Logo, company name

Description automatically generated

**ICT619 Artificial Intelligence**

**Assignment 02**

**Music Recommendation system based on mood.**

**(MOOSIC)**

**Submitted By:**

**Manish Shakya (34509621)**

**Renuka limbu (34591492)**

**Mohammad Israt E Hasan** (**34700886**)

Table of Contents

[Abstract 1](#_Toc163251183)

[Introduction 1](#_Toc163251184)

[Problem domain 2](#_Toc163251185)

[Motivation 2](#_Toc163251186)

[Related Works 2](#_Toc163251187)

[**Literature Review** 2](#_Toc163251188)

[Methodology 4](#_Toc163251189)

[Dataset Description 4](#_Toc163251190)

[Feature selection 5](#_Toc163251191)

[Exploratory Data Analysis 7](#_Toc163251192)

[Classifications 9](#_Toc163251193)

[Building a recommendation system 10](#_Toc163251194)

[Results and outcome 10](#_Toc163251195)

[References 12](#_Toc163251196)

# Abstract

Moosic is a ground-breaking method to customized music recommendation that takes use of the strong link between human emotions and musical preferences. This unique technology employs powerful machine learning algorithms to precisely anticipate a user's mood based on their interaction with music and other surrounding elements, resulting in a highly customized and emotionally resonant musical listening experience. It expands on existing recommendation systems by combining emotional intelligence, contextual awareness, and adaptive learning processes, guaranteeing that song recommendations are not just mood-appropriate but also contextually relevant. The system also includes innovative features like interactive mood maps for music discovery and social elements for community interaction.

# Introduction

Music, being a global language, powerfully connects with our emotions. In today's digital era, while people have access to an infinite number of music, choosing one that suits their mood remains difficult. This enormous selection might lead to decision paralysis. Furthermore, practically all systems recommend music based on the user's favourite songs or their overall listening habits, with little consideration given to the emotional component of song selection. To address this gap, our research introduces a Mood-Based Song Classifier and Recommendation System. We employ machine learning to determine users' emotions based on their previous song choices, and then personalize suggestions to their emotional state. This not only gives a personalized listening experience, but also sheds light on the complex interplay between emotions and musical choices. We want to transform users' daily experiences with music platforms by ensuring that each song suggestion reflects their emotional journey. By analysing elements such as valence, acousticness, danceability, energy, instrumentalness, liveness, loudness, tempo, mood our system discerns subtle emotional cues within songs. We will also introduce an intuitive user interface that captures real-time feedback, fine-tuning our machine learning models to align recommendations with dynamic music selections and different mood states. (Torabi, 2023)

Moreover, the integration of contextual data such as time of day, weather, and personal schedules allows for the recommendation of music that not only matches the current mood but also complements the listener's environment, promoting an immersive auditory experience. Our research also contributes to the field by compiling a unique dataset, blending musical features with psychological profiles derived from user interactions, filling a niche in emotion-centric music recommendation. (Rokach, 2015)

By bridging the gap between vast music libraries and the individual's emotional needs, we anticipate a shift in how users interact with music streaming services. Our system aspires to be more than a platform—it seeks to be an empathetic companion that understands and enhances the listener's emotional well-being through music.

# Problem domain

* **Understanding Users' Preferences and Context**: Personalization and context awareness are critical. To increase relevance and user happiness, the system must precisely model individual preferences and tailor recommendations to the user's present position or activity.
* **Music Content Analysis and Semantic Gap**: Accurate music classification and recommendation need effective feature extraction from music files, as well as bridging the semantic gap between technical music features and user-perceived attributes such as mood or emotion.
* **Diversity in Recommendations**: Striking a balance between matching user tastes and providing fresh, different musical options is critical. This eliminates over-specialization of suggestions, allowing consumers to discover new music and artists while expanding their listening experience.

# Background

## Motivation

We want to utilize ML to anticipate users' emotions based on past tracks and recommend relevant tunes. We will be focusing on understanding the users’ preferences by predicting the user’s current emotion or mood. This will improve the user experience, personalizes music consumption, and provides insights into the emotional effect of music. Recognizing the strong correlation between emotions and music preferences inspired us to create a practical solution that improves people's daily interactions with music.

## Literature review and Related Works

### Literature Review

* A music recommendation system using real time facial expression detections.

This paper describes the creation of a Mood-Based Music Recommendation System that employs machine learning to determine a user's mood based on real-time facial expression data. The system, which aims to increase consumer happiness by providing individualized song choices, is separated into two modules: mood recognition and recommendation.  
  
The mood detection module uses Keras' MobileNet model to categorize emotions from photos and works well with mobile devices due to its computational economy. The music recommendation module uses a song collection classified by mood and saved on Firebase to propose songs that fit the observed mood.

The Mood-Based Music Recommendation System was built utilizing machine learning techniques, notably the MobileNet model with Keras for real-time facial expression detection, and Java for Android app development. The solution deploys the trained ML model on mobile devices with TensorFlow Lite, uses Firebase for cloud storage and database services, and trains the model with datasets such as FER 2013 and MMA Facial Expression Recognition.

The suggested method detects mood with around 75% accuracy and offers music that may improve the listener's mood. Future enhancements might include adding physiological cues for increased accuracy and broadening suggestions beyond music to other types of entertainment such as movies or TV shows depending on identified mood. (Madik, A.)

* SVR-based music mood classification and context-based music recommendation by Seungmin Rho , Byeong-jun Han and Eenjun Hwang

The work "SVR-based Music Mood Classification and Context-based Music Recommendation" by Seungmin Rho, Byeong-jun Han, and Eenjun Hwang, published in October 2009, investigates a system meant to recommend music based on song mood categorization. This recommendation system combines many significant technology developments to propose music that is appropriate for the user's mood and surroundings.

The process involves classifying music moods using Support Vector Regression (SVR), which converts mood classification into a regression issue. The system classified music moods with an accuracy rating of 87.8% after utilizing SVR. The recommendation element makes use of both collaborative filtering and ontology technologies to determine the user's mood and situational environment before delivering suitable music recommendations. (Rho, S. (2009)).

Existing System

* **Lucyd**

It is a music recommendation engine created by four graduate students from UC Berkeley's Master of Information and Data Science (MIDS) program. Lucyd allows users to ask for music recommendations using their preferred phrases.

* **Reel time**

The AI system requires users to subscribe. Users may post photographs of huge groups, like retail malls, movie theaters, and restaurants. The algorithm distinguishes joyful and sad moods. The system distinguishes between joyful and sad expressions and assesses the scenario based on the faces of the participants.

* **Music.AI**

It suggests music based on user moods. It combines collaborative and content-based filtering approaches. Music selection considers emotion, time, atmosphere, and learning history.

# Methodology

## Dataset Description

We used Spotify Music data, which includes 686 songs from various artists and genres. All song features are retrieved using the Spotify API. These metrics, such as pace, energy, valence, and danceability, shed light on the qualities and atmosphere of songs. The features indicated in the table below are Raw Features.

We have performed some data pre-processing,

* Loading the dataset “data\_moods.csv” from <https://www.kaggle.com/datasets/musicblogger/spotify-music-data-to-identify-the-moods>
* The dataset has 686 rows and 19 features.

|  |  |
| --- | --- |
| Feature | Description |
| name | The title of the track |
| album | The album where the track is from |
| artist | The artist who performed the track |
| id | A unique identifier for the track |
| release\_date | The release date of the track |
| popularity | A measure of the track's popularity |
| length | The duration of the track in milliseconds |
| danceability | A measure of how suitable a track is for dancing |
| acousticness | A measure of the acousticness of a track |
| energy | A measure of intensity and activity |
| instrumentalness | Indicates the likelihood that a track contains no vocals |
| liveness | Detects the presence of an audience in the recording |
| valence | Describes the musical positiveness conveyed by a track |
| loudness | The overall loudness of a track in decibels |
| speechiness | A measure of the presence of spoken words in a track |
| tempo | The overall estimated tempo of a track in BPM |
| key | The key the track is in |
| time\_signature | An estimated overall time signature of a track |
| mood | The mood classification of the track |

* The dataset statistics descriptions

A screenshot of a calculator

Description automatically generated

* Check whether dataset has any missing values

A screen shot of a computer

Description automatically generated

There are no missing values in the dataset. To deal with Categorical features such as 'name', 'album', 'artist', and 'id', they are removed since they do not significantly add to the model's prediction.   
Numerical features are standardized to have a mean of 0 and a standard deviation of 1, which is critical for machine learning algorithms sensitive to feature scales, such as gradient descent, SVM, and k-nearest neighbours.   
Additionally, the category target label ‘mood' is label encoded. This is crucial because many algorithms, such as the SVM, ANN and XGB Classifier, need categorical target labels to be integer-encoded.

# Feature selection

A Random Forest classifier is used to evaluate the importance of each feature in predicting the target of variable ’mood’. Random Forest, being an ensemble of decision trees, inherently provides feature importance scores based on the average impurity reduction the feature.

A graph with green bars

Description automatically generated

Using the random forest classifier, we discovered that we do not require the dataset's time signature or key to potentially increase our model performance. We'll need to transform the category data into numerical numbers. To simplify, we shall encode the categorical characteristics using pandas' get dummies.

# Exploratory Data Analysis

The mood distribution of songs in dataset is shown below.

**A graph with green squares

Description automatically generated**

**Correlation mapping**

A screenshot of a graph

Description automatically generated

To analyse the correlation heatmap for numerical characteristics and discovered a positive link between danceability and valence. This shows that songs that are more suited for dancing have a more upbeat atmosphere. Similarly, Acoustic Ness and Energy are negatively correlated, implying that songs with greater acousticness ratings have lower energy.

**PCA & T-SNE**

**PCA**

A diagram of a diagram

Description automatically generated with medium confidence

**T-SNE**

**A diagram of a variety of colored dots

Description automatically generated with medium confidence**

First, the PCA and TSNE plots show that moods are clearly distinct (in different clusters), with the exception of energetic and cheerful tunes, which have a large degree of overlap. This was to be expected, given that most energetic tunes are also pleasant.

# Classifications

Most of the classification models are determined which are most significant for the dataset and calculated their accuracy, precision, and recall.

We utilized the following models, splitting the data into 70:20:10 for training, testing, and evaluation: Logistic regression, SGD classifier, Gaussian Naive Bayes, Decision Tree, Random Forest, SVM Linear, and XGB classifier.

Logistic Regression is a linear classification approach that calculates probabilities using a logistic function. We utilized the sigmoid function with cross-entropy loss and L2 penalty (Ridge). (Ryan, 2020)

SGD Classifier: An SVM linear classifier with Stochastic Gradient Descent, L2 regularization, and α = 0.0001.   
Gaussian Naive Bayes: A probabilistic classifier based on Bayes' theorem that assumes features follow a Gaussian distribution.

A Decision Tree is a tree structure with nodes representing characteristics and branches representing decision rules. Splits are calculated by reducing entropy, resulting in maximum information gain. Random Forest: A collection of decision trees trained on random selections of data and then aggregated for predictions.

XGB Classifier: A gradient boosting framework that employs decision trees and has been improved for speed. (Nadeem, 2022)

SVM: A classifier that uses hyperplanes in a high-dimensional space to segregate data. Linear: Separates using a single straight line. Polynomial: Uses polynomial curves to separate. RBF uses non-linear boundaries depending on distance from a central point. (Nadeem, 2022)

A table with numbers and letters

Description automatically generated

# Building a recommendation system

For an effective recommender system, there must be more data points to recommend songs, therefore we populated the dataset by extracting data from Spotify's API, which included essential musical attributes such as valence, danceability, acousticness, energy, instrumentalness, liveness and loudness

Using these features, we utilized XGBoost classifier, which gave us an accuracy of 82.01%. To categorize new songs by mood and add them as new features to the newly extracted dataset. The recommendation system novel approach is individualized music suggestions, and we have developed a novel way for recommending songs to users.

Firstly, user will be choosing 2 -3 songs by themselves and it will be as a data for us to serve our classifier

The algorithm then calculates the average values for 11 musical features, such as valence and danceability, from the user's current playlist. This average is then utilized as input for the XG-Boost classifier, which determines the user's current mood. As the user may be in numerous moods and will play both a peaceful and a sad tune during the session. Thus, averaging is quite significant. With the mood set, begin searching for songs for the user.  
We used a Euclidean search strategy to select songs that closely match the computed average feature values while guaranteeing that the mood of these songs coincides with the user's inferred mood in Classifier.

# Results and outcome

A screenshot of a computer

Description automatically generated

The average of all 11 characteristics is computed and used to assess the user's mood.

In the example above, we receive the mood as Energetic and a list of 5 songs that are recommended to the user based on the Euclidean distance.

We noticed that the elements we employed are quite successful in determining the mood of a music, and as a result, they can accurately represent the listener's mood. The XGB Classifier model produced a good result. We have developed a new method of song recommendation known as the "Moosic", which classifies the mood and analyzes users' music preferences before providing personalized song recommendations that are both sonically compatible and emotionally resonant. The recommender system is considerably good and provides.

# Future works

For the future works we will be integrating with larger dataset. Also, we will be developing a web app where a user can listen to a music and get recommendation as per their moods and behaviours.

# References

Madik, A. (2021). *Mood based music recommendation system*. [online] Available at: <https://www.researchgate.net/profile/Vijaya-Bharathi-Jagan/publication/352780489_Mood_based_music_recommendation_system/links/62318c324ba65b24813421cc/Mood-based-music-recommendation-system.pdf>.

Nadeem (2022). *Introduction to XGBoost Algorithm*. [online] Analytics Vidhya. Available at: <https://medium.com/analytics-vidhya/introduction-to-xgboost-algorithm-d2e7fad76b04>.

Rho, S. (2009). *SVR-based music mood classification and context-based music recommendation*. [online] Available at: <https://www.researchgate.net/publication/221571871_SVR-based_music_mood_classification_and_context-based_music_recommendation>

**Ricci, F., Rokach, L., & Shapira, B.** (2015). Recommender Systems: Introduction and Challenges. In *Recommender Systems Handbook* (pp. 1-34)

Torabi, N. (2023). *The Inner Workings of Spotify’s AI-Powered Music Recommendations: How Spotify Shapes Your Playlist*. [online] Medium. Available at: <https://neemz.medium.com/the-inner-workings-of-spotifys-ai-powered-music-recommendations-how-spotify-shapes-your-playlist-a10a9148ee8d>.

Ryan, A. (2020). *Linear Classification with Logistic Regression*. [online] Available at: <https://www.cs.princeton.edu/courses/archive/spring19/cos324/files/logistic-regression.pdf> [Accessed 5 Apr. 2024].