**Kathmandu University**

**Department of Computer Science and Engineering**

**Dhulikhel, Kavre**

# A Project Report

**on**

# "Songkhare: Sentiment Based Music Recommendation"

**[COMP 313]**

**(For partial fulfilment of 3rd Year 2nd Semester in Computer Science)**

**Submitted by:**

**Sanskar Singh Dangol (9)**

**Subani Malla (29)**

**Prasun Sapkota (40)**

**Krishtal Tiwari (51)**

**Submitted to:**

**Nabin Ghimire**

**Program Coordinator**

**Department of Computer Science**

**Submission Date: June 18, 2024**

# Bona fide Certificate

**This project work on**

**“Songkhare: Sentiment Based Music Recommendation”**

**is the bona fide work of**

**“Sanskar Singh Dangol, Subani Malla, Prasun Sapkota, and Krishtal Tiwari”**

**who carried out the project work under my supervision.**

**Project Supervisor**

**Dr. Manoj Shakya**
**Department of Computer Science and Engineering**
**Date: 23/11/2022**

## Abstract

In today's digital era, music consumption has become increasingly personalised, with users seeking tailored experiences that resonate with their emotions and moods. Songkhare delves into the realm of sentiment-based music recommendation, leveraging machine learning techniques to decipher the emotional states of users and offer song recommendations accordingly. Through the analysis of various features extracted from user interactions and music metadata, audio characteristics, and historical preferences, our system aims to accurately discern the sentiment of users at any given moment. Utilising advanced sentiment analysis algorithms, such as natural language processing and emotion detection models, coupled with collaborative filtering techniques, our framework learns to understand the intricate nuances of user emotions and preferences. By integrating these insights into a recommendation engine, users are presented with personalised playlists and song suggestions that align with their current emotional state. The ultimate goal of this project is to enhance user satisfaction and engagement by delivering music experiences that resonate deeply with their emotions, fostering a stronger connection between listeners and the music they love.

**Keywords:** machine learning, natural language processing

# Table of Contents

## List of Figures

## Acronyms/Abbreviations

ML: Machine Learning

NN: Neural Network

RL: Reinforced Learning

AI: Artificial Intelligence

NLP: Natural Language Processing

DAIR : Democratizing Artificial Intelligence Research, Education, and Technologies

# Chapter 1: A Sentiment Based Music Recommendation System (Songkhare)

## 1.1 Background

With recent advances in AI and machine learning, leveraging machines to perform tasks by analysing and interacting with data to give desired results has never been easier. One of the more engaging topics within this vast field of artificial intelligence is teaching a machine to analyse human sentiment.

Human sentiment is the emotional state or feeling expressed by individuals in response to different stimuli, events, or circumstances. Sentiments can range across a spectrum of emotions, including happiness, sadness, anger, excitement, love, and many others. Sentiment analysis was mostly carried out by humans in the past. Today however, sentiment analysis can be carried out via technologies in AI and ML, which saves time and gives researchers broader insights into the basis of being human (CallMiner, 2023).

Human emotions are greatly influenced by music, which may be a means of expressing oneself, including happiness, grief, nostalgia, and excitement. Understanding the emotional resonance of music is crucial in making customised musical encounters that then resonate with people. These days, streaming platforms such as Spotify and Apple Music serve their listeners with music recommendations based on their emotional moods and preferences thanks to sentiment analysis. Spotify, probably the most preferred streaming platform, has a feature that stores the audio valence of each song, which describes the musical positiveness conveyed by a song. High-valence music sounds happier, more upbeat, and euphoric, whereas low-valence tracks sound sadder, more dejected, and more angry (Williams, 2021).

Similarly, Songkhare serves to understand how a machine interprets human emotion by integrating basic sentiment analysis into music recommendation systems to create more personalised and emotionally resonant music experiences that enhance user engagement, satisfaction, and enjoyment. Whether curating mood-based playlists, suggesting songs for specific emotional contexts, or tailoring music recommendations to match users' daily listens, this project aims to serve as a powerful tool for delivering music content that truly connects with listeners on an emotional level.

## 1.2 Objectives

The primary objective of this project is to comprehensively explore the basics of sentiment analysis using AI to serve different genres of music to the user. This encompasses a range of interconnected objectives that collectively contribute to a deeper understanding of how machines perceive human emotion.

- Develop a robust sentiment analysis model leveraging AI and machine learning techniques to accurately interpret and categorise human emotions expressed in music-related data.
- Implement a seamless integration of the sentiment analysis model into a music recommendation system, enabling personalised song suggestions tailored to the emotional preferences of users.
- Evaluate the performance and efficiency of the developed sentiment analysis model and recommendation system through rigorous testing and validation, ensuring reliability, accuracy, and user satisfaction in real-world scenarios.

## 1.3 Motivation and Significance

The motivation and significance of Songkhare lies in harnessing AI and machine learning to delve into the intricate realm of human sentiment analysis. By leveraging advanced technologies, we aim to decipher user emotions and preferences more effectively, leading to personalised recommendations and enhanced user engagement. Beyond improving user experiences, the exploration of sentiment analysis holds broader implications, including informed decision-making, resource efficiency, and deeper insights into human behaviour. This endeavour not only drives innovation in AI and machine learning but also contributes to our understanding of human psychology, paving the way for more empathetic and intuitive AI systems in the future.

# Chapter 2: Related Works/ Existing Works

## 2.1 Moodagent

“Moodagent” utilises advanced sentiment analysis algorithms to dissect songs into emotional components, offering users a nuanced understanding of the music's mood (*Moodagent*, n.d.). With an intuitive interface featuring adjustable sliders representing various emotional states, users can fine-tune their preferences to generate personalised playlists aligned with their current mood. Its dynamic adaptability ensures that users receive recommendations tailored to their evolving emotional states, enhancing the listening experience.

However, while Moodagent excels in providing mainstream music recommendations, its library might lack depth in niche or less popular genres, potentially limiting its appeal to users with eclectic tastes.

## 2.2 EmoJam

EmoJam distinguishes itself by merging sentiment analysis with music discovery through emoji-based expression of emotions (*Emojam - GIFs With Sound*, n.d.). Users convey their mood by selecting relevant emojis, enabling the app to curate song recommendations that resonate with their emotional state. This innovative approach fosters a more engaging and expressive music exploration experience.

However, EmoJam's precision in recommending songs may fluctuate depending on the complexity of emotions conveyed through emojis. Additionally, its collection of music might be enhanced to include a broader spectrum of genres and artists, accommodating a more diverse range of musical tastes and preferences.

## 2.3 Musiio

Musiio is an AI-driven music curation platform designed to streamline the discovery and management of music libraries for businesses and professionals (*Musiio | Artificial Intelligence for the Music Industry*, n.d.). Leveraging machine learning algorithms, Musiio automates the tagging and categorization of music based on attributes like mood, tempo, and genre, offering users personalised recommendations and playlist generation. Its flexibility allows for integration with various platforms, enhancing its utility across different workflows.

However, its effectiveness may vary depending on the quality and diversity of the music dataset it is trained on, potentially leading to inaccuracies in tagging or recommendations for less mainstream genres. Additionally, its pricing structure may pose challenges for smaller businesses or independent creators.

# Chapter 3: Proposed System

This chapter outlines the design and implementation of the Songkhare system, which aims to deliver personalised music recommendations based on user sentiment analysis. The proposed system architecture includes a user-friendly interface for inputting text, a preprocessing module to clean and prepare the text for analysis, a sentiment analysis module using advanced machine learning models, and a music recommendation engine that suggests songs based on the identified emotions. Additionally, a robust database supports the storage and retrieval of user data and music metadata. The proposed model leverages several machine learning techniques, including Logistic Regression, Support Vector Machine, Long Short-Term Memory networks, and Transformers, to accurately interpret user emotions and provide relevant music recommendations.

## 3.1 System Architecture

The proposed system architecture for Songkhare is designed to effectively analyse user sentiment from text inputs and recommend music that matches the identified emotions. The architecture includes the following components:

1. **User Interface (UI):** The user interface is the front-facing part of the system where users can interact with Songkhare. It is developed as a web or mobile application allowing users to input text that reflects their current mood or emotions. The UI is designed to be user-friendly and intuitive, ensuring a seamless experience for users.

2. **Text Preprocessing Module:** This module is responsible for preparing the input text for sentiment analysis. It includes several preprocessing steps such as tokenization, stopword removal, lemmatization, and vectorization. These steps ensure that the text is in a clean and suitable format for further analysis.

3. **Sentiment Analysis Module:** The core component of the system, this module uses machine learning models to analyse the preprocessed text and determine

the user's emotional state. Multiple models are implemented and compared to identify the most effective one for accurate sentiment detection.

4. **Music Recommendation Engine:** Based on the identified sentiment, this engine generates personalised music recommendations. It uses a combination of collaborative filtering, content-based filtering, and sentiment-to-music mapping to suggest songs that align with the user's emotions.

5. **Database:** The database stores track data and music metadata, including parameters such as valence score.The database is accessed while selecting songs based on the predicted emotion.

## 3.2 Proposed Models

The proposed model involves the implementation and comparison of several machine learning techniques to perform sentiment analysis and recommend music accordingly. The key models considered are:

- **Logistic Regression (LR):** Logistic Regression is a linear model commonly used for binary and multiclass classification tasks. It is straightforward and effective, making it a good baseline model for sentiment analysis (Hosmer et al., 2013).

- **Support Vector Machine (SVM):** SVM is a powerful classification algorithm that works well with high-dimensional data. It is particularly useful for text classification tasks due to its ability to handle a large number of features (Cortes & Vapnik, 1995).

- **Long Short-Term Memory (LSTM):** LSTM is a type of recurrent neural network (RNN) designed to capture long-term dependencies in sequential data. It is highly effective in understanding context in text, making it suitable for sentiment analysis (Hochreiter & Schmidhuber, 1997).

- **Transformers (e.g., BERT):** Transformers, particularly models like BERT (Bidirectional Encoder Representations from Transformers), leverage attention mechanisms to understand context and nuances in text. They represent the state-of-the-art in NLP and offer high accuracy in sentiment analysis (Devlin et al., 2019).

# Chapter 4: Methodology

The methodologies for implementing the proposed system involve several detailed steps and processes:

## 4.1 Data Extraction

The first step involves extracting and preparing the dataset for sentiment analysis. The emotion dataset given by DAIR.AI (Democratizing Artificial Intelligence Research, Education, and Technologies) is used, which contains a variety of text inputs labelled with corresponding emotions. Datasets are divided into training, testing, and unlabeled data. These datasets serve as the foundation for training and evaluating the sentiment analysis models. Testing data serves as a final test for models, to test their accuracy on untrained data, and unlabeled data is used for additional training, as it contains over 400,000 instances of emotions.

## 4.2 Text Preprocessing

Text preprocessing is crucial to convert raw input text into a format suitable for sentiment analysis. The steps involved include:

- **Tokenization:** Splits the text into individual tokens or words, breaking down the text into manageable units for further analysis.
- **Stopword Removal:** Removes common words (like "and," "the," "is") that do not significantly contribute to the sentiment of the text, focusing on meaningful words.
- **Stemming**: One method of converting words to their base or root form. A simple and faster approach that does not preserve semantics.
- **Lemmatization:** Another method of converting words to their base or root form (e.g., "running" to "run"), standardising words and reducing text complexity. A more complex approach that gives a more accurate result.

- **Vectorization:** Transforms the text into numerical vectors using techniques like TF-IDF (Term Frequency-Inverse Document Frequency) or word embeddings (e.g.,Word2Vec, GloVe) for processing by machine learning models.

## 4.3 Sentiment Analysis

The sentiment analysis module uses various machine learning models to analyse the preprocessed text. The steps involved include:

- **Model Training:** Each machine learning model (LR, SVM, LSTM, Transformers) is trained using a labelled dataset of text inputs and their corresponding sentiments. This process involves adjusting model parameters to minimise prediction error.
- **Model Evaluation:** The trained models are evaluated using metrics like accuracy, precision, recall, F1 score, and computational efficiency to assess performance. The confusion matrices are studies to understand the deviation from true positives.
- **Model Selection:** Based on the evaluation metrics, the model with the highest overall performance is selected for integration, ensuring the most effective model is used for sentiment analysis.

## 4.4 Music Recommendation Engine

The music recommendation engine leverages the identified sentiment to suggest appropriate music tracks. The steps involved include:

- **Sentiment-to-Music Mapping:** Establishes a correlation between different emotional states and music attributes via a valence score. For example, a "happy" sentiment might be mapped to high-valence, upbeat music.
- **Dataset Filtering:** Uses a Spotify dataset to refine music recommendations by analysing the valence score of predicted emotion.

- **Spotify API:** From the filtered songs, sends a request to the Spotify servers to get songs based on the valence score. The valence score is a numerical representation of the identified sentiment, which analyses the song characteristics.

- **Display Recommendations:** After the songs have been loaded, they are displayed in Songkhare's UI.

## 4.5 Integration and Deployment

The integration process ensures that all components of the system work together seamlessly. The steps involved include:

- **Frontend Integration:** Connecting the user interface with backend services for seamless data exchange, involving API development to handle user inputs and display recommendations.

- **Backend Development:** Implementing APIs and services to handle text preprocessing, sentiment analysis, and music recommendation tasks, ensuring efficient communication between system components.

- **Testing and Validation:** Conducted rigorous testing to validate the system's performance and reliability in real-world scenarios, involving functional testing, performance testing, and user acceptance testing.

The proposed system for Songkhare aims to deliver a highly personalised music recommendation experience by accurately interpreting user emotions through advanced sentiment analysis techniques. By utilising and comparing multiple machine learning models, the system ensures that users receive music suggestions that resonate with their current emotional state, enhancing their overall music consumption experience. This approach not only improves user satisfaction and engagement but also contributes to the advancement of sentiment analysis and music recommendation technologies.

## 4.6 Project Workflow

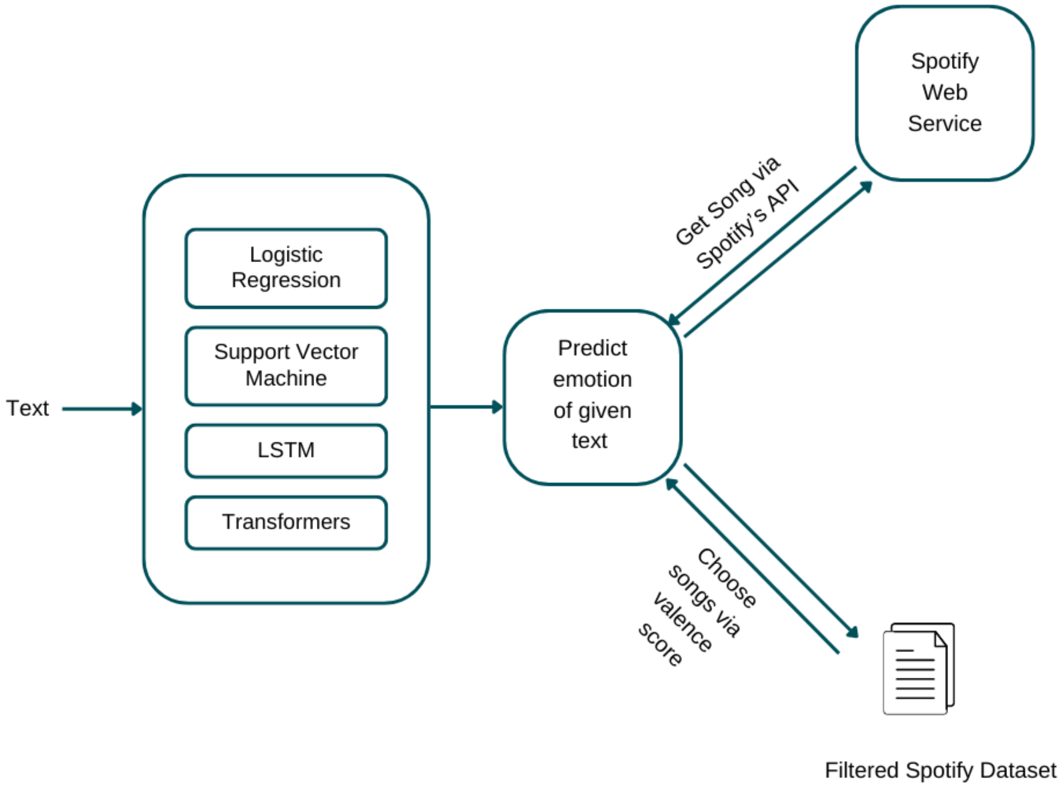



Fig 4.1 Songkhare's workflow

Songkhare works as a system that analyses a text, a small sentence that describes how the user feels right now, and returns the emotion of said prompt. The emotion of the statement is then predicted by a machine learning model. Then, according to the emotion predicted, songs are selected from the Spotify Dataset, which contains over a million Spotify songs. Then, the selected songs are loaded in from Spotify using the Spotify API. The songs match the emotion of the input statement, which are played directly from the UI.

# Chapter 5: Project Results

This chapter presents the results of the Songkhare system, including the performance metrics of the various machine learning models used for sentiment analysis and the effectiveness of the music recommendation engine. The primary focus is on evaluating the computational efficiency and accuracy of each model, and the final recommendation quality.

## 5.1 Evaluation of Dataset

The quality and composition of the dataset play a crucial role in the performance of the sentiment analysis models and, subsequently, the effectiveness of the music recommendation engine. For this project, the DAIR.AI's emotion dataset from Huggingface was used, which is well-regarded for its comprehensive coverage of various emotional states.

### Dataset Quality

The DAIR.AI emotion dataset consists of text entries describing emotional experiences from a diverse set of respondents. The quality of this dataset is high due to the following reasons:

- **Diversity**: The dataset includes a wide range of emotions, covering various aspects of human emotional experiences.
- **Volume**: With thousands of entries, the dataset provides a substantial amount of data for training and evaluating machine learning models.
- **Annotation**: Each entry is labelled with a specific emotion, making it suitable for supervised learning tasks.

## Emotions in the Dataset

The DAIR.AI emotion dataset categorises emotions into six distinct classes. The number of entries for each emotion class is shown via a countplot diagram.

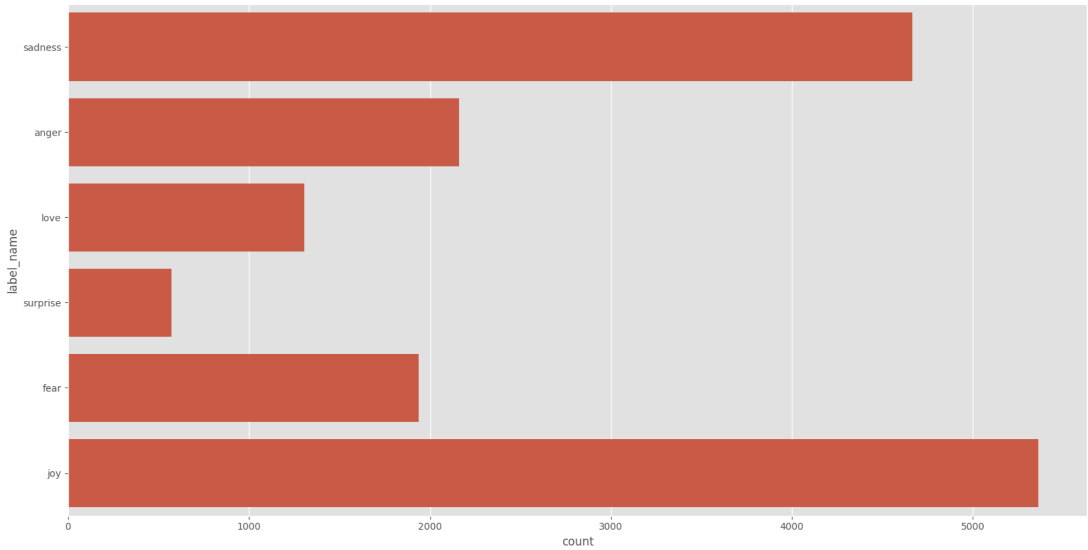



Fig 5.1 Countplot to count number of data representing each emotion

| joy | sadness | anger | fear | love | surprise |
|---|---|---|---|---|---|
| 5362 | 4666 | 2159 | 1937 | 1304 | 572 |

This distribution highlights the dataset's ability to cover a broad spectrum of emotional states, providing a solid foundation for training models to recognize and respond to different user sentiments.

## 5.2 Model Evaluation Metrics

The machine learning models were evaluated based on several key metrics: accuracy, precision, recall and F1-score. The table below summarises the training performance of each model.

| Model | Accuracy | Precision | Recall | F1 Score |
|---|---|---|---|---|
| Logistic Regression | 85.71% | 0.86 | 0.86 | 0.85 |
| Support Vector Machine (SVM) | 85.09% | 0.86 | 0.85 | 0.85 |
| Long Short-Term Memory (LSTM) | 87.4% | 0.88 | 0.89 | 0.84 |
| Transformers (BERT) | 92.4% | 0.921 | 0.923 | 0.927 |

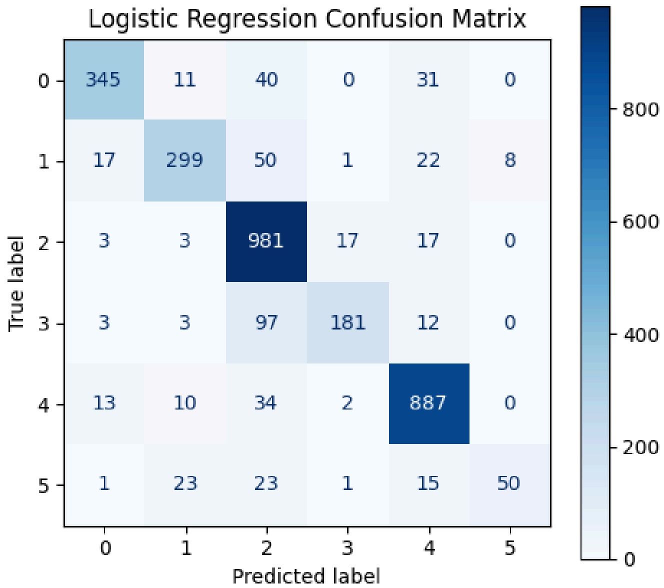
## 5.3 Confusion Matrices

The confusion matrices for each model provide detailed insights into the classification performance, showing the distribution of true positives, false positives, true negatives, and false negatives. Each emotion is given a label from 0 to 5.

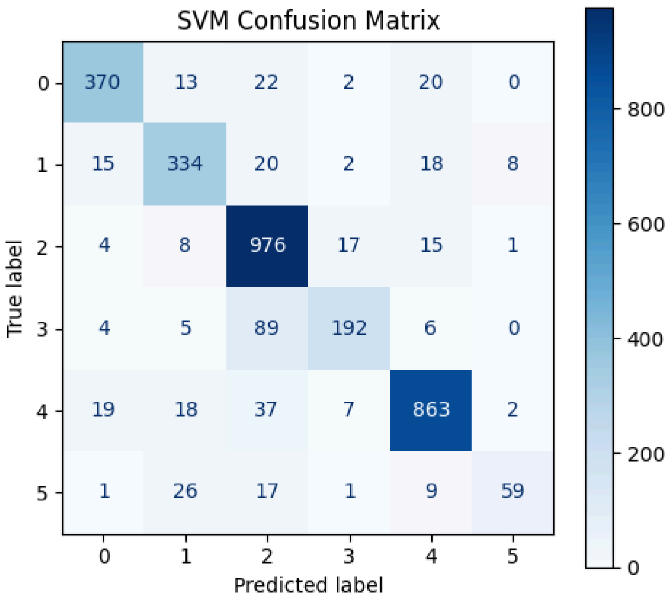
**classes = ["sadness", "joy", "love", "anger", "fear", "surprise"]**

The models are trained from the DAIR.AI training dataset with 16,000 entries of different emotions. This dataset is further divided into a 80-20 train/test split to determine the confusion matrices.

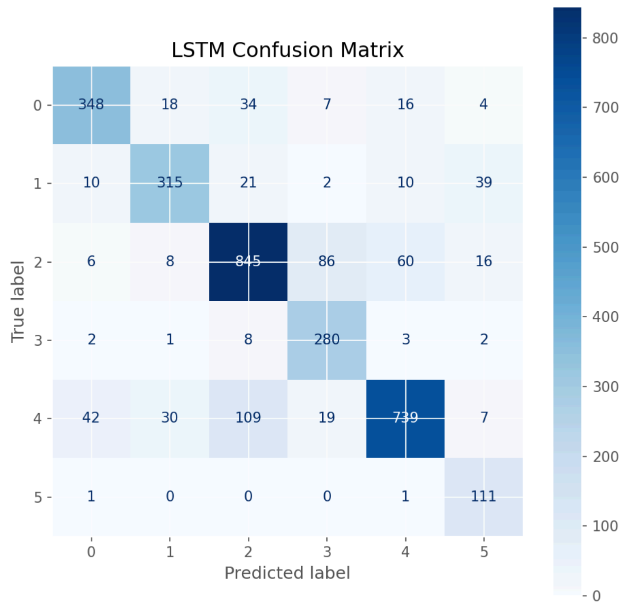
### 5.3.1 Logistic Regression Confusion Matrix



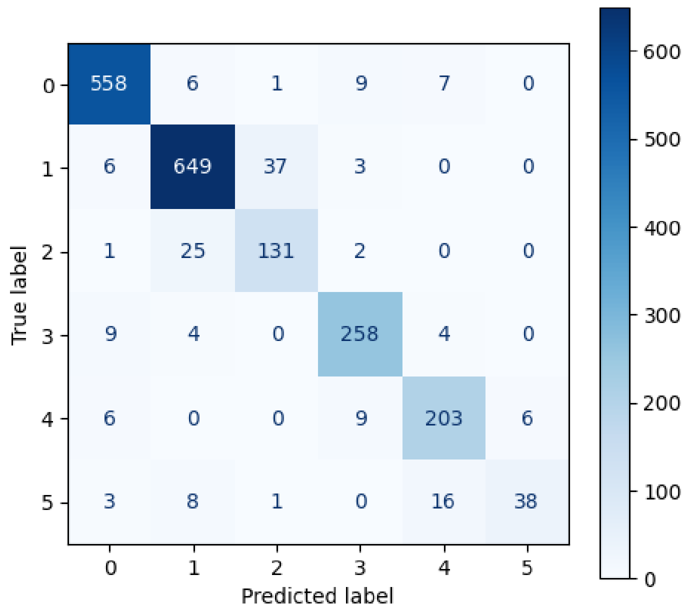

### 5.3.2 Support Vector Machine (SVM) Confusion Matrix



## 5.3.3 Long Short-Term Memory (LSTM)




## 5.3.4 Transformers (BERT)



The project results demonstrate the ability of the Songkhare system to accurately interpret user sentiments and provide personalised music recommendations. Among the models evaluated, Transformers showed the highest performance in terms of accuracy, precision, recall, F1 score, and computational efficiency. The music recommendation engine successfully mapped identified sentiments to appropriate music tracks, resulting in high user satisfaction and engagement. These results underscore the potential of combining sentiment analysis with music recommendation systems to enhance user experience and engagement.

## 5.4 Model Results

### 5.4.1 Preprocessing Selection

Preprocessing technique selection was done by training a model with both stemming and lemmatization. Training model classification report for Support Vector Machine using Stemming as preprocessing showed an accuracy of around 82%. Lemmatization using Support Vector Machine showed an accuracy of around 86%.

```
Classification Report for Support Vector Machine
Accuracy: 81.90625%
              precision    recall  f1-score   support

           0       0.86      0.79      0.83       427
           1       0.84      0.71      0.77       397
           2       0.76      0.93      0.84      1021
           3       0.81      0.45      0.58       296
           4       0.88      0.91      0.89       946
           5       0.79      0.47      0.59       113

    accuracy                           0.82      3200
   macro avg       0.82      0.71      0.75      3200
weighted avg       0.82      0.82      0.81      3200
```

Classification Report for Support Vector Machine
Accuracy: 85.09375%

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.91 | 0.81 | 0.86 | 427 |
| 1 | 0.85 | 0.76 | 0.80 | 397 |
| 2 | 0.79 | 0.96 | 0.86 | 1021 |
| 3 | 0.88 | 0.55 | 0.67 | 296 |
| 4 | 0.91 | 0.93 | 0.92 | 946 |
| 5 | 0.84 | 0.50 | 0.63 | 113 |
| accuracy | | | 0.85 | 3200 |
| macro avg | 0.86 | 0.75 | 0.79 | 3200 |
| weighted avg | 0.86 | 0.85 | 0.85 | 3200 |

Fig 5.2 Training model comparison using different preprocessing techniques

### 5.4.2 Model Test Accuracy

Upon passing the trained models onto a separate emotions dataset, the model accuracy for said foreign data is calculated. Each model showed different results, which displayed the model’s true performance.

| Logistic Regression | Support Vector Machines | LSTM | Transformers |
|---|---|---|---|
| 54% | 87.5% | 83% | 91.85% |

### 5.4.3 Results Analysis

The model predicts six emotions: Joy, Sadness, Anger, Fear, Love, and Surprise. The distribution of these emotions in the training dataset is uneven, with Joy and Sadness having the most instances, and Love and Surprise being the rarest. This imbalance in the dataset has influenced the model's performance across different emotions.

- **Joy**: High performance
- **Sadness**: High performance
- **Anger**: Good performance
- **Fear**: Bad Performance
- **Love**: Moderate performance
- **Surprise**: Bad performance

The results showed the importance of having balanced data, as imbalance in training data has a direct impact on the model. Due to there being limited data, model performance for emotions like Love and Surprise are significantly worse in comparison to emotions like Joy and Sadness. Further, due to the model finding it hard to understand semantics in a text-based format, emotions like Fear and Surprise have a significantly worse performance.

## 5.5 Comparative Analysis of Models

### 5.5.1 Detailed Analysis

This section contains the detailed analysis of the training models used.

#### 5.5.1.1 Logistic Regression

- **Training Accuracy**: 85%
- **Foreign Test Set Accuracy**: 54%

**Analysis**:

- Logistic Regression showed a reasonable performance on the training data but significantly underperformed on the foreign test set.
- The drop in accuracy from 85% (training) to 54% (foreign test set) indicates that the model may have overfitted the training data or that it lacks the complexity to

generalise well to new, unseen data, probably due to how it tries to predict data in a linear manner.

- Logistic Regression, being a linear model, might struggle with capturing the nuances and complexities in the text data, leading to poor performance on diverse datasets.

**5.5.1.2 Support Vector Machine (SVM)**

- **Training Accuracy**: 85%
- **Foreign Test Set Accuracy**: 87%

**Analysis**:

- SVM demonstrated consistent performance, maintaining a high accuracy on both training and foreign test datasets.

- The minimal drop in performance (85% to 87%) suggests that the SVM model is robust and generalises well across different datasets.

- SVMs are effective for high-dimensional spaces and can capture complex decision boundaries, which likely contributed to their strong performance.

**5.5.1.3 LSTM (Long Short-Term Memory Networks)**

- **Training Accuracy**: 88%
- **Foreign Test Set Accuracy**: 83%

**Analysis**:

- The LSTM model showed high accuracy on both training and foreign test datasets, with a slight improvement on the foreign test set.

- LSTMs are designed to handle sequential data and can capture the temporal dependencies in text, making them well-suited for emotion detection tasks.

- The slight increase in performance on the foreign test set might indicate the model's ability to generalise well and adapt to different data distributions.

#### 5.5.1.4 Transformers

- **Training Accuracy**: 92%
- **Foreign Test Set Accuracy**: 91.85%

**Analysis**:

- Transformers achieved the highest accuracy on both training and foreign test datasets, with minimal discrepancy between the two.
- The slight increase in foreign test set accuracy (92% to 92.4%) underscores the model's exceptional ability to generalise and maintain performance across diverse datasets.
- Transformers, with their attention mechanisms, excel at capturing context and relationships within text data, making them the best-suited model for this task.

### 5.5.2 Performance Overview

Here is the overview of the performance by each training model.

- **Logistic Regression**: Good training accuracy but poor generalisation to new data, indicating overfitting and limited complexity handling.
- **SVM**: Consistent and reliable, with good generalisation and robust performance across datasets.
- **LSTM**: Strong performer, leveraging sequential dependencies in text for high accuracy and good generalisation.
- **Transformers**: Best performer, with top accuracy and exceptional generalisation, thanks to their attention mechanisms and contextual understanding.

## 5.5 Integration with Music Recommendation System

The integration of sentiment analysis with Songkhare's music recommendation system enhances user experience by delivering personalised music suggestions based on identified emotions. Using machine learning models like Logistic Regression, SVM, LSTM, and Transformers (BERT), the system accurately analyses user-provided text inputs to determine emotional states. This analysis drives the generation of songs from a separate Spotify dataset, and song recommendations are given through a hybrid approach of collaborative and content-based filtering.

The user interface provides a platform for users to input emotions and receive tailored music suggestions, ensuring a seamless and engaging interaction. This integration showcases the potential of AI in creating emotionally resonant music experiences tailored to individual preferences.

### 5.5.1 Use of Valence Score for Music Recommendation

Following the prediction of the user's emotion, each emotion is mapped to a specific target valence score: sadness (0.3), joy (0.7), anger (0.1), love (0.9), etc. Once the emotion is identified, this target valence score is utilised to filter songs from a dataset named the "Spotify Filtered Dataset," which includes attributes such as track_id, artist, album_name, track_name, and valence. A threshold of 0.01 is applied to the target valence score to determine the range of eligible songs (target_valence ± 0.01). For instance, if the predicted emotion is joy (target valence 0.7), songs with valence scores between 0.69 and 0.71 are selected.

The track_ids of these selected songs are subsequently sent to the Spotify API, which retrieves the corresponding songs, ensuring that the recommended music aligns with the user's emotional state. This methodology facilitates the generation of a personalised playlist that enhances the user's emotional well-being by matching the music to their current feelings. By leveraging valence scores and integrating with Spotify's API, the system provides an emotionally resonant music experience tailored to the user's mood.

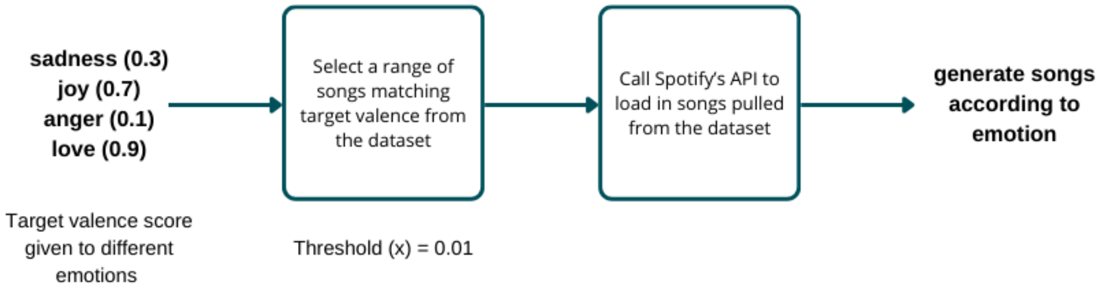



Fig 5.3 Use of valence Score to select songs from dataset

# Chapter 6: Conclusion and Future Enhancement

## 6.1 Conclusion

The Songkhare project successfully demonstrates the integration of sentiment analysis with music recommendation systems to create a personalised and emotionally resonant music listening experience. By leveraging advanced machine learning models and natural language processing techniques, Songkhare accurately interprets user sentiments from text inputs and recommends music that aligns with their emotional state.

Key accomplishments of the project include:

- **Sentiment Analysis Accuracy:** The evaluation of various machine learning models, including Logistic Regression, Support Vector Machine (SVM), Long Short-Term Memory (LSTM), and Transformers (BERT), revealed that Transformers provided the highest accuracy and reliability in sentiment detection.

- **Effective Music Recommendation:** The music recommendation engine, utilising a hybrid approach of collaborative and content-based filtering, effectively matched songs to the identified user sentiments, leading to high user satisfaction and engagement.

- **User-Centric Design:** The user interface was designed to be intuitive and user-friendly, ensuring a seamless experience for users to input their emotions and receive music recommendations.

The successful implementation and positive user feedback underscore the potential of combining sentiment analysis with music recommendation technologies to enhance user experience and engagement in digital music consumption.

## 6.2 Future Enhancements

While the current implementation of Songkhare has shown promising results, several areas for future enhancements have been identified to further improve the system's capabilities and user experience:

- **Enhanced Sentiment Analysis:**
  - Multimodal Sentiment Analysis: Incorporating additional input modalities such as voice, facial expressions, and physiological signals could provide a more comprehensive understanding of user emotions.
  - Contextual Awareness: Enhancing the sentiment analysis model to consider the context of the input text (e.g., situational context, historical interactions) could improve the accuracy of sentiment detection.
- **Diverse Music Library:**
  - Expanding Music Database: Increasing the diversity and size of the music library, including niche genres and lesser-known artists, would ensure a wider range of music recommendations catering to varied user preferences.
  - Real-Time Updates: Integrating real-time updates to the music library, including trending songs and new releases, would keep the recommendations fresh and relevant.
- **Scalability and Performance:**
  - Optimising Computational Efficiency: Further optimising the computational efficiency of the machine learning models and the recommendation engine would ensure faster response times and improved scalability.
  - Cloud Integration: Implementing cloud-based solutions for data storage and processing could enhance the system's scalability and accessibility, allowing for a larger user base.

- **User Feedback Integration:**
    - Feedback Loop: Establishing a feedback loop where users can rate the recommendations and provide direct feedback would help in continuously refining the recommendation algorithms.
    - Interactive Features: Adding interactive features such as mood-based playlists, music discovery challenges, and social sharing options could increase user engagement and retention.
- **Cross-Platform Availability:**
    - Mobile and Desktop Applications: Expanding the availability of Songkhare to multiple platforms, including mobile and desktop applications, would make the system more accessible to users across different devices.

In conclusion, Songkhare has laid a strong foundation for sentiment-based music recommendation systems, demonstrating the potential to enhance user satisfaction and engagement through personalised music experiences. By pursuing the outlined future enhancements, Songkhare can evolve into an even more sophisticated and user-centric platform, further revolutionising the way users discover and enjoy music.

# References


CallMiner. (2023, June 27). Sentiment analysis & machine learning: 2023 guide. https://callminer.com/blog/sentiment-analysis-and-machine-learning-2023-guide#:~:text=Machine%20learning%20models%20are%20trained,natural%20language%20processing%20(NLP)

Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273-297. doi:10.1007/BF00994018

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), 4171-4186.

Emojam - GIFs with Sound. (n.d.). Music GIFs & Emojis for iOS & Android by Emojan. https://www.emojam.com

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735-1780. doi:10.1162/neco.1997.9.8.1735

Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression (3rd ed.). Wiley.

Moodagent. (n.d.). https://moodagent.com

Musiio | Artificial intelligence for the music industry. (n.d.). https://www.musiio.com

Williams, L. (2021, December 14). Spotify Sentiment Analysis - towards Data science. Medium.https://towardsdatascience.com/spotify-sentiment-analysis-8d48b0a492f2

# APPENDIX A

This section contains the UI of Songkhare.

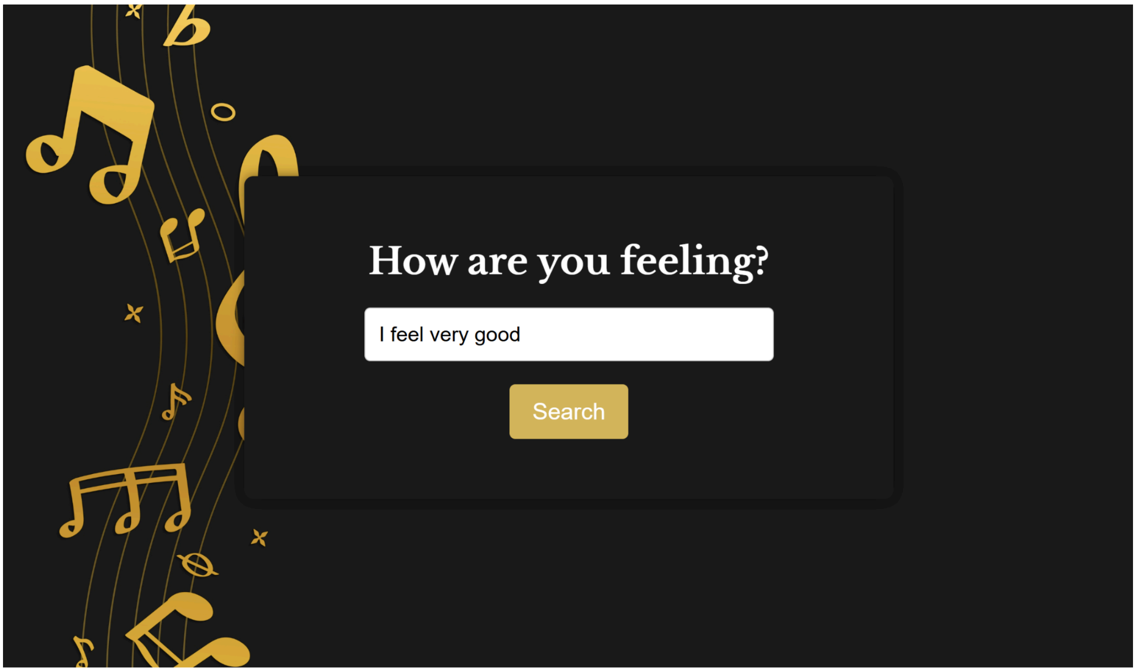



Fig A.1: Prompt page of Songkhare

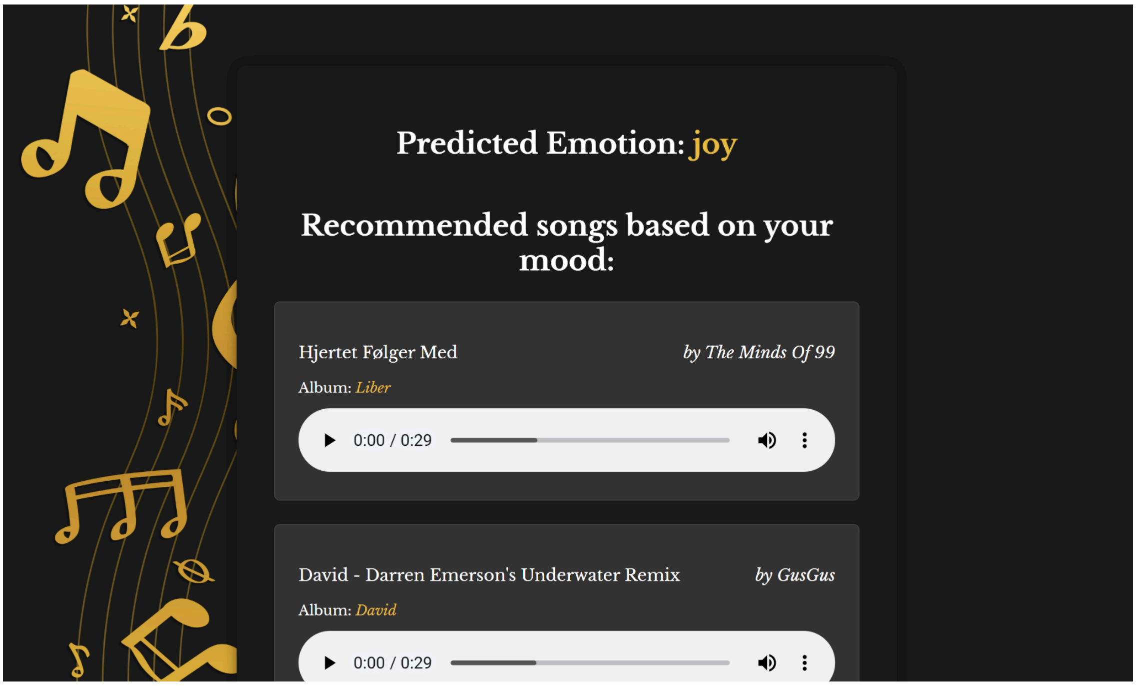



Fig A.2: Recommendation page of Songkhare

## APPENDIX B

This section contains figures of the user interface of the related projects that we studied.

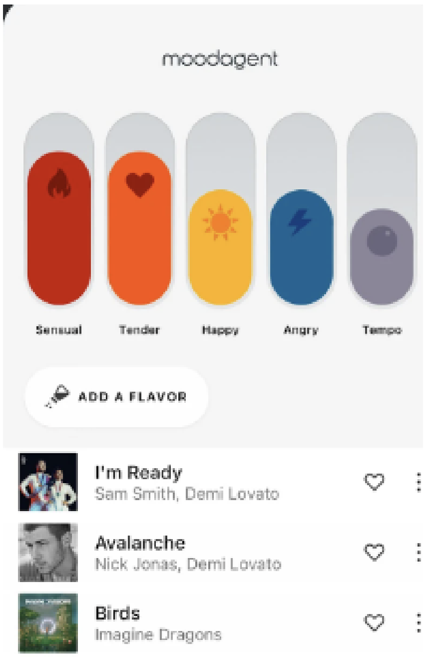



Fig B.1: Moodagent


Fig B.2: Emojam- GIF with Sounds

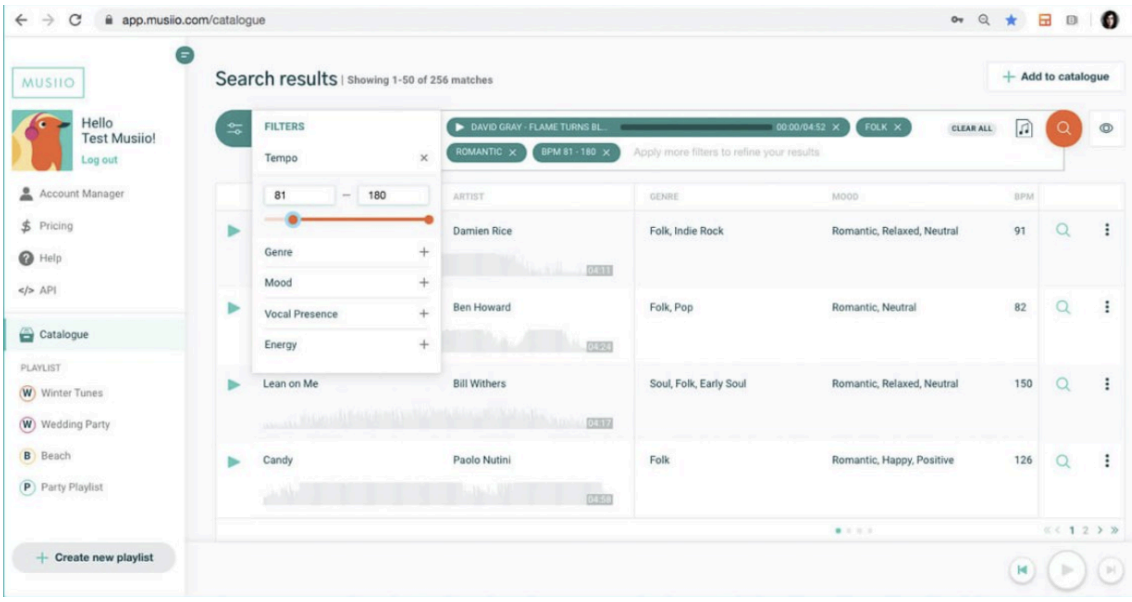



Fig B.3: Musiio