Improvising Spotify Music Recommendation System for Therapeutic Need

Group 7 – Nhat-Phuong H., Gia-Huy H., Mai-Anh N.

College of Engineering and Computer Sciences, VinUniversity

CECS1030 – Computational and Algorithmic Thinking

Instructor, Prof. Kok-Seng Wong

Teaching Assistant, Vu Duc Anh

June 1th, 2025

I. Executive Summary

A. Project Summary

As part of the Spotify product team, this project built a mood-aware music therapy feature that proactively guides users' emotional transitions by blending their dominant listening moods with target therapeutic states. Leveraging Spotify's track metadata and anonymized user sessions, we designed a system that detects emotional "personas" based on Carl Jung - inspired clusters (happy, calm, sad, angry) and generates personalized a 10-music playlists with adjustable blend ratios (e.g., 70/30, 60/40, or 50/50) to support mood shifts.

B. Methodology

Our approach began with integrating 1.18 million rows of user listening history with track audio features including valence, energy, acousticness, danceability, and more; and performing an exploratory data analysis (EDA) to address skewness and normalize feature scales. We determined four mood clusters using the elbow method for K-Means that achieved a silhouette score of 0.40 and a Davies-Bouldin index of 0.84, which confirmed coherent emotional groupings despite natural overlaps. Next, we classified user personas by mapping their recent mood distributions into one of the four clusters via a decision-tree model, enriching these classifications with descriptive taglines and feedback prompts. Finally, we deployed a content-based K-Nearest Neighbors algorithm to assemble blended playlists, allowing users to specify precise mood-transition ratios and hear seamless progressions from their current state to their desired state.

C. Key Insights

Our mood clusters reliably distinguish between emotional states, laying the groundwork for clear "therapeutic blend" pathways. In pilot testing with 43 VinUni students, around 70% of participants reported a positive shift in mood, and we observed engagement improvements with 100% playlist completion rates. Early testers specifically praised the persona cards and blend sliders, noting that these features empowered them to manage their moods independently. Moving forward, we will refine the project scope and integrate expert validation and behavioral data to improve how the music recommendation system aligns with therapeutic principles and objectively measures its impact.

II. Problem and Dataset Recap

A. Global Problem

The global music therapy market has experienced steady growth which was valued at approximately US$2.4 billion, with projections indicating a compound annual growth rate (CAGR) of 9.1% through the end of the decade [1]. This growth is primarily driven by rising global awareness of mental health challenges (stress, anxiety, chronic illnesses, etc.) alongside advancements in digital health technologies (mobile applications, AI/ML-based personalization, VR/AR). Music therapy is widely recognized for its physiological and psychological benefits, such as lowering cortisol levels, heart rate and blood pressure, making it a well-supported complementary treatment for conditions like depression, anxiety, and insomnia [2]. However, the field faces a significant bottleneck due to a global shortage of trained music therapists with an estimated shortfall of 185,000 professionals by 2024 and overall limited public awareness [1]. Moreover, although music therapy particularly fosters social connection, reduces isolation, and enhances overall wellness, yet each professional session costs $80–$120, making it inaccessible for many [3], signifying the urgent need for scalable, accessible digital solutions.

B. Spotify Problem

Spotify has long led the industry in personalized music experiences, from Discover Weekly to Spotify Wrapped which is their cultural moment that engaged with over 200 million users daily [4], showing how deeply people connect with music. At the same time, the platform is exploring features that enhance music discovery and artist promotion. For example, Spotify's recent "Share to TikTok" integration successfully promoted 84% [5] of Global 200 songs in 2024 that first went viral on TikTok, reinforcing that emotional context drive discovery and retention. Yet, a major untapped opportunity remains as mood-driven and therapeutic personalization. Despite Spotify users frequently using music to manage stress, energy, and sleep, the platform lacks tools to intentionally guide users through emotional transitions in a structured way. Therefore, our team decided to embed an emotional-transition model into Spotify's recommendation engine by analyzing song features to infer a user's current mood, thus, generates playlists that seamlessly guide listeners from negative states (anger, sadness) toward positive ones (calm, happiness), delivering on-demand, scalable music therapy.

C. Dataset Recap

Our analysis relies on two SQL-integrated datasets including:

- Spotify Tracks Dataset [6]: A tabular CSV of nearly 90,000 top-streamed tracks across 125 genres, featuring metadata of track ID, title, artist(s), primary genre, popularity score and audio descriptors: valence, energy, tempo, danceability, acousticness, instrumentalness, speechiness, loudness, key, and mode
- Streaming History [7]: This dataset features an anonymized log of user behavior, including their play records: endTime', 'artistName', 'trackName', 'msPlayed'.

For better understanding of the dataset, Exploratory Data Analysis (EDA) is conducted. We first examine the distributions of each audio feature to understand the overall landscape of the track dataset. According to Figure 1, danceability, energy, and valence all cluster between roughly 0.5 and 1.0, indicating that most of the tracks are upbeat, energetic, and emotionally positive. In contrast, acousticness and instrumentalness are heavily right-skewed, showing that most songs contain vocals and electronic elements (as opposed to being purely acoustic or instrumental). Speechiness shows a long‑tailed distribution, meaning that while most tracks have little to no spoken content, a small subset (such as rap or spoken‑word pieces) exhibits higher speechiness values. Then, to assess how these features interact, we turned to a correlation heatmap. As shown in Figure 2, which notably reveals a strong positive relationship between energy and loudness ($r = 0.76$) and clear negative correlations between acousticness and both energy ($r = –0.73$) and loudness ($r = –0.58$), then we confirmed the correlation by creating a scatterplot (Figure 3) for pairs having $|r| > 0.6$, which could be observed that louder, more energetic songs tend to be less acoustic.

III. Data Preprocessing Summary

To prepare the datasets for modeling, we implemented a structured data cleaning and transformation workflow. We began by removing duplicate and NaN entries in both the track and listening-history tables using a left inner join, preventing redundancy and ensuring consistency. Records missing any of the nine core audio features including valence, energy, acousticness, danceability, instrumentalness, speechiness, loudness, key, mode and so on, or those lacking valid timestamp data were excluded to maintain data integrity.

Categorical variables such as mode and genre were removed to ensure compatibility with machine learning algorithms.

IV. MODEL DEVELOPMENT

A. Objective & Model Selection

With the goal to identify coherent emotional clusters in the audio‑feature space, evaluate each method's performance using both hard‑cluster and fuzzy‑cluster validity indices, and compare their suitability for playlist recommendation in a therapeutic context, we decided to use K-Means, Gaussian Mixture Models (GMM), and Fuzzy C-Means (FCM) as each aligns well with specific requirements of mood-based music clustering while balancing scalability, interpretability, and the scalability within Spotify's infrastructure.

K-Means was chosen as our primary clustering method due to its computational efficiency on large audio datasets and its ability to assign to the nearest cluster centroid, which makes the resulting clusters easy to interpret by examining their means [8]. K-means tends to produce compact, well-separated clusters when the number of clusters is chosen appropriately and can produce relatively high silhouette scores on music features (for example 0.48 for 2 clusters [9], or 0.465 for 6 clusters [10]) indicating distinct, low-overlap clusters. Its time complexity is **O(nki)** with n songs, k clusters, i iterations [9], so it remains tractable on large audio datasets.

Meanwhile, GMMs build on K-means by treating data as mixtures of Gaussians, providing soft cluster assignments [11] and allowing a song to have partial membership across clusters, which reflects the fuzzy nature of musical emotions [12]. The same study also indicates that clustering into Thayer's four mood quadrants with a GMM yielded high on average classification accuracy of about 86.3% on a 250-song dataset [12]. Besides, FCM is also a soft clustering version of K-means where each track has membership degrees for all clusters, modeling overlaps and benefits a song to belong to multiple clusters [9]. FCM minimizes a fuzzy objective (measured by the Fuzzy Partition Coefficient - FPC), balancing compactness and overlap; a study reported a 0.76 FPC for three clusters, indicating well-defined fuzzy groupings [9]. Together, these three models collectively address the primary requirements of mood-based music clustering by balancing efficiency, interpretability, and the ability to model overlapping emotional content.

All clustering models used the preprocessed feature whose 3 standardized audio descriptors (valence, energy, acousticness) were described in Section III. These features were selected based on both exploratory data analysis (EDA) and previous research suggesting their strong relevance in music emotion classification [13].

B . Model Development

a. K-Means

We first applied K-Means clustering, a popular partitioning technique that minimizes intra-cluster variance by assigning points to the nearest centroid [14]. To determine the optimal number of clusters, we applied the Elbow Method and identified a clear "elbow" at k = 4 (Figure 4), suggesting 4 as the ideal cluster count.

We configured the model with four clusters (n_clusters=4), running the algorithm multiple times (n_init=10) to ensure stability and setting a maximum of 300 iterations (max_iter=300) for convergence. A fixed random seed (random_state=42) was used to guarantee reproducibility of results. Based on the cluster means, the data grouped naturally into four moods - angry, calm, sad, and happy - each distinguished by characteristic levels of energy, acousticness, and valence.

Evaluation using Silhouette Score [15] and Davies-Bouldin Score [16] yielded values of 0.4015 and 0.8389, respectively. These indicate reasonably well-separated and compact clusters, supporting the effectiveness of the selected features.

b. Gaussian Mixture Models (GMM)

Next, we applied Gaussian Mixture Models (GMM), a probabilistic clustering method that models the data as a mixture of Gaussian distributions [17]. Unlike K-Means, GMM provides soft assignments by estimating the probability that each data point belongs to each cluster, allowing for more nuanced membership.

To determine the optimal number of components, we relied on both the Bayesian Information Criterion (BIC) [18] and Akaike Information Criterion (AIC) [19], which both indicated 4 components - in line with the K-Means results (Figure 5).

The clusters produced by GMM corresponded closely to the same mood categories identified earlier: Angry, Sad, Happy, and Calm. The model showed high confidence in its assignments, with an average cluster membership probability of 0.9092.

However, evaluation metrics revealed weaker cluster separation: the Silhouette Score [14] dropped to 0.1102, suggesting considerable overlap between clusters, and the Davies-Bouldin Score [15] increased to 2.0537, indicating less compact and less distinct clusters compared to K-Means.

c. Fuzzy C-Means Clustering

Fuzzy C-Means [20] allows each data point to belong to multiple clusters with varying degrees of membership, aligning well with the nature of emotional states in music. Using the skfuzzy package, we identified the optimal number of clusters based on the Fuzzy Partition Coefficient (FPC), which was highest at 0.73 for 2 clusters (Figure 6) and remained reasonably high at 0.54 for 4 clusters. For consistency with prior models and interpretability, we selected 4 clusters.

The resulting clusters aligned with core mood categories, and performance metrics showed a Silhouette Score [14] of 0.3992 and a Davies-Bouldin Score [15] of 0.8631, closely matching K-Means and outperforming GMM. This suggests that while fuzzy clustering allows for emotional ambiguity, it still forms coherent groupings.

C. Mood-based Recommendation

To support mood-based music recommendations for an Awareness Month campaign, we developed a mood transition model that identifies a user's dominant mood cluster and maps it to a target mood to promote emotional well-being, such as "calm" or "happy." The model transitions from negative moods like "angry" to "calm" or "sad" to "happy," while retaining the current state if the dominant mood is already "happy" or "calm." Recommendations are generated using a k-nearest neighbors (KNN) algorithm, which retrieves songs from the target mood cluster that are closest in feature space (e.g., energy, acousticness, valence) to the centroid of the user's original mood cluster. This approach ensures personalized, mood-regulating song suggestions by leveraging similarity in musical features.

Given the high dimensionality of musical features, mood clusters may overlap. KNN is well-suited for this task as it relies on distance metrics (e.g., Euclidean distance) to identify similar songs, effectively handling overlapping clusters. Additionally, since the Awareness Month campaign involves weekly mood assessments, KNN's computational efficiency is advantageous, as it requires minimal retraining and scales well with small-to-medium datasets, making it easier to generate timely recommendations.

Validation Methodology

To validate the model, we employed both qualitative and quantitative methods to compare the effectiveness of personalized versus non-personalized playlists.

Qualitative Validation

A group of five students was selected for qualitative testing, each agreeing to share their Spotify listening history for research purposes. We exported their Spotify data, analyzed their listening history, and generated personalized recommendation playlists based on the mood transition model. Due to the lack of direct Spotify API access, playlists were manually created and shared with participants. After one day of listening, we collected feedback on their experience, focusing on the perceived relevance and emotional impact of the recommended songs. This qualitative approach provides insights into user satisfaction and the model's ability to deliver mood-regulating recommendations.

Quantitative Validation

For quantitative validation, we conducted a survey within an Arts Appreciation class [21]. A non-personalized playlist, targeting a dominant mood cluster (e.g., "happy" or "calm"), was provided to the entire class (n=41) for two sessions within one week. Participants ranked the effectiveness of the playlist in regulating their mood on a standardized scale. We compared these results with the qualitative feedback from the personalized playlist group to test the hypothesis that personalized recommendations yield a higher positive response rate than non-personalized ones. The results were analyzed to assess the model’s performance and validate its effectiveness in delivering mood-based recommendations.

This dual validation approach ensures a comprehensive evaluation of the system, combining user experience insights with measurable outcomes to refine and improve the mood-based recommendation model.

V. Evaluation Result

A. Model Comparison and Final Selection

To evaluate the performance of our clustering model, we employed two widely used internal validation metrics: the Silhouette Score [14] and the Davies-Bouldin Index[15]. These metrics are particularly suited for unsupervised learning tasks where no ground truth labels exist, making them appropriate for clustering emotional features in music data.

The Silhouette Score [14] measures the degree of separation between clusters by comparing the mean intra-cluster distance with the mean nearest-cluster distance for each data point. The score ranges from -1 to 1, where values near +1 indicate well-separated, dense clusters, values around 0 suggest overlapping clusters, and negative values imply incorrect cluster assignment. A commonly accepted benchmark is that scores above 0.5 indicate meaningful clustering, while scores around 0.2–0.5 reflect moderate structure [14].

The Davies-Bouldin Index [15] evaluates clustering by measuring the average similarity between each cluster and its most similar counterpart, based on intra-cluster and inter-cluster distances. Lower scores denote better clustering, with values closer to 0 indicating compact, well-separated clusters. Typically, a Davies-Bouldin Index below 1.0 is considered acceptable, and values below 0.5 suggest high-quality clustering [21].

Below is a summary of model performance:

Algorithm	Silhouette Score	Davies-Bouldin Score
K-Means	0.4015	0.8389
Fuzzy C-Means	0.3992	0.8631
GMM	0.1102	2.0537

K-Means performed best overall, creating the most distinct and compact clusters, while Fuzzy C-Means showed competitive performance and theoretical advantages in modeling emotional overlap. Although GMM provided valuable probabilistic insights, its practical clustering quality was the weakest.

We selected K-Means as our final model due to its superior evaluation scores, intuitive results, and strong alignment with therapeutic music grouping goals. These results provide a solid basis for generating personalized, mood-driven music recommendations.

B. Recommendation

The personalized recommendation system outperforms the group non-personalized setting:

Quantitative:

- 53.5% of participants agree that the mood-based recommendation system evoked positive emotions
- 23.3% of participants disagree with the method, showing a preference for artist-based recommender

Qualitative:

- 4/5 participants enjoy the customized therapeutic experience
- 5/5 participants would share the persona cards on their social media
- 4/5 agreed that the social impact of such campaigns will be beneficial to Spotify and users in general

VI. Business Interpretation

The successful first phase (focused on prototyping and iterative scope definition) validates the concept and capitalizes on current trends in user behavior, particularly the growing demand for personalized, wellness-oriented digital experiences. By addressing users' emotional and mental wellness needs through tailored music experiences, Spotify can strengthen its brand identity while simultaneously contributing to broader public health and well-being initiatives.

A. Branding Identity

The collected evidence suggests a more pronounced mediated effect of emotionally intelligent music experiences in personalized settings compared to group contexts. This underscores the growing necessity for mood-stabilizing digital interactions tailored to individual emotional states. Such an approach presents a distinctive value proposition for Spotify, particularly as it seeks to differentiate itself in an increasingly competitive digital streaming landscape dominated by platforms like Apple Music and YouTube Music. These competitors are actively leveraging their large user bases to introduce adjacent products, making innovation in emotional personalization a critical avenue for competitive advantage.

Spotify can strategically capitalize on this opportunity by enhancing its recommendation algorithms to deliver emotionally intelligent, user-specific music experiences. This is especially relevant for its core demographic who frequently turn to music as a method of emotional regulation and stress relief. This user segment often shares common pain points with users of other platforms, such as algorithmic fatigue, lack of emotional resonance, and insufficient personalization.

By addressing these needs, Spotify is well-positioned to implement a Blue Ocean Strategy that redefines market boundaries and generates uncontested value in the wellness-oriented digital experience sector. The growing cultural momentum around mental health awareness further supports this initiative. Early positive feedback from mental wellness campaigns suggests that users are both receptive to and actively seeking tools that enhance emotional well-being.

B. Social Responsibility

As one of the four core pillars within Spotify's Equity Impact Framework, Mental Health Awareness represents a foundational principle of the company's broader social impact strategy. This project has the potential to significantly elevate Spotify's social responsibility profile by addressing the accessibility gap between high-cost, clinician-led music therapy and the more scalable, at-home experiences offered through digital streaming. By leveraging its platform to support emotional well-being, Spotify can advance its mission of social equity, creating a strong alignment between the objectives of this initiative and the company's ongoing commitment to inclusive, wellness-centered innovation.

The results from the initial phase of this project validate the conceptual framework and confirm the presence of a viable market opportunity. The findings align with current trends in user behavior, particularly the increasing demand for personalized and emotionally responsive digital services. This reinforces the potential for Spotify to lead innovation in emotionally intelligent music experiences that support mental wellness and user satisfaction.

VII. Limitations and Future Works

A. Limitation

First, although the specific focus of our project has allowed us to obtain a portrait of the effects of music recommendation on mental well-being of VinUniversity students, the result of our project can be difficult to generalize to situations where individuals have different mental stages, listening behavior, and preferences. The mood-based recommendation system is designed to provide convenient therapeutic experience regardless of demographics biases. Although the selected individuals for the qualitative interview are representing the largest proportion of Spotify demographic, they are individuals with medium-high listening frequency, tech-savvy, and low stress response may not be representative enough for the population of Spotify.

Second, the therapeutic effects of music recommendations typically unfold over time and are influenced by listening frequency and context. Our study was conducted as a one-time survey rather than through longitudinal or laboratory-controlled methods. This limited timeframe and lack of experimental manipulation may affect the reliability and stability of our findings. Additionally, while the mixed-mood recommendation model aims to balance emotional responses, it could potentially lead to adverse effects (i.e., a "reverse effect") where users experience a deterioration in mood rather than improvement.

Third, the effectiveness of the recommendation model was assessed based on users' self-reported, subjective experiences via surveys. Participants may have refrained from giving extreme ratings, which could lead to a skewed interpretation of the model's effectiveness. The absence of high or low scores suggests a possible overestimation of the model's therapeutic efficacy.

B. Future work

For future research and development, we will focus on refining the project scope, enhancing product validation methods, and integrating more sophisticated data-driven strategies for evaluating user engagement and mental well-being outcomes.

First, we propose to engage domain experts to supervise and validate the therapeutic efficacy of the proposed music recommendation system. Given the integration of music therapy principles within a Spotify-like digital platform, substantial expert consultation is necessary to ensure that the design and implementation of system features align with established therapeutic frameworks and accurately reflect the intended psychological outcomes.

Second, the current study relies primarily on self-reported data obtained through surveys, which may primarily capture subjective experiences. Such data may lack objective behavioral evidence required for a comprehensive evaluation of the system's efficacy. In future research, we intend to supplement self-report measures with behavioral metrics, including click-through rates, playback durations, and interaction patterns. These metrics will help reduce reliance on participant introspection and minimize the risk of response bias, particularly those arising from extreme or socially desirable judgments.

Furthermore, with behavioral data collection in place, subsequent iterations of the project will incorporate a hypothesis-driven feedback loop to facilitate agile and adaptive understanding of user behavior. This will enable the deployment of controlled A/B testing frameworks to rigorously evaluate the effects of different recommendation strategies on users' mental and emotional states. By adopting agile development methodologies, we aim to establish iterative refinement processes that inform the recommendation logic, user interface, and interaction flow—tailored to diverse demographic and psychographic segments.

Reference List

[1] SkyQuest Technology Consulting Pvt. Ltd. (2022b, September 21). Music Therapy Market to Generate Revenue of $4.42 Billion by 2028 | Over 50% Business in Music Therapy are Less than 5 Years Old | SkyQuest Technology. GlobeNewswire News Room. https://www.globenewswire.com/news-release/2022/09/21/2520318/0/en/Music-Therapy-Market-to-Generate-Revenue-of-4-42-Billion-by-2028-Over-50-Business-in-Music-Therapy-are-Less-than-5-Years-Old-SkyQuest-Technology.html

[2] de Witte, M., Pinho, A., Stams, G. J., Moonen, X., Bos, A. E. R., and van Hooren, S. (2022b). Music therapy for stress reduction: a systematic review and meta-analysis. Health Psychol. Rev. 16, 134–159. doi: 10.1080/17437199.2020.1846580

[3] Hookway, J. (2025, February 16). The hidden costs of music therapy you need to know. Brainwave Watch. https://brainwave.watch/the-hidden-costs-of-music-therapy-you-need-to-know/

[4] Velardo, V. (2021, December 7). Spotify's Discover Weekly explained — Breaking from your music bubble or, maybe not? Medium. https://medium.com/the-sound-of-ai/spotifys-discover-weekly-explained-breaking-from-your-music-bubble-or-maybe-not-b506da144123

[5] Marshall, T. (2025b, February 17). Does TikTok drive viral music? RouteNote Blog. https://routenote.com/blog/does-tiktok-drive-viral-music/#:~:text=TikTok%20and%20Luminate's%20'Music%20Impact%20Report'&text=Key%20findings%20from%20the%20report,average%20short%2Dform%20video%20consumer.

[6] Spotify Tracks Dataset. (2022, October 22). Kaggle. https://www.kaggle.com/datasets/maharshipandya/-spotify-tracks-dataset?fbclid=IwY2xjawKlT_JleHRuA2FlbQIxMQABHsngqYYDkKPmSwe-9orrgKP8Bw381SE0glWwB0KDIL5CH0b-TTOD_tKo9wxr_aem_epDtEXddoIfB8XoyrvvLtw

[7] Metarex21 (n.d.). SpotXtract: The Ultimate Spotify Streaming History Extractor GitHub. https://github.com/metarex21/SpotXtract/blob/main/json/StreamingHistory6.json

[8] GeeksforGeeks. (2025, May 13). K means Clustering – Introduction. GeeksforGeeks. https://www.geeksforgeeks.org/k-means-clustering-introduction/

[9] Zhang, T., Liu, X., Guo, Z., & Tian, Y. (2024b). Adaptive music recommendation: Applying machine learning algorithms using low computing device. *Journal of Software Engineering and Applications*, *17*(11), 817–831. https://doi.org/10.4236/jsea.2024.1711045

[10] Mohammad Taief, A. (2024). Application of LLMs and Embeddings in Music Recommendation Systems. *The Arctic University of Norway*. https://munin.uit.no/bitstream/handle/10037/34168/thesis.pdf?sequence=2&isAllowed=y#:~:text=pace,inside%20each%20cluster%20and%20increase

[11] GeeksforGeeks. (2025b, May 16). *Gaussian Mixture Model*. GeeksforGeeks. https://www.geeksforgeeks.org/gaussian-mixture-model/

[12] Griffiths, D., Cunningham, S., Weinel, J., & Picking, R. (2021). A multi-genre model for music emotion recognition using linear regressors. *Journal of New Music Research*, *50*(4), 355–372. https://doi.org/10.1080/09298215.2021.1977336

[13] Dalida, M. R., Aquino, L. B., Hod, W. C., Agapor, R. A., Huyo-A, S. L., & Sampedro, G. A. (2022). Music mood prediction based on Spotify's audio features using logistic regression. *2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, 1–5. https://doi.org/10.1109/hnicem57413.2022.10109396

[14] MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, 1, 281–297. https://projecteuclid.org/euclid.bsmsp/1200512992

[15] Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, 20, 53–65.

[16] Davies, D. L., & Bouldin, D. W. (1979). A Cluster Separation Measure. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-1(2), 224–227.

[17] Reynolds, D.A. (2009) Gaussian Mixture Models. In: Li, S.Z. and Jain, A., Eds., Encyclopedia of Biometrics, Springer, Boston, MA, 659-663. https://doi.org/10.1007/978-0-387-73003-5_196

[18] Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

[19] Schwarz, G. (1978). Estimating the dimension of a model. The Annals of Statistics, 6(2), 461–464. https://doi.org/10.1214/aos/1176344136

[20] Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. Computers & Geosciences, 10(2-3), 191–203. https://doi.org/10.1016/0098-3004(84)90020-7

[21] Halkidi, M., Batistakis, Y., & Vazirgiannis, M. (2001). On clustering validation techniques. Journal of Intelligent Information Systems, 17(2-3), 107–145. https://doi.org/10.1023/A:1012801612483

[22] Huy, H., Anh, N., Phuong, H. (2025), Unpublished raw data on the effect of mood-based music recommendation on VinUni students. VinUniversity

Appendix

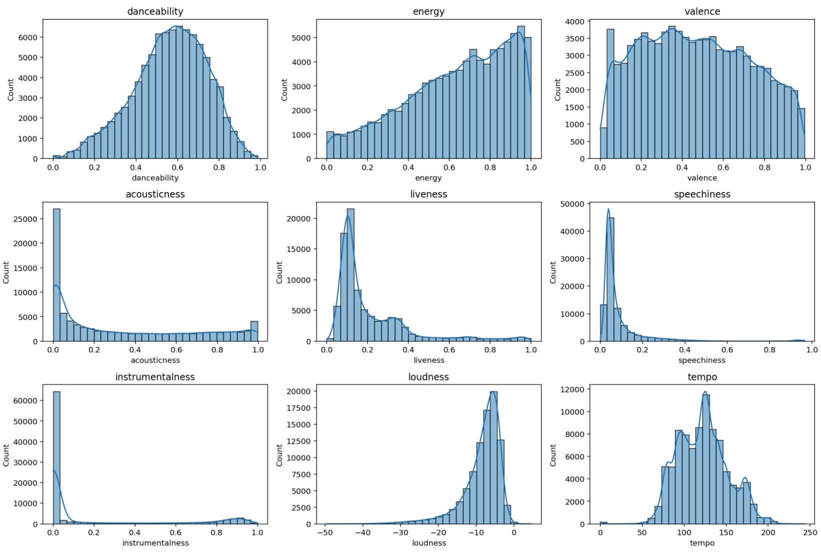

Figure 1: Histogram of each feature

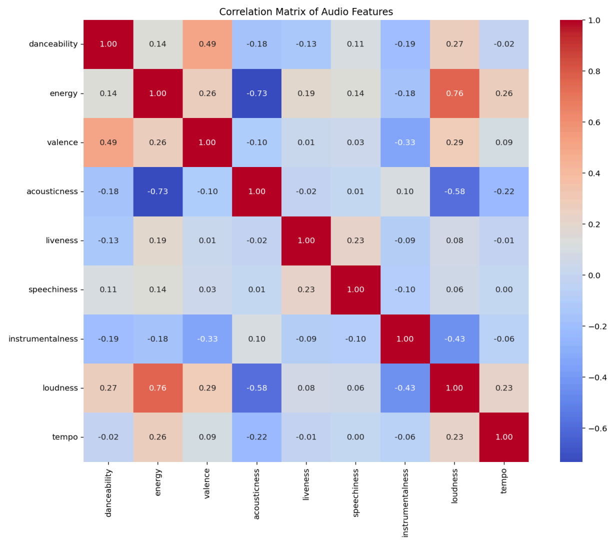

Figure 2: Correlation Matrix

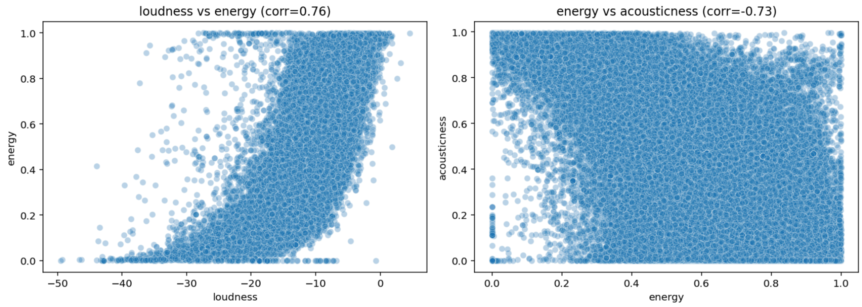

Figure 3: Scatterplot of highly correlated features (>0.6)

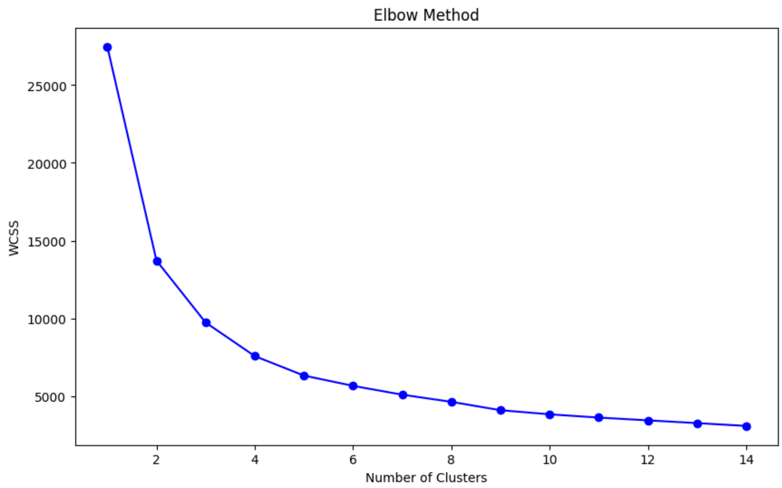

Figure 4: Elbow Method of KMeans method

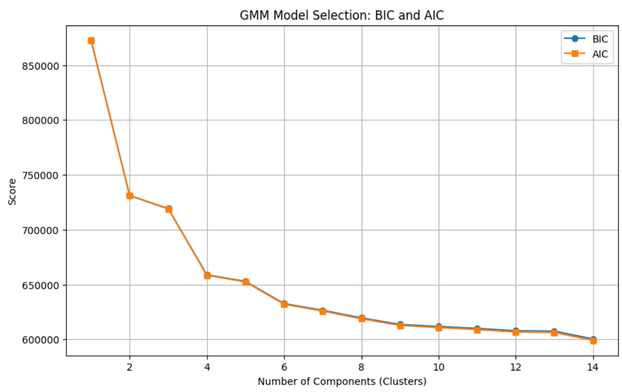

Figure 5: Elbow Method of GMM method

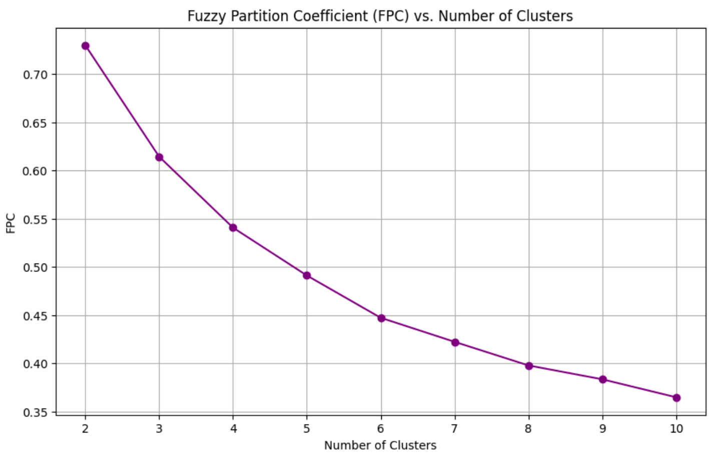

Figure 6: Elbow Method of Fuzzy C-Means method

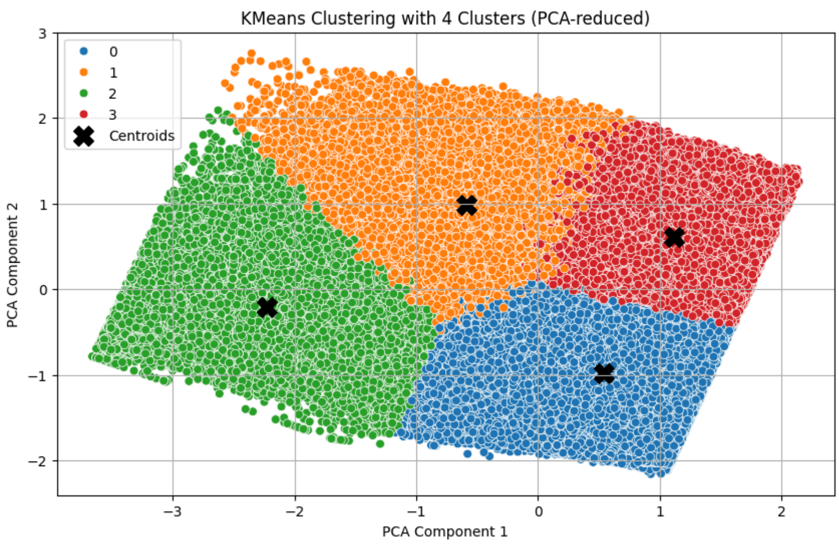

Figure 7: KMeans Clustering Result

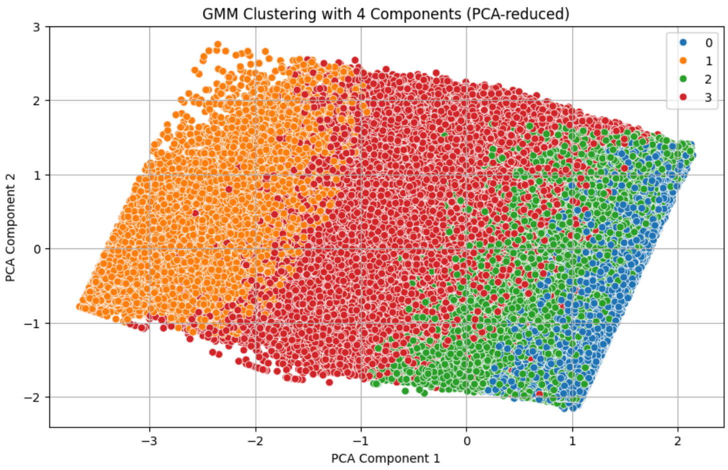

Figure 8: GMM Clustering Result

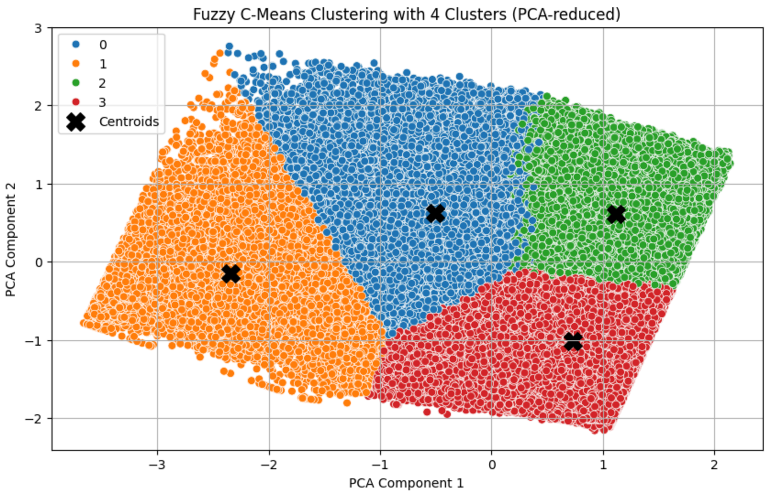

Figure 9: Fuzzy C-Means Clustering Result

Figure 10: Participants' mood before music experience (n=23)

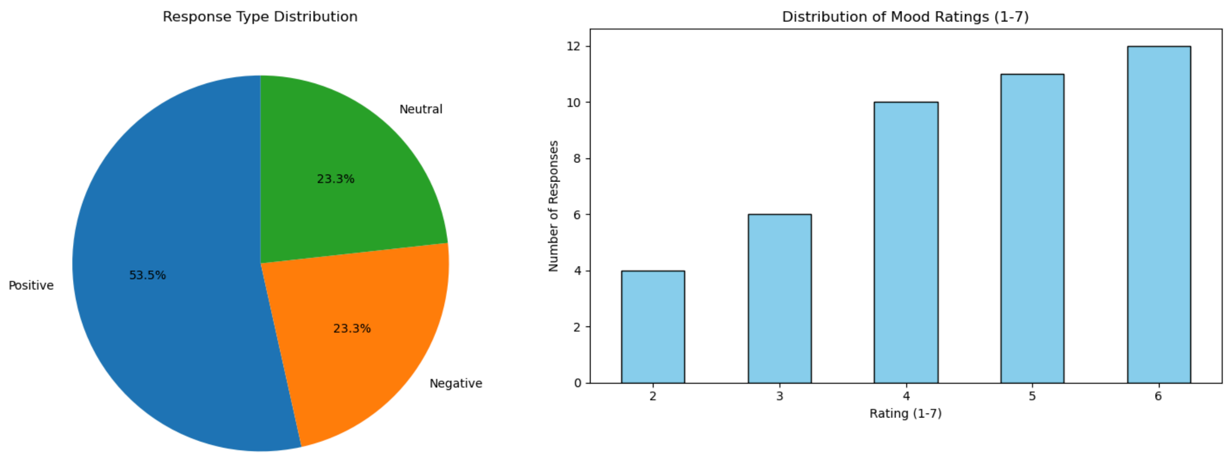

Figure 10: Survey on Non-personalized recommendation system within Arts Appreciation class (n=41)