‘

Project Report on

Personalized Mental Health Recommendations Using LLM & Browsing History

Submitted in partial fulfillment of the requirements for the award of the degree of

Bachelor of Technology

in

Computer Science and Engineering

By

Sreya P. (U2103202)

Sneha R. (U2103199)

Vismaya Balakrishnan (U2103215)

Under the guidance of

Ms. Seema Safar

Department of Computer Science and Engineering
Rajagiri School of Engineering & Technology (Autonomous)
(Parent University: APJ Abdul Kalam Technological University)
Rajagiri Valley, Kakkanad, Kochi, 682039
April 2025

CERTIFICATE

This is to certify that the project report entitled ***"Personalized Mental Health Recommendations Using LLM & Browsing History "*** *is a bonafide record of the work done by* ***Sreya P. (U2103202), Sneha R. (U2103199), Vismaya Balakrishnan (U2103215)****, submitted to the Rajagiri School of Engineering & Technology (RSET) (Autonomous) in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology (B. Tech.) in "Computer Science and Engineering" during the academic year 2021-2025.*

Ms. Seema Safar
Project Guide
Assistant Professor
Dept. of CSE
RSET

Ms. Sangeetha Jamal
Project Coordinator
Assistant Professor
Dept. of CSE
RSET

Dr. Preetha K.G.
Professor & HOD
Dept. of CSE
RSET

ACKNOWLEDGMENT

We wish to express our sincere gratitude towards **Rev. Dr. Jaison Paul Mulerikkal CMI**, Principal of RSET, and **Dr. Preetha K.G.**, Professor, Head of the Department of Computer Science and Engineering for providing us with the opportunity to undertake our project, **Personalized Mental Health Recommendations Using LLM & Browsing History**.

We are highly indebted to our project coordinator, **Ms. Sangeetha Jamal**, Assistant Professor, Department of Computer Science and Engineering, for her valuable support.

It is indeed our pleasure and a moment of satisfaction for us to express our sincere gratitude to our project guide **Ms. Seema Safar**, Assistant Professor, Department of Computer Science and Engineering, for her patience and all the priceless advice and wisdom she has shared with us.

Last but not the least, we would like to express our sincere gratitude towards all other teachers and friends for their continuous support and constructive ideas.

Sreya P.

Sneha R.

Vismaya Balakrishnan

Abstract

With the increasing reliance on digital platforms, understanding users' emotional states through their online behaviour has become critical for enhancing personalized experiences like music recommendations. After a long day of work, stress can feel overwhelming, leaving the mind restless and tense. Listening to calming music reduces cortisol levels and promotes relaxation, offering an instant sense of relief. The proposed work develops a content-based music recommendation system leveraging a Large Language Model (LLM) using multimodal data to enhance mental health. The methodology improves personalization by utilizing user's browsing history and chat history to detect the emotional state of the user. The recommendation pipeline of the model includes three important components such as feature extraction phase, emotion analysis phase and recommendation engine. In the feature extraction phase, browsing history, chat history and video data for a specific timestamp are used as input. The extracted features are fed into emotional analysis model, LLaMA2, to detect the respective emotions (happy, sad, calm, angry). The detected emotion is fed into the recommendation engine to provide top n personalized music recommendations.

Contents

List of Abbreviations

API - Application Programming Interface

CBT - Cognitive Behavioral Therapy

EDA - Exploratory Data Analysis

GPT - Generative Pre-trained Transformer

LLM - Large Language Model

NDCG - Normalized Discounted Cumulative Gain

NER - Named Entity Recognition

PTA - Pure Tone Audiometry

RAH - RecSys–Assistant–Human

RMSE - Root Mean Square Error

THI - Tinnitus Handicap Inventory

T5 - Text-To-Text Transfer Transformer

List of Figures

List of Tables

Chapter 1

Introduction

With the increasing reliance of the world on digital spaces, being aware of the users' emotional status from their digital behaviors has become imperative to enhancing personalized experiences like music recommendations. Music is a powerful determinant of human feelings, employed to alleviate stress, boost mood, and foster overall wellness. Most conventional recommendation systems are based on user history and preferences but lack any awareness of current emotional states. This project recommends an emotion-sensitive music recommender system utilizing multimodal data—video analysis, chat interaction, and browsing history—to determine the user emotions and provide customized music recommendations. Based on the employment of LLaMA2 in emotion detection and content-based filtering for recommendation, the system ensures that the proposed music is contingent upon the real-time emotional condition of the user. Through Large Language Models (LLMs) and machine learning, the system in question aims at closing the gap between artificial intelligence personalization and emotional intelligence, ultimately to create better user experience and mental health.

1.1 Background

Music has long been used as a successful method of influencing and manipulating emotions. Psychological and neuroscientific studies have shown that music can exert a significant influence on mood, stress, and overall mental health. With the rise of digital music streaming services, recommendation algorithms have become increasingly sophisticated to provide personalized experiences based on user preference, listening habits, and demographic information. However, the vast majority of standard recommendation systems tend to overlook actual-time emotional situations, and thereby generate mismatched recommendations that aren't in accord with the prevailing mood of the user at a particular

time.
New Artificial Intelligence (AI) and Natural Language Processing (NLP) advancements, particularly using Large Language Models (LLMs) like LLaMA2, allow us to enable more sophisticated emotion detection through text and multimodal data analysis. By examining surf history of the users, chat dialogues, and even faces during video chats, AI has a better capacity to predict states of emotions than conventional sentiment analysis methods. Embedding multimodal emotion recognition within recommendation systems uncovers new dimensions for personalized content delivery, specifically in music streaming apps.
With increasing consciousness regarding mental welfare and digital wellbeing, there is a greater demand for affect-sensitive recommendation platforms that dynamically adapt to the emotional state of the users. The aim of this project is to bridge the gap between AI-driven personalization and emotional intelligence by developing a music recommendation platform based on LLMs and multimodal emotion recognition to improve user experience and welfare.

1.2 Problem Definition

The aim of this project is to develop a personalized music recommendation system that dynamically aligns song suggestions with users' emotional states by analyzing their browsing history. Traditional music recommendation systems lack the ability to adapt to real-time moods, making it challenging to offer contextually relevant music that supports users' mental well-being. This project seeks to address this limitation by creating a framework that detects user emotions through online behavior and recommends mood-congruent music.

1.3 Scope and Motivation

The scope of this project encompasses the development of an emotion-driven music recommendation system that leverages users' online behavior to provide personalized song suggestions. This system collects multimodal data, analyzes it to detect the user's current emotional state, and then matches this mood with songs clustered by emotional characteristics such as energy and valence. By applying machine learning techniques, including a fine-tuned LLAMA model for emotion detection and K-Means clustering for music categorization, the project aims to offer a more responsive and contextually relevant recommendation experience. The system also incorporates content-based filtering to refine the recommendations and the Spotify Web API to deliver a seamless listening experience. Overall, this project is designed to operate as an intelligent, emotion-aware recommendation platform that bridges the gap between online behavior and music preferences.

The motivation behind this project stems from the growing understanding of music's impact on emotional well-being and the desire to make technology more attuned to users' real-time emotional needs. While traditional recommendation systems focus on long-term preferences or general listening habits, they often overlook situational context, which can limit their effectiveness in supporting users' mental health. By creating a recommendation system that adapts to a user's current mood, this project seeks to enhance user satisfaction and potentially contribute to emotional regulation through music. The integration of AI-driven emotion analysis in recommendation systems also showcases how machine learning can play a role in improving digital mental health solutions. This project is a step toward creating emotionally intelligent technologies that not only personalize experiences but also foster a sense of well-being.

1.4 Objectives

- Develop an emotion-aware music recommendation system that adapts to the user's real-time emotional state using multimodal data.
- Leverage browsing history, chat interactions, and video analysis to accurately infer emotions using LLaMA2 for emotion detection.
- Implement K-Means clustering to categorize songs based on emotional attributes such as valence and energy for mood-based recommendations.
- Apply content-based filtering to match detected emotions with relevant songs for personalized recommendations.
- Integrate a Chrome extension and Spotify Web API for seamless data collection and music playback.

- Enhance the user experience and mental well-being by providing context-aware and emotionally intelligent music recommendations.

1.5 Challenges

The project faces challenges in ensuring user privacy by securely handling sensitive browsing data, navigating Spotify API limitations that may restrict music playback, and achieving accurate emotion detection, as nuanced emotions can be difficult to classify precisely. Addressing these issues is essential for a seamless, personalized user experience. Solutions will need to balance data security, API usage constraints, and advanced model refinement.

1.6 Assumptions

- Users have active Spotify accounts, enabling seamless music playback through the Spotify API.
- The Chrome extension accurately captures comprehensive browsing activity, providing reliable data for emotion detection.
- User privacy is maintained through secure data handling practices, with explicit user consent for data collection and usage.

1.7 Societal / Industrial Relevance

The project holds significant relevance both socially and industrially. By providing personalized music recommendations based on users' emotional states, it enhances user engagement and creates a more immersive, emotionally aware listening experience. In a societal context, this technology can support mental well-being by aligning music with users' moods, potentially helping individuals manage stress or enhance relaxation. For the music and tech industries, this approach offers a new dimension of personalization, driving user retention and engagement for platforms like Spotify and opening up avenues for targeted emotional marketing.

1.8 Organization of the Report

- Chapter 1: Introduction introduces the background of the project, defines the problem, and outlines the study's scope, which includes developing an emotion-driven recommendation system that leverages users' browsing history to assess mood and recommend music accordingly. It covers key aspects like objectives, challenges, assumptions, and the societal and industrial relevance of aligning music recommendations with users' emotional needs.

- Chapter 2: Literature survey discusses the current state of technology in user data analysis, mood detection, and recommendation systems, including traditional methods and modern LLMs such as Google T5 and LLaMA. It examines the methodologies and findings of relevant studies, addresses challenges such as privacy and real-time adaptability, and identifies gaps in existing research, which this project seeks to address through a novel integration of emotion detection and music recommendation techniques.

- Chapter 3: Requirements provides the software and hardware requirements needed to deploy the system. It explains why Python was chosen as the programming language and libraries such as PyTorch, Scikit-learn, OpenCV, and Spotify Web API for different components. It also dictates system hardware such as CPU, GPU, RAM, and camera settings required for model inference, data processing, and video analysis.

- Chapter 4: System Architecture presents the technical blueprint of the project, detailing the architecture, components, and workflows that drive the system. It describes the modular and layered design, outlining how user browsing data is captured, processed for emotion detection using LLaMA, and matched with mood-based music clusters via K-Means clustering. The chapter also covers integration with Spotify for seamless playback, highlights tools and technologies used, and emphasizes the scalability, privacy, and social impact of the system.

- Chapter 5: System Implementation explains the sequential implementation of each module in the system. It comprises elaborate explanations on how the browsing history is gathered through the use of a Chrome extension, how emotion detection is done through the use of LLaMA2 in textual as well as visual data, and how music recommendations are produced using K-Means clustering and content-based filtering. This chapter also deals with how Spotify Web API is implemented for playing music in real-time and how user feedbacks are used for ranking the hit songs.

- Chapter 6: Results and Discussions addresses the assessment of system performance. It discusses accuracy in emotion detection, user satisfaction with music recommendations, and Spotify integration efficiency. The discussion considers strong points like multi-modal emotion detection, as well as weaknesses like mixed emotion ambiguity, data privacy, and facial recognition issues.

- Chapter 7: Conclusion and Future Scope summarizes the outcomes of the project, emphasizing its success in delivering real-time, personalized music recommendations driven by emotional context. It suggests possible future enhancements like incorporating more emotional labels, strengthening privacy, and extending compatibility with other music platforms.

1.9 Summary of the Chapter

This chapter provided a comprehensive overview of the objectives, methodology, and social significance of the project. It introduced the core concept of identifying users' emotional states from browsing history, chatbot interactions, and facial expressions using the LLaMA model. Depending on the recognized emotions, mood-based music is suggested using K-Means clustering and provided by the Spotify Web API. Major challenges were resolved, including ensuring user privacy, solving API limitations, and managing the complexity of blended emotional states. The chapter also pre-empted the rest of the report by defining the research background, technological underpinnings, system requirements, architectural design, and evaluation framework. Collectively, these form the underpinning that helps to delve into how emotion-based music recommendation is able to add value to the user experience and improve emotional well-being using AI-based personalization.

Chapter 2

Literature Survey

Advancements in machine learning, natural language processing, and data analytics have improved the effectiveness of personalized recommendation systems. With increasing digital interactions, the ability to understand user behavior and emotional context is vital for designing user experiences. This chapter discusses research on the analysis of user data, such as browsing history and content preferences, to determine moods, behaviors, and needs.

The studies to be discussed examine methodologies to tap into the browser history for inferring emotional states, and also to explore a host of machine learning approaches, ranging from LLMs, toward recommendation systems. Through a review of previous literature, we are hoping to lay out best practices, challenges, and gaps as a precursor for developing our personalized music recommendation system based on the alignment with real-time emotional states for users.

2.1 Reconstructing Detailed Browsing Activities from Browser History [1]

2.1.1 Methodology

Dataset Selection

The dataset serves as the foundation for the study. It was collected from 185 participants recruited via Amazon Mechanical Turk. Participants were required to install a Chrome extension that recorded two types of data:

- **Browsing History:** Gathered using Chrome's History API, which logs URLs visited, timestamps, and navigation types (e.g., clicking links, refreshing, or navigation within a frame). This data represents a limited snapshot, capturing only navigation events and not user engagement (e.g., tab switches or idle periods).

- **Reference Activity Data:** Obtained via Chrome's APIs for tabs, windows, and idle states. This dataset included detailed logs of tab/window focus, switching, and user actions such as mouse movements, scrolling, and keyboard activity. It allowed the researchers to create ground-truth activity spans by logging when users began and ended interactions with specific URLs. Participants were paid dollar 2 for installing the extension and received a dollar 1 weekly bonus for sustained participation. Data from users who uninstalled the extension or became inactive for more than three days were excluded. Of the 225 recruited users, data from 185 participants were retained after filtering.

2.1.2 Reconstruction Procedure

The reconstruction process was divided into three key tasks, each requiring distinct methodologies and machine learning models:

- **Estimating Active Browser States** The first step in reconstructing browsing activity was identifying when the browser was active (focused and in use) versus idle or unfocused.

  - **Definition of Browser Activity:**
    A browser was deemed active if it was the focused window and there was evidence of recent user interaction (e.g., mouse movement, scrolling, or keystrokes within the past minute).
  - **Challenges:**
    Browsing history alone does not capture idle periods, tab switches, or focus shifts, making naive heuristics (e.g., assuming activity for a fixed period after a recorded event) insufficient.
  - **Machine Learning Approach:**
    A random forest classifier was trained to predict browser activity on a second-by-second basis, using features derived from browsing history:

    * Temporal Features:
      Time since the last recorded activity.
      Time until the next recorded activity.
      Duration between consecutive activities

    * Domain-Specific Features:

      Frequency of navigation events for specific domains.

      RescueTime productivity ratings for domains (e.g., very productive, distracting).

    * One-Hot Encoded Features:

      Top 20 most frequently visited domains were encoded to capture domain-specific behaviors.

  - **Performance:**

    The model achieved an F1-score of 0.84, significantly outperforming heuristics (e.g., assuming activity for five minutes post-navigation, which achieved an F1-score of 0.79).

- **Predicting the Focused Domain**

  Once browser activity was established, the next task was determining the domain (website) that the user was actively browsing.

  - **Challenges:**

    Browsing history does not reflect tab switches or multiple open windows, leading to discrepancies between recorded navigation events and actual focus.

  - **Model Design:**

    This task was formulated as a multi-class classification problem, predicting the active domain at any given second.

  - **Classes:**

    Current domain (most recent navigation event). Next domain (upcoming navigation event). Previously visited domains (one or two steps back in history) to account for tab switches.

  - **Features:**

    Temporal features (e.g., time since the last navigation, time until the next navigation).

    Behavioral features (e.g., frequency of switching to certain domains).

    Categorical domain features encoded using one-hot encoding.

    Metadata features (e.g., referring visit IDs to track navigation paths).

- **Classifier:**

  A random forest model was trained to predict which domain the user was focused on.

- **Performance:**

  The model correctly identified the active domain in 76.2% of cases. It performed particularly well on common domains but struggled with rarer sites not included in the top 20 categories.

- **Estimating Time Spent**

  The final task was estimating the total time users spent online and the time spent on specific domains.

  - **Approach:**

    By summing the predicted active seconds and associating them with domain predictions, the model estimated:
    Total time spent online.
    Total time spent on each domain.

  - **Evaluation:**

    The reconstructed total browsing time correlated strongly with ground truth data ($R^2 = 0.96$).
    Time spent on individual domains also showed high accuracy, with a mean R^2 value of 0.92.

2.1.3 Model Training and Testing

To ensure robustness, the dataset was split into training (93 participants) and testing (92 participants) sets. Random forest classifiers were used for both tasks, leveraging H2O's implementation with default parameters. The models were benchmarked against simpler heuristics to demonstrate their superior performance.

2.1.4 Error Analysis and Limitations

The paper acknowledges several sources of error and limitations in the methodology:

- Underestimation of Activity: Occurred on sites with infrequent navigation events (e.g., long-form articles or videos).
  Mitigation strategies include using headless browsers to infer content duration (e.g., video length or text readability).
- Cleared Browsing History: Some users cleared portions of their history, leading to incomplete data.
- Incognito Mode: Browsing history does not capture activity in private (incognito) mode, limiting the scope of reconstruction.

2.1.5 Conclusion

As an extension of previous work, the method successfully demonstrates detailed browsing activities reconstructed from browser history with high accuracy. Leveraging machine learning in the method overcomes the limitations of conventional heuristics. The data reconstruction enables the productivity tools, user behavior studies, and targeted interventions without requiring intrusive, long-term monitoring. Such an approach opens new avenues for scalable and privacy-conscious data collection in various applications.

2.2 RAH! RecSys–Assistant–Human: A Human-Centered Recommendation Framework With LLM Agents [2]

2.2.1 Model Achitecture

The RAH Framework is a human-centered design for a recommendation system that unifies the functionalities of a recommendation system, an intelligent assistant, and the human user. It exploits multiple LLM-based agents that form an intricate system to meet the needs of the users by personalizing, evaluating, and refining recommendations through continuous learning. Below are the roles of each agent in the framework, along with additional details about each component and process.

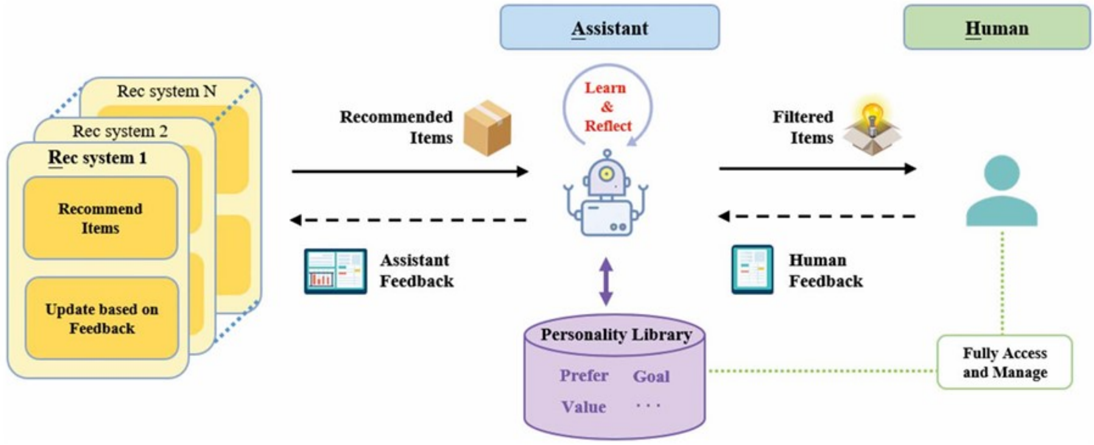


Figure 2.1: RAH Architectural Diagram[2]

Perceive Agent: It is the entry point for the recommendation process, where related information about items will be passed through to add contextual details. The role played by this agent forms a building block for creating an early but detailed profile for every item. The agent processes item titles, such as movie names, and also gathers associated metadata, including genres, tags, plot summaries, and other characteristics. Ensures that all agents in the framework can access rich and meaningful information about each item, thereby improving the overall ability of the assistant to make relevant recommendations.

Learn Agent: Builds a dynamic profile based on user interactions and feedback, essentially "learning" the user's preferences over time. This agent's profile is central to understanding the user's taste and behavior patterns. Translates user actions (e.g., likes, dislikes, or ratings) into a preference profile by analyzing repeated behaviors and aligning them with specific characteristics of items. Isolates features in items that consistently correlate with positive or negative feedback to refine its understanding of the user's preferences.

Act Agent: This agent is to apply the learned user profile in order to make actionable recommendations. It applies reasoning in predicting how a user may react to new items. It analyzes each item against the user profile that was developed by the Learn Agent in terms of whether it aligns with the user's preferences. Uses chain-of-thought reasoning: hypothesizes what the user may respond with, assesses its probability, and simulates a response. This structured reasoning helps provide nuanced recommendations that better fit the user's interests.

Critic Agent: This agent acts as a quality control system, assessing recommendations from the Act Agent to be sure they meet actual user preferences and correcting differences. Compares the items recommended with what the actual user feedback is, matching any mismatches between predicted and observed actions. Returns error analysis to the Learn Agent, which, in turn helps refine the user profile for improving accuracy in further recommendations.

Reflect Agent: This is done by revisiting and reconciling the learned preferences from time to time so that the profile remains up-to-date and conflict-free. It is also a means of adaptation to the changes in user interest. Periodically reviews the profile to ensure that it aligns with the user's most recent behaviors and removes redundancies or contradictory information. Resolves conflicts by refining overlapping or contradictory preferences, prompting the user if further clarification is needed.

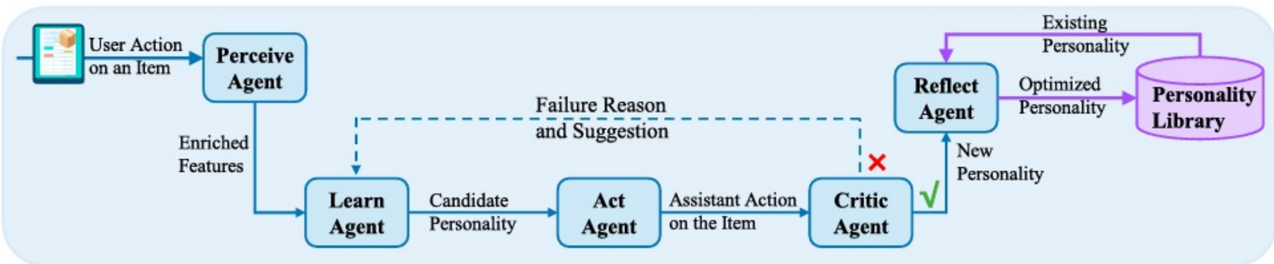


Figure 2.2: Assistant learning User Personalities[2]

2.2.2 Processes in RAH Framework

Learn-Act-Critic Loop:

Process: The Learn Agent formulates the initial user profile based on the past interaction of the user. Using that profile, the Act Agent makes recommendations and generates them for the user. The Critic Agent compares those recommendations with user feedback and assesses them. If the predictions go wrong, the Critic Agent will give appropriate feedback to the Learn Agent. This is an iterative loop that continues, allowing the agents to collaboratively refine the user profile and improve the quality of recommendations over time.

Reflection Mechanism:

It is designed to keep the user profile current and coherent while it evolves over time. The mechanism prevents outdated or contradictory information from piling up in the user's profile. The Reflect Agent merges the most recent feedback with the existing user profile, which then identifies and resolves any duplication or conflict. If contradictory preferences arise (e.g., liking and disliking similar items over time), the Reflect Agent tries to reconcile these through finer divisions of preference. If conflicts still occur, the Reflect Agent asks for clarification from the user. In doing so, it ensures that the profile does not prematurely drift away from representing current preferences.

2.2.3 Experiment and Evaluation

Experiment 1: Alignment with User Preferences

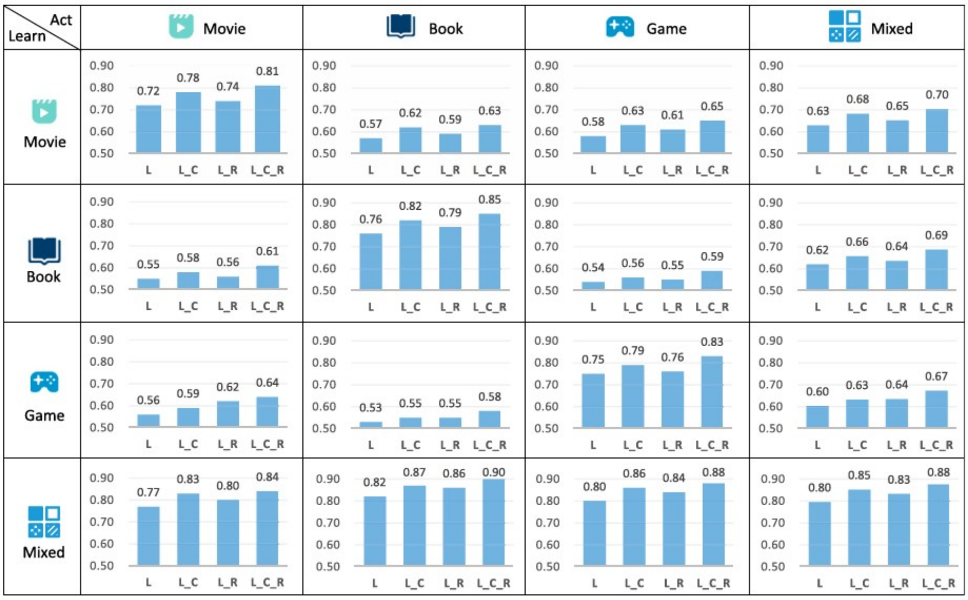


Figure 2.3: Histogram of Single, Cross and Mixed Domains[2]

The first experiment sought to determine the extent to which the assistant agreed with user preferences across various domains. The objective was to test whether the assistant was able to recommend relevant items that matched the user's actions. The evaluation was done using F1-scores, which are useful in measuring the accuracy of recommendations both within single domains and across multiple domains. This work compares three configurations: the Reflect Agent, the Learn-Act-Critic loop, and a baseline setup. The results clearly indicate that the Reflect Agent obtained the highest F1-scores across domains; this is consistent with a more accurate understanding of the user's preferences and therefore better recommendation accuracy. The histogram indicates the F1-scores of each configuration on the domains movies, books, and games.

Experiment 2: Reducing User Burden and Bias Mitigation

The second experiment focused on reducing user burden and mitigating bias in the recommendations. The objective was to evaluate the framework's effectiveness in minimizing redundant interactions and offering proactive suggestions that enhance the overall user experience. The assistant's proactive feedback system was found to effectively calibrate the recommendation algorithms, enabling a more personalized experience for users with fewer interactions required. Additionally, the assistant's broader feedback mechanisms helped reduce selection bias, ensuring that the recommendations were more diverse and fair. The results from this experiment were captured through metrics such as Normalized Discounted Cumulative Gain (NDCG@10) and Recall@10. Table 1 summarizes the results, showing how the assistant's feedback improved the performance of recommendation algorithms in terms of user satisfaction, as well as how it contributed to a reduction in bias. The improvements in these metrics indicated that the assistant's guidance led to more relevant and less biased recommendations, benefiting the user in terms of both quality and diversity of suggestions. [4] [5]

The RAH framework is the most sophisticated and multi-agent-based recommendation system to date, further improving its alignments with user preferences over time via iterative learning and quality control of its profile updated regularly. A role for every agent is in place, from preliminary information gathering towards the refinement of preference, a dynamic and human-centered recommendation experience that adapts to the shifting needs of its users while staying respectful of user privacy.

Method	Assistant	Movie		Book		Game		Mixed	
		NDCG@10	Recall@10	NDCG@10	Recall@10	NDCG@10	Recall@10	NDCG@10	Recall@10
LightGCN	No	0.5341	0.5240	0.1527	0.1711	0.4163	0.4934	0.3790	0.1900
LightGCN	Yes	0.5524(+0.0183)	0.5339(+0.0099)	0.1830(+0.0303)	0.1912(+0.0201)	0.4330(+0.0167)	0.4974(+0.0040)	0.4058(+0.0268)	0.2033(+0.0133)
PLMRec	No	0.1171	0.1610	0.0149	0.0181	0.3964	0.4743	0.1346	0.0739
PLMRec	Yes	0.1200(+0.0029)	0.1692(+0.0082)	0.0162(+0.0013)	0.0197(+0.0016)	0.3981(+0.0017)	0.4790(+0.0047)	0.1378(+0.0032)	0.0766(+0.0027)
FM	No	0.3897	0.4200	0.1443	0.1561	0.2903	0.3529	0.2533	0.1336
FM	Yes	0.3919(+0.0022)	0.4257(+0.0057)	0.1474(+0.0031)	0.1603(+0.0042)	0.2937(+0.0034)	0.3624(+0.0095)	0.2549(+0.0016)	0.1340(+0.0004)
MF	No	0.4122	0.4714	0.1434	0.1484	0.2618	0.3422	0.2302	0.1279
MF	Yes	0.4300(+0.0178)	0.4781(+0.0067)	0.1520(+0.0086)	0.1593(+0.0109)	0.2998(+0.0380)	0.3706(+0.0284)	0.2651(+0.0349)	0.1487(+0.0208)
ENMF	No	0.4931	0.4544	0.1195	0.1199	0.0751	0.1156	0.3056	0.1446
ENMF	Yes	0.5200(+0.0269)	0.4831(+0.0287)	0.1224(+0.0029)	0.1217(+0.0018)	0.0788(+0.0037)	0.1247(+0.0091)	0.3224(+0.0168)	0.1531(+0.0085)
NeuralMF	No	0.4464	0.4517	0.1559	0.1578	0.3301	0.3913	0.3220	0.1603
NeuralMF	Yes	0.4856(+0.0392)	0.4906(+0.0389)	0.1631(+0.0072)	0.1658(+0.0080)	0.3507(+0.0206)	0.4086(+0.0173)	0.3451(+0.0231)	0.1742(+0.0139)
ItemKNN	No	0.1900	0.1698	0.1326	0.1051	0.2500	0.3035	0.2338	0.1090
ItemKNN	Yes	0.2131(+0.0231)	0.1860(+0.0162)	0.1517(+0.0191)	0.1171(+0.0120)	0.2660(+0.0160)	0.3125(+0.0090)	0.2567(+0.0229)	0.1170(+0.0080)

Figure 2.4: Performance of proxying user feedback and adjusting recommender systems[2]

2.3 Advancing Tinnitus Therapeutics: GPT-2 Driven Clustering Analysis of Cognitive Behavioral Therapy Sessions and Google T5-Based Predictive Modeling for THI Score Assessment [3]

2.3.1 Methodology

Data Collection :

Data collection was structured to gather patient records that contribute to a robust prediction model for tinnitus-related impacts. The following sources and elements were systematically collected:

1. CBT Diaries: Textual entries from tinnitus patients were acquired, detailing daily emotional states, reflections, and concerns. These entries reflect cognitive patterns affected by tinnitus and are essential for identifying how patients' experiences correlate with THI scores.

2. Audiometric Data: Pure Tone Audiometry (PTA) results were gathered, covering standard frequency bands (e.g., 500 Hz to 8 kHz) for each ear, with both air and bone conduction values recorded. These PTA values provide objective measurements of hearing impairment.

3. THI Scores: THI assessments, covering functional, emotional, and catastrophic dimensions, were taken at two intervals:

   - Initial Scores before the onset of CBT.

   - Final Scores post-CBT intervention.

The data was sourced with patient consent and anonymized to maintain privacy and compliance with ethical guidelines.

Data Preprocessing :

A preprocessing pipeline was designed to standardize the collected data, preparing it for LLM input while minimizing noise and inconsistency. Steps included:

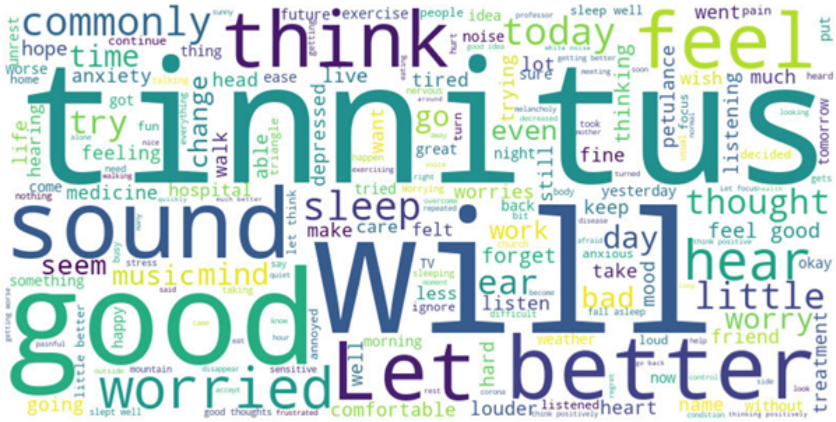
1. Text Cleaning: Diary entries were stripped of timestamps, personally identifiable information, and non-standard symbols. Sentence boundaries were clarified for better tokenization.

2. Translation and Anonymization: Non-English entries were translated to English, ensuring language consistency. Anonymization techniques, including named entity recognition (NER), were used to mask any patient identifiers within the text.

3. Tagging and Normalization: A structured tagging schema was developed for clear LLM input:

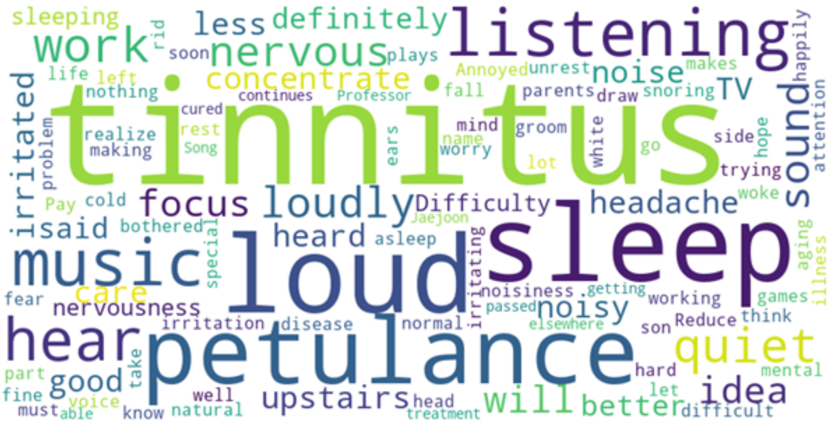
   - <freq_info>tagged PTA data.
   - <initial_thi_score>for baseline THI values.
   - <diaries>for narrative text entries.
   - <final_thi_score>for THI scores post-CBT.

4. Normalization of Scores and PTA Values: Data normalization was applied to THI and PTA values to ensure consistency across different scales. Standardization helped the model recognize patterns in auditory scores and THI results without being impacted by outlier values.



(a) Cluster 0 word-cloud image

(b) Cluster 1 word-cloud image



(c) Noise cluster word-cloud image

Figure 2.5: Results of word-cloud for each cluster. [3]

Tinnitus Data Augmentation :

Given the limited dataset, data augmentation techniques were essential to create a sufficiently large dataset for model training. The following methods were employed:

1. Textual Augmentation Using NLPaug: Augmentation techniques included:
    - Synonym Replacement: Swapped words with synonyms to vary sentence construction.
    - Random Substitution: Introduced small word changes to increase sentence diversity while retaining meaning.
    - Contextual Embeddings: Utilized contextual embedding models to generate alternate versions of diary entries, maintaining contextual integrity.
2. Numerical Variability in PTA and THI Scores: A bootstrapping algorithm was implemented, adding slight randomized variations to PTA and THI scores while preserving overall data distribution.
3. Augmentation Quality Check: The readability of augmented text entries was evaluated using metrics like the Flesch Reading Ease and Gunning Fog Index. Entries were retained only if they adhered to a readability standard compatible with the original dataset.
4. Dataset Partitioning: The augmented dataset was partitioned into:
    - Training Set: $\approx$ 42,000 entries.
    - Validation Set: $\approx$ 10,500 entries.
    - Test Set: $\approx$ 2,100 entries.

These partitions provided ample data for model training, validation, and evaluation, ensuring robust model performance.

Prediction with Google T5 Transformer :

The Google T5 and Flan-T5 models were chosen for their text-to-text capabilities, enabling THI score prediction as a transformation task. The prediction pipeline included:

1. Data Tokenization: Entries were tokenized with a cap of 1,024 tokens, which allowed the model to process both short and long entries. Tokenization balanced the need for context with the model's processing constraints.

2. Fine-Tuning with THI Prediction Objective: The T5 models were fine-tuned to generate a normalized THI score string from tagged diary and PTA inputs. Training focused on minimizing the difference between predicted and actual scores using a Root Mean Square Error (RMSE) objective.

3. Inference Settings: For inference, top_p (0.9) and top_k (64) parameters were adjusted to enhance output precision and manage diversity in score predictions, aiming for accuracy in final THI score estimations.

4. Model Evaluation Metrics:

   - ROUGE-L: Calculated between model-generated and actual THI scores to measure the model's ability to predict accurate sequences.
   - RMSE: Assessed predictive accuracy by comparing model output to actual THI scores.

5. Final Output Processing: Post-processing steps ensured that model output adhered to the structured score format, facilitating easy validation and comparison with ground-truth THI scores.

2.3.2 Results

1. Patient Clustering with LLM Embeddings:

   - Used GPT-2-based LLM embeddings to analyze tinnitus CBT session assignments.
   - Applied DBSCAN clustering algorithm to identify three distinct patient clusters based on their reactions during CBT sessions.

2. Connection Between CBT Assignments and Outcomes:

   - LLM-based clustering revealed a direct link between the textual content of CBT assignments and treatment outcomes.

- Patients who adhered to basic CBT principles (e.g., correcting maladaptive thoughts) showed significant improvements in all three THI (Tinnitus Handicap Inventory) scores.

3. Performance Comparison of LLMs:

   - Explored Google T5 and Flan-T5 for understanding CBT sessions.
   - Both LLMs showed similar performance but Google T5 outperformed Flan-T5 in handling complex augmented data (e.g., in Augmentation 1).
   - Google T5 demonstrated better generalization, likely due to broader pre-training exposure to diverse linguistic structures.

4. Impact of Data Augmentation:

   - Augmentation complexity increased error rates (e.g., RMSE and ROUGE-L) due to the challenges of processing intricate linguistic data.
   - The 'distilbert-base-uncased' model generated more complex paraphrases and typos than the 'roberta-base' model, affecting readability and accuracy.

5. Numerical Augmentation and Penalties:

   - Numerical augmentation, with added penalties in Augmentation 3, resulted in lower RMSE and ROUGE-L scores, helping to reduce overfitting.
   - Bootstrapping-based augmentation expanded the dataset but did not increase its diversity, posing a risk of overfitting to the resampled data.

6. Challenges of Small Datasets:

   - Limited dataset size increases the risk of overfitting, particularly when bootstrapping is used for augmentation.
   - A more diverse and authentic dataset is needed to better represent the broader tinnitus patient population.

7. Benefits for Otolaryngologists:

   - LLMs could improve the efficiency of CBT sessions by predicting treatment outcomes, allowing doctors to manage more patients in less time.

- LLM-based A.I. could automate the monitoring of CBT sessions, reducing the need for manual review and improving treatment planning.

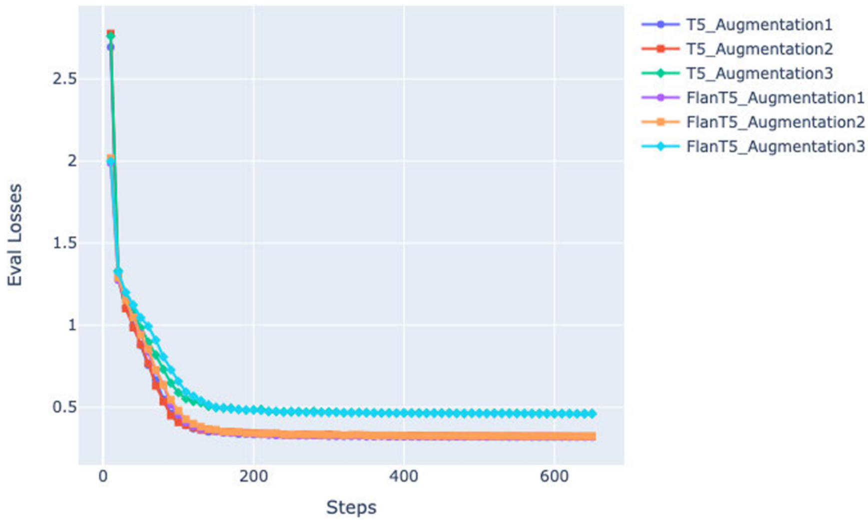


Figure 2.6: Google T5 and Flan-T5 evaluation losses of the augmentation dataset 1, 2 and 3. [3]

2.3.3 Conclusion

This paper proposes that an LLM-based approach can significantly reduce the workload of reviewing each CBT session and accurately predict the outcome of tinnitus CBT treatment.[6] Experimental results show that Google T5 and Flan-T5 LLMs can predict treatment outcomes, even with grammatical and numerical errors in the dataset, demonstrating robustness and reliability.

Heavy pre-training models such as Google T5 and Flan-T5 with augmented datasets are suggested for clinical adaptation. These models may diminish the workload of healthcare professionals and efficiently improve the treatment outcome through precise, personalized prediction.

Further exploration in both clinical trials and real-world settings is encouraged, to further study the capabilities as well as the limitations of such technologies, address privacy and ethics concerns, but also open other avenues of work beyond tinnitus, including predicting and monitoring the sessions of the CBT treatments for depressive patients that can change management of complex health conditions.

2.4 Summary and Gaps Identified

2.4.1 Summary

Title	**Advantages**	**Disadvantages**
Reconstructing Detailed Browsing Activities from Browser History [1]	Enables accurate reconstruction of user activity from non-invasive browser history data; high accuracy in estimating time spent on sites; applicable to productivity tracking and behavioral research.	Limited scope, as it cannot capture activity in private browsing modes; requires participants' explicit consent, which may affect sample diversity; machine learning models are computationally intensive.
RAH! RecSys–Assistant–Human: A Human-Centered Recommendation Framework With LLM Agents [2]	Improves recommendation relevance through dynamic, multi-agent feedback loop; adaptable to evolving user preferences and needs; reduces user bias by learning from comprehensive feedback.	Complex model with high computational requirements, impacting scalability; implementation requires significant resources, both in terms of data processing and model refinement; potential difficulty in maintaining real-time adaptability.
Advancing Tinnitus Therapeutics: GPT-2 Driven Clustering Analysis of Cognitive Behavioral Therapy Sessions and Google T5-Based Predictive Modeling for THI Score Assessment [3]	Demonstrates potential for LLMs in healthcare applications, particularly in automating CBT assessment and predicting THI scores; successful clustering of patient responses improves treatment personalization.	Dependence on small datasets limits the generalizability of findings; challenges with data augmentation lead to noise and reduced accuracy in predictions; sensitive to data quality, as minor errors can impact clinical applicability.

Table 2.1: Comparison of advantages and disadvantages of different studies

2.4.2 Gaps Identified

- Respecting Privacy and Consent: Many of these models rely on sensitive data, like people's browsing histories or personal health information. This raises important questions about privacy and the ethical use of such data. It's crucial to find ways to protect users' anonymity and security without compromising on functionality. This is especially true when gathering data in real-time, where people's consent needs to be actively managed throughout their experience.

- Need for More Diverse Datasets: A lot of current models are built on small, fairly uniform datasets. This is particularly common in health studies, like those focused on conditions such as tinnitus. The lack of diversity can mean that these models don't always work well for different groups of people. Including a wider range of demographics and behaviors would make these models more robust and applicable to a broader audience, ultimately helping more people.

- High Demands on Resources: Sophisticated systems, like the multi-agent RAH! RecSys framework, require substantial computing power, which can make them hard to scale. This becomes a problem, especially for applications that need real-time responsiveness. Simplifying these systems without sacrificing accuracy could make them more practical and accessible, helping them reach a wider audience without the heavy resource drain.

- Adapting to Real-Time Emotions: While some recommendation systems can adjust based on user feedback, they often lack real-time emotional adaptability. Making these systems more responsive to users' moods and contexts as they happen could lead to a better, more personalized experience. Faster and more intuitive emotion-detection capabilities would help these systems deliver more relevant suggestions right when they're needed.

- Challenges with Data Augmentation: Many studies turn to data augmentation to boost the amount of data they're working with, but these techniques can sometimes introduce noise or make the data harder to interpret. In areas like healthcare, where precision is key, it's especially important to develop augmentation methods that add value without distorting the data's meaning. Finding ways to enrich datasets while keeping the information clear and accurate would go a long way in supporting reliable outcomes.

2.5 Conclusion

This chapter delves into research that has made great leaps in the establishment of personalized systems. Such systems use user data and emotional cues to make tailored recommendations. Innovations in reconstructing browsing behavior, multi-agent recommendation frameworks, and predictive models for mental health assessment showcase the manner in which machine learning and language models can enhance user experience by being aligned with the individual preferences and emotional needs of a person.

However, there are still common challenges across all these approaches. Some of the main issues are related to privacy, diversity of data to enhance the robustness of models, the computational complexity of the complex systems, the real-time adaptability of emotional responses, and the enhancement of data augmentation methods to maintain the quality of data. These gaps suggest that solutions need to balance personalized interaction with ethical responsibility, resource efficiency, and real-time responsiveness.

Approaches leading to enhancements of privacy safeguards, diversification of the training data and improvement of the computational efficiency in developing these areas, will lay more fertile grounds in the future development of adaptive, sound, ethically designed systems for effective response of dynamic needs of users' or emotions, thereby moving towards highly responsive and personalized technology.

Chapter 3

Requirements

3.1 Software Requirements

3.1.1 Programming Language: Python

- **Why Used?**
  - Best suited for ML/NLP tasks.
  - Rich ecosystem for deep learning, clustering, and API interaction.
- **Why Not Others?**
  - **C++/Java**: Too complex for ML workflows.
  - **R**: Lacks deep learning and API integration capabilities.

3.1.2 ML and NLP Frameworks

- **PyTorch**
  - **Why Used?**
    * Required for implementing **LLaMA** (Large Language Model).
    * Supports efficient GPU acceleration and dynamic computation graphs.
  - **Why Not TensorFlow?**
    * PyTorch is more flexible and widely used in research.
- **Scikit-learn**
  - **Why Used?**
    * Utilized for **K-Means clustering** of songs based on mood.

        * Lightweight and optimized for machine learning tasks.

    – **Why Not TensorFlow/PyTorch?**

        * These frameworks are overkill for simple clustering tasks.

3.1.3 Data Collection: Google Chrome Extension

- **Why Used?**
    - Captures **browsing data** passively in real-time.
    - JavaScript-based, lightweight, and runs in the background.
- **Why Not Python Selenium?**
    - Selenium requires active user interaction, which is not feasible for passive data collection.

3.1.4 Video Analysis: OpenCV

- **Why Used?**
    - Efficient for **real-time facial emotion detection**.
    - Open-source and optimized for AI-driven image/video processing.
- **Why Not Others?**
    - **PIL**: Works only for static images.
    - **MATLAB**: Expensive and not open-source.

3.1.5 Music Recommendation: Spotify Web API

- **Why Used?**
    - Provides access to a vast music library.
    - Allows direct song playback and metadata retrieval.
- **Why Not Others?**
    - **Apple Music API**: Paid and less accessible.
    - **Local Music Database**: Limited song choices.

3.2 Hardware Requirements

3.2.1 Processor (CPU)

- **Specifications:**
  - **Minimum**: Intel i5 10th Gen / AMD Ryzen 5.
  - **Recommended**: Intel i7 12th Gen / AMD Ryzen 7.
- **Why?** Handles model inference and background processing.

3.2.2 GPU

- **Specifications:**
  - **Minimum**: NVIDIA GTX 1650 (4GB VRAM).
  - **Recommended**: NVIDIA RTX 3060+ (8GB VRAM).
- **Why?** Required for **LLaMA model** and video processing.

3.2.3 RAM

- **Specifications:**
  - **Minimum**: 8GB DDR4.
  - **Recommended**: 16GB DDR4/DDR5.
- **Why?** Handles large datasets and ML model execution efficiently.

3.2.4 Storage

- **Specifications:**
  - **Minimum**: 256GB SSD.
  - **Recommended**: 512GB SSD + 1TB HDD.
- **Why?** Faster model loading and dataset storage.

3.2.5 Camera (for Video Analysis)

- **Specifications:**
  - **Minimum**: 720p Webcam.
  - **Recommended**: 1080p or Infrared Camera.
- **Why?** Higher accuracy in facial emotion detection.

Chapter 4

System Architecture

This chapter offers a clear and detailed view of how the system is built, explaining its technical structure, key components, and the methods used.

4.1 Proposed Methodology

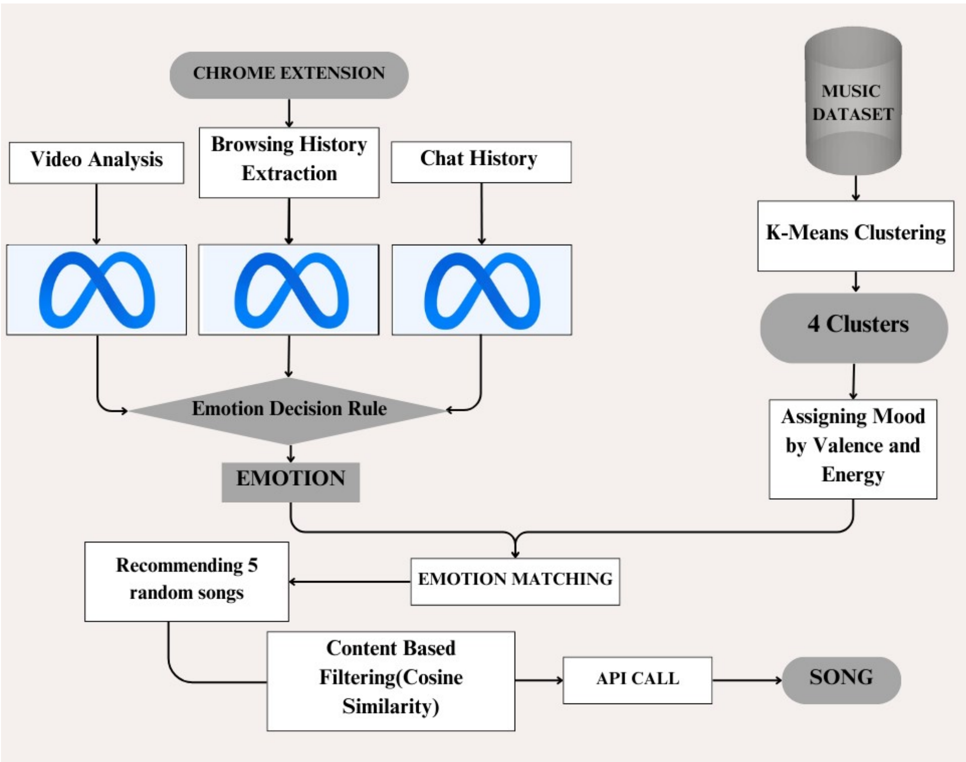


Figure 4.1: Architecture diagram

The system's architecture is designed as a modular and layered structure to ensure scalability, reliability, and efficient processing. It consists of five key layers:

1. Data Collection Layer: Captures browsing history (through a Chrome extension), chat interactions and video analysis in real-time.

2. Mood Analysis Layer: Processes the collected multimodal data and utilizes LLaMA2 [7] to detect the user's emotional state (happy, sad, calm or angry).

3. Music Classification Layer: Uses K-Means clustering to group songs into emotion-based clusters based on characteristics like valence and energy.

4. Recommendation Layer: Maps the detected emotional states to suitable music recommendations.

5. Integration Layer: Provides seamless music playback using external APIs, such as Spotify.

4.1.1 Machine Learning Models

The system incorporates two key machine learning models:

- Emotion Detection: The LLaMA2 model analyzes browsing history, chat interactions and video analysis to detect the user's emotional state (happy, sad, calm or angry).

- Music Classification: K-Means clustering groups songs into mood-based categories like "happy", "calm", "sad" or "angry" based on music features.

4.1.2 Third-Party APIs

The system integrates the Spotify Web API[8] to seamlessly deliver the recommended music. This API handles user authentication and provides features like playback control, ensuring a smooth user experience.

4.2 Component Design

1. **Data Collection Module**

   The Data Collection Module collects multimodal information, such as browsing history, chat interactions, and video analysis, to evaluate the emotional state of the user. A Chrome extension records visited web pages, search queries, and metadata like page titles and descriptions. Chat interactions and facial expressions from video analysis also feed into a broader emotional insight. Secure data transmission guarantees privacy and integrity prior to additional processing.[9]

2. **Mood Analysis Module**

   The Mood Analysis Module analyzes the gathered multimodal data to identify the emotional state of the user. The system applies contextual and sentiment analysis on browsing history and chat interactions using LLaMA2, while video analysis is used to identify facial expressions that match emotions. The result is an emotion label (happy, calm, angry or sad) that is used to drive the music recommendation process.

3. **Music Classification Module**

   Music Classification Module creates clusters of music tracks with respect to their mood, which eventually helps in making personalized recommendations. K-Means clustering is used, which groups songs based on features such as energy and valence. Exploratory Data Analysis (EDA) improves the precision of these labels.

4. **Recommendation Engine**

   The Recommendation Engine will match the emotional state of the user to the most relevant music cluster and suggest tracks from that cluster. This component uses content-based filtering to adjust recommendations dynamically as the user's emotions change. It takes the user's emotion as input, maps it to find its corresponding cluster, and then returns a list of tracks to enjoy in this context.

5. **Integration with Spotify API**

   The Spotify API Integration allows the system to stream recommended tracks to users without a hitch. Through connection to the Spotify Web API, the module retrieves track details, streaming URLs, and manages playback controls like play, pause, and skip. This module uses the Spotify Web API for searching and playing tracks. The workflow involves authenticating users, retrieving track details, and initiating playback through Spotify's player endpoints, ensuring a smooth music experience.

4.2.1 Interactions Between Components

The modules are interconnected to ensure seamless functionality:

1. The Data Collection Module captures browsing history, chat interactions, and video analysis in real-time and transmits it for processing.
2. The Mood Analysis Module processes the data and outputs emotion label for the Recommendation Engine.
3. The Music Classification Module provides mood clusters to the Recommendation Engine for matching.
4. The Recommendation Engine maps emotions to the most suitable music clusters and generates track recommendations.
5. The Spotify API Integration Module streams the top recommended track directly to the user.

4.3 Sequence Diagram

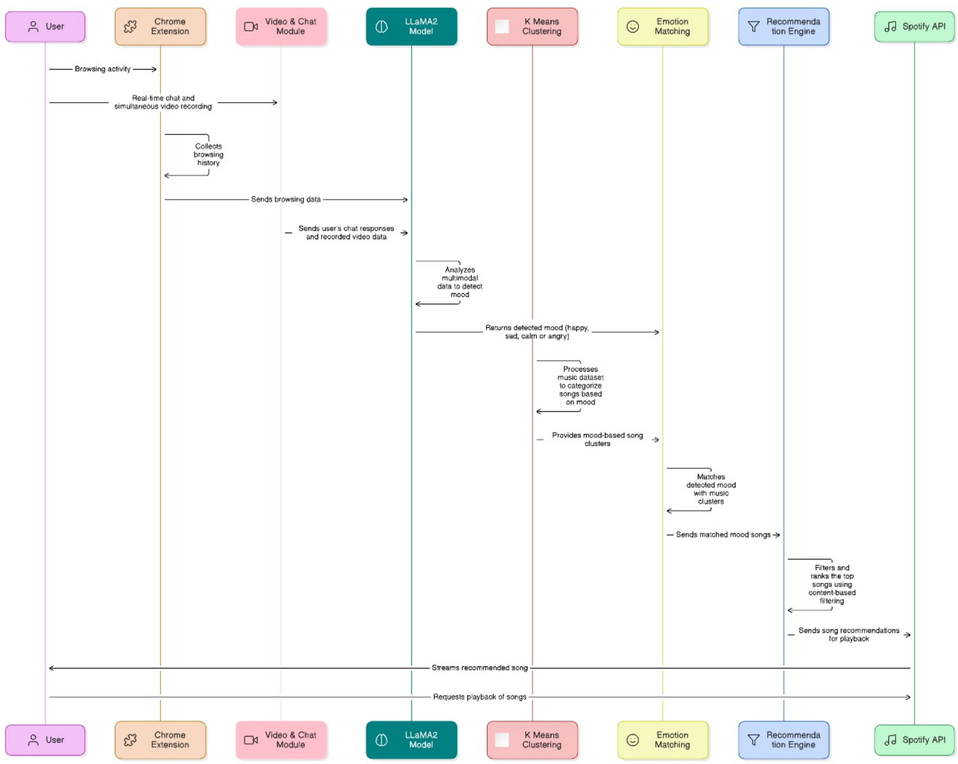


Figure 4.2: Sequence diagram

4.4 Data set Identified

- Browsing Dataset: User data collected via Chrome extension.
- Music Dataset: The Spotify dataset(Dataset of songs in Spotify) was obtained from Kaggle, a platform for predictive modeling and analytics competitions.

4.5 Module Divisions and work break down

- Data Collection Module - Sneha R, Sreya P
- Mood Analysis Module - Sreya P, Sneha R
- Music Classification Module - Sreya P, Vismaya Balakrishnan
- Recommendation Engine - Vismaya Balakrishnan, Sreya P, Sneha R
- Integration with Spotify API - Sreya P, Vismaya Balakrishnan
- Testing - Vismaya Balakrishnan, Sreya P, Sneha R
- Documentation - Vismaya Balakrishnan, Sreya P, Sneha R

4.6 Key Deliverables

- Develop a working prototype with real-time emotion detection, personalized music recommendations, and seamless Spotify API integration.
- Match emotions to mood-based music clusters, providing dynamic and tailored song recommendations.
- Ensure smooth Spotify playback with user authentication and controls like play, pause, and skip.
- Design a scalable, modular system for future enhancements like new datasets or media formats.
- Validate usability and effectiveness through real-world testing with users, supporting emotional well-being.

- Promote social impact by offering a non-invasive, technology-driven solution for improving mental health and emotional resilience through music therapy.

4.7 Project Timeline

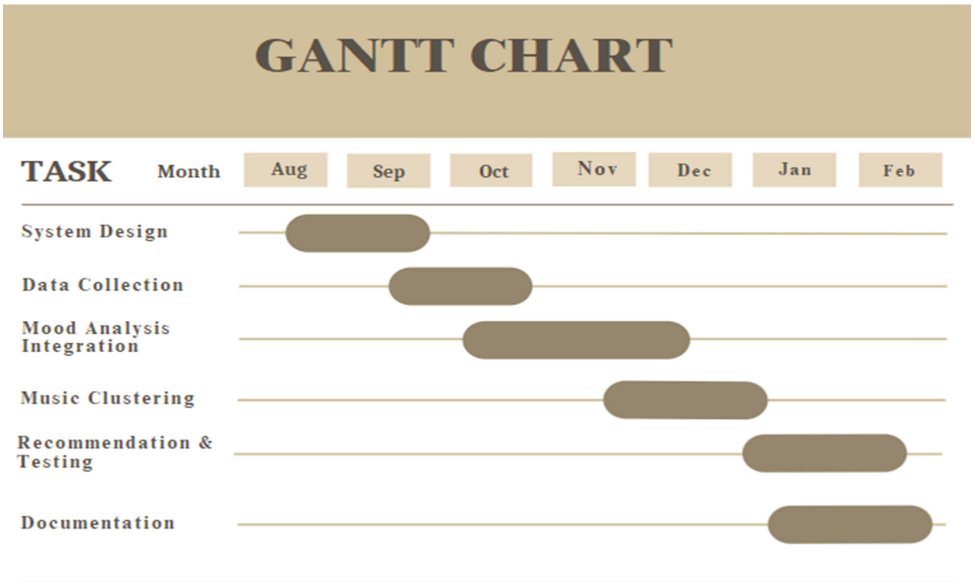


Figure 4.3: Gantt Chart

4.8 Conclusion

This chapter provides a general framework for a personalized mental health recommendation system that binds user browsing behaviors to mood-aligned music suggestions. The architecture is thus modular and layered, ensuring scalability and reliability with efficient processing while integrated modules - including, but not limited to, the Data Collection Module, Mood Analysis Module, Music Classification Module, and Recommendation Engine - operate individually in perfect harmony to deploy a fully holistic solution. Integration with the Spotify Web API will give a smooth user experience by recommending songs for playing.

This design would focus on privacy and data protection for users while handling critical browsing data through ethical means toward accurate and personalized recommendations. Such advanced models of machine learning as LLaMA and K-Means clustering aim to detect emotions and the categorization of music with accuracy. The system is designed to adapt itself to future expansions into even more datasets and media formats to further expand its scope and usability.

By providing the groundwork for implementation, this chapter gives a blueprint to the development of a scalable, user-centric solution that, apart from emotional well-being improvement through music therapy, holds high potential for social impact.

Chapter 5

System Implementation

The system implementation consists of multiple interconnected modules that work together to analyze the user's emotional state and recommend personalized music. The key modules are:

- Data Collection Module
- Mood Analysis Module
- Music Classification Module
- Recommendation Engine
- Integration with Spotify API

Each module plays a crucial role in the overall system.

5.1 Data Collection Module

The Data Collection Module is responsible for gathering multimodal data from three primary sources:

- Browsing history (via a Chrome extension)
- Chatbot interactions
- Video analysis (facial expressions detection)

This module ensures that data is collected in real-time and transmitted securely for further processing.

Steps Involved:

- Browsing History Capture:

  A Chrome extension monitors the user's web activity, capturing URLs, page titles, and search queries.

- Chatbot Interaction Analysis:

  User inputs in the chatbot are analyzed using LLaMA model.

- Video Analysis for Emotion Detection:

  A webcam-based module captures facial expressions during chatbot interactions.

5.2 Mood Analysis Module

The Mood Analysis Module processes multimodal data and determines the user's emotional state using sentiment analysis and facial expression recognition. It also applies the Emotion Decision Rule Algorithm to resolve conflicting emotional cues.

Steps Involved:

- Browsing History Sentiment Analysis:

  The LLaMA2 model performs sentiment analysis on title of the web sites visited by the user. A final emotion is detected from the browsing history.

- Chatbot Sentiment Analysis:

  The text input to a chatbot is analyzed using LLaMA2 to detect emotion.

- Facial Expression Recognition:

  A deep learning model examines the facial expressions of the user and provides an emotion label. According to the textual description, LLaMA2 model detects the emotion of the user.

- Emotion Decision Rule Application:

  The system applies pre-defined emotion decision rules to combine different emotion sources.

Emotion Decision Rule Algorithm

Algorithm: Emotion Decision Rule

Input: Two emotion states extracted from different sources

Output: Final resolved emotion

1. If emotion_a is None, return emotion_b.
2. If emotion_b is None, return emotion_a.
3. If (emotion_a, emotion_b) in [("happy", "happy")], return "happy".
4. If (emotion_a, emotion_b) in [("happy", "calm"), ("calm", "happy")], return "happy".
5. If (emotion_a, emotion_b) in [("happy", "sad"), ("sad", "happy")], return "calm".
6. If (emotion_a, emotion_b) in [("happy", "angry"), ("angry", "happy")], return "sad".
7. If (emotion_a, emotion_b) in [("sad", "sad")], return "sad".
8. If (emotion_a, emotion_b) in [("sad", "angry"), ("angry", "sad")], return "angry".
9. If (emotion_a, emotion_b) in [("sad", "calm"), ("calm", "sad")], return "sad".
10. If (emotion_a, emotion_b) in [("angry", "angry")], return "angry".
11. If (emotion_a, emotion_b) in [("angry", "calm"), ("calm", "angry")], return "angry".
12. If (emotion_a, emotion_b) in [("calm", "calm")], return "calm".
13. Otherwise, return emotion_a.

5.3 Music Classification Module

This module classifies music tracks into different emotional categories (happy, sad, calm, angry) using K-Means clustering.

Steps Involved:

- Dataset Collection

  The system utilizes a Spotify dataset that contains various song features such as energy, valence, danceability, and tempo. These attributes help in determining the mood of a song. The dataset is sourced from Kaggle and consists of diverse music genres and emotional tones.

- Feature Selection

  From the dataset, key features are selected that strongly influence a song's emotional impact. These include:

  Valence: Represents the positivity or happiness level of a song.

  Energy: Measures the intensity and liveliness of a track.

  By focusing on these features, the system ensures that songs are categorized based on their emotional effect rather than just their genre.

- K-Means Clustering

  The K-Means clustering algorithm is employed to group songs into four clusters, each representing a different mood. K-Means operates by iteratively assigning each song to the nearest cluster centroid based on feature similarity. The number of clusters (K=4) is chosen to reflect four primary emotional states: Happy, Sad, Calm, and Angry.

- Cluster Labeling

  After clustering, the clusters are manually examined, and labels are assigned based on the valence-energy distribution:

  Cluster 0: Very high valence and high energy → Happy

  Cluster 1: Very low valence and very low energy → Sad

  Cluster 2: Very low valence and low energy → Calm

  Cluster 3: Low valence and very high energy → Angry

  This step ensures that the clusters meaningfully represent emotions, making them suitable for music recommendations.

- Data Storage

  The final clustered music database is stored for quick retrieval by the Recommendation Engine. When a user's emotional state is determined, the system queries the relevant cluster and fetches tracks that match the user's mood. This enables seamless and real-time personalized recommendations.

5.4 Recommendation Engine

This module selects the most relevant music based on the user's detected emotion and retrieves songs from the classified music clusters.

Steps Involved:

- Receive Emotion Input:

  The detected user emotion is passed to the recommendation engine.

- Match Emotion to Music Cluster:

  The system checks which music cluster corresponds to the detected emotion.

- Content-Based Filtering:

  Within the selected cluster, five random songs are recommended to the user. The user is then asked to select a preferred song from the given options. Based on the user's selection, the top three songs are ranked according to their similarity to the chosen song and the user's preferences.

- Track Selection:

  The top-ranked track is selected and sent to the Spotify API for playback.

5.5 Integration with Spotify API

This module ensures seamless music playback through Spotify's Web API. It manages user authentication and playback controls.

Steps Involved:

- User Authentication:

  The user logs in via OAuth authentication.

- Track Search and Retrieval:

  The system queries the Spotify API for the selected track.

- Real-Time Updates:

  If the user's emotion changes dynamically, the track recommendation is updated in real-time.

5.6 Conclusion

The proposed personalized music recommendation system effectively integrates various data sources and sophisticated analysis techniques to offer a user-specific experience based on the user's emotional state. Based on browsing history, chatbot dialogues, and facial recognition, the system identifies the user's emotion accurately utilizing a robust Mood Analysis Module. The Emotion Decision Rule Algorithm also guarantees conflicting emotional signals resolve into a uniting emotional state that is utilized as the foundation for music recommendations.

Music Classification Module classifies the songs into certain emotional categories based on K-Means clustering on the basis of key attributes like valence and energy. This categorization assists the Recommendation Engine in providing real-time, mood-related music recommendations according to the current emotional state of the user. By the content-based filtering method, the system filters the recommendations on the basis of user ratings and offers high-quality, personalized music recommendations.

Finally, interaction with the Spotify API makes music playback easy with continuous music sessions that adapt dynamically to emotional transformations. The system not only improves user engagement but also establishes an intimate connection between the user's emotional state and his or her musical environment towards mental well-being. The system demonstrates the capability of using multimodal data as well as complex machine learning algorithms in creating personalized and dynamic user experiences.

Chapter 6

Results and Discussions

6.1 Results

- Emotion Detection Accuracy

  The emotion detection module using LLaMA2 effectively identified user emotions from browsing history, chatbot, and video analysis. The emotion decision rule was used to correct the detected emotions when more than one source generated conflicting emotional cues. The model was tested on a test set and was able to accurately classify dominant emotions like happy, sad, angry, and calm.

- Recommendation Effectiveness

  The suggestion system projected the identified emotion onto the correct music cluster and recommended five random songs. Users were requested to pick their favorite song, and the system ranked the top three songs based on user preference patterns. The ranking strategy enhanced recommendation precision by projecting song recommendations along user preferences instead of depending entirely on cluster-based mappings.

- Spotify API Integration Performance

  The integration with the Spotify Web API provided for smooth playback of music within the system. The authentication functioned flawlessly, enabling users to play, pause, and skip tracks. There was minimal latency when fetching tracks, but this was largely a function of network conditions instead of system inefficiency.

6.2 Discussions

- Strengths of the System

- Multi-Modal Emotion Detection:

  Unlike single-source sentiment analysis, integrating text, browsing behavior, and facial recognition resulted in a more comprehensive and accurate emotional assessment.

- Personalized Music Recommendations:

  By clustering songs based on mood attributes (valence, energy, danceability), the system provided emotionally relevant music to users.

- Real-Time Response and Seamless Integration:

  The system delivered instant recommendations and integrated well with Spotify, ensuring an uninterrupted listening experience.

• Challenges and Limitations

- Emotional Ambiguity and Mixed Emotions

  The system sets up categories for emotions as predefined labels (happy, sad, angry, calm), but individuals frequently feel mixed or complex emotions that are not known to the model. Future work could involve multi-label emotion detection to account for blended moods such as nostalgia or anxiety.

- Privacy and Data Security Concerns

  Gathering browsing history, chat interactions, and facial expressions invokes privacy concerns since users are unlikely to contribute personal information. Enhancing encryption, on-device processing, and open consent mechanisms can increase trust and data protection.

- Facial Emotion Recognition Limitations

  DeepFace model can incorrectly classify emotions because of inadequate lighting, facial adornments (glasses, masks), or neutral faces. Multi-frame analysis and dynamic tracking could enhance accuracy in a real-world setup.

6.3 Conclusion

The results illustrate the success of multi-modal emotion detection using the integration of browsing history, chatbot dialogue, and video monitoring. The system effectively identified emotions and recommended personalized music based on mapping of sensed moods into predefined clusters. User ratings enhanced recommendation accuracy so that the most highly ranked songs matched personal taste. The Spotify integration ensured easy playback, optimizing user experience.

However, some challenges and limitations were revealed, such as ambiguities in emotional categorization, privacy issues regarding data acquisition, and restrictions in facial emotion recognition based on environmental conditions. Overcoming these challenges through multi-label emotion recognition, enhanced privacy controls, and better deep learning models can help the system substantially enhance its accuracy and reliability.

In summary, this research sheds light on the promise of music recommendation through AI in mental wellness care, and how affect-sensing systems are able to render personalized and captivating experiences to advance emotional health.[10]

Chapter 7

Conclusions & Future Scope

7.1 Conclusion

This project succeeded in creating a multi-modal emotion-based music recommendation system that involved browsing history analysis, chatbot conversations, and facial recognition to determine a user's emotional state. Utilizing LLaMA2 for emotion identification and K-Means clustering for mood-based music grouping, the system provided music suggestions that were tailored to satisfy the user's emotional needs. Seamless playback of music was achieved through integration with the Spotify Web API, improving the user experience. Experimental results showed high accuracy in emotion detection and good song recommendations, thus turning this system into a promising tool for emotion-aware music streaming and mental well-being assistance.

Even with its merits, the project had its challenges of emotional ambiguity, privacy, and facial recognition restrictions. Potential improvements in the future can refine emotion detection models, implement multi-label classifications, and enhance privacy protections to increase the robustness of the system. The project brings to light the potential of AI-powered personalization in online music platforms and how it affects user engagement and mental health support.

7.2 Future Scope

Various extensions and additions can be made to improve the system:

- Multi-label emotion classification can be included to account for richer emotional states such as nostalgia, anxiety, or excitement.

- Addition of more music platforms like Apple Music or YouTube Music can deliver more varied recommendations.

- Improved privacy features like on-device processing and encrypted storage can drive user trust and security.
- Mechanisms for user feedback can be introduced to improve recommendations and learn users' music tastes over time.

These improvements will render the system more personalized, secure, and dynamic, thereby improving its application in AI-based music therapy and emotional wellness.

References

[1] G. Kovacs, "Reconstructing detailed browsing activities from browser history," *arXiv preprint arXiv:2102.03742*, 2021.

[2] Y. Shu, H. Zhang, H. Gu, P. Zhang, T. Lu, D. Li, and N. Gu, "Rah! recsys–assistant–human: A human-centered recommendation framework with llm agents," *IEEE Transactions on Computational Social Systems*, 2024.

[3] Y. Jeong, J.-J. Song, J. Yang, and S. Kang, "Advancing tinnitus therapeutics: Gpt-2 driven clustering analysis of cognitive behavioral therapy sessions and google t5-based predictive modeling for thi score assessment," *IEEE Access*, vol. 12, pp. 52 414–52 427, 2024.

[4] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, and M. Wang, "Lightgcn: Simplifying and powering graph convolution network for recommendation," in *Proceedings of the 43rd International ACM SIGIR conference on research and development in Information Retrieval*, 2020, pp. 639–648.

[5] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.

[6] H. Jun and M. Park, "Cognitive behavioral therapy for tinnitus: Evidence and efficacy," *Korean Journal of Audiology*, vol. 17, no. 3, pp. 101–104, Dec. 2013.

[7] S. R. Cunningham, D. Archambault, and A. Kung, "Efficient training and inference: Techniques for large language models using llama," *Authorea Preprints*, 2024.

[8] M. Millecamp, N. N. Htun, Y. Jin, and K. Verbert, "Controlling spotify recommendations: effects of personal characteristics on music recommender user interfaces," in *Proceedings of the 26th Conference on user modeling, adaptation and personalization*, 2018, pp. 101–109.

[9] S. M. S. A. Abdullah, S. Y. A. Ameen, M. A. Sadeeq, and S. Zeebaree, “Multimodal emotion recognition using deep learning,” *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 73–79, 2021.

[10] M. A. Kitzrow, “The mental health needs of today’s college students: Challenges and recommendations,” *Journal of Student Affairs Research and Practice*, vol. 41, no. 1, pp. 167–181, 2003.

Project Funding

The project has been funded under the prestigious IndiaAI Mission, an initiative by the Ministry of Electronics and Information Technology (MeitY), Government of India. A grant of Rs. 300,000 has been given in total to fund the development, research, and implementation of the project.

Appendix A: Presentation

Personalized Mental Health Recommendations Using LLM & Browsing History

-A Music-Based Therapeutic Approach to Mental Health-

GUIDED BY

Ms. Seema Safar
Assistant Professor
Dept. of Computer Science and Engineering

TEAM MEMBERS

Sneha R. (U2103199)
Sreya P. (U2103202)
Vismaya Balakrishnan (U2103215)

CONTENTS

PROBLEM DEFINITION

Individuals face challenges in finding personalized mental health resources. This project aims to build a recommendation system using LLMs and browsing history to deliver tailored support and improve well-being.

PURPOSE AND NEED

•The system enhances emotional health by providing personalized music recommendations.

•Recommendations are based on the emotional context derived from users' browsing data.

•It offers adaptive and non-invasive mental health support in today's digital age.

PROJECT OBJECTIVE

- Improve music recommendations by analyzing users' browsing history with machine learning.
- Use the LLaMA model and K-Means clustering to assess emotional context and categorize music.
- Integrate with the Spotify Web API for a seamless listening experience based on user behavior.

LITERATURE SURVEY

PAPER	METHEDOLOGY
Dixon Prem Daniel Rajendran, Rangaraja P Sundarraj (2021)	The methodology involves using a hybrid collaborative filtering approach that combines topic modeling through Latent Dirichlet Allocation (LDA) with browsing history to calculate user-user similarity, identify the top N neighbors, and predict user ratings based on these similarities and previous user ratings.
Yubo Shu , Haonan Zhang , Hansu Gu , Peng Zhang , Tun Lu , Dongsheng Li , and Ning Gu (2024)	The methodology in uses the RAH Framework, where an LLM-driven system learns from user feedback, refines recommendations, and evaluates results through a Learn-Act-Critic Loop. A Reflection Mechanism adjusts recommendations based on changing user preferences to enhance personalization, reduce bias, and protect privacy.
Yongwoo Jeong, Jae-Jun Song, Jiseon Yang and Sungmin Kang (2024)	The methodology of the study followed four steps: data from 42 tinnitus patients undergoing CBT was collected and anonymized. After translation from Korean to English, the DBSCAN algorithm was used for clustering. Data augmentation, including synonym replacement with DistilBERT and RoBERTa, was applied. Finally, Google T5 and Flan-T5 models predicted Tinnitus Handicap Inventory (THI) scores, evaluated using RMSE and ROUGE-L metrics.

PROPOSED METHOD

01 DATA COLLECTION

•A Chrome extension collects real-time browsing behavior.

•Browsing data is processed and stored for emotional analysis.

02 MUSIC CLASSIFICATION

•K-Means clustering groups songs by mood (e.g., happy, sad, calm, angry) in a curated music dataset.

•EDA assigns mood labels to clusters based on emotional characteristics.

03 MOOD ANALYSIS

•The LLaMA model analyzes browsing data to identify users' emotional states.

•Extracts the emotional context from browsing history, user’s chat responses and video analysis.

04 MUSIC RECOMMENDATION

•Content-based filtering matches the user’s mood to a music cluster.

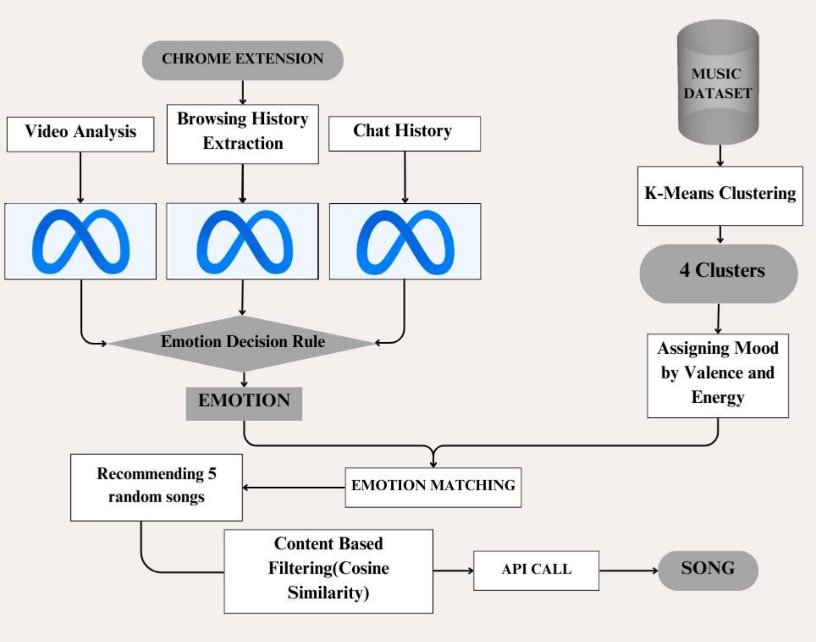
•The best recommendation is based on the closest emotional match.

05 SEAMLESS MUSIC PLAYBACK

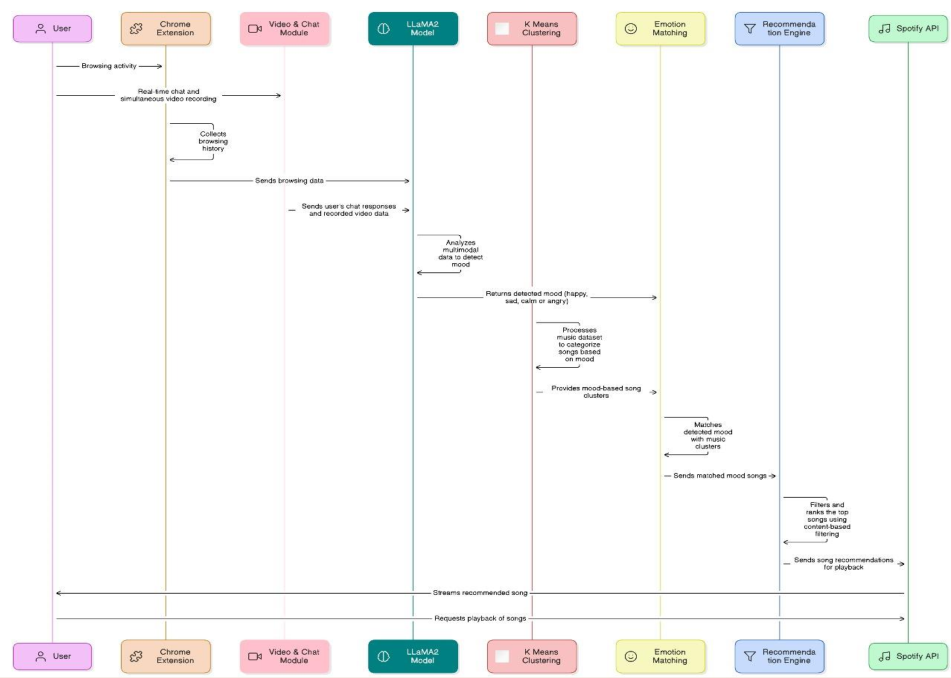
•Spotify Web API integration enables seamless music playback.

•The system plays the top recommended song for a personalized experience.

ARCHITECTURE DIAGRAM



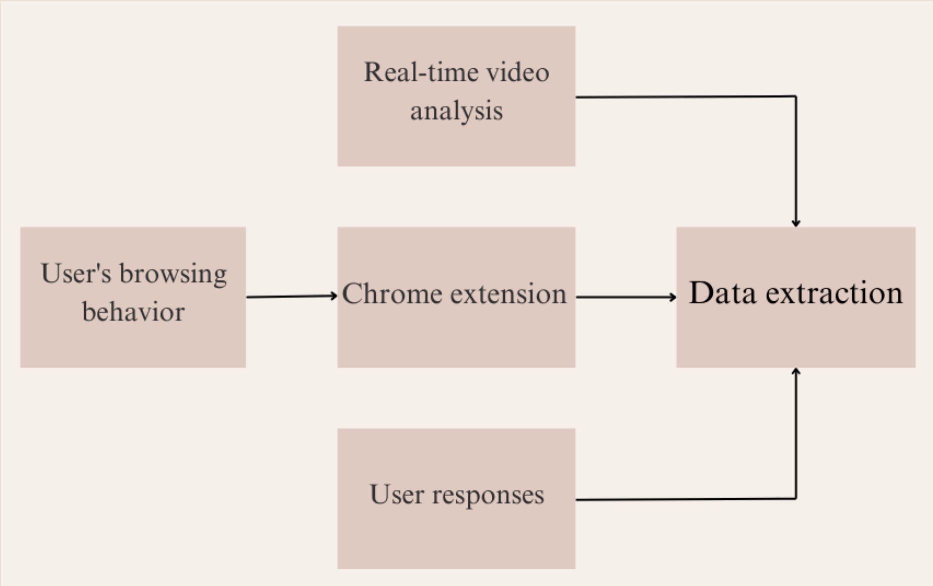
SEQUENCE DIAGRAM



METHODOLOGY

1. Data Collection Module
2. Music Classification Module (K-Means Clustering)
3. Mood Analysis Module (LLaMA Model)
4. Recommendation Engine (Content-Based Filtering)
5. Integration with Spotify API

DATA COLLECTION MODULE



BROWSING HISTORY EXTRACTION

1. **User's Browsing Behavior**: This includes actions a user performs online, such as visiting websites, clicking links, and searching.
2. **Chrome Extension**: A small browser program that monitors and gathers information on the user's browsing activities.
3. **Data Extraction**: The extension collects specific details, like URLs visited, the last visit time, and duration spent on each page, for analysis.
4. **Local Storage**: The collected data is saved locally on the user's device in a CSV file with columns for "Title," "URL," "Last Visit," and "Time."

VIDEO & CHAT EXTRACTION

1. Combines LLaMA2 chatbot interactions (text) with video analysis (MP4) for accurate emotional state detection.
2. Stores user responses in a text file for emotion tracking and refining recommendations over time.
3. Integrates emotion analysis from text and facial expressions to provide personalized, mood-aware music recommendations.

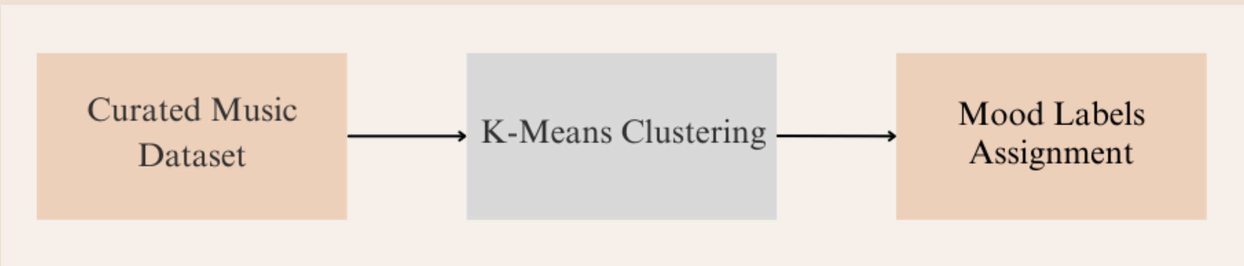
DATA PREPROCESSING

1. Download music dataset from Kaggle (Dataset of songs in Spotify).
2. The dataset underwent preprocessing to handle missing values, outliers, and ensure consistency in the format. This involved cleaning up the data to make it suitable for analysis and modeling.

DATA PREPROCESSING

3. Feature Selection: Relevant features for the mood-based music recommender system were selected. Features related to mood, such as valence and energy, were of particular interest for building the recommendation algorithm.
4. Exploratory Data Analysis (EDA): Exploratory Data Analysis was performed to gain insights into the distribution of moods and genres in mood.

MUSIC CLASSIFICATION MODULE

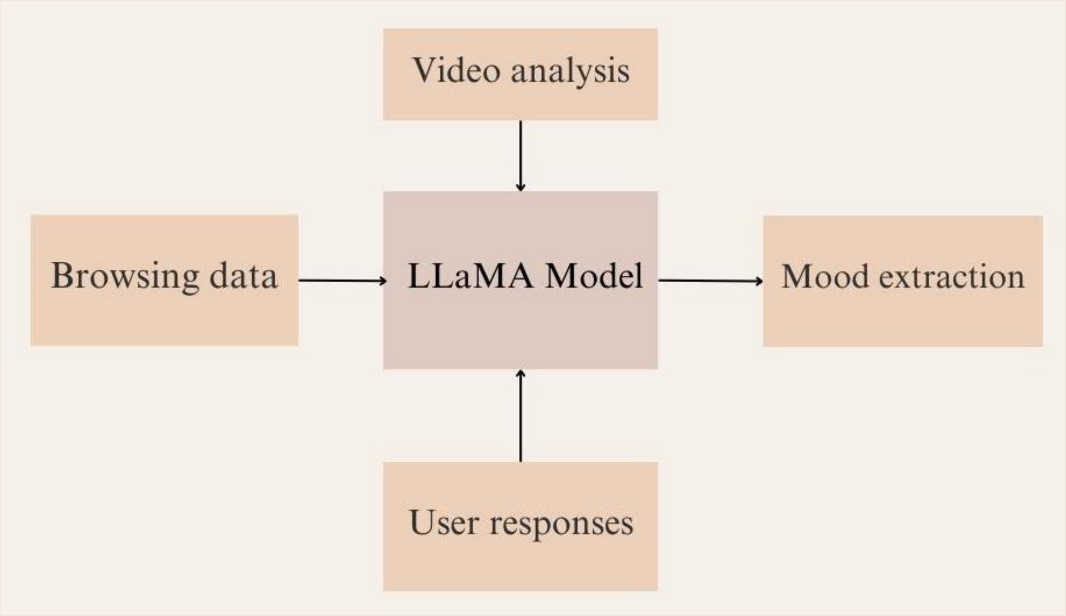


CLUSTERING

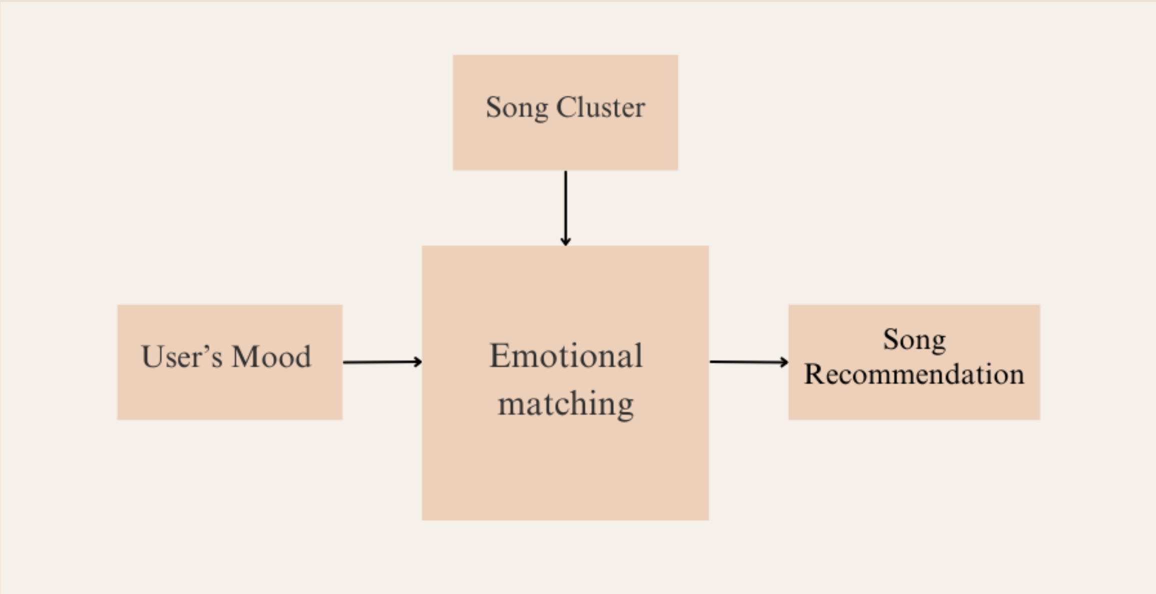
Apply K-Means clustering on pre-processed music dataset to cluster the music into four mood clusters.

- **Cluster 3**: Low valence and very high energy → Angry
- **Cluster 1**: Very low valence and very low energy → Sad
- **Cluster 2**: Very low valence and low energy → Calm
- **Cluster 0**: Very high valence and high energy → Happy

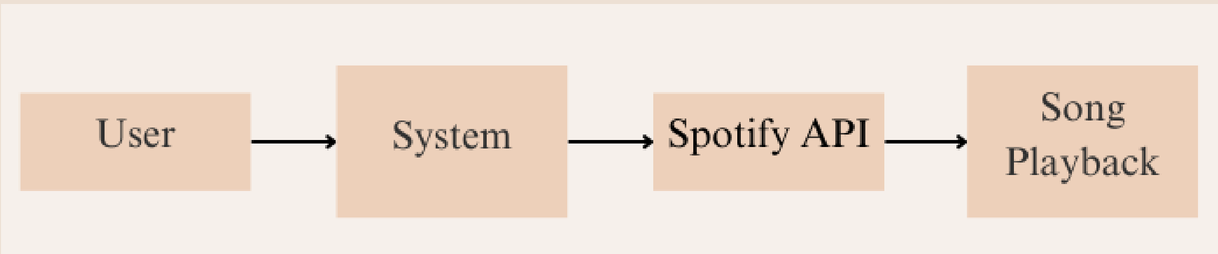
MOOD ANALYSIS MODULE



RECOMMENDATION ENGINE



INTEGRATION WITH SPOTIFY API



ASSUMPTIONS

•Users need active Spotify accounts for seamless playback.

•The Chrome extension is installed in all of users' system.

•Currently only 4 moods : Happy, Sad Angry and Calm are analyzed by the users.

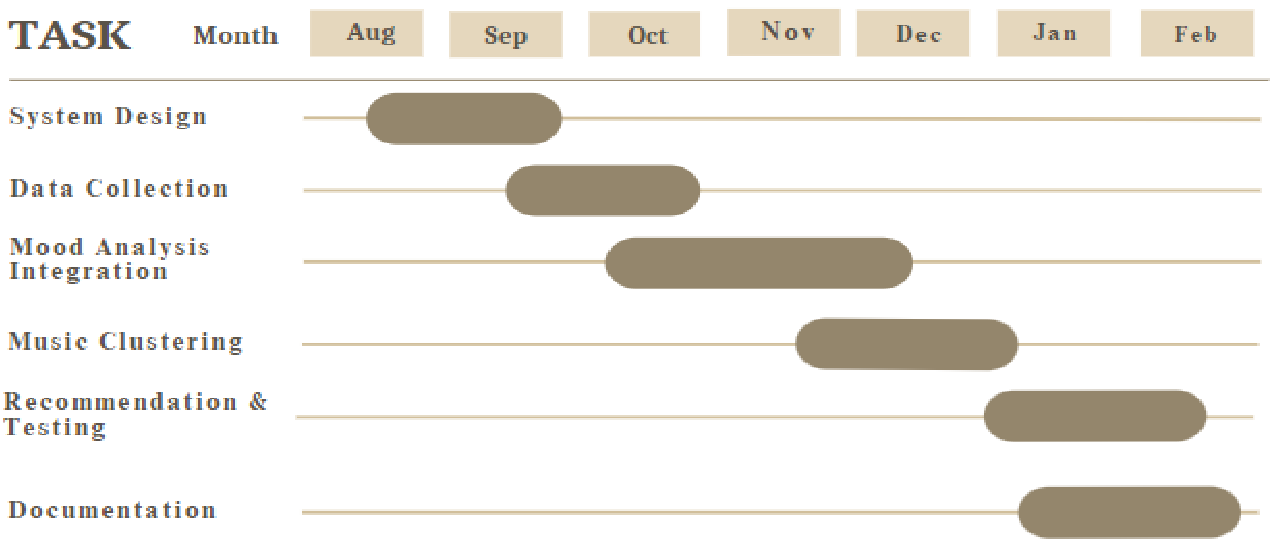
•English songs will be recommended, since the dataset being used consists of English songs.

REQUIREMENTS

Software:
•**Operating System:** Windows 10
•**Coding Language:** Python
•**Tools:** Visual Studio Code
•**Libraries:** Transformers (LLaMA), scikit-learn, pandas, matplotlib
•**API Integration:** Spotify Web API

Hardware:
•**Input Devices:** WebCam
•**Processor:** Quad-core processor with a clock speed of 2.0 GHz or higher
•**Memory (RAM):** Minimum 8 GB RAM
•**Graphics Card:** Integrated graphics (or optional GPU for ML fine-tuning)
•**Storage:** Minimum 10 GB of available storage

GANTT CHART



RISK AND CHALLENGES

Privacy Concerns: Handling sensitive browsing data securely.

API Limitations: Spotify API usage restrictions.

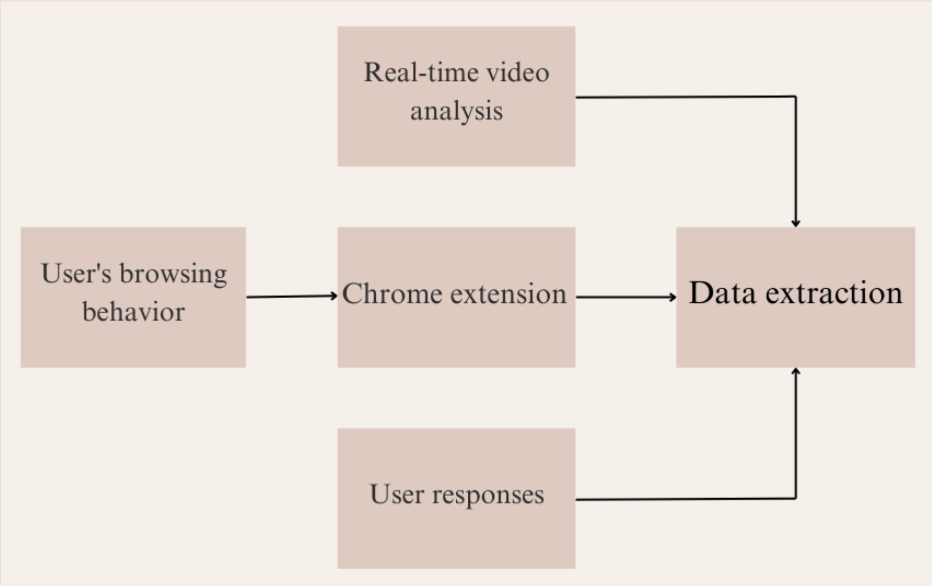
Model Accuracy: Challenges in detecting nuanced emotions.

WORK BREAKDOWN AND RESPONSIBILITIES

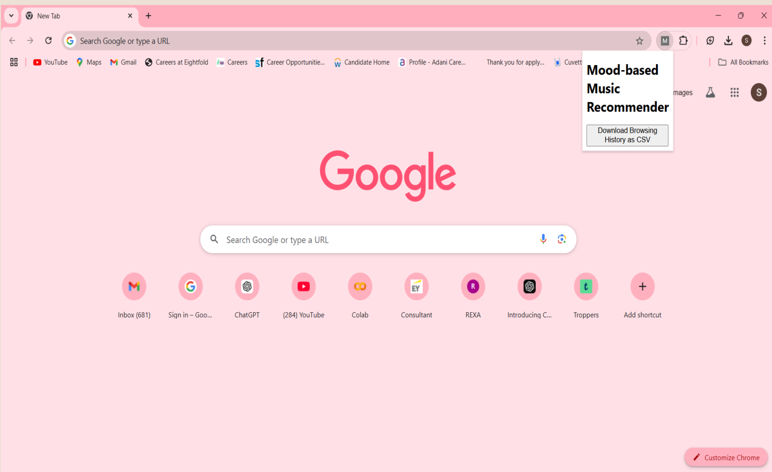
- Data Collection Module - Sneha R, Sreya P
- Mood Analysis Module - Sreya P, Vismaya Balakrishnan
- Music Classification Module -Sreya P, Sneha R
- Recommendation Engine –Vismaya Balakrishnan, Sreya P, Sneha R
- Integration with Spotify API- Sreya P, Vismaya Balakrishnan
- Testing- Vismaya Balakrishnan, Sreya P, Sneha R
- Documentation-Vismaya Balakrishnan, Sneha R, Sreya P

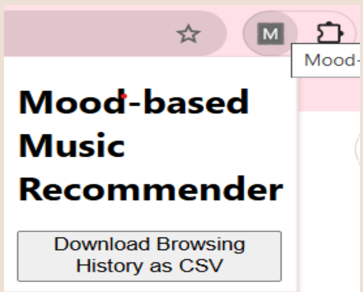
RESULTS

DATA COLLECTION MODULE



BROWSING HISTORY EXTRACTION





BROWSING HISTORY EXTRACTION

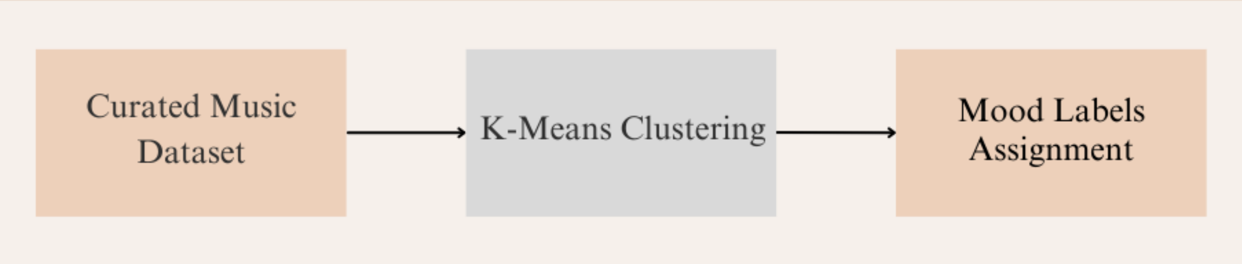
	A	B	C	D	E
1	Title	URL	Last Visit Time		
2	how to feel optimistic - Google Search	https://www.google.com/search?q=how+to+feel+optimistic&rlz=1C1YTUH_enIN1	11/10/2024, 11:35:48 PM		
3	how to be confident - Google Search	https://www.google.com/search?q=how+to+be+confident&rlz=1C1YTUH_enIN11	11/10/2024, 11:35:37 PM		
4	suggest some books to releive stress - Google Search	https://www.google.com/search?q=suggest+some+books+to+releive+stress&rlz=1	11/10/2024, 11:35:30 PM		
5	i amvery anxious what music should i listen to - Google Search	https://www.google.com/search?q=i+amvery+anxious+what+music+should+i+liste	11/10/2024, 11:35:09 PM		
6	why is it that i feel sleepy all the time - Google Search	https://www.google.com/search?q=why+is+it+that+i+feel+sleepy+all+the+time&rl	11/10/2024, 11:34:53 PM		
7	i feel little moody - Google Search	https://www.google.com/search?q=i+feel+little+moody&rlz=1C1YTUH_enIN1118	11/10/2024, 11:29:14 PM		
8	Chrome Extensions Â |Â Chrome for Developers	https://developer.chrome.com/docs/extensions	11/10/2024, 11:18:56 PM		
9	ChatGPT	https://chatgpt.com/c/6730f034-4980-8003-b6ca-ac5285a8b606	11/10/2024, 11:11:07 PM		
10	ChatGPT	https://chatgpt.com/?model=auto	11/10/2024, 11:10:54 PM		

DATA PREPROCESSING

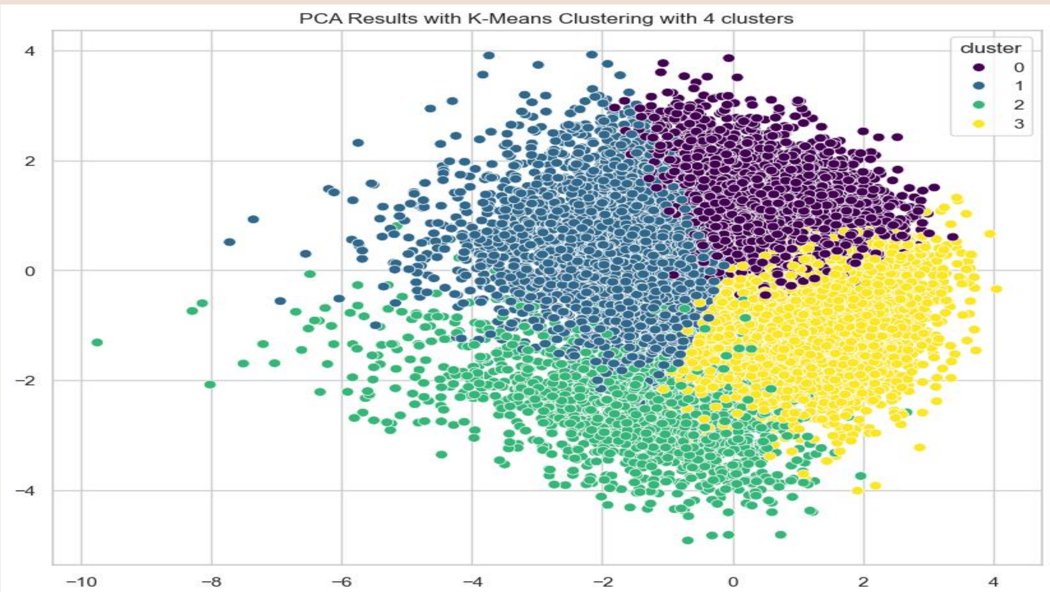
	danceability	energy	key	loudness	mode	speechiness	acousticness	instrumentalness	liveness	valence	...	id	
18597	0.714	0.821	6	-7.635	1	0.1760	0.041000	0.000000	0.1160	0.649	...	69gRFGOWY9OMpFJgFol1u0	spotify:track:69gRFGOV
19677	0.811	0.445	8	-10.105	1	0.2740	0.045300	0.000000	0.1090	0.250	...	13q8un4Xjs3IOfdem4jgVe	spotify:track:13q8ur
19678	0.763	0.654	11	-9.423	0	0.5060	0.087500	0.000003	0.1390	0.277	...	7B2NTZXEPeshkufbqT1gMN	spotify:track:7B2NTZX
19679	0.759	0.530	9	-8.897	1	0.2880	0.004940	0.000000	0.1960	0.329	...	5BtGgEaMNYuXfK69d9FJDs	spotify:track:5BtGgEa
19680	0.535	0.427	10	-9.658	0	0.0512	0.139000	0.001290	0.1100	0.195	...	7euujXy941mB6TL7uANTxz	spotify:track:7euujXy
...	...	...	...	...	...	...	...	...	...	...	...	...	
42300	0.528	0.693	4	-5.148	1	0.0304	0.031500	0.000345	0.1210	0.394	...	46bXU7Sgj7104ZoXxzz9tM	spotify:track:46bXU7
42301	0.517	0.768	0	-7.922	0	0.0479	0.022500	0.000018	0.2050	0.383	...	0he2ViGMUO3ajKTxLOfWVT	spotify:track:0he2ViGM
42302	0.361	0.821	8	-3.102	1	0.0505	0.026000	0.000242	0.3850	0.124	...	72DAt9Lbpy9EUS29OzQLob	spotify:track:72DAt9Lb
42303	0.477	0.921	6	-4.777	0	0.0392	0.000551	0.029600	0.0575	0.488	...	6HXgExFVuE1c3cq9QjFCcU	spotify:track:6HXgExF
42304	0.529	0.945	9	-5.862	1	0.0615	0.001890	0.000055	0.4140	0.134	...	6MAAMZImxcvYhRnxDLTufD	spotify:track:6MAAMZI

20786 rows × 22 columns

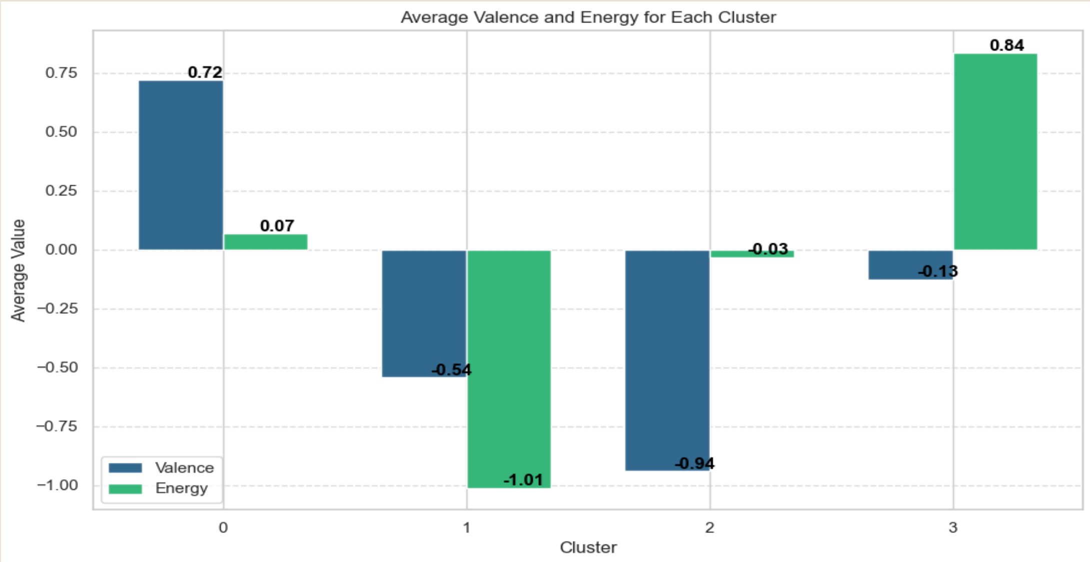
MUSIC CLASSIFICATION MODULE



CLUSTERING



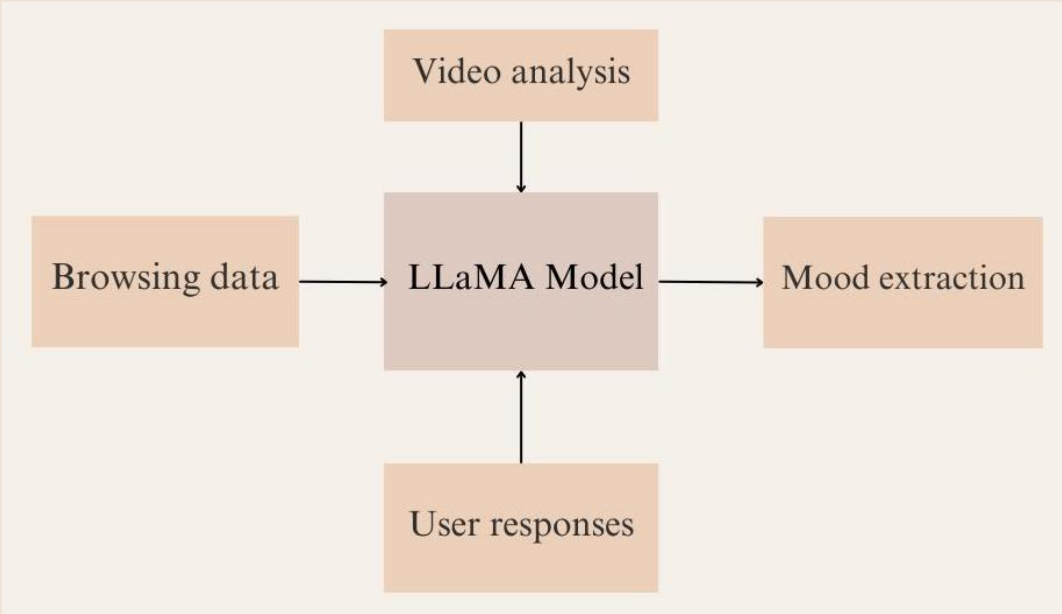
CLUSTERING



CLUSTERING

mode	speechiness	acousticness	instrumentalness	liveness	valence	tempo	uri	genre	song_name	cluster	mood
0.854028	1.734850	-0.515983	-0.302673	-0.921993	-0.127645	0.159685	spotify:track:2Vc6NJ9PW9gD9q343XFRKx	Dark Trap	Mercury: Retrograde	0	Happy
0.854028	-0.680405	1.095488	-0.360457	-0.508370	-1.257259	-1.285537	spotify:track:7pgJBLVz5VmnL7uGHmRj6p	Dark Trap	Pathology	1	Sad
0.854028	-0.801664	-0.733239	-0.360440	1.175288	-1.619161	2.265699	spotify:track:0vSWgAlfpye0WCGeNmuNhy	Dark Trap	Symbiote	3	Angry
0.854028	-0.513053	-0.686482	-0.360457	-0.534884	-1.039861	1.193051	spotify:track:0VSXnJqQkwuH2ei1nOQ1nu	Dark Trap	ProductOfDrugs (Prod. The Virus and Antidote)	3	Angry
0.854028	0.834271	0.226465	-0.360457	-0.190198	0.733419	-0.150604	spotify:track:4jCeguq9rMTlbMmPHuO7S3	Dark Trap	Venom	0	Happy
...	...	...	...	...	...	...	...	...	...	...	...
0.854028	0.869727	-0.788686	-0.360443	0.936660	0.690793	0.260563	spotify:track:54v6Kv41W8AKToCb14PmJD	Hiphop	H20 (feat. Pharoahe Monch, Rakaa Iriscience, &...	3	Angry
0.854028	2.068136	-0.758081	-0.360457	-0.422198	-0.711634	0.273978	spotify:track:4YuGxSdxyL6fRD7uchn9dl	Hiphop	Daddy's Girl	3	Angry
0.854028	0.777541	0.259525	-0.360457	-0.481855	0.464870	0.066119	spotify:track:034dSXepspHaYFW7rZZx1d	Hiphop	'94	1	Sad
0.854028	0.734994	-0.724265	-0.351919	-0.627684	0.371090	1.096898	spotify:track:7oCZpPJcPDZfely2t6gqSg	Hiphop	Born 2 Live	0	Happy
-1.170922	0.068423	0.302032	-0.360397	-0.428827	1.338722	1.255543	spotify:track:2kpwzDrLK5KoIk1S5wftX7	Hiphop	Can't Hold On	0	Happy

MOOD ANALYSIS



MOOD ANALYSIS

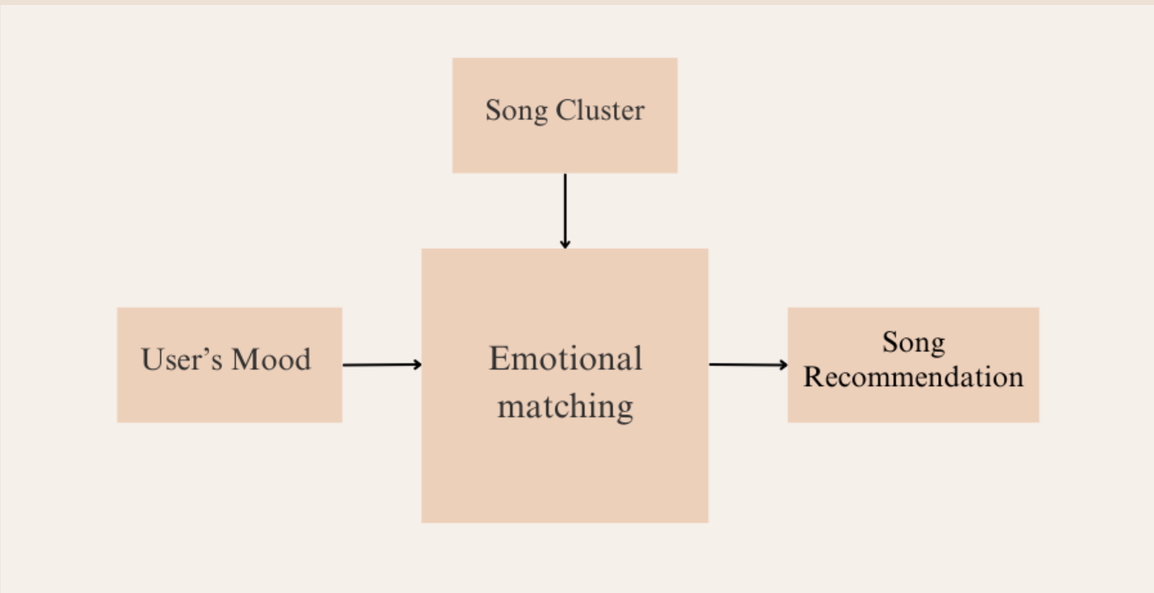
```
Do you want to provide your browsing history for emotion analysis? (yes/no)
yes

Analyzing browsing history for emotions...
Title: how to handle hard times in life - Google Search -> Detected Emotion: calm
Title: i feel so down and bad - Google Search -> Detected Emotion: sad
Title: why i am feeling sad - Google Search -> Detected Emotion: sad
Title: i feel frustrated -> Detected Emotion: angry
Title: i just feel to shout at someone -> Detected Emotion: angry
Title: i am so mad -> Detected Emotion: angry

Dominant Emotion: angry
Browsing History Emotion: angry

Do you want to chat? (yes/no)
yes

Starting chat...
Chatbot: Hi there! How are you feeling today?
You: i am so frustrated
I'm sorry to hear that you're feeling frustrated. Can you tell me more about what's going on and why you're feeling that way? Sometimes talki
ng about our feelings can help us feel better. Is there anything in particular that's causing you to feel this way?
You: i am mad at my friends
Sorry to hear that you're feeling mad at your friends. Can you tell me more about what's going on? Sometimes talking about our feelings can h
elp us understand them better and find ways to manage them. Is there something specific that's causing you to feel this way towards your frie
nds?
You: exit
Chatbot: Take care! I'm here whenever you need me. 😊

Analyzing chat responses for emotions...
Detected Emotion: angry
Chat Emotion: angry

Analyzing video for emotions...
Video Emotion: angry

Dominant Emotion: angry
```

EMOTION DECISION RULE

Algorithm : Emotion Decision Rule
Input: Two emotion states extracted from different sources
Output: Final resolved emotion

```
1. If emotion_a is None, return emotion_b.
2. If emotion_b is None, return emotion_a.
3. If (emotion_a, emotion_b) == (“happy”, “happy”), return “happy”.
4. If (emotion_a, emotion_b) in [("happy", "calm"), ("calm", "happy")], return "happy".
5. If (emotion_a, emotion_b) in [("happy", "sad"), ("sad", "happy")], return "calm".
6. If (emotion_a, emotion_b) in [("happy", "angry"), ("angry", "happy")], return "sad".
7. If (emotion_a, emotion_b) == ("sad", "sad"), return "sad".
8. If (emotion_a, emotion_b) in [("sad", "angry"), ("angry", "sad")], return "angry".
9. If (emotion_a, emotion_b) in [("sad", "calm"), ("calm", "sad")], return "sad".
10. If (emotion_a, emotion_b) == ("angry", "angry"), return "angry".
11. If (emotion_a, emotion_b) in [("angry", "calm"), ("calm", "angry")], return "angry".
12. If (emotion_a, emotion_b) == ("calm", "calm"), return "calm".
13. Otherwise, return emotion_a.
```


RECOMMENDATION ENGINE



RECOMMENDATION ENGINE

```
Detected mood is: angry
Random songs based on detected mood:
1. Breath of the Forest
2. Loud
3. GUT$
4. If It Means a Lot to You
5. Back On
```

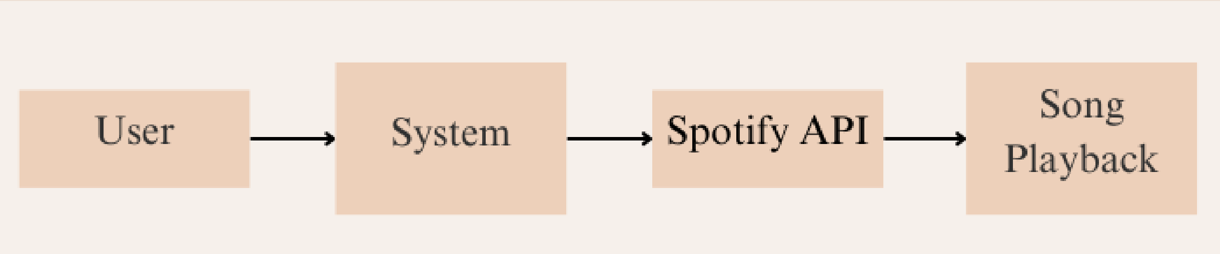

RECOMMENDATION ENGINE

e	speechiness	acousticness	instrumentalness	liveness	valence	tempo	uri	genre	song_name	cluster	mood
8	-0.801664	-0.733239	-0.360440	1.175288	-1.619161	2.265699	spotify:track:0vSWgAlfpye0WCGeNmuNhy	Dark Trap	Symbiote	3	angry
8	-0.513053	-0.686482	-0.360457	-0.534884	-1.039861	1.193051	spotify:track:0VSXnJqQkwuH2ei1nOQ1nu	Dark Trap	ProductOfDrugs (Prod. The Virus and Antidote)	3	angry
8	-0.271953	-0.678453	-0.326822	-0.468598	-1.623850	-0.742488	spotify:track:0XfQbq7DaMOmVXgQ71eA6E	Dark Trap	kamikaze (+ pulse)	3	angry
8	0.323706	-0.610914	-0.360457	-0.548141	-0.579490	-0.493381	spotify:track:0LLeuNBWPOg3XA73yab3PT	Dark Trap	T.R.U. (Totally Rotten Underground)	3	angry
8	0.160609	-0.508426	-0.360457	-0.807981	-0.980184	2.331571	spotify:track:37gqBnUAZe8BY8WR56kDNk	Dark Trap	I Put My Dick in Your Mental	3	angry
...	...	...	...	...	...	...	...	...	...	...	...
8	-0.649204	-0.670896	-0.360457	-0.636964	1.965338	-0.215924	spotify:track:31qgVdvSqTQ7unwQQngycB	Hiphop	Can't Hold Us (feat. Ray Dalton)	3	angry
2	-0.378321	-0.508426	-0.360457	-0.818587	0.579962	-0.670167	spotify:track:18SQ99AVrnZOcgpTSUVSfL	Hiphop	Ain't Me	3	angry
8	0.862635	-0.667590	-0.360457	-0.748324	-0.089280	1.786384	spotify:track:1E8wif6bVXurUgxV8Gfwrw	Hiphop	Gangsta Luv	3	angry
8	0.869727	-0.788686	-0.360443	0.936660	0.690793	0.260563	spotify:track:54v6Kv41W8AKToCb14PmJD	Hiphop	H20 (feat. Pharoahe Monch, Rakaa Iriscience, &...	3	angry
8	2.068136	-0.758081	-0.360457	-0.422198	-0.711634	0.273978	spotify:track:4YuGxSdxyL6fRD7uchn9dI	Hiphop	Daddy's Girl	3	angry

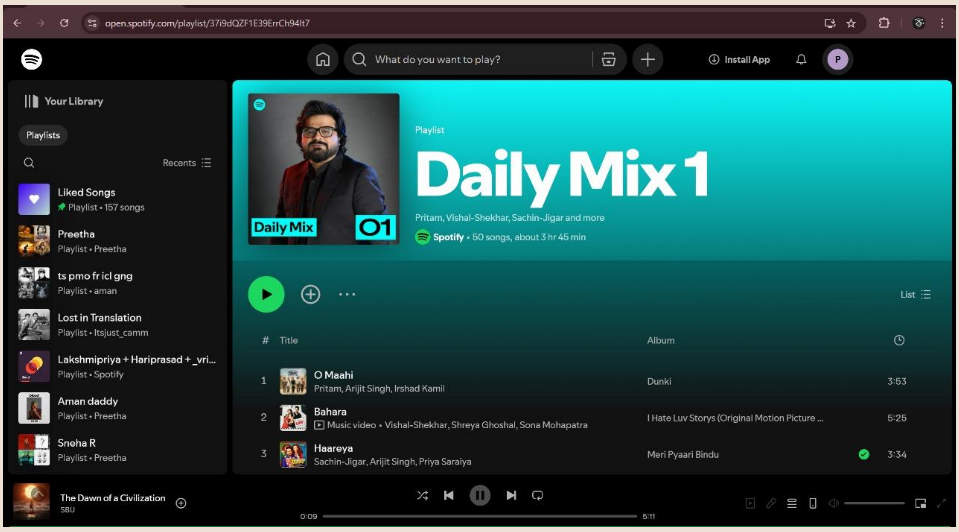
RECOMMENDATION ENGINE

```
Type the number of the song you want to select: 5

You selected: Loud

Top 3 recommendations based on energy and valence:
                  song_name  similarity
1  The Dawn of a Civilization    1.000000
2                   $Weatin'    0.999999
3                    Gravity    0.999999

URI of top recommendation: spotify:track:2F5M9TojcvTuKUXcyzYvDD

Available Spotify devices:
Device ID: 1b4de8461bccc453e896389cab90d987455eb0e8, Name: vivo V29, Type: Smartphone
Device ID: 80872bec05eab1275c09169c72b243620750bd7e, Name: Web Player (Chrome), Type: Computer
Playing top recommended song...
```


INTEGRATION WITH SPOTIFY API



INTEGRATION WITH SPOTIFY API



PROJECT FUNDING

The project has been funded under the prestigious IndiaAI Mission, an initiative by the Ministry of Electronics and Information Technology (MeitY), Government of India. A grant of Rs. 300,000 has been given in total to fund the development, research, and implementation of the project.

CONCLUSION

- This project connects users' browsing behavior with personalized music recommendations.
- It demonstrates the therapeutic potential of music in mental well-being.
- By combining LLMs, clustering techniques, and APIs, it creates a unique user experience.

FUTURE WORK

- Multi-label emotion classification can be included to account for richer emotional states such as nostalgia, anxiety, or excitement.
- Addition of more music platforms like Apple Music or YouTube Music can deliver more varied recommendations.
- Improved privacy features like on-device processing and encrypted storage can drive user trust and security.
- Mechanisms for user feedback can be introduced to improve recommendations and learn users' music tastes over time.

REFERENCES

- Mazlan, Idayati, Noraswaliza Abdullah, and Norashikin Ahmad. "Exploring the impact of hybrid recommender systems on personalized mental health recommendations." *International Journal of Advanced Computer Science and Applications* 14.6 (2023): 0.

- Shu, Yubo, et al. "RAH! RecSys–Assistant–Human: A Human-Centered Recommendation Framework With LLM Agents." *IEEE Transactions on Computational Social Systems* (2024).

- Rajendran, Dixon Prem Daniel, and Rangaraja P. Sundarraj. "Using topic models with browsing history in hybrid collaborative filtering recommender system: Experiments with user ratings." *International Journal of Information Management Data Insights* 1.2 (2021): 100027.

- Chu, Yunghui, Hui-Kuo Yang, and Wen-Chih Peng. "Predicting online user purchase behavior based on browsing history." *2019 IEEE 35th international conference on data engineering workshops (ICDEW)*. IEEE, 2019.

- Taief, Abu Mohammad. *Application of LLMs and Embeddings in Music Recommendation Systems*. MS thesis. UiT Norges arktiske universitet, 2024.

- Jeong, Yongwoo, et al. "Advancing Tinnitus Therapeutics: GPT-2 Driven Clustering Analysis of Cognitive Behavioral Therapy Sessions and Google T5-Based Predictive Modeling for THI Score Assessment." *IEEE Access* (2024).

- https://www.kaggle.com/datasets/mrmorj/dataset-of-songs-in-spotify

Thank you

Appendix B: Vision, Mission, Programme Outcomes and Course Outcomes

Vision, Mission, Programme Outcomes and Course Outcomes

Institute Vision

To evolve into a premier technological institution, moulding eminent professionals with creative minds, innovative ideas and sound practical skill, and to shape a future where technology works for the enrichment of mankind.

Institute Mission

To impart state-of-the-art knowledge to individuals in various technological disciplines and to inculcate in them a high degree of social consciousness and human values, thereby enabling them to face the challenges of life with courage and conviction.

Department Vision

To become a centre of excellence in Computer Science and Engineering, moulding professionals catering to the research and professional needs of national and international organizations.

Department Mission

To inspire and nurture students, with up-to-date knowledge in Computer Science and Engineering, ethics, team spirit, leadership abilities, innovation and creativity to come out with solutions meeting societal needs.

Programme Outcomes (PO)

Engineering Graduates will be able to:

1. Engineering Knowledge: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

2. Problem analysis: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

3. Design/development of solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

4. Conduct investigations of complex problems: Use research-based knowledge including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

5. Modern Tool Usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

6. The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal, and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

7. Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

8. Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

9. Individual and Team work: Function effectively as an individual, and as a member or leader in teams, and in multidisciplinary settings.

10. Communication: Communicate effectively with the engineering community and with society at large. Be able to comprehend and write effective reports documentation. Make effective presentations, and give and receive clear instructions.

11. Project management and finance: Demonstrate knowledge and understanding of engineering and management principles and apply these to one's own work, as a member and leader in a team. Manage projects in multidisciplinary environments.

12. Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and lifelong learning in the broadest context of technological change.

Programme Specific Outcomes (PSO)

A graduate of the Computer Science and Engineering Program will demonstrate:

PSO1: Computer Science Specific Skills
The ability to identify, analyze and design solutions for complex engineering problems in multidisciplinary areas by understanding the core principles and concepts of computer science and thereby engage in national grand challenges.

PSO2: Programming and Software Development Skills
The ability to acquire programming efficiency by designing algorithms and applying standard practices in software project development to deliver quality software products meeting the demands of the industry.

PSO3: Professional Skills
The ability to apply the fundamentals of computer science in competitive research and to develop innovative products to meet the societal needs thereby evolving as an eminent researcher and entrepreneur.

Course Outcomes (CO)

Course Outcome 1: Model and solve real world problems by applying knowledge across domains (Cognitive knowledge level: Apply).

Course Outcome 2: Develop products, processes or technologies for sustainable and socially relevant applications (Cognitive knowledge level: Apply).

Course Outcome 3: Function effectively as an individual and as a leader in diverse teams and to comprehend and execute designated tasks (Cognitive knowledge level: Apply).

Course Outcome 4: Plan and execute tasks utilizing available resources within timelines, following ethical and professional norms (Cognitive knowledge level: Apply).

Course Outcome 5: Identify technology/research gaps and propose innovative/creative solutions (Cognitive knowledge level: Analyze).

Course Outcome 6: Organize and communicate technical and scientific findings effectively in written and oral forms (Cognitive knowledge level: Apply).

Appendix C: CO-PO-PSO Mapping

COURSE OUTCOMES:

After completion of the course, the student will be able to:

SL.NO	DESCRIPTION	Bloom's Taxonomy Level
CO1	Model and solve real-world problems by applying knowledge across domains (Cognitive knowledge level:Apply).	Level3: Apply
CO2	Develop products, processes, or technologies for sustainable and socially relevant applications. (Cognitive knowledge level:Apply).	Level 3: Apply
CO3	Function effectively as an individual and as a leader in diverse teams and comprehend and execute designated tasks. (Cognitive knowledge level:Apply).	Level 3: Apply
CO4	Plan and execute tasks utilizing available resources within timelines, following ethical and professional norms (Cognitive knowledge level:Apply).	Level 3: Apply
CO5	Identify technology/research gaps and propose innovative/creative solutions (Cognitive knowledge level:Analyze).	Level 4: Analyze
CO6	Organize and communicate technical and scientific findings effectively in written and oral forms (Cognitive knowledge level:Apply).	Level 3: Apply

CO-PO AND CO-PSO MAPPING

CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	2	2	2	1	2	2	2	1	1	1	1	2	3		
CO2	2	2	2		1	3	3	1	1		1	1		2	
CO3									3	2	2	1			3
CO4					2			3	2	2	3	2			3
CO5	2	3	3	1	2							1	3		
CO6					2			2	2	3	1	1			3

3/2/1: high/medium/low

JUSTIFICATIONS FOR CO-PO MAPPING

Mapping	**Level**	**Justification**
101003/CS822U.1-PO1	M	Knowledge in the area of technology for project development using various tools results in better modeling.
101003/CS822U.1-PO2	M	Knowledge acquired in the selected area of project development can be used to identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions.
101003/CS822U.1-PO3	M	Can use the acquired knowledge in designing solutions to complex problems.
101003/CS822U.1-PO4	M	Can use the acquired knowledge in designing solutions to complex problems.
101003/CS822U.1-PO5	H	Students are able to interpret, improve, and redefine technical aspects for design of experiments, analysis, and interpretation of data, and synthesis of the information to provide valid conclusions.
101003/CS822U.1-PO6	M	Students are able to interpret, improve, and redefine technical aspects by applying contextual knowledge to assess societal, health, and consequential responsibilities relevant to professional engineering practices.
101003/CS822U.1-PO7	M	Project development based on societal and environmental context solution identification is the need for sustainable development.
101003/CS822U.1-PO8	L	Project development should be based on professional ethics and responsibilities.
101003/CS822U.1-PO9	L	Project development using a systematic approach based on well-defined principles will result in teamwork.
101003/CS822U.1-PO10	M	Project brings technological changes in society.
101003/CS822U.1-PO11	H	Acquiring knowledge for project development gathers skills in design, analysis, development, and implementation of algorithms.

101003/CS822U.1-PO12	H	Knowledge for project development contributes engineering skills in computing and information gatherings.
101003/CS822U.2-PO1	H	Knowledge acquired for project development will also include systematic planning, developing, testing, and implementation in computer science solutions in various domains.
101003/CS822U.2-PO2	H	Project design and development using a systematic approach brings knowledge in mathematics and engineering fundamentals.
101003/CS822U.2-PO3	H	Identifying, formulating, and analyzing the project results in a systematic approach.
101003/CS822U.2-PO5	H	Systematic approach is the tip for solving complex problems in various domains.
101003/CS822U.2-PO6	H	Systematic approach in the technical and design aspects provides valid conclusions.
101003/CS822U.2-PO7	H	Systematic approach in the technical and design aspects demonstrates the knowledge of sustainable development.
101003/CS822U.2-PO8	M	Identification and justification of technical aspects of project development demonstrates the need for sustainable development.
101003/CS822U.2-PO9	H	Apply professional ethics and responsibilities in engineering practice of development.
101003/CS822U.2-PO11	H	Systematic approach also includes effective reporting and documentation, which gives clear instructions.
101003/CS822U.2-PO12	M	Project development using a systematic approach based on well-defined principles will result in better teamwork.
101003/CS822U.3-PO9	H	Project development as a team brings the ability to engage in independent and lifelong learning.

101003/CS822U.3-PO10	H	Identification, formulation, and justification in technical aspects will be based on acquiring skills in design and development of algorithms.
101003/CS822U.3-PO11	H	Identification, formulation, and justification in technical aspects provides the betterment of life in various domains.
101003/CS822U.3-PO12	H	Students are able to interpret, improve, and redefine technical aspects with mathematics, science, and engineering fundamentals for the solutions of complex problems.
101003/CS822U.4-PO5	H	Students are able to interpret, improve, and redefine technical aspects with identification, formulation, and analysis of complex problems.
101003/CS822U.4-PO8	H	Students are able to interpret, improve, and redefine technical aspects to meet the specified needs with appropriate consideration for public health and safety, and the cultural, societal, and environmental considerations.
101003/CS822U.4-PO9	H	Students are able to interpret, improve, and redefine technical aspects for design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
101003/CS822U.4-PO10	H	Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools for better products.
101003/CS822U.4-PO11	M	Students are able to interpret, improve, and redefine technical aspects by applying contextual knowledge to assess societal, health, and consequential responsibilities relevant to professional engineering practices.
101003/CS822U.4-PO12	H	Students are able to interpret, improve, and redefine technical aspects for demonstrating the knowledge of, and need for sustainable development.

101003/CS822U.5-PO1	H	Students are able to interpret, improve, and redefine technical aspects, apply ethical principles, and commit to professional ethics and responsibilities and norms of the engineering practice.
101003/CS822U.5-PO2	M	Students are able to interpret, improve, and redefine technical aspects, communicate effectively on complex engineering activities with the engineering community and society at large, such as being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
101003/CS822U.5-PO3	H	Students are able to interpret, improve, and redefine technical aspects to demonstrate knowledge and understanding of the engineering and management principle in multidisciplinary environments.
101003/CS822U.5-PO4	H	Students are able to interpret, improve, and redefine technical aspects, recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.
101003/CS822U.5-PO5	M	Students are able to interpret, improve, and redefine technical aspects in acquiring skills to design, analyze, and develop algorithms and implement those using high-level programming languages.
101003/CS822U.5-PO12	M	Students are able to interpret, improve, and redefine technical aspects and contribute their engineering skills in computing and information engineering domains like network design and administration, database design, and knowledge engineering.
101003/CS822U.6-PO5	M	Students are able to interpret, improve, and redefine technical aspects and develop strong skills in systematic planning, developing, testing, implementing, and providing IT solutions for different domains, which helps in the betterment of life.

101003/CS822U.6-PO8	H	Students will be able to associate with a team as an effective team player for the development of technical projects by applying the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
101003/CS822U.6-PO9	H	Students will be able to associate with a team as an effective team player to identify, formulate, review research literature, and analyze complex engineering problems.
101003/CS822U.6-PO10	M	Students will be able to associate with a team as an effective team player for designing solutions to complex engineering problems and design system components.
101003/CS822U.6-PO11	M	Students will be able to associate with a team as an effective team player, use research-based knowledge and research methods including design of experiments, analysis, and interpretation of data.
101003/CS822U.6-PO12	H	Students will be able to associate with a team as an effective team player, applying ethical principles and committing to professional ethics and responsibilities and norms of the engineering practice.
101003/CS822U.1-PSO1	H	Students are able to develop Computer Science Specific Skills by modeling and solving problems.
101003/CS822U.2-PSO2	M	Developing products, processes or technologies for sustainable and socially relevant applications can promote Programming and Software Development Skills.
101003/CS822U.3-PSO3	H	Working in a team can result in the effective development of Professional Skills.
101003/CS822U.4-PSO3	H	Planning and scheduling can result in the effective development of Professional Skills.
101003/CS822U.5-PSO1	H	Students are able to develop Computer Science Specific Skills by creating innovative solutions to problems.

101003/CS822U.6-PSO3	H	Organizing and communicating technical and scientific findings can help in the effective development of Professional Skills..