

Mood Based Classification of Bangla Music Analyzing lyrical data using Bert Algorithm

*Capstone project report to be submitted in partial fulfillment of the
requirements for the degree*

of

Bachelor of Science in Computer Science and Engineering

by

**Fahimul Islam Shakil
171014035**

Maliha Mahajebin	Hafizur Rahman Fahim
171014031	171014054

Under the supervision of

Mohammad Rifat Ahmed Rashid



**COMPUTER SCIENCE AND ENGINEERING
UNIVERSITY OF LIBERAL ARTS BANGLADESH**

SPRING 2021

© *Fahimul Islam Shakil, Maliha Mahajebin, and Hafizur Rahman Fahim*
All rights reserved

DECLARATION

Project Title Mood Based Classification of Bangla Music Analyzing lyrical data using Bert Algorithm
Authors Fahimul Islam Shakil, Maliha Mahajebin, and Hafizur Rahman Fahim
Student IDs 171014035, 171014031, and 171014054
Supervisor Mohammad Rifat Ahmed Rashid

We declare that this capstone project report entitled *Mood Based Classification of Bangla Music Analyzing lyrical data using Bert Algorithm* is the result of our own work except as cited in the references. The capstone project report has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

shakil

Fahimul Islam Shakil
171014035

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

maliha

Maliha Mahajebin
171014031

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

fahim

Hafizur Rahman Fahim
171014054

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Date: June 14, 2021

CERTIFICATE

This is to certify that the capstone project report entitled **Mood Based Classification of Bangla Music Analyzing lyrical data using Bert Algorithm**, submitted by **Fahimul Islam Shakil** (Student ID: 171014035), **Maliha Mahajebin** (Student ID: 171014031) and **Hafizur Rahman Fahim** (Student ID: 171014054) are undergraduate students of the **Department of Computer Science and Engineering** has been examined. Upon recommendation by the examination committee, we hereby accord our approval of it as the presented work and submitted report fulfill the requirements for its acceptance in partial fulfillment for the degree of *Bachelor of Science in Computer Science and Engineering*.



Mohammad Rifat Ahmed Rashid
Assistant Professor

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Dr. Syed Akhter Hossain
Professor and Head

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Place: Dhaka

Date: June 14, 2021

ACKNOWLEDGEMENTS

We would like to express our deep and sincere gratitude to our research supervisor, *Mohammad Rifat Ahmed Rashid*, for giving us the opportunity to conduct research and providing invaluable guidance throughout this work. His dynamism, vision, sincerity and motivation have deeply inspired us. He has taught us the methodology to carry out the work and to present the works as clearly as possible. It was a great privilege and honor to work and study under his guidance.

We are greatly indebted to our honorable teachers of the Department of Computer Science and Engineering at the University of Liberal Arts Bangladesh who taught us during the course of our study. Without any doubt, their teaching and guidance have completely transformed us to the persons that we are today.

We are extremely thankful to our parents for their unconditional love, endless prayers and caring, and immense sacrifices for educating and preparing us for our future. We would like to say thanks to our friends and relatives for their kind support and care.

Finally, we would like to thank all the people who have supported us to complete the project work directly or indirectly.

Fahimul Islam Shakil, Maliha Mahajebin and Hafizur Rahman Fahim

University of Liberal Arts Bangladesh

Date: June 14, 2021

To my mother
FA Mother name
and father
FA Father name

–Fahimul Islam Shakil

To my mother
SA Mother name
and father
SA Father name

–Maliha Mahajebin

To my mother
TA Mother name
and father
TA Father name

–Hafizur Rahman Fahim

ABSTRACT

Music mood classification is very emergent for the musician, singer, listener. But along with audio(MP3) and Video(MP4), music lyrics are also vital components. Many researches and works have already been done for English, Hindi music mood, and mood base recommendation. There are few works done with Bangla music mood and recommendation system based on Bangla lyrics. The objective of this capstone project is to classify the Bangla Music Mood and Recommendation system based on Mood along with the genre. We are classifying 5 different Bangla music categories like Romantic, sad, happy, etc. We are using a dataset consisting of 4088 Bangla song lyrics that have different Moods and genres. Our proposed model is to do multiclass classification using Bangla BERT(Bidirectional Encoder Representation from Transformers) base. The NPL(Natural Language Processing) BERT model is open-source and it helps to achieve good accuracy, that makes it great for building models.

Keywords: sentiment analysis, Deep learning,Bert, Bangla music mood.

Contents

1	Introduction	1
1.1	Problem Statement:	1
1.2	Aims and Objectives:	1
1.3	Motivation and contribution:	2
2	Literature Review	3
2.1	Sentiment Analysis:	3
2.2	Music Based Sentiment Analysis:	3
2.3	Lyrics Sentiment Analysis:	4
2.4	Recommendation System:	5
3	Project Specification	6
3.1	Dataset description	6
3.2	Deep Learning Approach	7
3.3	Application specification	8
3.4	Specific Requirements	8
3.4.1	Software Requirements (for user)	8
3.4.2	Hardware requirements (for user)	9
3.4.3	System Attributes	9
4	System Design	10
4.1	Exploratory Data Analysis	10
4.2	Use case Diagram	12
4.3	Data Flow Diagram (DFD)	13
4.4	Sequence Diagram	14

5	Methodology	15
5.1	Overview	15
5.2	Deep Learning Architecture	15
5.3	Software Development Process	16
5.3.1	Incremental Mode	16
5.3.1.1	Requirement analysis	17
5.3.1.2	Design & Development	17
5.3.1.3	Testing	17
5.3.1.4	Implementation	17
5.4	Design Pattern	18
5.4.1	Model-view-controller	18
5.4.1.1	Model	18
5.4.1.2	View	18
5.4.1.3	Controller	19
5.5	Technology Used	19
6	Project Contribution	20
6.1	Overview:	20
6.2	Proposed System:	20
6.3	User Interface:	20
6.4	Results:	23
6.4.1	Deep Learning Model Results:	23
6.4.2	Application Outputs:	24
7	Result and Analysis	26
7.1	Deep Learning Model Performance Analysis	26
7.2	Security Analysis	28
7.2.1	Security	28
7.2.2	Privacy	29
7.2.3	Integrity	29
7.3	Evaluation	29
7.3.1	Interference Time	29
7.3.2	Application performance (Time Complexity)	30
7.4	Cost Analysis	30
7.4.1	Hardware	30

7.4.2	Software cost	31
7.5	Social Legal & Ethical Issues	31
8	Conclusion and Future Works	32
8.1	Scope for Future Work	32
8.2	Conclusion	32
8.3	Gantt Chart	33
	Bibliography	34

List of Figures

3.1	picture of genre categories in our dataset	6
4.1	graph of mood categories in our dataset	10
4.2	graph of token count in our dataset	11
4.3	use case diagram of our project	12
4.4	Data flow diagram of our project	13
4.5	Sequence diagram of our project	14
5.1	BERT Architecture	16
5.2	Architecture of our full application process.	17
5.3	Software Development Process	18
6.1	Home page of our website	21
6.2	Five Mood categories in our website	21
6.3	Songs are showing in cards based on Mood	22
6.4	Songs are showing in cards based on Mood along with audio	22
6.5	Details view of card with lyrics and audio clip.	23
6.6	Showing post and get data from API	24
6.7	Showing Json format data from our REST API	25
7.1	Training and validation accuracy graph	27
7.2	Confusion Matrix	28
8.1	Gantt Chart of our project	33

List of Tables

7.1	suggestions with the purpose of fine-tuning.	26
7.2	Classification Report	27

Chapter 1

Introduction

1.1 Problem Statement:

Music is very important in our day to day life. Many researches have already been done for the classification of english songs by using mood. On the other hand, few works done for classifying Bangla music based on mood. Bangla music has a very rich music culture but very few noticeable works are done for Bangla music classification based on mood. As much research has been done on Bangla audio songs but not based on Lyrics. So we decided to work on detecting the mood of Bangla songs based on lyrics and build a software that will help people to listen to songs according to their mood.

1.2 Aims and Objectives:

Music plays a very important rule for all people around the world. Music is the only language that everyone can understand. We can't imagine a party, movie, or event without music. Researchers find that music enhances brain functioning. . This is like when you have a great relationship with our mood to be happy, you can listen to happy songs and when you are in a sad mood or losing someone, this is helpful to listen to sad songs that make you feel much better. When we are listening to sad or happy songs many applications recommend similar mood type songs. This feature is growing and gaining popularity far and wide. The aim of this project is to create an application that classifies and predicts songs mood by analyzing the lyrics and suggest the similar type of genre and mood songs.

1.3 Motivation and contribution:

Our Main motivation of this project is to build a software that will classify Bangla songs mood based on lyrics as well as people can listen to songs depending on their mood. We tried to classify both audio and lyrics based on songs. In this project we are proposing a multi class classification model by using BERT(Bidirectional Encoder Representation from Transformers), a very popular NLP(Natural Language Processing) sequence-to-sequence model. We are using deep learning library keras to classify our five different moods such as Romantic, sad, happy, religious,patriotism.Google made BERT model as an open source to improve the efficiency of NLP and the efficiency is much higher the the other classification model like naive bayes , logistic regression. In our project we are using the BERT model,TensorFlow hub, pre-training, small BERT's, Kaggle, Adam(Adaptive moments) used for fine tuning.

Chapter 2

Literature Review

2.1 Sentiment Analysis:

Sentiment analysis is a popular topic in this era. Many researches have been done and still going on this topic. A huge survey has been done on sentiment analysis's recently updated algorithm by some students of Ain Shams University in 2014. They analyzed 54 articles and cites. They give a result by analysing that for open field research, SC and FS algorithms have more enhancements. Machine Learning algorithm that most frequently used to solve SC problems are- Native Bayes, Support Vector Machines. They also analysed most of the papers are on English language, but the research on other languages are growing up. [1]

Another survey on the challenges on Sentiment analysis was done by Hussein, et al. in 2016. This survey was based on 47 papers and they discussed, the importance and effects on of sentiment 191 analysis challenges in sentiment evaluation which is based on two com- 192 parisons. The result of this analysis shows another essential factor to recognize the 196 sentiment challenges. This is domain-dependence. In all type of reviews the popular challenge is the negation challenge. [2]

2.2 Music Based Sentiment Analysis:

In this era many works have been done on Music analysis. Some have been done on the basis of audio signal, some on the basis of Lyrics and the classification have been done on various topics. Such as artists, categories, mood and so on.

Some work had been done on Bangla song Review to research on the acceptance of a young star's song. They collected the comments from the comment section of a Youtube music video of a Bangla young star and researched on this if this that if the comments were positive or negative. They used a lexicon based backtracking approach on each sentence of the dataset.[3]

Some students of Daffodil International University Bangladesh had also done some work on Bangla Song. They had done Bangla song genre classification using a neural network. In that paper, they proposed a deep learning model by using the Sequential using Sequential Model of the deep learning library Keras to classify 6 different Bangla music genres. They used audio signals for that purpose and collected MP3 dataset of Bangla songs.[4]

2.3 Lyrics Sentiment Analysis:

Automatic mood classification using tf*idf based on lyrics was the earliest work done by van Zaanen, M.; Kanters, P.H.M. in 2010. They divided the mood of songs into different classes such as sad, happy, angry, relaxed and so on. They used tf*idf which helped to emphasize on the emotional words of lyrics of a song and automatically describe the mood.[5] Mood Based Classification of Music by Analyzing Lyrical Data Using Text Mining was done by Amity University Uttar Pradesh, Noida – INDIA. They also used Native Bayes classifier for text mining and getting the mood of the song. They also used some audio features of the song.[6] Sebastian Raschka had done some work on Predicting the mood of music from song lyrics using Bayes classifier. They developed a web app to perform mood prediction by giving the artist name and song title. They collected a huge number of songs from the 'Million song dataset' and trained the data using Bayes theorem and classified the songs in two moods, happy and sad. The results have shown that a naive Bayes model can predict the positive class with high precision. To filter a large music library for happy music with a low false positive rate, this model is useful.[7] Similar work is done on a number of Bangla song lyrics by some students of United International University Bangladesh. They have worked on Music mood classifiers on Bangla song lyrics using Native Naive Bayes classifier. They also divided the songs into two moods, sad and happy. They showed an optimal result

that the sad songs are increasing day by day as peoples' emotion is going down with the era.[8]

2.4 Recommendation System:

Recommendation System (RS) is a popular topic having a major research interest. These researches aims to help users by suggesting them to find items online according to their interest.

A recent paper which was published on 2021 which gives a brief Overview,showing the Research Trends and suggests Future Directions of recommendation system.In this paper it is analysed that,till now most of the researches in RS is focused on CF and knowledge-based approach.China is in top among all the countries who have done research related to recommendation system.Recommendation system related most of the papers are published in IEEE.During the period of 2013–2014, most of these researches have been done.[9]

Some other Survey of Music Recommendation Systems have been done by some people in 2012. In their paper,basic metadata-based model is explained. They also described about two popular music recommendation approaches which are collaborative filtering and content-based model.This models have some limitations too. By using a hybrid model, using the advantage of this two models,the result can be improved. [10]

Chapter 3

Project Specification

3.1 Dataset description

: Our dataset is a collection of 4000 Bangla songs from different Bangla music websites. I scraped dataset from various Bangla song's lyrics websites and labeled the dataset. Our dataset is in .CSV format. There are four columns in the dataset. They are title, category, lyrics and mood. Title is the song's name published by the copyright author. There are 20 plus categories in the data set; those are mainly the genre of the songs. Fig 3.1. In the lyrics column, there is the full version of the

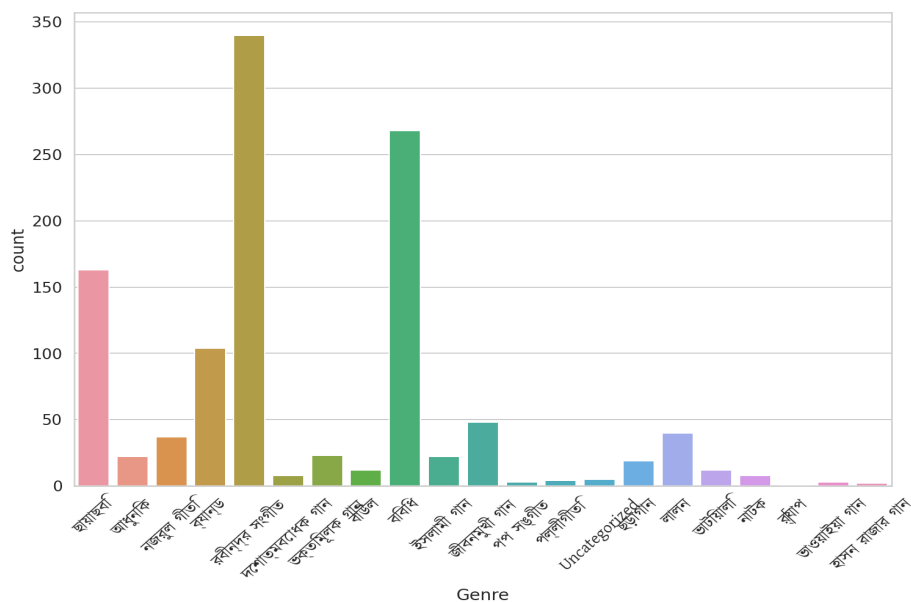


Figure 3.1: picture of genre categories in our dataset

song's lyrics. Mood column is the important part of this dataset. There are four Bangla songs moods we identify from various Bangla and English songs mood identification research done previously and the moods are romantic, sad, happy and relaxation[11]. There is some unbalance in the data-set because of in Bangla music there are little amount of relaxation songs we find out from Bangla music history

3.2 Deep Learning Approach

BERT (Bidirectional Encoder representation from Transformers) is a new paper distributed by specialists at Google computer based intelligence Language. It made a ruckus to AI group of people introducing best in it's class brings about various NLP(natural language processing) task, including question answering, Normal Language Surmising (MNLI),next sentence pr edition,Musk language identification and others. [12]

The BERT key is dedicated to development by applying the Bi-directional representation of transformer, a mainstream considered model, to the visualization of language.It has no past effort to see a series of books while preparing from left to right or left to right, right to left. The results of this paper show that bidirectionally prepared language models can have a deeper sense of language settings and flows than single-heading language models. In this treatise, experts elaborate on a new method called Covered LM (MLM) that allows bi-directional preparation in an already anomalous model.

How it works: BERT uses transformer, which is a consideration system that can learn the logical relationship between words or sub-words in learning content. In transformer's vanilla structure, the transformer contains two components — an encoder that examines the substance data and a decoder that conveys a gauge for the task. Because BERT goes likely create a model, simply the encoder system is fundamental. Google described Transformer's feature-by-feature in a document.[13] Instead of directional model that continuously reads content information (from left to right or from right to left), the Transformer encoder examines the entire gathering of words right away. This way is considered bidirectional, But it is more accurately say that is omnidirectional. This brand permits the model sample to get used to adjusting a word dependent at the entirety from its environmental factors like

left and right of the word. The schematic below is undeniable level portrayal from Transformer encoder. The information is a consecutive sequence of token counts, first inserted into a vector and then processed by the nervous system. The throughput is a series of H-size vectors, where each vector is compared to an information token with similar files.

3.3 Application specification

“Bangla Music” is a collection of historical and modern songs web applications we are developing. This website is based on music moods we identify and selected though our research those are romantic, sad, happy and relaxation. There is a small amount of work and research done for Bangla music. There is no research we find on Mood base Bangla song’s lyrics identification. However, there are many applications for Bangla music or lyrics. But rare to find audio along with lyrics in a single application. For this we decided to develop this application where users can find audio songs whatever mood they are in and also can read the lyrics. There are mainly three parts in this application. Our website design, development, and creating a machine learning model. Creating a machine learning model we used BERT base uncased pre-trained model for training and validating our dataset. For developing APIs, we are using Fastapi. Fastapi is based on standard python types hints and it’s a modern, faster and web framework for building APIs using python (ref: Fastapi website). Through this admin can post, delete, get, update Bangla music songs and lyrics. All of the data is stored in a database. For our design purpose we are using React library and ES7. We are fetching data from the APIs depending on lyrics mood and showing them into our website along with audio.

3.4 Specific Requirements

3.4.1 Software Requirements (for user)

This is a very simple application for the user to use. There is no need for additional software to run this web application. Users just need a browser like chrome, Firefox and Microsoft edge etc and internet connection. where they can browse the website. Otherwise there is no need to install any software for the user to run, hear or read the music and lyrics from this web application.

3.4.2 Hardware requirements (for user)

This is a simple web application. Users can use this application on mobile, computer and also any other devices that support the html5 version. There are many browsers like internet explorer version4 doesn't support html5 version in case they need version 9 to run html5. Users can browse this Bangla Music application by any kind of android mobile or tv, iOS devices.

3.4.3 System Attributes

Every good software has some attributes mainly the Acceptability, Dependability and security, Efficiency and Maintainability. Our web application almost meets all the standard attributes. This web application is very simple to use and has no complex functionality. Bangla music application is designed for all platforms of users. Mobile and computer users can hear and read lyrics very easily by just a couple of clicks. This is very similar to other web applications that we are using in our everyday life. This is a dependable software because if unfortunately, our system fails there is no physical and economical damage. We are always making a backup database and API system for any kind of system failure and it will support us to stabilize our website. This is a fully private database and API so malicious users cannot get access to the database or any kind of information, only the admin can access the information. The efficiency of our software is much better. This is a web application so there is no need for physical memory for use. When users access the website data is shown immediately because this is a very light weight website. Everything we are showing to the user from an API for this it works very faster and the user will get their desired music. The maintainability of this software is simply not so critical. Through user needs we can change any kind of features easily and the most important thing is that the admin can add, delete, update and post Bangla music and lyrics very easily by the need of user demand.

Chapter 4

System Design

4.1 Exploratory Data Analysis

Due to the regulation of dopamine in the brain, sounds like music and noise have a major impact on our mood and emotions. Dopamine in the brain is a neurotransmitter that actively participates in emotional behavior and mood regulation. This is the fact that we choose to work in a musical mood.

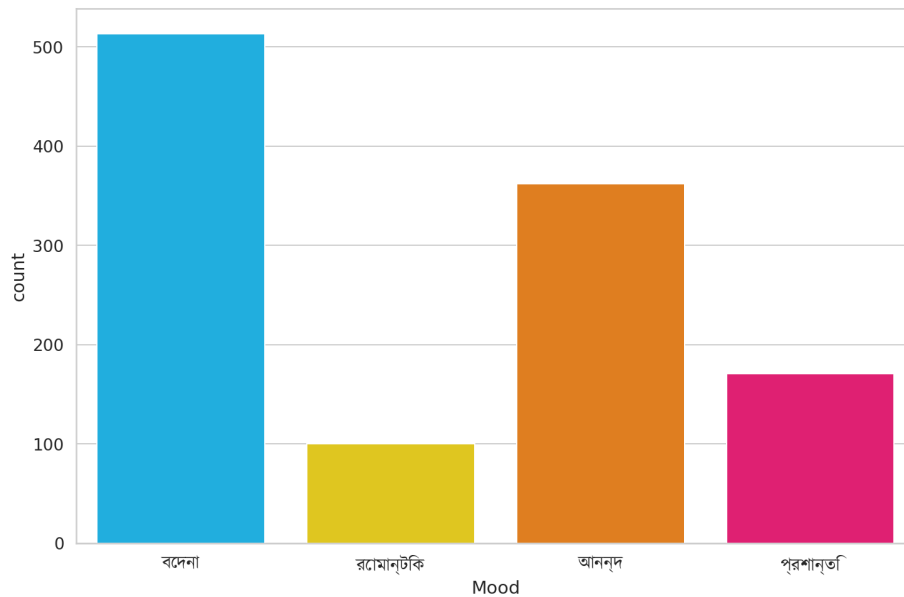


Figure 4.1: graph of mood categories in our dataset

There are 21 categories in our dataset. In our dataset we select four kinds of mood. romantic, sad, happy and relaxation. There is a huge unbalance in our dataset. Fig 4.1 showing the amount of songs we collect from various Bangla songs websites by scrape. sad(513),romantic (362), happy (168), relaxation (100) are amount of songs we labeled for training purpose. sad and romantic (362) are the highest amount of data we collect. But this is difficult to find happy (168), relaxation (100) songs lyric from Bangla songs documentary. We are still trying to increase the dataset and balance the dataset for getting a good performance. Token count defines the length of the individual data. We set the maximum length 512.

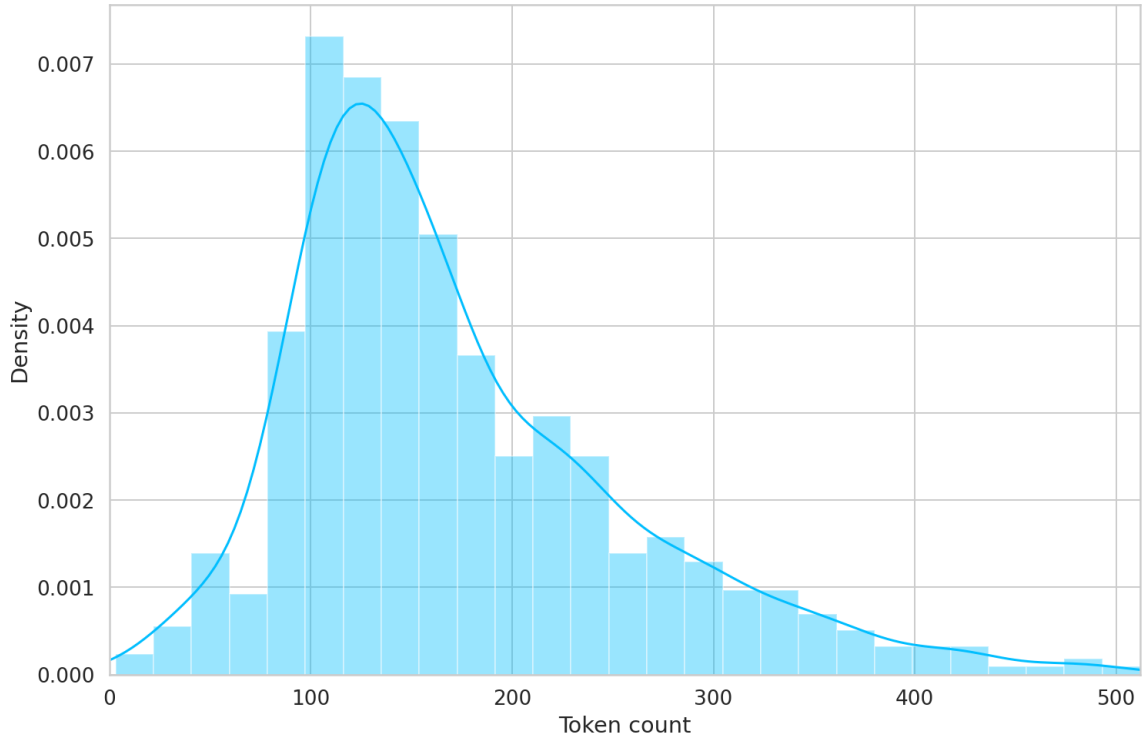


Figure 4.2: graph of token count in our dataset

4.2 Use case Diagram

In this application only the admin can post lyrics. After posting lyrics using FastAPI the ML model classifies the mood and stores the data into a database. If any change is needed in the application admin can update (edit, delete) the data and fetch the data using the get method and make a connection between the API and interface. Users can access the web application only. Web applications contain categories, lyrics and audio clips. There are five categories: sad, romance, happy, relaxation and mixed songs. Fig:4.3 showing the use case diagram of our project.

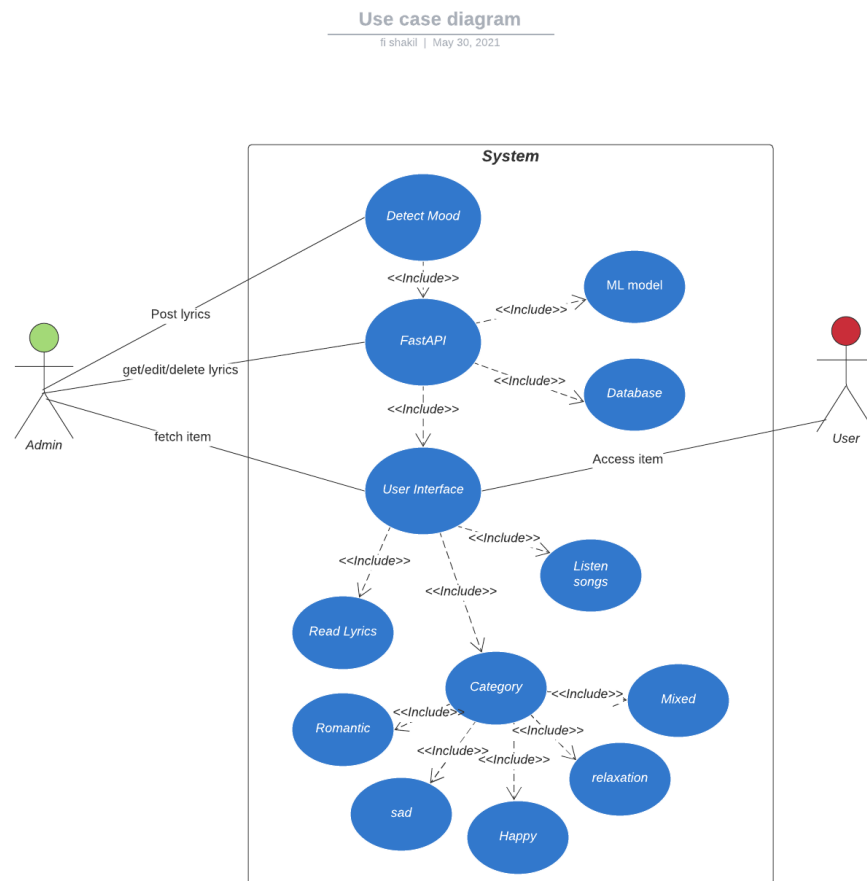


Figure 4.3: use case diagram of our project

4.3 Data Flow Diagram (DFD)

Admin is posting data using FastAPI. Using FastAPI we connect with the database and return the data on FastAPI. For classifying songs mood ML model is processing lyrics mood and pushing the data again into the database. Using the get method admin fetching the data from the database also updates the data if any need for edit and delete. User accessing the data from the user interface. User interface is showing the data by fetching the data from the database. Fig:4.5 showing the DFD of our project

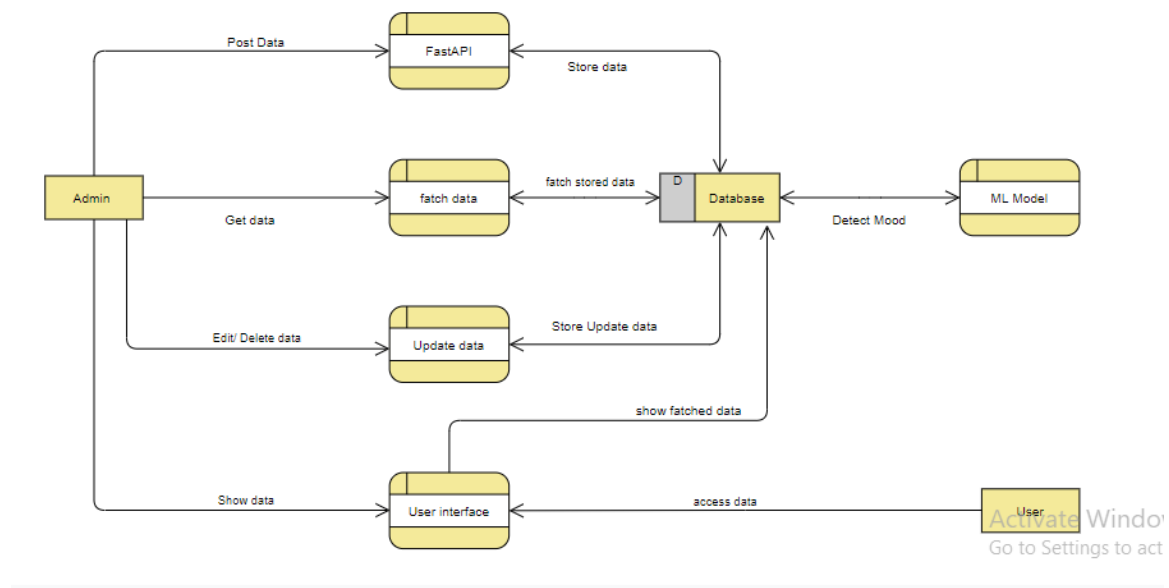


Figure 4.4: Data flow diagram of our project

4.4 Sequence Diagram

In our application, at the first time period admin has to post the data using FastAPI. Posted data are classifying by the ML model we created earlier. Then return the classified data in FastAPI. Also admin can get, put, delete the data using FastAPI. Admin pushes the data into the database after classifying the mood and also updates the data if any need. Then return the data from the database. User interface is fetching the data form API and showing the data into the interface. From the user interface, users can find categories. Can read lyrics and listen to audio songs. Fig:4.5 showing the sequence diagram of our project.

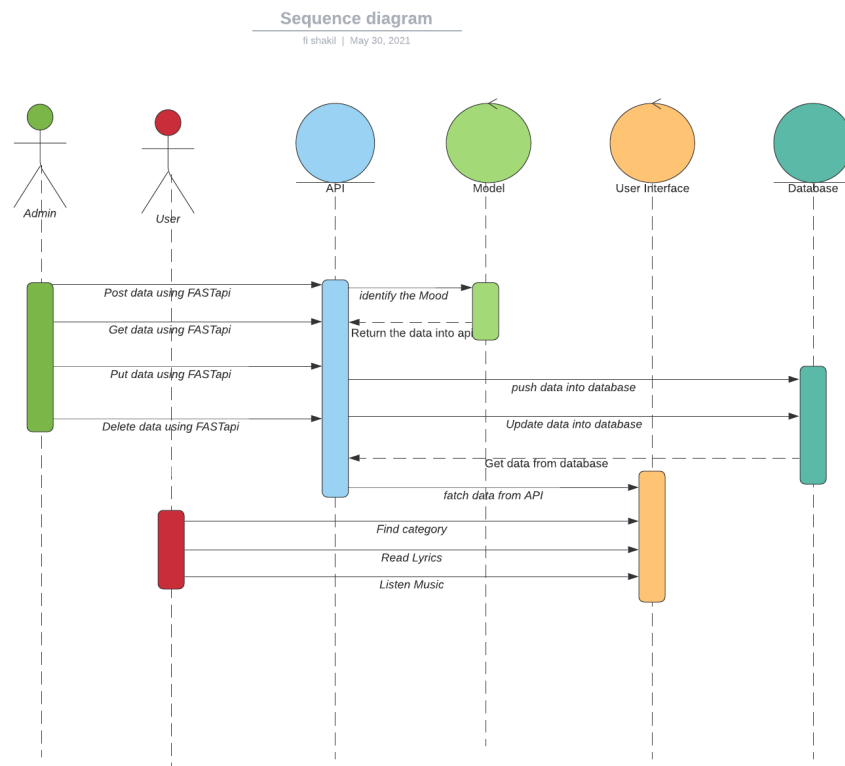


Figure 4.5: Sequence diagram of our project

Chapter 5

Methodology

5.1 Overview

We are going to describe here what method we used to develop our software. The methodology we are applying is briefly described in this chapter. We are going to describe our architecture. How our software developed etc.

5.2 Deep Learning Architecture

We used Bangla BERT BASE Uncase Model as the algorithm of our project. Here we are describing the BERT Model first. We are giving description of the NLP framework according to the research team behind BERT Model: “BERT stands for Bidirectional Encoder Representations from Transformers. It is designed to pre-train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context. As a result, the pre-trained BERT model can be fine-tuned with just one additional output layer to create state-of-the-art models for a wide range of NLP tasks.” BERT is a “deeply bidirectional” model. Bidirectional means learning information from both the left and the right side of a token’s context during the training phase. The bidirectionality of a model is important for truly understanding the meaning of a language. The BERT architecture builds on top of Transformer. We currently have two variants available: 1.BERT Base 2.BERT Large

Fig:5.1 is the Architecture of BERT model.

We used BERT Base uncase in our model. This model contain 12 encoder layers, 768 feed forward networks and 12 attention heads. A special token [CLS] is

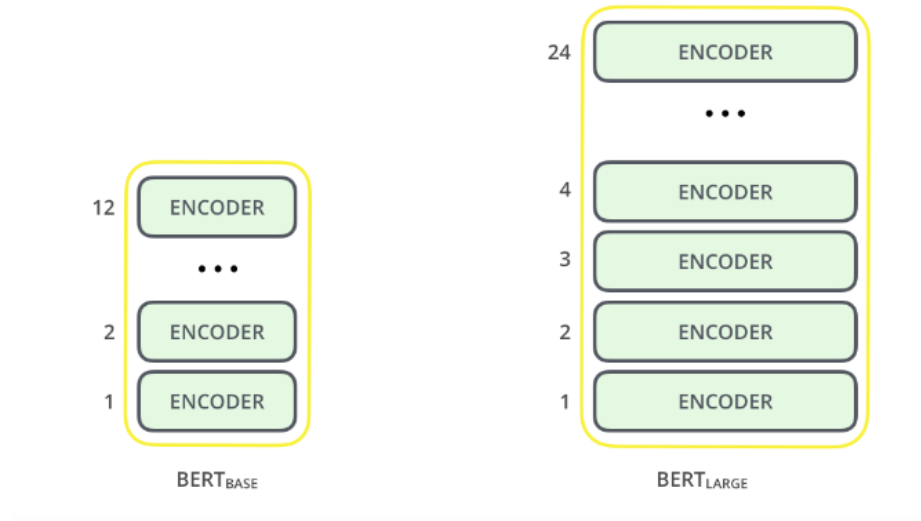


Figure 5.1: BERT Architecture

provided for the first input token, the reason for which will become apparent later. CLS stands for classification here.

In this architecture Fig:5.2 we basically described our work flow. First of all we collected the data and labeled the dataset. There is around 4000 songs in our dataset. We labeled all the song's mood. Then in 2nd step we do pre-processing of data which include data cleaning, tokenizing and padding.

After that we trained the data using our Bangla BERT Base Model. After training data we do the evaluation of it. We tested the data.

Lastly, we used the API to make our website. We post and get data from API and showed it to the website. Thus, finally our website shows our expected result.

5.3 Software Development Process

5.3.1 Incremental Mode

We used an Incremental model to develop our software. It is a process where the to-do list of project is divided into multiple standalone modules of SDLC. Each module goes through a process which includes some phases- 1. Requirements 2. Design 3. Implementation 4. Testing After each module is completed, it is added to the previous release. Until the complete system is achieved, the process continues.

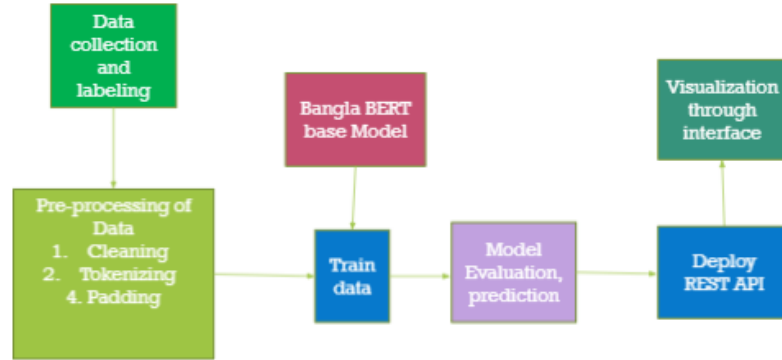


Figure 5.2: Architecture of our full application process.

5.3.1.1 Requirement analysis

In the first phase of the incremental model, we analysed the problem. We gather information , read papers and research on Sentiment analysis to understand the topic. Then we thought about our idea, what things are required for our project, such as algorithm, model etc.

5.3.1.2 Design & Development

We started to design the UI of our software in this phase. We thought about how it could be more user friendly. After the design of the system functionality, the development method is finished with success. We developed our website which detect mood of songs and recommends songs also.

5.3.1.3 Testing

The testing phase is used to check the performance of each existing function. We had a dataset of around 4088 songs. We divided the songs and took 20% of the songs to test how our website works.

5.3.1.4 Implementation

In this phase, coding is done of the development system. Two things are done here-the final coding that design in the designing and development phase as well

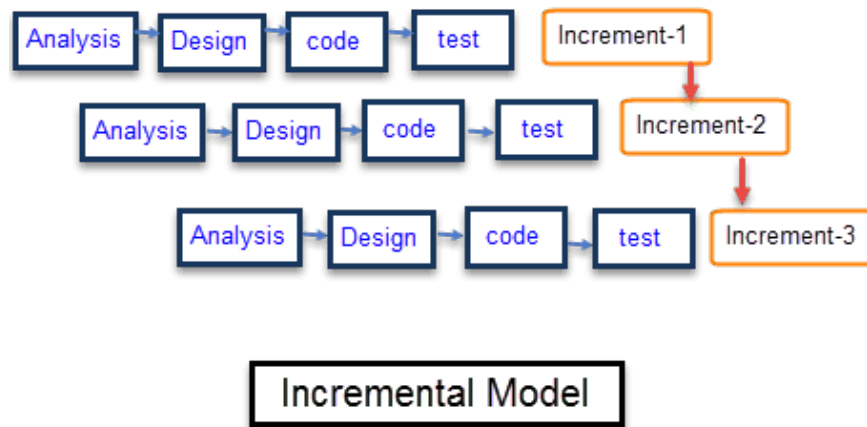


Figure 5.3: Software Development Process

as tests the functionality in the testing phase. The final upgraded product we get after finishing this process. So at the end we have done our implementation finally.

5.4 Design Pattern

5.4.1 Model-view-controller

The Model-View-Controller (MVC) is a structural pattern that consists of three interconnected parts which are model ,view and controller . It provides the basic structure for designing an application for desktop or mobile and also for web applications.

5.4.1.1 Model

it is the specific representation of the information with which the system operate .logic ensures the integrity of the data and allows to derive it.

5.4.1.2 View

Represents the Model in a suitable format to interact and access the data, usually called “user interface”.

5.4.1.3 Controller

controller is a pattern which manipulates data using the Model component and works with the Views to render the final output.

5.5 Technology Used

For developing this project, we are using a transformer library by hugging face. Transformers (some time ago known as pytorch-transformers and pytorch-pretrained-bert) gives universally useful designs (BERT, GPT-2, RoBERTa, XLM, DistilBert, XLNet...) for Natural Language Understanding (NLU) and Natural Language Generation (NLG) with over 32+ pretrained models in over 100 dialects and profound interoperability between Jax, PyTorch and TensorFlow. From the transformer we are using BERT model architectures for developing our desired model. We design our user interface using the Redux and React.js library. For website development purposes we make an API using FastApi and store our data into PostgreSQL.

Chapter 6

Project Contribution

6.1 Overview:

Here we are going to describe how our project contributes to the research of music mood analysis. We will also discuss how it is user friendly, how the UI works. We are also showing the result outputs in this chapter.

6.2 Proposed System:

We are going to build an application which detects the mood of the song either it is audio or lyrics based. Our software will be used by people to listen to songs according to their moods. This system will help people to know the mood of a song and get suggestions of songs according to that mood. They will have a website where they will go and put a song, the site will automatically detect the mood. This will be also helpful to know the statistics of moods of Bangla songs.

6.3 User Interface:

The interface of Bangla Gaan is a user -friendly website and very simple to use. There is no need to Sign up just browse the website in Google and read, listen to your favorite Bangla song.fig:6.1

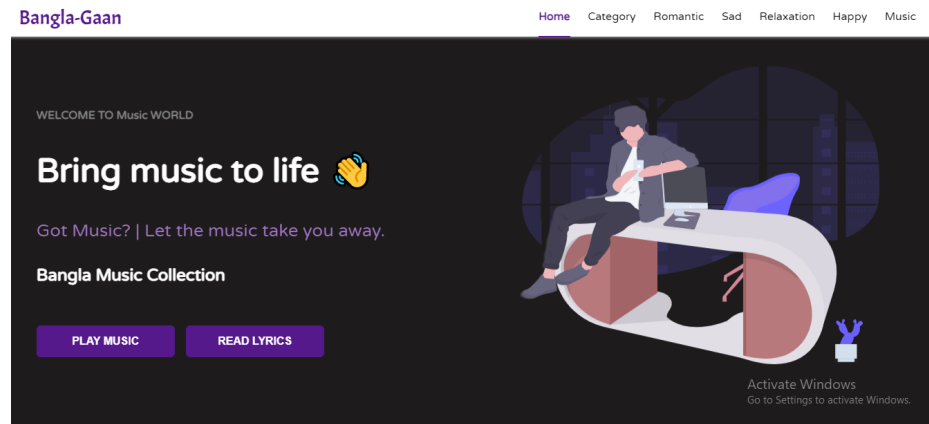


Figure 6.1: Home page of our website

In this website we have five categories of songs are romantic, sad, happy, relaxation and mixed songs.fig:6.2

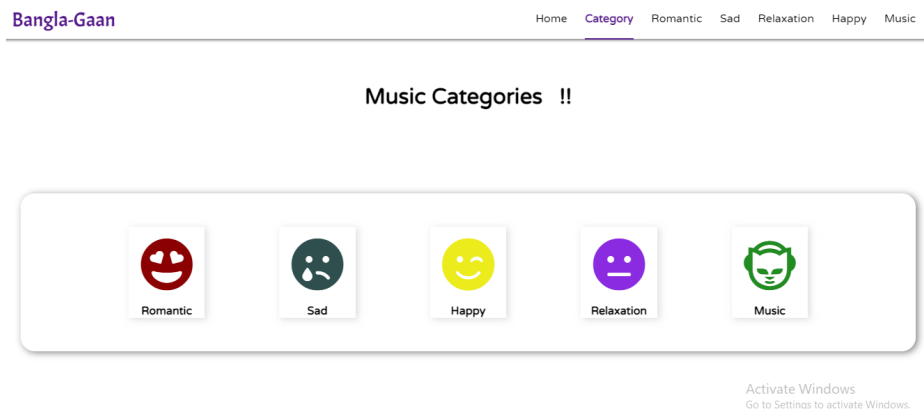


Figure 6.2: Five Mood categories in our website

Users will get all categories of songs in the navigation bar and also in the categories page. In the main page all categories contain some song cards. Those cards contain the title, genre and mood of the songs.fig:6.3&6.4

Romantic Songs



Figure 6.3: Songs are showing in cards based on Mood

Relaxation Songs

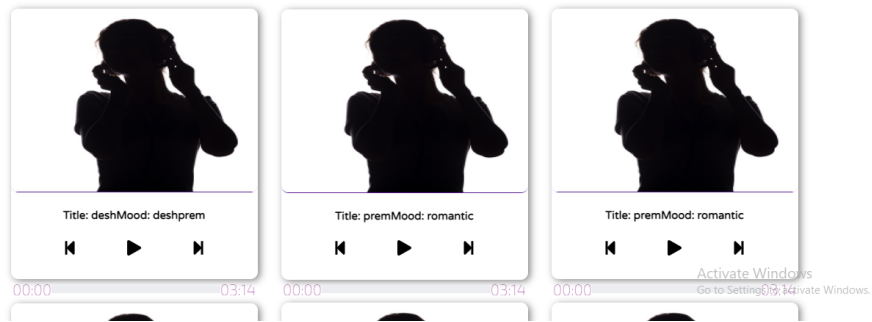


Figure 6.4: Songs are showing in cards based on Mood along with audio

A details page will open by clicking in the card. This details page contains lyrics, audio, genre, title. At the end of the page there are some suggested songs of the same mood. fig:6.5

Users can also play audio music on the cards. All categories have similar functionality. This is very easy to use and needs a couple of clicks to get the songs.

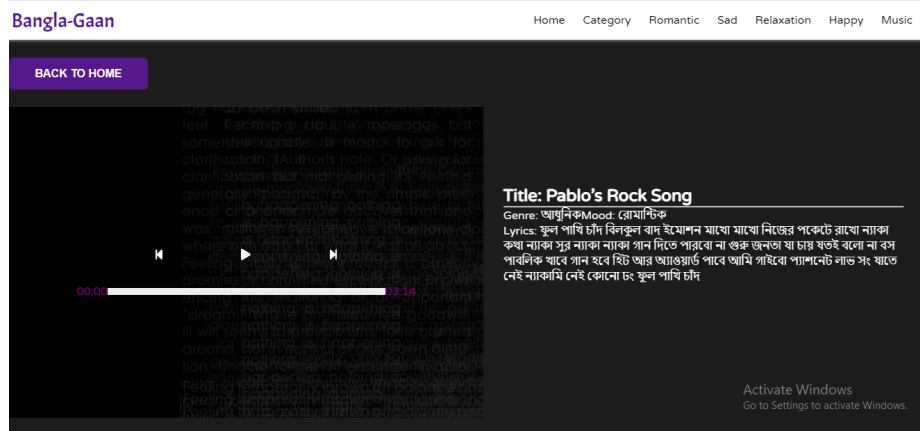


Figure 6.5: Details view of card with lyrics and audio clip.

6.4 Results:

6.4.1 Deep Learning Model Results:

For building this model we prepare our dataset and label it into four categories. We build the model based on the moods of a song. For creating the model first we are using a classifier that can classify the lyrics moods. We used the BertModel and built our lyrics classifier that can identify the moods of a lyrics. Our classifier appoints a large portion of the truly difficult work for BertModel. We use a dropout class for certain regulations and an affiliate class only for our benefit. Note that the last layer returns a rough yield because the unfortunate operation of cross-entropy in PyTorch is what it takes to work. we Apply a softmax function to the output to get the predicted probabilities of the trained model. In order to recreate the learning process from the BERT document, we use Hugging Face's AdamW optimizer. Correct the weight loss so that it looks like the original document. We will also use a linear programmer without a heating step. For training and validation, we set the number of epochs, batch size and learning rate . When the model is created this can identify the mood of a song's lyric.

6.4.2 Application Outputs:

Our application is successfully showing the data from the API we built for this project. When an admin posts data using the FastAPI, this data is sent to the model we created and this will identify the mood of the songs and store the data into a database.

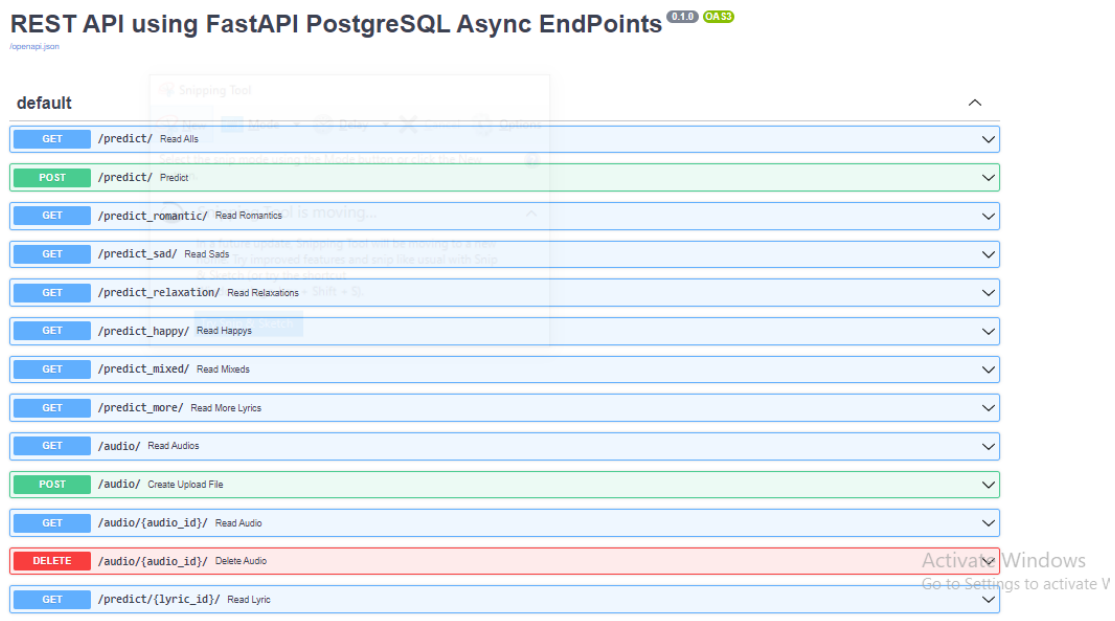


Figure 6.6: Showing post and get data from API

Our application fetches the data from the API (json format fig:6.7) and shows the outputs in a website. Outputs contain the lyrics, title, genre, mood, audio clip. By browsing the website, users can listen and read Bangla songs.

```

    confidence: 0.5405130386352539,
  },
- {
  id: 7,
  sentiment: "রোমান্টিক",
  text: "ভাড়ার টাকা না দিতে পারায় ঢাকার কেরানীগঞ্জে এক প্রতিবন্ধী নারীকে বাস থেকে ফেলে দেওয়ার ঘটনায় সমালোচনা মুখর নেটিজেনরা",
  title: "sad",
  genre: "sad",
  confidence: 0.5405130386352539,
},
- {
  id: 8,
  sentiment: "বেদনা",
  text: "শীতের রাতে আঁধার টুঁয়ে ছুটছি আমি ভাসছি নাও এ, কোন কথা তে তোমার আমার এ অভিমান প্রহর গুনে, চাঁদের আলো মেঘের দলে যে  

  লিখছি তোমায় গানের সুরে একলা হওয়া মধ্যরাতে, ভাবছি শুধুই তোমার আমার এ গল্প-গান।  

  শীতের রাতে আঁধার টুঁয়ে দাগ কেটেছে মন কাগজে, তোমার আমার এ অভিমান রোজ সুর লিখে যায় পাতার ভাঁজে, সাঁঝের বাতি নিভলো ভোরে  

  লিখছি তোমায় গানের সুরে একলা হওয়া মধ্যরাতে, ভাবছি শুধুই তোমার আমার এ গল্প-গান।",
  title: "লিখছি তোমায় ",
  genre: "আধুনিক",
  confidence: 0.9980863332748413,
},
- {
  id: 9,
  sentiment: "বেদনা",
  text: "প্রতিটি ঘরে আজ ঢল নেমেছে ঢল নেমেছে আকাশে ইশান কোলে মেঘ জমেছে মেঘ জমেছে যে মেঘে ধ্বনি তোলে বৈশাখের বৈশাখেরি যে মেঘের  

  title: "প্রতিটি ঘরে আজ ঢল",
  genre: "আধুনিক",
  confidence: 0.9953462481498718,
},
- {
  id: 10,
  sentiment: "রোমান্টিক",
  text:

```

Figure 6.7: Showing Json format data from our REST API

Chapter 7

Result and Analysis

7.1 Deep Learning Model Performance Analysis

For Model Performance we train our dataset and analyze our model by validation dataset. We create the helper function for training and evaluate our model. The BERT author made various suggestions with the purpose of fine-tuning [14]:

Table 7.1: suggestions with the purpose of fine-tuning.

Fine-Tune				
Batch size	8	16	32	64
Learning rate (Adam)	8e-5	5e-5	3e-5	2e-5
Number of epochs	40	100	200	300

For training and validation, we set the epochs 100. We set our batch size 8. If you increase the batch size, it will significantly reduce the time of training the dataset, however will reduce the training accuracy. There is a problem in google Colaboratory when we use the GPU. When we increase the batch size google colab gets out of memory. For this we used a lesser batch size for training the model. Model training should be familiar with two exceptions: every time a package is loaded into the model, the scheduler is called. We use clipgrad_norm to cut the gradient of the model to avoid gradient explosion. At the training loop we are storing the training history. After completing the training and validation process we are storing the better model represented by the rate of highest value of validation accuracy. After 7 epoch our training accuracy starts to approach 100 %. But our validation accuracy is much lower than the training accuracy. We need to fine tune

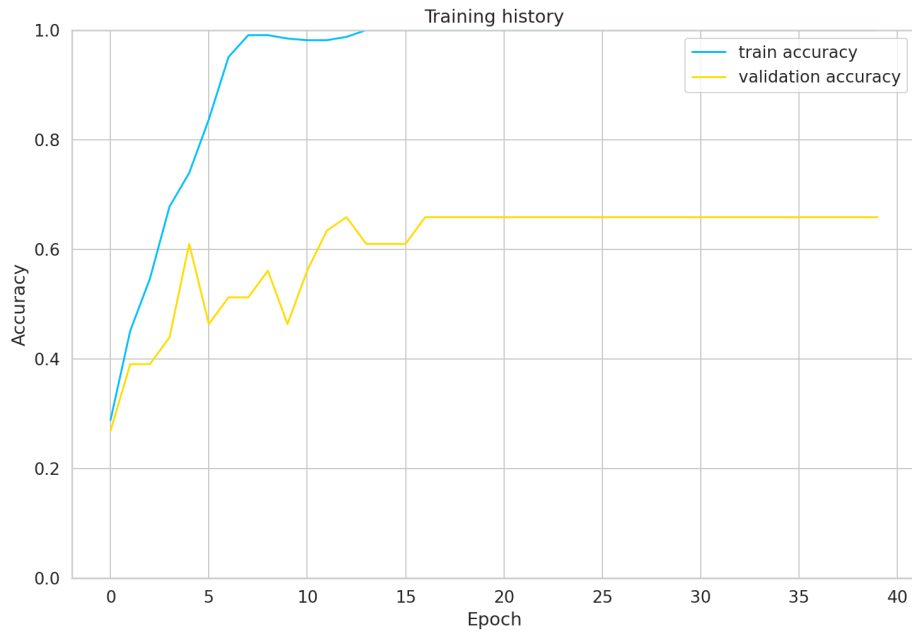


Figure 7.1: Training and validation accuracy graph

the parameter to get an average accuracy. However, this problem is occurring because of our data unbalance when training the model and for the batch size. When we increase the batch size our google colab is getting out of memory. This is a common issue I found out every moment. When I decrease the batch size everything is working perfectly but it's made in our model accuracy. We got an accuracy of 65% based on validation accuracy for our model. This is the classification report of the model. Looks like it's very hard to identify the romantic, sad , and relaxation. Be-

Table 7.2: Classification Report

Classification report.	precision	recall	f1-score	support
Sad	0.39	0.50	0.44	14
Romantic	0.62	0.38	0.48	13
Happy	0.62	0.72	0.67	18
Relaxation	0.50	0.40	0.44	10
Accuracy			0.53	55
macro avg	0.53	0.50	0.51	55
weighted avg	0.54	0.53	0.52	55

cause there is some similarity in lyrics between those classes. This confusion will

be cleared by the confusion matrix. In confusion matrix class sad and romantic at

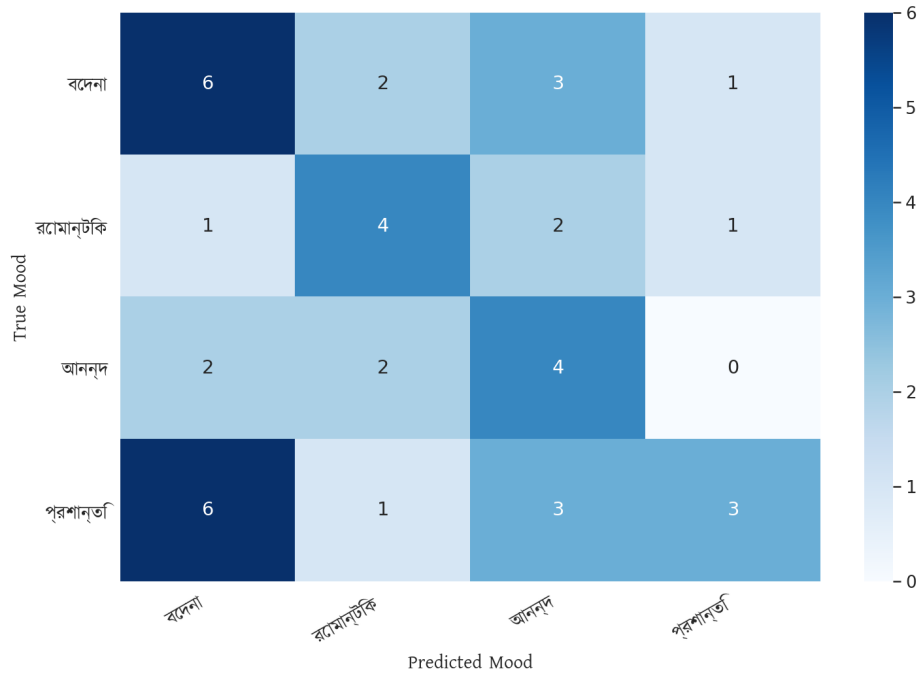


Figure 7.2: Confusion Matrix

a roughly equal frequency. Also sad and romantic class has some equal frequency of happy class. By confusion matrix we confirm that our model is having difficulty classifying the romantic and relaxation songs mood.

7.2 Security Analysis

Security Analysis is a process which measures security at the application level that helps to prevent data within the application from being stolen or intruded; not only that it also prevents any user from leaking any data or being a victim to the intruder.

7.2.1 Security

This application ensures all users that any data will not be stolen or leaked and admin will control whole system and prevents any unauthorized access. Thus, this application is safe and secure and recommended to use.

7.2.2 Privacy

If an user uses an application that collects personal information then the user must need a Privacy Policy . Personal data can be many types. It could be the users name, mail address, phone number It can also be IP addresses and any information collected through cookies .Moreover, This application doesn't collect or leak any user's information which an user can be rest assured.

7.2.3 Integrity

Integrity is crucial as it illustrates the safety, security, and maintainability of your program or application . An Application Integrity serves clients and ensures their organizations for the safety and security. As this application's Integrity, helps to enforce an organization's data integrity goals. They can be used to defend denotation of user errors, to secure against unauthorized changes in application code, to help prevent flaws or vulnerabilities and to defend malicious code from entering to this application.

7.3 Evaluation

7.3.1 Interference Time

Our web application is too fast to load. It just takes a few seconds to load and the user can use this application. There are two types of users. Admin and normal user. Interfaces are different for admin and user. Admin interface takes much time at the time of data post and update. Because when a data post by admin using FastApi is loaded the classify model restarts the full process of data preprocessing and text classification and then classify the Mood of a song's lyric. After classifying the mood is also load the database and push the data into the data table. This full process is time consuming. Another feature of this application takes a huge time to process. Audio to text conversion. This process takes approximately 8-10 minutes to process the data and classify the Mood.

7.3.2 Application performance (Time Complexity)

Our application performance is average and not so good. The main fact for the average performance is the data unbalance. Our model accuracy is much lesser. But it can classify any kind of Bangla song's mood correctly. This process is a time consuming process. Our application is bidirectional. We said Bidirectional because it can identify songs mood from lyrics and also from audio clips. Based on lyrics it takes 1-2 minutes to classify the lyrics mood. This process is faster than the audio classification process. For audio clips we are using google converter. We divide the audio clips into many chunks. Bases on length of the audio clip chunks will be more than 20-30. This process is a very time consuming process. We convert the audio to text, preprocess the text and classify the text. But our model can classify the text correctly from audio clips. Everything we are posting into the API is stored in a database. Then we are fetching the data using the get method and showing it to our web application. This process is a faster process. No need to wait to load the songs.

7.4 Cost Analysis

7.4.1 Hardware

In this project we are using this hardwares given below-

- Laptop with good configuration
- Good internet connection
- Processor
- Memory
- Display
- Server

7.4.2 Software cost

- Windows 10
- Web browser
- Visual studio
- Google Colab
- FastAPI
- PostgreSQL
- Microsoft office

7.5 Social Legal & Ethical Issues

This project has a great impact on our society. As it gives statistics how the moods of songs are in an era. As in our project we can see a category having more songs. So, it indicates the mental growth of the generation. Analysing our dataset we can tell people like to listen "sad" songs more than others songs. In the case of Ethical issues we should obviously care about the security of the users who are logged in to our system. We should ensure that people's private data should not be disclosed. As our project is song related we should also ensure there is no plagiarism of songs. We have tried to maintain all the ethical issues to build our project.

Chapter 8

Conclusion and Future Works

8.1 Scope for Future Work

The application has a very wide distant future scope which is to track users activities or moods which varies person to person .This application will track users mood and recommend music depending on their mood through a recommendation system . Basically there are three different types of functions in the recommendation system. These are - User-Product Relationship ,Product-Product Relationship ,User-User Relationship. A user who has preference to romantic music will be recommended by this recommender system and the same method works with a user who prefers sad or happy music.

We will develop a better model of this application whose accuracy will be better than the current model . However , we will collect more Bengali songs in our dataset .We will develop an android app for this application too and will publish it on google play store .Thus, a journal will be published on this project.

8.2 Conclusion

We are working on the mood classification of Bangla songs lyrics. We have built up a Web application or either both which will identify the mood of the Bangla songs Based on the lyrics. It will also suggest similar mood songs' lyrics. Music industry is huge and a lot of people work with music. Our application will help those people to know what type of Bangla songs are related to peoples' emotions nowadays. It will also help the people to understand normal peoples' emotions and that will make it easier for them to make songs according to this. It will also be

a psychological research on peoples' emotions. What type of songs are increasing according to the mood of people will let us know in what way the emotions are changing of people. So we think our work on mood analysis of songs based on lyrics will be useful work indeed. We are hoping to publish a journal on this work. Also after completing the application we will exhibit it to different competitions as well as we will try to upload this on Google Play Store.

8.3 Gantt Chart

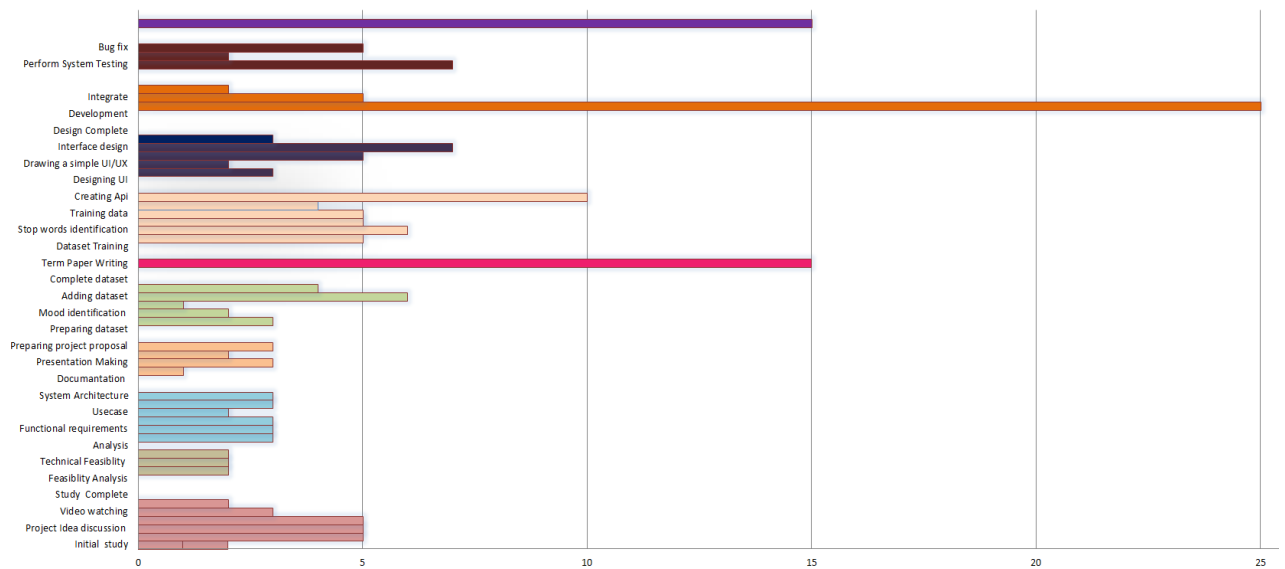


Figure 8.1: Gantt Chart of our project

Bibliography

- [1] Walaa Medhat, Ahmed Hassan, and H. Korashy. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5:1093–1113, 2014.
- [2] D. M. E. M. Hussein. A survey on sentiment analysis challenges. *Journal of King Saud University: Engineering Sciences*, 30:330–338, 2016.
- [3] Tapasy Rabeya, Narayan Ranjan Chakraborty, Sanjida Ferdous, Manoranjan Dash, and Ahmed Al Marouf. Sentiment analysis of bangla song review- a lexicon based backtracking approach. In *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, pages 1–7, 2019.
- [4] Md. Afif Al Mamun, Imamul Kadir, AKM Shahariar Azad Rabby, and Abdullah Al Azmi. Bangla music genre classification using neural network. In *2019 8th International Conference System Modeling and Advancement in Research Trends (SMART)*, pages 397–403, 2019.
- [5] M. Zaanen and P. Kanter. Automatic mood classification using tf*idf based on lyrics. In *ISMIR*, 2010.
- [6] Nirbhay Kashyap, Tanupriya Choudhury, Dev Kumar Chaudhary, and Roshan Lal. Mood based classification of music by analyzing lyrical data using text mining. In *2016 International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE)*, pages 287–292, 2016.
- [7] Sebastian Raschka. Musicmood: Predicting the mood of music from song lyrics using machine learning. *CoRR*, abs/1611.00138, 2016.

- [8] Nazia Parvin Urmi, N. Ahmed, Md. Habibur Rahman Sifat, S. Islam, and Abu Shafin Mohammad Mahdee Jameel. Banglamusicmood: A music mood classifier from bangla music lyrics. 2021.
- [9] Pradeep Kumar Singh, Pijush Kanti Dutta Pramanik, Avick Kumar Dey, and Prasenjit Choudhury. Recommender systems: an overview, research trends, and future directions. *International Journal of Business and Systems Research*, 15(1):14–52, 2021.
- [10] Y. Song, S. Dixon, and M. Pearce. A survey of music recommendation systems and future perspectives. 2012.
- [11] Erion Çano and M. Morisio. Music mood dataset creation based on last.fm tags. 2017.
- [12] J. Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *NAACL-HLT*, 2019.
- [13] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. Huggingface’s transformers: State-of-the-art natural language processing, 2020.
- [14] Chi Sun, Xipeng Qiu, Yige Xu, and Xuanjing Huang. How to fine-tune bert for text classification?, 2020.