Mapping Emotional Dynamics Through Music Consumption Patterns

Mattia Billa
Dept. of Physics,
Informatics and Mathematics
Univ. of Modena and Reggio Emilia
Modena, Italy
mattia.billa@unimore.it

Marco Furini
Dept. of Communication and Economics
Univ. of Modena and Reggio Emilia
Reggio Emilia, Italy
marco.furini@unimore.it

Manuela Montangero
Dept. of Physics,
Informatics and Mathematics
Univ. of Modena and Reggio Emilia
Modena, Italy
manuela.montangero@unimore.it

***Abstract*—Emotions significantly influence human behavior, impacting decisions, social interactions, and overall mental well-being. Understanding emotional dynamics over time is crucial not only for clinical applications, such as early detection of psychological distress, but also for enhancing user-centered services like recommendation systems and personalized advertising. Traditional emotion recognition methods rely on physiological sensors, facial expressions, or textual analysis—approaches that, while effective in controlled scenarios, suffer from issues of invasiveness, scalability, and privacy. In this paper, we propose a novel, non-invasive approach to infer users' emotional flow over time by analyzing their music listening history. Music is deeply intertwined with emotions: individuals often select songs based on their current mood, and conversely, music can induce emotional changes. Leveraging this connection, we identify recurring listening habits using a spatio-temporal clustering technique that jointly considers musical similarity and listening time. Each habit is then associated with a corresponding emotion, enabling us to reconstruct a daily emotional profile over a typical week for each user. We validate our method through an experimental study involving 14 real users who provided their listening histories. Results show the system's ability to capture individual emotional trends, highlighting its potential for both psychological monitoring and commercial applications in emotion-aware computing.**

***Index Terms*—Clustering, Recommendation System, Music Playlists**

I. Introduction

Emotions play a central role in individuals' daily lives, deeply influencing decisions, behaviors, and interpersonal relationships. The literature in psychology and neuroscience has extensively demonstrated that emotions drive a wide range of human activities, from decision-making processes related to purchases to the choice of travel destinations, the management of social relationships, and the regulation of psychophysical well-being [1], [2]. A person's emotional state is closely linked to their mental and physiological balance, and understanding its evolution over time can offer valuable insights both for the prevention of psychological disorders (such as stress, anxiety, and depression) and for optimizing marketing strategies and service personalization.

In recent years, several approaches have been proposed to recognize users' emotions, relying on physiological signals, facial expressions, body posture, or language analysis. The use of biosensors (e.g., electroencephalography, GSR sensors) allows for direct and reliable measurements of emotional states, but it presents significant limitations in terms of invasiveness, cost, and scalability [3]. Other approaches focus on textual language analysis, for example via social media or instant messaging, with promising results in specific contexts, but with challenges related to privacy, data discontinuity, and linguistic variability [4], [5]. More recently, models based on facial expressions have gained popularity, often supported by computer vision techniques [6], [7]. While promising, these methods are also intrusive and subject to environmental constraints (e.g., lighting conditions, camera position, etc.).

In parallel, numerous studies have highlighted the strong connection between music and emotions. Music not only reflects the emotional state of the listener, but can also profoundly influence it, stimulating the production of neurotransmitters associated with pleasure and mood regulation [8], [9]. Research has classified the mechanisms through which music evokes emotions, including emotional contagion, episodic memory, and rhythmic entrainment [10], [11]. Moreover, music is often chosen based on one's current mood, creating a privileged and natural channel for inferring personal emotions. This connection has also been investigated in the context of streaming platforms, where listening history can provide meaningful cues about users' psychological well-being.

In light of these considerations, there is a growing need for a mechanism capable of reconstructing a user's *emotional flow* over time, not just classifying discrete emotional states, but tracing a continuous and coherent emotional trajectory. Such a flow could be crucial for detecting early signs of psychological distress, enhancing personalized recommendation systems (in music, media, or advertising), and supporting real-time adaptive interventions [12]. A dynamic and temporal view of emotions thus represents a fundamental step toward a deeper understanding of user experience.

In this context, we propose a non-invasive method to infer users' daily emotional flow based on their *listening history*, on a weekly schedule. Listening history is a time-coded list of songs played by the user, usually enriched with metadata such as playback duration, artist, and genre [13]. Unlike sensor- or camera-based methods, our approach relies solely on data already available from music streaming platforms, preserving user privacy and enabling easy integration and scalability. A *listening habit* is a collection of similar tracks that have been played by a user at the same time slot on the same day of the week, representing a pattern in the listening history [14].

The method involves retrieving a set of analytical *features* (extracted from both audio content and metadata) for each song listened to, identifying listening habits through clustering, and associating each habit with the dominant emotion of the songs it contains. This process enables the generation of a daily emotional report for the user, organized in a weekly schedule, and analyzable at different levels of temporal granularity.

The validity of our approach was assessed through an experimental phase involving 14 real users who anonymously shared their listening history. The results are consistent with previous large-scale research by Heggli et al. [15] on diurnal listening trends and indicate that the method is able to accurately identify individual emotional trends, demonstrating potential for both clinical and commercial applications.

The rest of the paper is organized as follows: in Section II we report on related work; in Section III we present our approach to listening histories analysis for user mood classification; in Section IV we report and discuss experimental results; finally, in Section V we conclude by also highlighting future research directions.

II. Related Work

Emotions are complex internal states resulting from the interplay between cognitive evaluations and physiological responses to stimuli, whether external (events, social interactions) or internal (memories, bodily sensations). They are inherently subjective and influence both behavior and physiological activity, making their classification a central task in affective science.

A. Theoretical Models of Emotion

Several theoretical frameworks have been proposed to describe and categorize emotions:

- **Ekman's discrete emotion theory** [16] identifies six basic, universally recognized emotions (happiness, sadness, anger, fear, disgust, surprise), widely applied in cross-cultural psychology and facial expression recognition.
- **Russell's circumplex model** [17] organizes emotions along two continuous axes — valence (positive vs. negative) and arousal (low vs. high) — making it particularly relevant in affective computing and sentiment analysis.
- **The PAD model** (Pleasure–Arousal–Dominance) [18] situates affective states in a three-dimensional space, with applications in consumer research and interactive design.
- **Plutchik's wheel of emotions** [19] defines eight primary emotions arranged in four opposing pairs, providing a dynamic view of affective relationships.
- The **OCC model**, frequently used in artificial intelligence and robotics, focuses on cognitive appraisal mechanisms underlying emotional responses.

B. Emotion Models in Music Analysis

In music-related affective computing, dimensional frameworks such as the Circumplex, PAD, and Plutchik's model are often preferred. They enable the mapping of affective variations to quantifiable musical attributes, including tempo, tonal mode, timbre, and dynamic range, thus supporting emotion-aware music recommendation and more personalized listening experiences.

A growing body of research investigates the relationship between musical content and emotional perception. Chang et al. [20] employed recurrent neural networks to predict emotional content from audio signals, showing potential for therapeutic applications and adaptive recommendations. Dutta and Mookherjee [21] analyzed large-scale Spotify data to explore valence trends and genre preferences over time. Chen et al. [22] applied Transformer Encoder architectures to detect the affective tone of musical tracks. Su and Zhou [23] adopted heuristic clustering to uncover latent relationships among musical features and self-reported listener emotions. Gujar and Reha [24] leveraged machine learning to correlate mood and listening behavior, discussing applications in mental health monitoring.

C. Our Contribution

Unlike approaches that infer emotional states from raw acoustic features, our method exploits tracks annotated with human-assigned emotional tags. These labels are used to train a classifier to assign moods to songs, and then the classifier is used to infer the temporal distribution of user emotions within the individual recurrent patterns in the listening history. This process enables the reconstruction of an *emotional flow* over time, providing valuable input for both psychological well-being assessment and personalized content delivery.

III. Methods

A. Data

This study relies on two types of data: the users' listening histories that are to be analyzed, and an emotion-based labeled song dataset.

1) Listening Histories: Every listening history is enriched through metadata (e.g., artist, album, title, duration, playback time, etc.) and through an audio feature vector $x_s \in \mathcal{R}^f$, where f is the song-feature dimension. Specifically, each song is described with 9 features (tempo, loudness, acousticness, energy, valence, danceability, instrumentalness, liveness, speechiness).

For each user, the listening history of the last two months is considered. This comes from the assumption that listening habits might change over time and seasons (e.g., summer vs. winter). We consider only songs that have been listened to for at least 30 seconds, assuming that shorter plays mean that the user did not like the song.

2) Mood Dataset: To identify the mood of each song, the Moodify dataset has been used[1]. The Moodify dataset consists of 1,200 songs collected and systematically labeled

[1] https://www.kaggle.com/datasets/abdullahorzan/moodify-dataset/

for emotion-based music classification and recommendation. The songs were retrieved from Spotify's most popular mood playlists using the Spotify API. Each song is categorized into one of four primary emotions (sad, happy, energetic, calm) based on Robert Thayer's traditional mood model [25], with 300 songs per emotion category to ensure statistical balance.

B. Mining Listening Habits

The first goal of this work is to extract the listening habits of users from their listening data. In the context of this paper, listening habits are defined as collections of similar tracks that a user tends to play during roughly the same time period. We make the following assumptions:

1) Listening habits are periodic; in particular, we assume that they follow a weekly pattern, reflecting the weekly schedule of everyday life. Therefore, our analysis only considers times of the day and days of the week.
2) The number of listening habits is not known in advance.
3) Listening habits do not follow fixed time slots.
4) Both temporal proximity and song similarity should be considered when identifying a user's listening habits.

The proposed approach is developed with these assumptions in mind. Furthermore, since listening habits are inherently personal and listening data can vary significantly from user to user, the entire approach is designed to be applied independently for each user.

To extract a user's listening habits, we first reduce the dimensionality of the audio feature space using Principal Component Analysis (PCA) while retaining the most relevant musical information. We then apply a spatio-temporal clustering method that jointly considers musical similarity and temporal proximity, identifying groups of songs that reflect distinct listening habits. This process yields a set of clusters, each corresponding to one recurring habit in the user's weekly listening pattern.

1) Dimensionality Reduction: Since clustering techniques and distance-based methods in general are negatively affected by the curse of dimensionality, we first compute a lower-dimensional embedding of the songs listened to by each user. To this end, we apply a PCA to the set of song features, excluding the time attribute. Starting with nine audio features, we retain only the principal components with the highest eigenvalues such that the cumulative explained variance reaches at least 80%. In practice, we found that for most users, four principal components are sufficient to meet this threshold.

2) Clustering: To identify listening habits, we focus on density-based clustering techniques, such as DBSCAN, because they do not require the number of clusters to be specified in advance and are more robust to outliers.

Indeed, our second assumption states that the number of listening habits is not known *a priori*, thus the use of clustering methods that require this information, such as *k*-means, can be problematic unless combined with additional heuristics to estimate the number of clusters. Moreover, methods such as *k*-means assume clusters to be roughly spherical and of similar size, which may not hold true for listening habits. Other approaches, such as Gaussian Mixture Models, share similar limitations due to assumptions on cluster shapes and the need to specify the number of clusters in advance.

In addition, listening histories may contain outliers, for example, when a user plays a song they do not typically listen to (e.g., suggested by a friend). Such cases must be accounted for during clustering.

Moreover, in this scenario, as highlighted in our fourth assumption, a key challenge is that listening habits depend on both temporal proximity and song similarity. Songs played at similar times, but with very different musical features, or similar songs played during unrelated periods, should not be grouped into the same habit. To address this, we employ a spatio-temporal extension of DBSCAN, called ST-DBSCAN [26]. This method distinguishes spatial and temporal features by applying separate thresholds (ε_1 for spatial similarity and ε_2 for temporal proximity) when determining neighborhoods. In practice, the implementation used in this work [27] preprocesses the spatial distance matrix for standard DBSCAN by identifying pairs of samples with a temporal distance greater than ε_2 and setting their spatial distance to a value exceeding ε_1, thereby preventing them from being grouped together.

In our setting, we adapt this approach by encoding the time feature as the number of minutes elapsed since Monday 00:00, while treating the song features as the spatial dimensions.

The main challenge in applying DBSCAN, and particularly ST-DBSCAN, lies in selecting appropriate hyperparameters: the minimum number of points required to form a cluster (MinPts), the spatial threshold (ε_1), and the temporal threshold (ε_2). Unlike other clustering methods, for which a simple grid search with a standard score (e.g., the Silhouette score) would be sufficient, tuning these parameters for DBSCAN is more delicate. Since DBSCAN can label varying numbers of points as outliers, standard scores can favor configurations that overestimate outliers. Alternative methods, such as identifying the knee point in the k-distance graph, may not always yield clear results due to multiple or ambiguous knee points.

In this work, we adopt the following heuristics to select the hyperparameters:

- For the minimum number of points (MinPts) required to form a cluster, we use the average number of songs per hour in the listening history, computed as:

$$\mathrm{MinPts} = \max\left(5,\ \mathrm{round}\Big(\frac{N}{24 \times 7}\Big)\right)$$

  where N is the total number of listened songs in the user's history, ensuring a minimum of 5 songs per cluster.
- For the temporal threshold ε_2, we assume that songs within a habit should be listened to within at most 30 minutes of each other.
- For the spatial threshold ε_1, we perform an iterative search starting from a small value (0.01) and gradually increasing it until fewer than 10% of the samples are classified as outliers.

To associate new listening events with an existing user habit, it is sufficient to compute the spatial distance (i.e., the distance based on the song features) to the songs already assigned to each habit. For each new song, the closest song in the feature space is identified. If both the spatial distance and the temporal distance to this closest song are below the corresponding thresholds (ε_1 and ε_2), the new point is assigned to the same habit. If either condition is not satisfied, the next closest song is considered, and the process is repeated until either a matching habit is found or no match is possible under the defined constraints.

C. Temporal Mood Profiling

The previous section described a method for identifying users' listening habits from their listening histories using clustering. Our next goal is to profile these habits in order to track users' mood patterns throughout the week. To this end, we first train a classifier to assign moods to songs, and then the trained classifier is used to predict the mood of each song in the users' listening habits. For each habit, the most frequent predicted mood is assigned as its representative mood.

TABLE I
COMPARATIVE RESULTS OF DIFFERENT CLASSIFIERS ON THE MOOD CLASSIFICATION DATASET.

Model Name	Mean Accuracy (%)	Std. Dev.
Logistic Regression	80.1	3.6
Decision Tree	73.1	4.9
Random Forest	80.2	4.6
XGBoost	80.2	4.2
Support Vector Machine (SVM)	79.2	3.4
Neural Network (MLP)	80.1	4.8

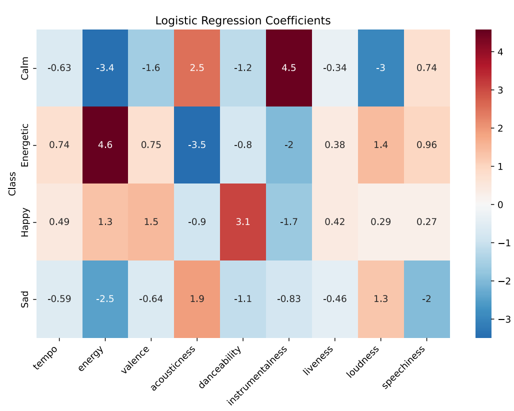

Fig. 1. Logistic Regression coefficients.

We train the mood classifier on the *Moodify* labeled dataset, which, we recall, consists of 1200 songs annotated with one of four mood categories: *sad*, *energetic*, *happy*, or *calm*. Several off-the-shelf classifiers were evaluated using 10-fold cross-validation; the results are reported in Table I. Notably, logistic regression achieves competitive accuracy comparable to less interpretable models such as random forests and XGBoost. Due to its interpretability, we select logistic regression as our final classifier, as it allows us to inspect and interpret the influence of each covariate. Figure 1 illustrates, for example, the learned coefficients for each feature, revealing intuitive patterns: for instance, *calm* songs tend to have higher values of *instrumentalness* and *acousticness*, and lower *energy* and *loudness*, while *happy* songs are more *danceable*, and *energetic* songs actually have more *energy* than the other classes.

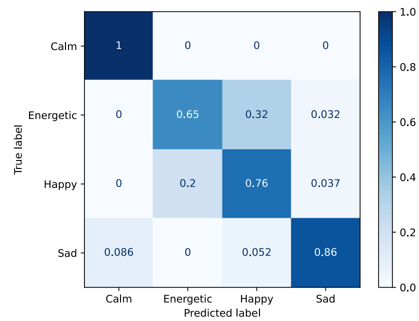

Fig. 2. Confusion matrix of the classifier over the mood classes.

Figure 2 presents the confusion matrix of the classifier, showing that *happy* and *energetic* songs are often confused: happy songs are confused with energetic songs 20% of the time, while the opposite happens 32% of the time. This confusion appears consistently across all tested models, indicating that it is an inherent characteristic of the dataset.

D. Graphical Weekly Mood Representation

Since a user habit may contain a mix of moods, it is important to communicate the level of purity or uncertainty associated with the assigned mood. We draw inspiration from the graphical representation of sequence conservation in bioinformatics, known as the *sequence logo* [28], to derive a graphical representation for the user's weekly mood. In a sequence logo, the height of each character is proportional to its frequency, quantified using Shannon entropy. In our approach, we adapt this concept by varying the opacity of the habit's mood color according to its uncertainty: the higher the entropy, the lower the opacity, indicating that the mood is less clearly defined.

For a categorical distribution with k possible outcomes, the entropy is given by:

$$H(X) = -\sum_{i=1}^{k} p_i \log_2 p_i,$$

where p_i is the probability of the i-th category. In our case, $k = 4$ (corresponding to the four moods), and p_i is estimated by the frequency of the i-th mood within the habit. Since the maximum entropy for four categories is 2, we define the opacity $O(X)$ as:

$$O(X) = 1 - \frac{H(X)}{2},$$

which ensures the opacity ranges from 0 (maximum uncertainty) to 1 (complete certainty).

IV. Experimental Results and Discussion

We collected the listening histories of 14 Spotify volunteer users contacted through our university teaching platforms. On average, 1200 samples were retrieved for each user (min. 108, max. 4358).

We applied the previously described method to the listening histories and, among the 14 users included in this study, an average of 23.5 listening habits were identified per user (min. 2, max. 40). The median number of habits is 24.5, with an interquartile range (IQR) of 22 to 31.25.

Each habit contains a variable number of listening events: the median number of songs per habit is 21 (IQR: 12–46), while the mean is 44.7. This large difference between mean and median indicates a strong right skew, with some habits encompassing hundreds of songs (maximum: 506). This wide range reflects the high variability in individual listening patterns and the diversity of music consumption during the two-month observation period.

It is important to note that habits are only detected when recurring patterns are actually present. For example, the user with the fewest detected habits (two) had listened to 405 songs in total, yet their listening history lacked sufficient temporal and musical regularity to produce more clusters.

A. Examples of Weekly Listening Habits

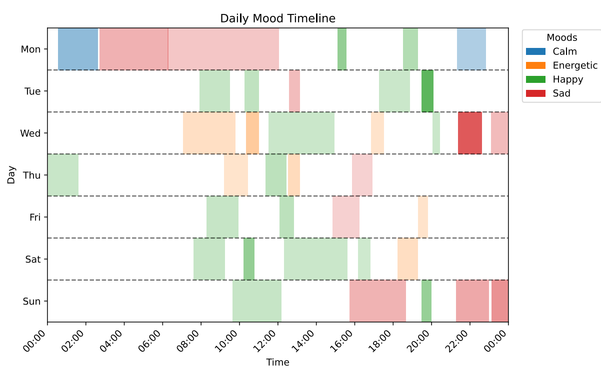

Fig. 3. Weekly mood timeline of User 1, with 36 habits identified and an average of 25 songs per habit (min: 6, max: 103).

We present the weekly emotion timeline of two representative users. Other users in the dataset exhibit broadly similar patterns, though many are less distinct or less diverse. Presenting all users would be redundant and exceed space constraints, so we focus on these two illustrative cases.

Figure 3 shows the weekly mood timeline for User 1, where stronger colors represent less uncertainty of the classified mood. This user tends to listen to calm and sad music during Monday mornings and late nights on Wednesdays and Sundays, with Wednesday nights showing a particularly consistent and marked habit. We can also observe the listening of energetic music on Wednesday and Thursday mornings, probably indicating some workout activities.

A similar trend can be observed in Figure 4 for User 2, who also listens to calm and sad music on weekday mornings

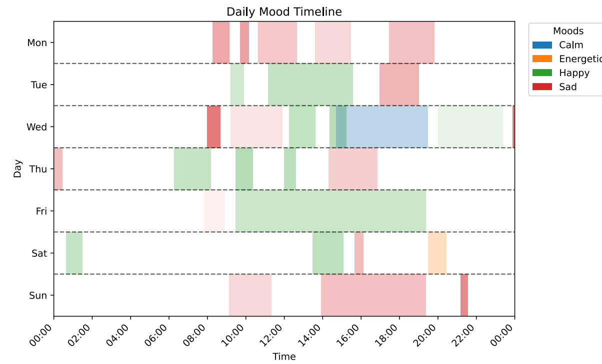

Fig. 4. Weekly mood timeline of User 2, with 28 habits identified and an average of 49.5 songs per habit (min: 9, max: 234).

and late nights. Notably, both User 1 and User 2 show a distinct tendency to listen to energetic music on Saturday nights, possibly indicating a social or leisure-related habit.

B. Overall Mood Trends Across Users

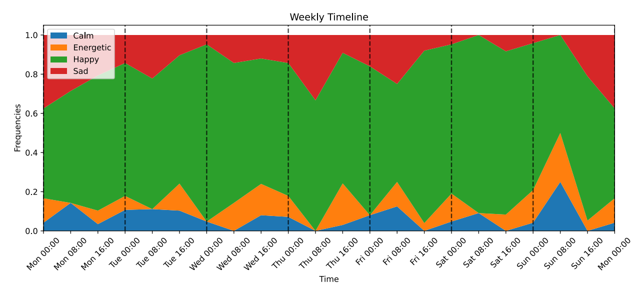

Fig. 5. Aggregate distribution of moods throughout the week for all users.

Aggregating the habits of all 14 users into a single dataset reveals broader mood trends throughout the week. Figure 5 shows that sad moods are most prevalent at the beginning of the workweek and on Sunday afternoons, consistent with the idea of a "Sunday evening blues" effect.

In contrast, calm and energetic moods are more common on weekend mornings and evenings, respectively. These trends align with expectations that weekends involve a greater mix of relaxing and social activities, reflected in more upbeat or quiet listening patterns.

Figure 6 shows the relative differences in the audio feature values of users' listening habits across the day. For each user's habit cluster, the median audio feature values were computed; then, for each time segment, the average was calculated and compared to the overall mean across all time segments.

The resulting daily pattern is consistent with previous large-scale research by Heggli et al. [15], who analyzed over two billion Spotify listening events and found clear diurnal trends: instrumentalness and acousticness tend to increase during the day, while energy, valence, and tempo decrease toward nighttime. Despite the much smaller and localized scope of our dataset, the similarity in these daily fluctuations supports

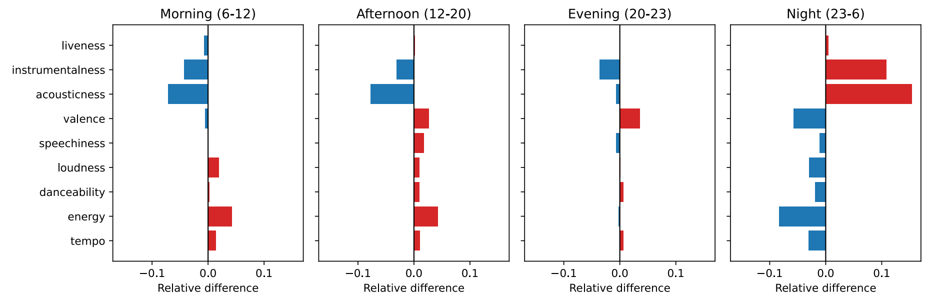

Fig. 6. Relative difference in audio feature values during times of the day. Blue indicates a relative decrease, and red indicates a relative increase.

the validity of our habit detection approach. Together, these results suggest that even at the individual level, mood-linked listening habits reflect the broader diurnal musical preference patterns observed globally.

V. Conclusions

In this paper, we present a method to derive a day-by-day weekly emotional flow of a user based on music listening history. To this end, we applied both supervised and unsupervised machine learning techniques. We tested the proposed approach on the anonymously shared listening histories of 14 real users, obtaining results consistent with previous related studies and suggesting applications in clinical monitoring and music recommendation.

Future work could explore non-linear dimensionality reduction methods, such as UMAP, alternative clustering algorithms like OPTICS with spatio-temporal extensions, and techniques that adapt to evolving user behavior over time to improve habit detection. Finally, direct validation with users or practitioners (e.g., psychologists) would provide valuable insight into the method's real-world effectiveness.

Acknowledgment

The authors wish to thank Alex Billa for his preliminary work on this project during his bachelor's thesis.

References

[1] J. S. Lerner, Y. Li, P. Valdesolo, and K. S. Kassam, "Emotion and decision making," *Annual review of psychology*, vol. 66, no. 1, 2015.

[2] C. Achar, J. So, N. Agrawal, and A. Duhachek, "What we feel and why we buy: the influence of emotions on consumer decision-making," *Current Opinion in Psychology*, vol. 10, pp. 166–170, 2016.

[3] G. Pei, H. Li, Y. Lu, Y. Wang, S. Hua, and T. Li, *Affective computing: Recent advances, challenges, and future trends*. AAAS, 2024, vol. 3.

[4] S. Zad, M. Heidari, H. James Jr, and O. Uzuner, "Emotion detection of textual data: An interdisciplinary survey," in *2021 IEEE World AI IoT Congress (AIIoT)*. IEEE, 2021, pp. 0255–0261.

[5] S. M. Mohammad and S. Kiritchenko, "Using hashtags to capture fine emotion categories from tweets," *Computational Intelligence*, vol. 31, no. 2, pp. 301–326, 2015.

[6] A. I. Siam, N. F. Soliman, A. D. Algarni, F. E. Abd El-Samie, and A. Sedik, "Deploying machine learning techniques for human emotion detection," *Computational intelligence and neuroscience*, p. 8032673, 2022.

[7] B. C. Ko, "A brief review of facial emotion recognition based on visual information," *sensors*, vol. 18, no. 2, p. 401, 2018.

[8] K. K. Jakupčević, I. R. Ercegovac, and S. Dobrota, "Music as a tool for mood regulation: the role of absorption vs. mindfulness," *Primenjena psihologija*, vol. 14, no. 2, pp. 229–248, 2021.

[9] V. N. Salimpoor, M. Benovoy, K. Larcher, A. Dagher, and R. J. Zatorre, "Anatomically distinct dopamine release during anticipation and experience of peak emotion to music," *Nature neuroscience*, vol. 14, no. 2, pp. 257–262, 2011.

[10] M. Barbaresi, D. Nardo, and S. Fagioli, "Physiological entrainment: a key mind–body mechanism for cognitive, motor and affective functioning, and well-being," *Brain Sciences*, vol. 15, no. 1, p. 3, 2024.

[11] J. Cespedes-Guevara and N. Dibben, "The role of embodied simulation and visual imagery in emotional contagion with music," *Music & Science*, vol. 5, p. 20592043221093836, 2022.

[12] R. Cavicchioli and M. Furini, "Listening to emotions: Inferring user mood through music consumption patterns," in *2025 IEEE Symposium on Computers and Communications (ISCC)*. IEEE, 2025.

[13] M. Furini and F. Fragnelli, "Social music discovery: an ethical recommendation system based on friend's preferred songs," *Multimedia Tools and Applications*, pp. 1–15, 2024.

[14] M. Furini and M. Montangero, "Understanding users music listening habits for time and activity sensitive customized playlists," in *2023 IEEE Consumer Commun. & Networking Conf. (CCNC)*, 2023, pp. 485–488.

[15] O. A. Heggli, J. Stupacher, and P. Vuust, "Diurnal fluctuations in musical preference," *Royal Society open science*, vol. 8, no. 11, p. 210885, 2021.

[16] P. Ekman, "Are there basic emotions?" *Psychological Review*, 1992.

[17] J. A. Russell, "A circumplex model of affect." *Journal of personality and social psychology*, vol. 39, no. 6, p. 1161, 1980.

[18] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Current Psychology*, vol. 14, pp. 261–292, 1996.

[19] R. E. Plutchik and H. R. Conte, *Circumplex models of personality and emotions.* American Psychological Association, 1997.

[20] X. Chang, X. Zhang, H. Zhang, and Y. Ran, "Music emotion prediction using recurrent neural networks," *arXiv preprint arXiv:2405.06747*, 2024. [Online]. Available: https://arxiv.org/abs/2405.06747

[21] S. Dutta and S. Mookherjee, "Exploring the emotional landscape of music: An analysis of valence trends and genre variations in spotify music data," *arXiv preprint arXiv:2310.19052*, 2023.

[22] Y. Chen, Z. Ma, M. Wang, and M. Liu, "Advancing music emotion recognition: A transformer encoder-based approach," in *Proceedings of the 6th ACM International Conference on Multimedia in Asia*. Association for Computing Machinery, 2024.

[23] J. Su and P. Zhou, "Machine learning–based modeling and prediction of the intrinsic relationship between human emotion and music," *ACM Trans. Appl. Percept.*, vol. 19, no. 3, 2022.

[24] S. S. Gujar and A. Y. Reha, "Exploring relatonship between music and mood through machine learning technique," in *Proc. of the 5th International Conference on Information Management & Machine Intelligence*. Association for Computing Machinery, 2024.

[25] R. E. Thayer, *The biopsychology of mood and arousal*. OUP, 1990.

[26] D. Birant and A. Kut, "St-dbscan: An algorithm for clustering spatial–temporal data," *Data & knowledge engineering*, vol. 60, no. 1, 2007.

[27] E. Cakmak, M. Plank, D. S. Calovi, A. Jordan, and D. Keim, "Spatio-temporal clustering benchmark for collective animal behavior," in *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Animal Movement Ecology and Human Mobility*, 2021.

[28] T. D. Schneider and R. M. Stephens, "Sequence logos: a new way to display consensus sequences," *Nucleic acids research*, vol. 18, no. 20, pp. 6097–6100, 1990.