of MER

For data collecting we firstly made a list of categorized song like happy, sad, romantic, angry and so on. Then we collected the required categorized songs. After that we gathered the lyrics and then filtered them. In filtering state we cleaned the song lyrics. We deleted the verse, chorus and digit as well. Finally we stored the data for experiment.

Experiment 3:

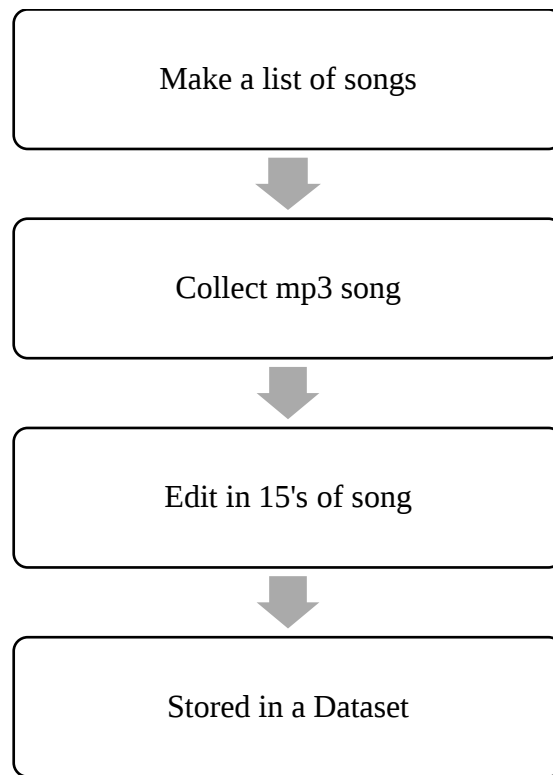


Figure 3.3: Data Collection Process of MPT

As we were working on two different sector, that's why we collected data in different way. For the personality detection we created 20 different question that can find out the personality. Another one, we created a web site along with 80 songs that will find out the musical interest.

3.4 Statistical Analysis:

Experiment 2:

As our experiment we have taken 200 different emotionally categorized songs. From those we have taken 160 data for training data and 40 for the test data. The categorized are: Happy, Sad, Angry, Love, Pain, Surprised, Enjoyment, Fear, Anguish, Excited. From these 10 categorizes, we have taken 20 songs for each one.

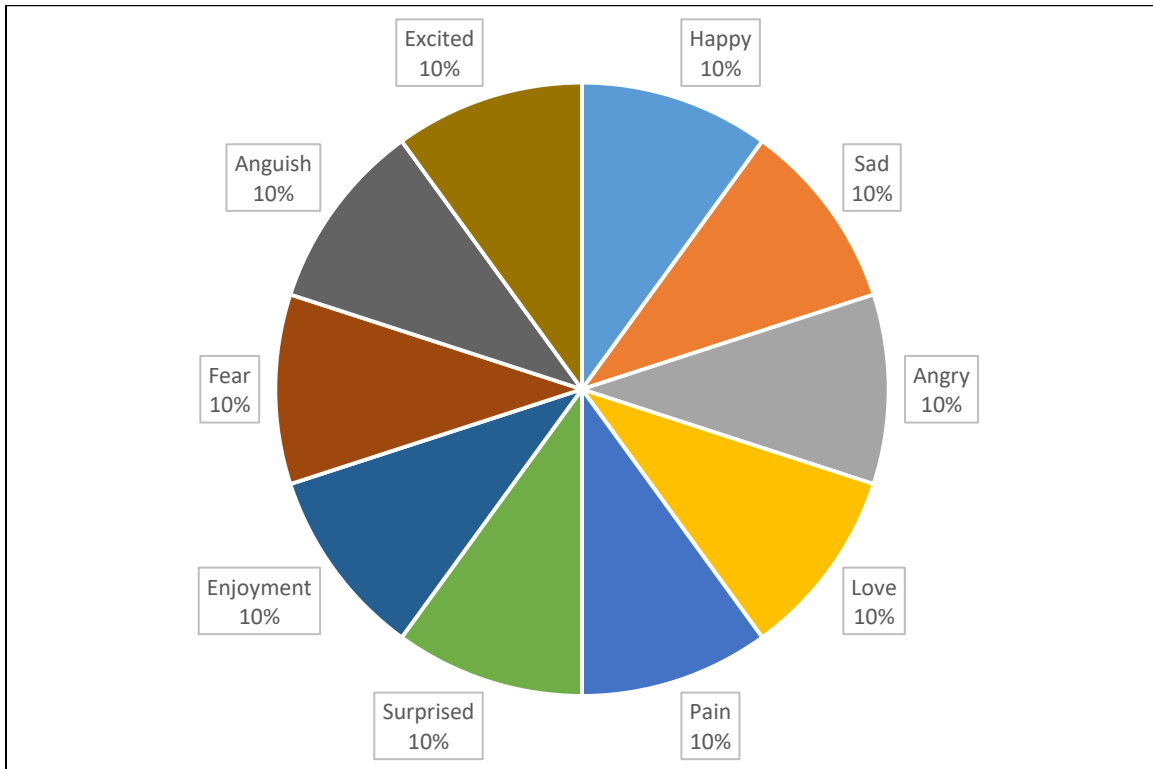


Figure 3.4: Song Category Based on Emotion

Experiment 3:

Here we are doing statistical analysis only for musical interest as we do not need any statistical data for personality detection. We collect the song according to the genre. They are:

- Classical
- Rabindra Songit
- Nazrul giti
- Adhunik
- Folk
- Boul
- Instrumental and
- Modern music.

Here we have taken 10 songs from each genre.

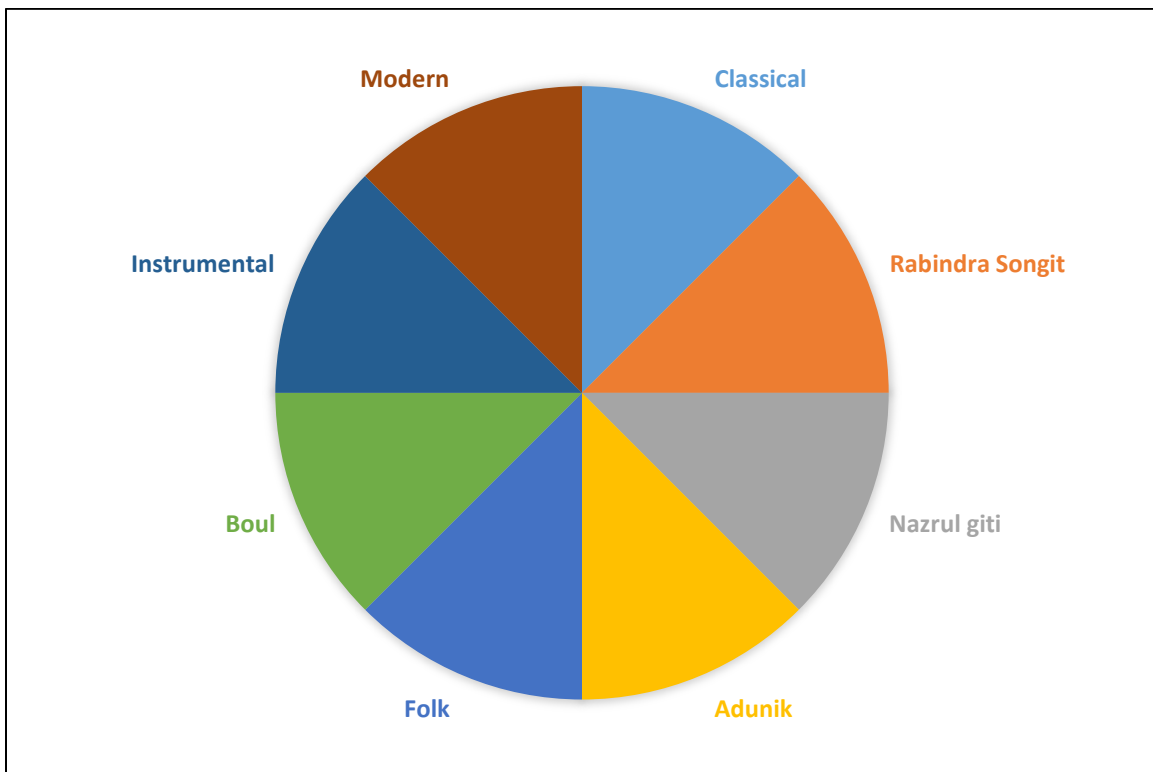


Figure 3.5: Song Category Based on Genre

3.5 Implementation Requirement:

Experiment 1:

To implement the experiment we mainly used Python. For this we were needed some third party packages.

- **NLTK:** A Natural language toolkit or more generally NLTK is a staff of libraries and program for python.
- **Stop_Words:** In natural language processing meaningless word or data are defined as stop word.
- **Gensim:** Gensim is a python library for topic modeling

Experiment 2:

To perform this experiment here we also used python. To obtain the gain, we were needed some third party packages like NLP libraries, BOW libraries, POS NEG dataset. With the help of this library and dataset we extracted 8 features. Then we make a custom dataset from those feature.

- NLP Libraries
- Bag of Words
- Positive Negative Dataset

Experiment 3:

We were performed on two experiment, so our requirement were also different.

- Google Form
- Html, CSS
- PHP, Ajax
- MySQL server
- Domain & Hosting

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction:

For music title recommendation we have tried to recommend an efficient title to a song. We also have tried to give a proper title that bear the depth interpretation of the corresponding song. For this purpose we have chosen better methodology from different algorithms, as we explained it clearly in the previous chapter (Research Methodology).

Music emotion recognition is an experiment where we have tried to recognize the emotion of a song. To get an efficient output, we applied some methodology as discussed in previous chapter.

Our focused was to generate a customize dataset. One was to detect human's personality and another one is for human's music interest. We surveyed both experiment on around 100 persons.

4.2 Experimental Results:

Experiment 1:

Form the experiment using Latent Dirichlet allocation (LDA) we have got different accuracy for different word of the lyrics. We have stored the best accuracy and chosen, that word contained sentence as our title.

```
>>> print(ldamodel.print_topics(num_topics=3, num_words=3))  
['0.141*health + 0.080*brocolli + 0.080*good', '0.060*eat +  
059*brother']
```

Figure 4.1: Result of LDA

We have also developed a graphical user interface (GUI) for title recommendation system. User can input the lyrics and it will show them an expected title.

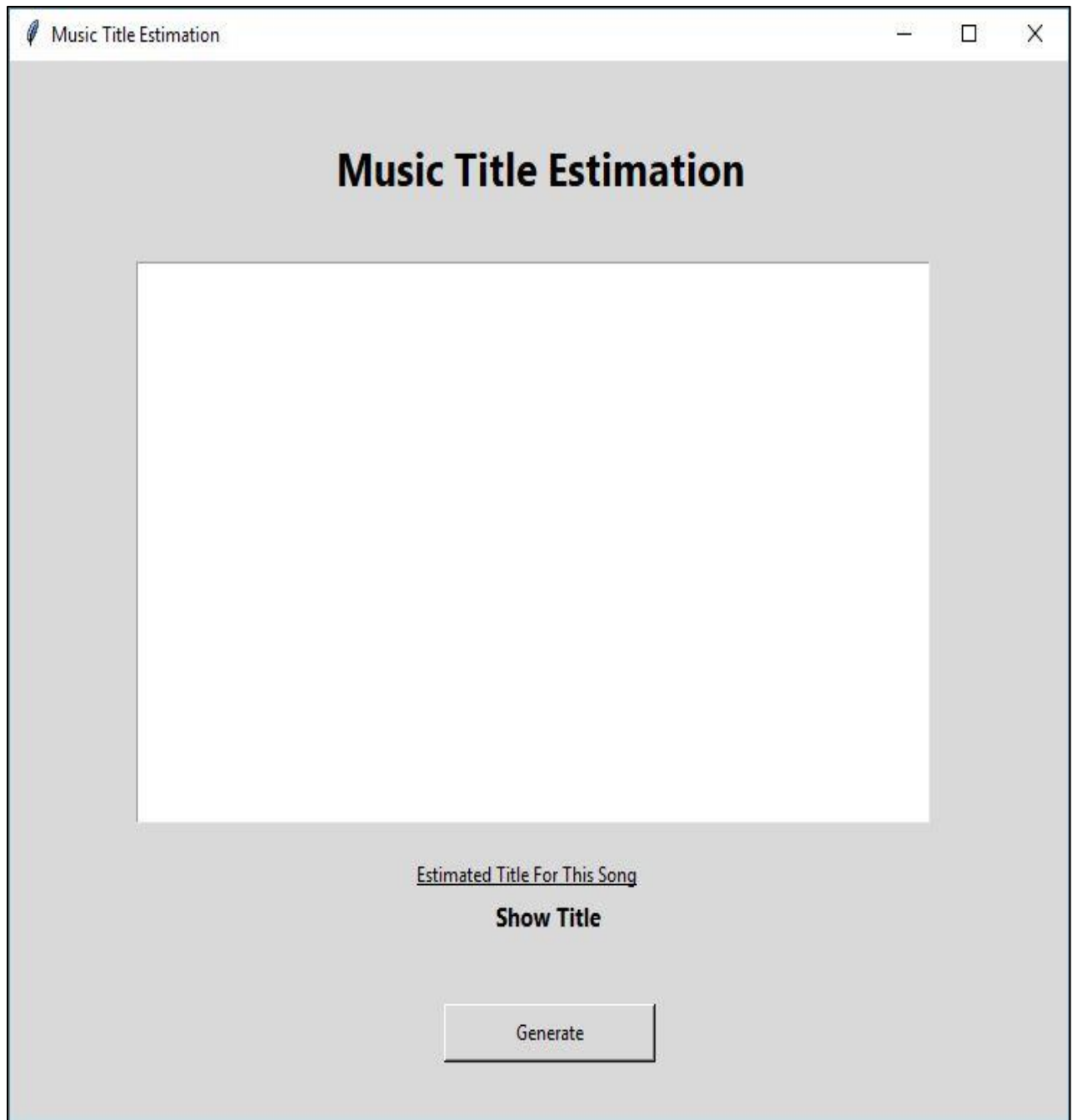


Figure 4.2: User Interface of STR

Experiment 2:

Table 4.1: Features Output

Name of the Feature	Instance
First line of the lyrics	A pillow of tears
Full uppercase word	0
Average length of word in each line	4 6 4 5 5 5 1 1 7 6 3 1 4 6 4 5 6 5 1 1 3 6 3 5 5 5
Average parts of speech	{'DT': 0.019, 'NN': 0.087, 'IN': 0.135, 'VBZ': 0.038, 'VBD': 0.019, 'NNP': 0.087, 'JJ': 0.009}
Average number of verb	20
Positive Negative word counting	Pos- 4 Neg – 2
Bag of words	'back': 6, 'you': 5, 'of': 4, 'Is': 4, 'A': 3, 'tears': 3, 'know': 3, 'not': 3, 'come': 3, 'Come': 3, 'to': 3, 'darling': 3,
Special Keywords	2

Experiment 3:

The experiment which we have finished that gave us the outcome like –

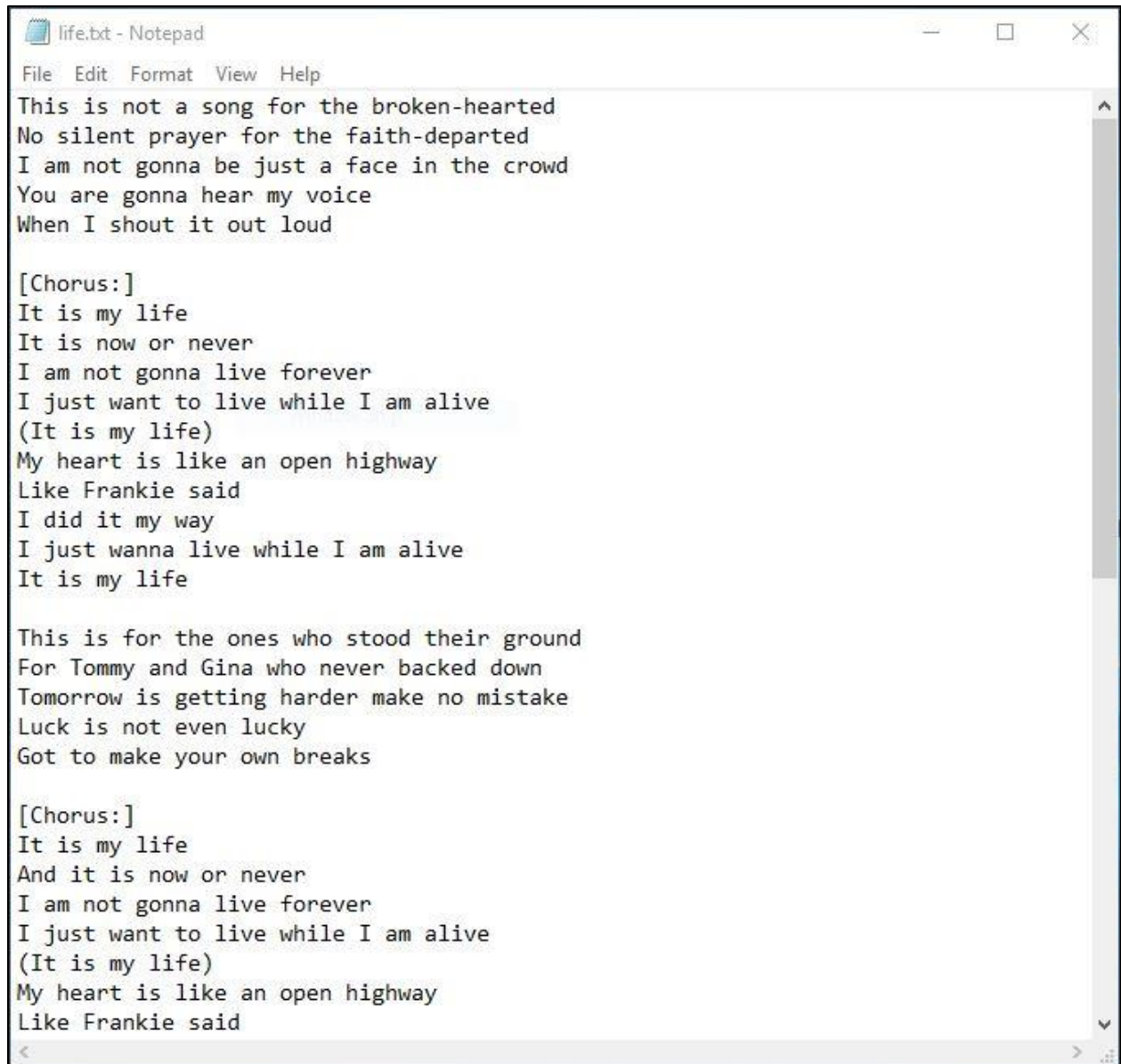
SN	ID	Openness	Conscientiousness	Extraversion	Agreeableness	Neutoricism
1	47	11	76	73	42	73
2	44	11	28	58	21	56
3	32	85	66	92	38	34
4	25	24	48	42	76	73
5	35	52	73	50	38	52
6	48	66	76	42	69	79
7	50	34	66	50	44	56
8	49	16	24	16	42	52
9	36	73	48	73	52	66
10	33	34	38	42	52	69
11	43	50	38	79	52	34
12	37	24	21	42	5	66
13	19	34	48	38	44	21
14	18	62	76	28	60	58
15	17	66	56	88	34	28
16	16	85	28	97	69	50
17	38	8	40	56	44	66
18	31	34	92	66	76	28
19	26	42	79	58	42	42
20	42	34	21	42	42	50
21	39	40	85	79	11	16
22	34	5	52	88	71	60

Figure 4.3: Dataset of Personality test

4.3 Descriptive Analysis:

Experiment 1:

Here is our experimental procedure and a scenario. It will clarify the conception of entire experiment. As our experimental data we have taken the song lyrics “**It’s my life**” by Bon Jovi.



```
life.txt - Notepad
File Edit Format View Help
This is not a song for the broken-hearted
No silent prayer for the faith-departed
I am not gonna be just a face in the crowd
You are gonna hear my voice
When I shout it out loud

[Chorus:]
It is my life
It is now or never
I am not gonna live forever
I just want to live while I am alive
(It is my life)
My heart is like an open highway
Like Frankie said
I did it my way
I just wanna live while I am alive
It is my life

This is for the ones who stood their ground
For Tommy and Gina who never backed down
Tomorrow is getting harder make no mistake
Luck is not even lucky
Got to make your own breaks

[Chorus:]
It is my life
And it is now or never
I am not gonna live forever
I just want to live while I am alive
(It is my life)
My heart is like an open highway
Like Frankie said
```

Figure 4.4: Raw data of song

After taking input we have tokenize the data. Then we removed the meaningless word using stopword from the data. After tokenize and removing meaningless data we merged the total remaining data. That is called stemming. To create an LDA model we have construct a document term matrix using gensim. After we have applied the LDA model and got the result.

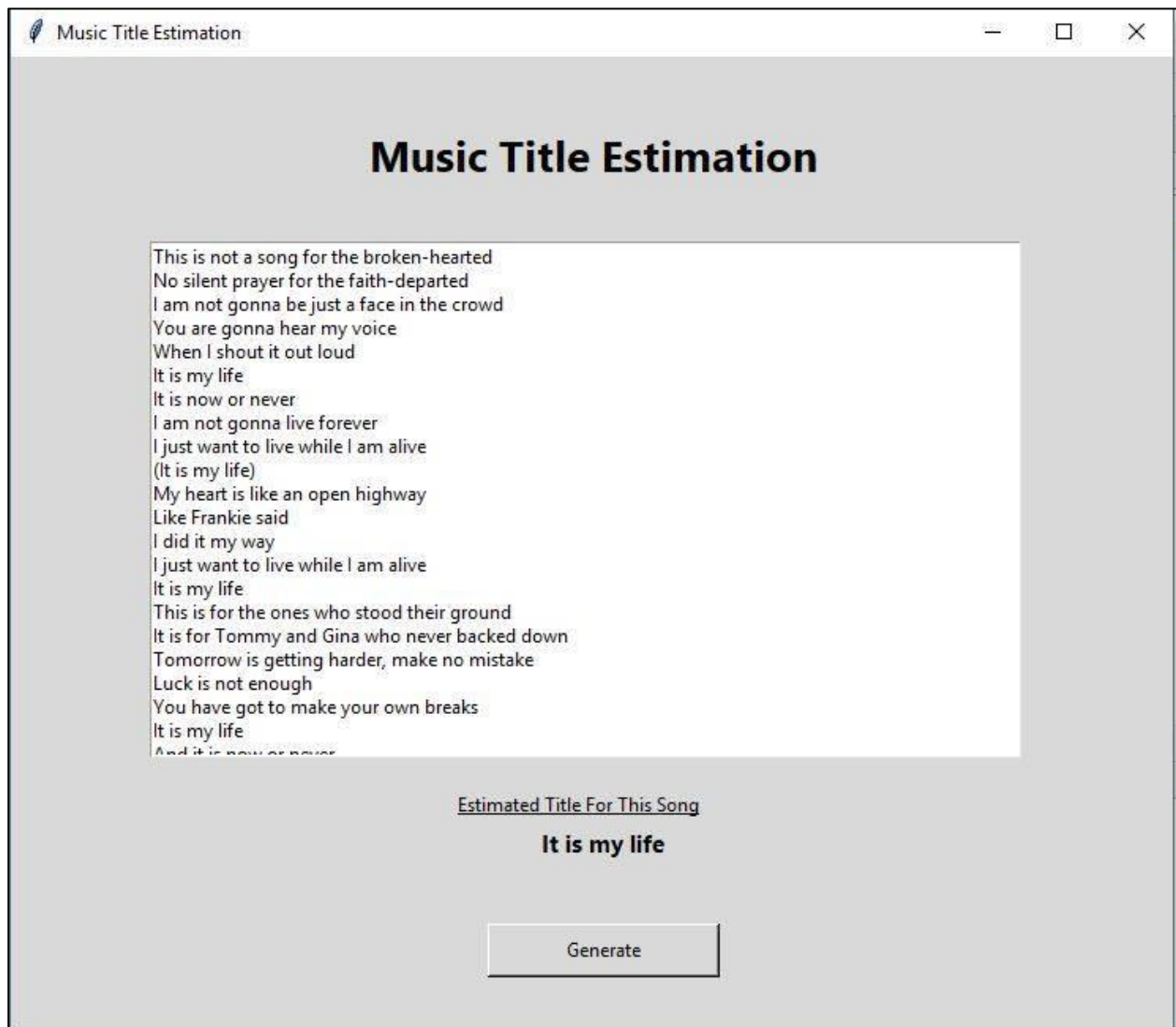


Figure 4.5: Result of UI of STR

Here the most accuracy word is “**Life: 0.047%**”. So we determined the song this about life. Line that includes the word “Life” would be our estimated title.

Song Title: **It is my life**

Experiment 2:

As we mentioned before we have collected data according to the category. Then we filtered the data and extracted 8 features from the filtered data. Let consider some songs as our tested data.

```
@data
0,0,4.863636,21,4,0,0.336735,0.071429,0.153061,0.22449,0.040816,0.142857,0.030612,0
0,0,5.6,25,2,1,0.175439,0.254386,0.035088,0.280702,0.105263,0.061404,0.061404,0
0,0,3.454545,19,1,1,0.314286,0.185714,0.085714,0.271429,0.071429,0.071429,0,0
0,0,3.107143,17,2,6,0.306667,0.173333,0.106667,0.24,0.053333,0.04,0.08,0
0,0,4.6,28,17,1,0.264151,0.132075,0.132075,0.273585,0.075472,0.066038,0.056604,0
0,0,4.392857,21,1,2,0.283333,0.225,0.033333,0.216667,0.058333,0.133333,0.041667,0
0,0,2.888889,20,2,1,0.23301,0.15534,0.116505,0.203883,0.145631,0.106796,0.019417,0
0,0,5.606061,40,10,9,0.206704,0.206704,0.03352,0.26257,0.067039,0.156425,0.067039,0
12,2,4.759259,93,3,7,0.235529,0.179641,0.035928,0.221557,0.139721,0.143713,0.031936,0.001996
0,0,4.516129,100,1,5,0.245098,0.191176,0.04902,0.264706,0.139706,0.083333,0.026961,0
0,0,5.111111,19,11,6,0.213483,0.191011,0.089888,0.224719,0.157303,0.078652,0.044944,0
4,0,4.928571,40,46,3,0.321168,0.145985,0.058394,0.175182,0.105839,0.127737,0.043796,0
7,0,7.76,136,11,1,0.127273,0.27013,0.041558,0.371429,0.033766,0.137662,0,0
0,0,5.777778,70,9,1,0.16996,0.280632,0.031621,0.343874,0.063241,0.102767,0.007905,0
0,0,5.090909,22,0,0,0.163462,0.240385,0.057692,0.269231,0.086538,0.115385,0.067308,0
0,0,4.409091,17,0,2,0.272727,0.284091,0.011364,0.193182,0.068182,0.136364,0.034091,0
1,0,3.521739,28,6,1,0.227848,0.113924,0.037975,0.35443,0.164557,0.075949,0.025316,0
0,0,3.785714,16,3,4,0.37,0.13,0.02,0.17,0.14,0.14,0.03,0
0,0,3.708333,17,5,2,0.361446,0.204819,0.036145,0.228916,0.072289,0.048193,0.048193,0
0,0,3.535714,21,1,5,0.258065,0.225806,0.075269,0.236559,0.075269,0.064516,0.053763,0
0,0,3.75,42,2,1,0.183673,0.244898,0.013605,0.340136,0.129252,0.034014,0.054422,0
0,0,3.538462,26,5,10,0.055556,0.288889,0.033333,0.388889,0.1,0.1,0.033333,0
0,0,3.228571,30,3,3,0.163636,0.145455,0.009091,0.318182,0.154545,0.181818,0.027273,0
0,0,2.756757,34,30,1,0.135417,0.34375,0.0625,0.364583,0.03125,0.052083,0.010417,0
0,0,4.954545,33,6,2,0.1,0.33,0.02,0.37,0.04,0.13,0.01,0
0,0,2.657143,24,10,3,0.212766,0.255319,0.021277,0.287234,0.095745,0.085106,0,0
0,1,3.653846,40,5,7,0.245902,0.20765,0.054645,0.240437,0.114754,0.081967,0.043716,0
0,0,5.5,39,1,2,0.194595,0.205405,0.07027,0.254054,0.075676,0.145946,0.037838,0
0,0,3.535714,21,2,5,0.258065,0.225806,0.075269,0.236559,0.075269,0.064516,0.053763,0
0,0,3.725,41,5,1,0.17931,0.248276,0.013793,0.337931,0.131034,0.034483,0.055172,0
0,0,3.538462,26,3,10,0.055556,0.288889,0.033333,0.388889,0.1,0.1,0.033333,0
0,0,3.228571,30,3,3,0.163636,0.145455,0.009091,0.318182,0.154545,0.181818,0.027273,0
```

Figure 4.6: Dataset of 8 feature

After getting 8 features we merge the entire data. Then we compared the test data with our train data. Then we gain the accuracy of songs category. We picked the best one and considered the emotion of the song.

Experiment 3:

First of all, assume that we are surveying on a subject name “Suddam”. As we have to survey form, we have given the subject music interest survey form first. Then the subject has to fill up the form with its name, age, and gender. Then an id will be generated automatically. Using this id we request the subject to sign in a consent letter and the subject has to remember the id.



Figure 4.7: Music Survey

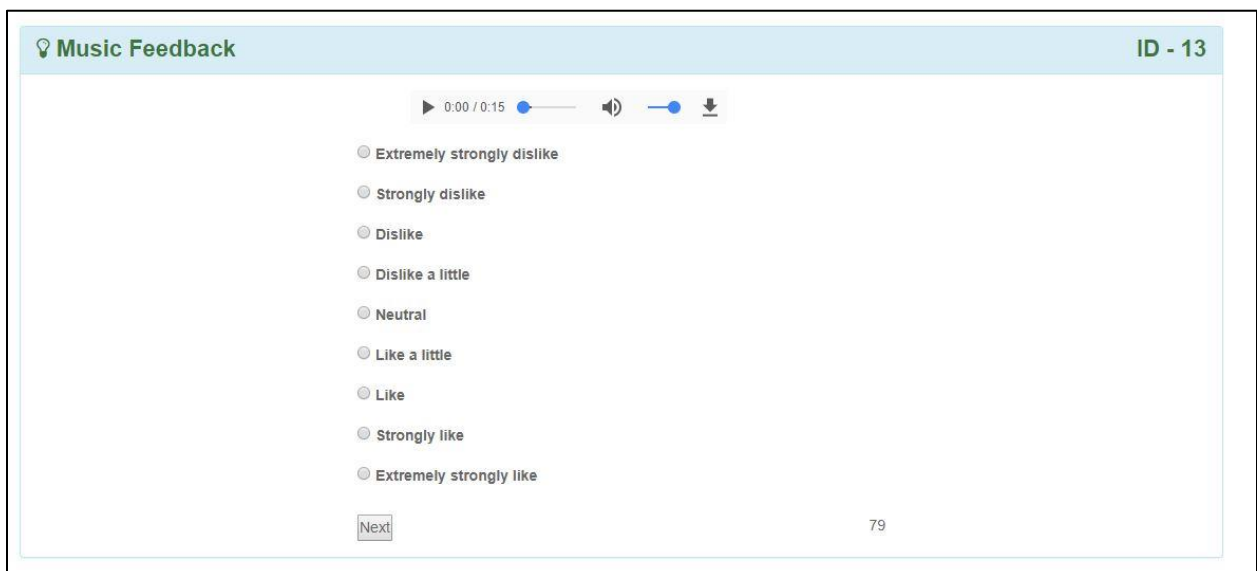


Figure 4.8: Survey form with ID

After that the subject listened all the songs and gave its feedback by choosing an option. Then we provide the subject another surveying form. Then the subject fill up the id as it generated in previous survey form. Then the subject fills up the full form by answering the questions. The id is saved because we need to merge the music interest form data and personality test form data for the same subject.

RealityAboutMe: A personality measuring questionnaire.

This questionnaire is based on the Mini-IPIP (International Personality Item Pool). based on this 20 interesting questions your personality could be reflected. Please try to give the answers honestly as these data will be used for solely research purpose.
Note: No personal information (name, StudentID, Address etc.) will not be stored in this dataset.
Request: Please do attend the another survey for "BanglaMusicInterest".
Footer: All the phot used in this survey are properly referenced to the concern publishers.

* Required

What is your UserID? (Ask for your UserID to the volunteers) *

13

I am the life of a party *



Figure 4.9: Personality form with ID

4.4 Summary:

Our experiment on music title recommendation is tremendous. We have also tried on some untitled songs. Most of the time we got relevant result for those song as expected. It is helpful for the user to obtain a proper title.

Our experiment on music emotion recognition, we have trained 160 songs and remaining 40 songs is for testing. 160 trained songs are categorized and 40 was uncategorized. It shows an excellent accuracy.

According to the personality of the subject what will be the music interest. It's the thing we have done. The customize dataset categorized those song and personality. When a subject choose a music it predict how is its personality and vise-versa.

CHAPTER 5

SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH

5.1 Summary of the Study:

Music is a form of art. When dissimilar kinds of sounds are structured a new sound which is pleasurable to the human beings, it is called music. Music is not just a source of diversion in our life. It is something more than it. We have worked to music. Firstly we have work to music title recommendation system. Secondly music emotion reorganization and third one is music personality traits. In music title recommendation system we have tried to recommend a well-organized title to a song. For title recommendation system, firstly we have collected song lyrics than filtering them and generate a preprocessed dataset. We have used topic model algorithm Latent Dirichlet allocation (LDA) and we have got different accuracy for different word of the lyrics. We have stored the best accuracy and chosen, that word contained sentence as our title. . It is supportive for the user to find a proper title. In music emotion reorganization, we have tried to recognize the emotion of music. For this experiment firstly we have collected song lyrics from different song category (happy, sad, angry, joyful etc.). Then we generate a customize dataset. We filtered the data and extracted 8 features from the filtered data. In this experiment we have trained 160 songs and rest of the songs is for testing. After getting 8 features we merge the entire data. Then we compared the test data with our train data. Then we get the accuracy of songs category. We chose the best one and considered the emotion of the song. In music personality traits, we have to generate a customize dataset. Firstly we created two survey from. One is interesting of music and another one is 20 question of human personality. We provide two survey from to the subject where we generate an id to merge the music interest form data and personality test form data for the same subject. The customize dataset categorized those song and personality. When a subject choose a music it predict how is its personality and vise-versa.

5.2 Conclusions:

Music is most important part of human life. Title should come after the song is completed and should reproduce a culmination of the lyrical content. In music title recommendation system, we have recommended a proper title based on song lyrics using topic model algorithm Latent Dirichlet allocation (LDA). Topic modeling provides methods for embryonic knowledge finding, finding relationships among data, understanding, and summarizing huge electronic archives. We believe that, this experiment is most helpful to generate the proper title from song lyrics. Emotion detection is an important research field in affective computing over the last decade. Emotion detection from song lyrics should be focused as a significant research issue in affective computing. This experiment used emotion ontology and machine learning methods to study the process of emotion classification. We have extracted 8 features from song lyrics and classify the music emotion. When we are listening a song, we fell various emotions there. Users use this information without classification in their search for songs, but the proposed system needs to meet their needs. We trust that researchers in associated fields may find the survey presented in this experiment helpful and advance the discussion in emotion reorganization from lyrics. In music personality traits, we have to generate a customize dataset.

5.3 Recommendations:

We recommend,

On experiment 1, produces satisfactory result. Therefore, we may use this module to recommend song title from lyrics stimuli.

On experiment 2, Music emotion recognizer will be able to find out the emotions of different songs of different genres. Finding this issue will help the system to predict the overall emotional state of a user.

On experiment 3, Music personality traits could be useful to find out the personality measurements of any user.

5.4 Implication for Further Study:

- Applying neural network for better recommendation.
- Applying pattern recognition algorithms to recognize the music emotion properly.
- Music personality traits will be useful to find personality measures which will lead to personalized recommender systems.

REFERENCES

1. Y. Hu, X. Chen , D. Yang, “Lyric-based Song Emotion Detection with Affective Lexicon and Fuzzy Clustering Method”, Proceedings of the 10th International Conference on Music Information Retrieval (ICMIR), 2009.
2. Hu, X., Downie, J.S, “Improving Mood Classification in Music Digital Libraries by Combining Lyrics and Audio”, In: Proceedings of Joint Conference on Digital Libraries (PCDL), 2010.
3. Laurier, C., Grivolla, J., Herrera, “Multimodal Music Mood Classification using Audio and Lyrics”, In: Proceedings of the International Conference on Machine Learning and Applications. ICMLA'08. Seventh International Conference (ICMLA), 2008.
4. Big Five personality traits, Link: https://en.wikipedia.org/wiki/Big_Five_personality_traits, [6th April 2018, 3.32 AM]
5. Jonathan Chang, David Blei , “ Relational Topic Models for Document Networks”, Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics, PMLR 5:81-88 (ICAIS), 2009.
6. Daniel Ramage, Evan Rosen, Jason Chuang, Christopher D. Manning and Daniel A. McFarland, “Topic Modeling for the Social Sciences”,2009.
7. David Newman , Edwin V. Bonilla , Wray Buntine ,” Improving topic coherence with regularized topic models”,2011.
8. Ricardo Malheiro,Renato Panda,Paulo Gomes,Rui Pedro Paiva. “Music Emotion Recognition from Lyrics: A Comparative Study”, 2013.
9. Dan Yang, Won-Sook Lee, “Music Emotion Identification from Lyrics”, 2009.
10. Edward Choa-Chun Kao, Chun-Chieh Liu, Ting-Hao Yang, “Towards Text-based Emotion Detection A Survey and Possible Improvements”, 2009.
11. Minh Kim, Hyuk-Chul Kwon, “Lyrics-Based Emotion Classification Using Feature Selection by Partial Syntactic Analysis”, 2011.

12. Ferwerda, Tkalcic, Schedl, “Personality Traits and Music Genre Preferences: How Music Taste Varies Over Age Groups”, In: Temporal Reasoning in Recommender Systems / [ed] Maria Bielikova, Veronika Bogina, Tsvi Kuflik, Roy Sasson, 2017.
13. Tomas Chamorro-Premuzic Adrian Furnham, “Personality and music: can traits explain how people use music in everyday life?”, 2010.
14. Use of emotions helps manager, Link: <https://www.mt.nl/management-team/gebruik-van-emoties-helpt-manager/4470>, [6th April 2018, 3.30 AM]

APPENDIX

Appendix A: Research Reflection

The purpose of this appendix is about research reflection. From summer 2017 semester we started our journey to research in this field. The main purpose of our research is song title recommendation, music emotion recognition and music personality traits. Also using our song title recommendation system module user can easily generate song title from lyrics, Music emotion recognition module will help the system to predict the overall emotional state of a user. Music personality traits could be useful to find out the personality measurements of any user.

A dataset that contain 200 song lyrics.

Link: [https://github.com/rafayetrafi/Music-Data-set-contains-200-song-](https://github.com/rafayetrafi/Music-Data-set-contains-200-song-lyrics)

PLAGIARISM REPORT

Checked By: www.plagramme.com

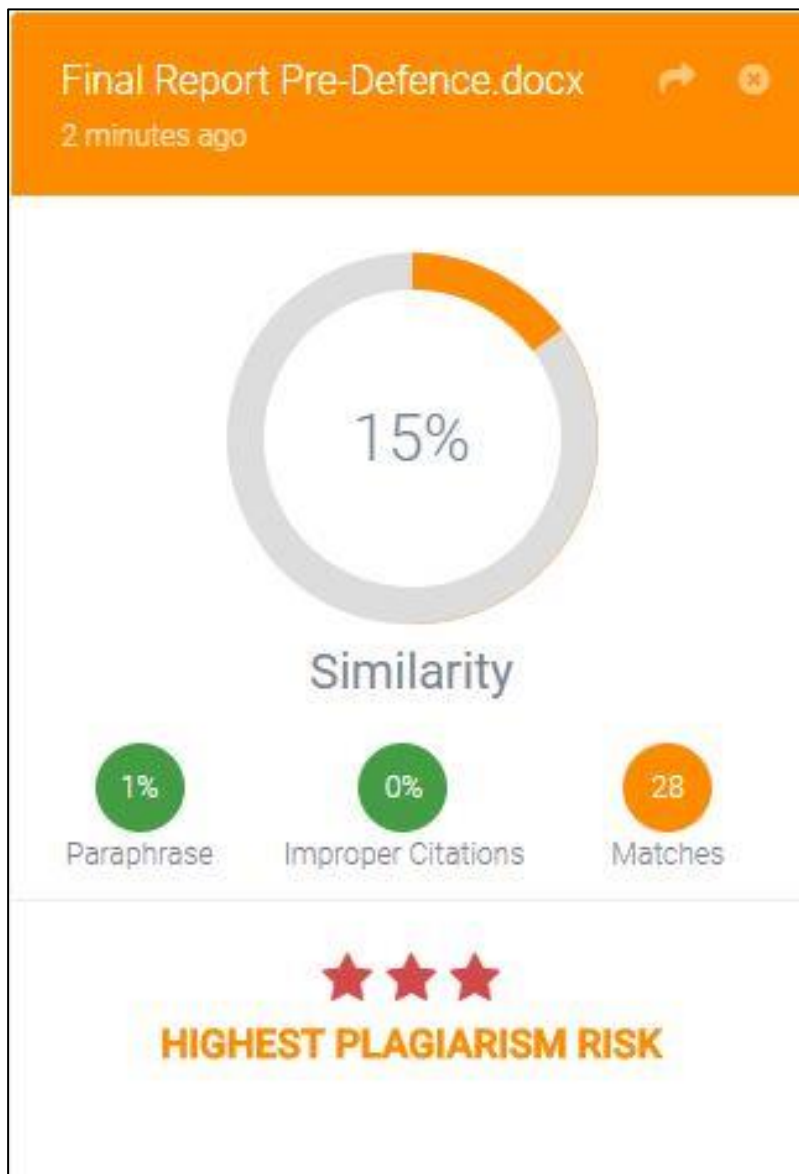


Figure: Plagiarism Report