Article

Not peer-reviewed version

A Continuous Music Recommendation Method Considering Emotional Change

Se In Baek and Yong Kyu Lee *

Posted Date: 20 March 2025

doi: 10.20944/preprints202503.1504.v1

Keywords: emotional changes; emotion prediction; multiple regression analysis; Markov chain; personalized mental model; cosine similarity; recommendation system

Article

A Continuous Music Recommendation Method Considering Emotional Change

Se In Baek [1] and Yong Kyu Lee [2,*]

[1] Knowledge Science Research Center, Dongguk University, Seoul, Republic of Korea
[2] Division of AI Software Convergence, Dongguk University, Seoul, Republic of Korea
* Correspondence: yklee@dongguk.edu

Abstract: Music, movies, books, pictures, and other media can change a user's emotions, which are important factors in recommending appropriate items. As users' emotions change over time, the content they select may vary accordingly. Existing emotion-based content recommendation methods primarily recommend content based on the user's current emotional state. In this study, we propose a continuous music recommendation method that adapts to a user's changing emotions. Based on Thayer's emotion model, emotions were classified into four areas, and music and user emotion vectors were created by analyzing the relationships between valence, arousal, and each emotion using a multiple regression model. Based on the user's emotional history data, a personalized mental model (PMM) was created using a Markov chain. PMM was used to predict future emotions and generate user emotion vectors for each period. A recommendation list was created by calculating the similarity between music emotion vectors and user emotion vectors. To prove the effectiveness of the proposed method, the accuracy of the music emotion analysis, user emotion prediction, and music recommendation results were evaluated. To evaluate the experiments, the PMM and the modified mental model (MMM) were used to predict user emotions and generate recommendation lists. The accuracy of the content emotion analysis was 87.26%, and the accuracy of user emotion prediction was 86.72%, an improvement of 13.68% compared with the MMM. Additionally, the balanced accuracy of the content recommendation was 79.31%, an improvement of 26.88% compared with the MMM. The proposed method can recommend content that is suitable for users.

Keywords: emotional changes; emotion prediction; multiple regression analysis; Markov chain; personalized mental model; cosine similarity; recommendation system

---

1. Introduction

Human emotions appear in various ways, even in response to identical external stimuli, influenced by individual standards, life patterns, personality, and way of thinking [1]. Emotionally intelligent computing refers to computers equipped with emotion recognition capabilities that learn and adapt to process human emotions, facilitating more effective interactions between humans and computers [2]. Recent advancements in information technology have stimulated active research on the interaction between emotions and technology. Numerous studies have explored human emotions and much effort has been made to find a link between technology and emotional states. However, emotion analysis outcomes are influenced by various factors, and it is difficult to confirm whether the emotion at a specific moment has been accurately determined. Moreover, individual variations in emotional trends can reduce satisfaction with emotion recognition capabilities.

To increase satisfaction with emotion recognition, it is necessary to consider individual-specific factors rather than general emotion recognition. Various studies have explored emotion-based recommendation methods [3,4] aiming to recommend diverse content types, such as music, movies, books, and art, while considering emotions. In the development of services and products, these recommendation technologies emphasize that user emotions play a pivotal role in service usage and item purchase decisions. Therefore, emotions significantly influence the recommendation of suitable items to users, and the users' choice of items may vary depending on their emotional state. As user's emotions evolve over time, the intensity and duration of emotions can vary even when experiencing the same emotion.

Studies on content recommendations that consider existing emotions [5-10] typically base their recommendations on users' present emotional state. Human emotions change regularly over time for each individual [11,12], and individuals often experience complex emotions rather than singular ones [13,14]. Therefore, recommendation results based solely on current emotions have limitations in terms of accuracy and diversity. To address these challenges, it is necessary to consider the complexity of emotions and their evolution over time.

Among emotional stimuli, images and music, which are visual and auditory, can evoke emotions quickly and persist as long-term memories [1]. Music, also called the language of emotions, is considered one of the greatest stimuli affecting user emotions [15,16] and has the capacity to evoke positive or negative emotions [17]. Differences in emotions expressed through music have been attributed to factors such as gender, language ability, and socioeconomic background [18-20]; however, there is no method to express general human emotions toward music. Recent studies have indicated that it is consistent across cultures [21,22]. Therefore, this study focuses on leveraging music data as a representative medium for influencing emotional changes among various forms of content.

Many studies are being conducted to analyze content and user emotions. Emotion refers to the degree or intensity with which the five senses are stimulated from the outside and react to it [23]. Emotional models have been studied to express these emotions using the concept of dimensions [24]. This two-dimensional approach has polarities representing valence and arousal [25]. The valence axis represents the degree of positivity and negativity of emotions and the arousal axis represents the intensity of arousal. Valence and arousal are correlated with emotions, and emotions can be expressed based on them [22,26,27]. Representative emotion models include those proposed by Russell [28] and Thayer [29]. Studies on emotion classification primarily use Thayer's model [30-32]. In this study, Thayer's emotion model was used to express emotions by dividing them into four emotion areas.

Among content type, lyrics and audio information, which are elements of music, are used to analyze the emotions of music [30,32-36]. There are also studies analyzing the emotions evoked by paintings [37] or movies [38]. To analyze a user's emotions, research has been conducted to infer emotions using social factors, facial expressions, voices, and brain wave data [39-41] and to predict emotions based on other people's emotions or actions[26,42,43]. Existing research on predicting emotions [26] predicted other people's emotions through a mental model integrated with a questionnaire on emotional transitions. This model is limited because it cannot account for personalized emotional changes, applying instead the average value of the emotional transition probability for the entire user group. To address this, this study proposes a personalized mental model. We attempted to classify music emotions using a multiple regression model based on music emotion data. Additionally, we aimed to predict changes in the user's emotions using Markov chain and interpolation methods based on the user's emotional history data.

In this study, we propose a continuous music recommendation method that considers emotional changes. First, we created an emotion vector for the music using valence and arousal values, elements of Thayer's emotion model. Using a multiple regression model, we analyzed the relationships between valence, arousal, and the four emotions to generate emotion vectors for the music. Second, we generated the user's emotion vector based on their emotional changes. The rate of emotional change for each user perspective was used to create these vectors. Each time, we generated an individual's emotional state transition matrix using a Markov chain, defining the user's individual mental model. The user's emotion vector was then generated by predicting the emotional state at time t_0+n, where t_0 is the current time. Third, we calculated the similarity between the music emotion vector and the user emotion vector to generate a recommendation list according to the emotional state at time t_0+n. For continuous recommendation, interpolation was used to estimate the user's

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2025 doi:10.20944/preprints202503.1504.v1

emotional state between times t_0 and t_0+n, creating a continuous recommendation list based on user emotions. We expect the proposed method to recommend content that is more suitable for users.

The structure of this paper is as follows. Chapter 2 describes existing emotion-based recommendation methods, emotion models, and emotion analysis methods as related research to this paper. Chapter 3 details the generation of music emotion vectors proposed in this paper, Chapter 4 describes the creation of user emotion vectors according to emotional changes, and Chapter 5 presents the continuous music recommendation method. Chapter 6 discusses the experiment and evaluation results of the proposed method. Finally, Chapter 7 provides the conclusion.

2. Related research

2.1. Emotion-Based Recommendations

Recently, with the growth of the OTT market and the production of various types of content, technologies that recommend customized content to users have been advancing. As user participation in service and product development increases alongside recommendation technology, the emotions experienced by users become crucial factors in their service utilization and purchasing decisions [4]. Consequently, research in emotional engineering, which analyzes and evaluates the relationship between human emotions and products, is actively progressing. Emotional engineering embodies engineering technologies aimed at faithfully reflecting human characteristics and emotions, with emotional design, emotional content, and emotional computing serving as integral research fields within this discipline [4]. In Moreover, with the proliferation of mobile devices and the increased use of social networking services, individuals can share their thoughts and opinions anytime and anywhere. Consequently, user emotions are analyzed through social media platforms, and numerous studies are underway based on recommendations based on user emotions.

Content such as music, movies, and books has a significant influence on changing users' emotions. Therefore, emotions are an important factor in content recommendation. Existing emotion-based recommendation studies analyze content and user emotions to suggest content that aligns with the user's emotional state.

The music recommendation method that considers emotions [5,44-46] recognizes facial expressions to derive the user's emotions and recommends music suitable for the user. As a recommendation method, content-based filtering (CSF), collaborative filling (CF), and similarity techniques were used to recommend music suitable for the user. In addition, a study recommending music by estimating emotions through brainwave data, which is a user's biosignal [47], a study recommending music by analyzing the user's music listening history data [48], and a study recommending music by analyzing the user's music listening history data [48], and even studying the user's keyboard input and mouse click patterns to analyze and recommend music [49]. Research on music recommendations that consider existing emotions mainly recommends music based on the user's current emotional state.

The emotion-based movie recommendation method [6] defines colors as singular emotion and recommends movies based on the color selected by the user. This approach combines content-based filtering, collaborative filtering, and emotion detection algorithms to enhance the movie recommendation system by incorporating the user's emotional state. However, because colors are defined as representing singular emotions and only consider the user's emotion corresponding to the selected color, the accuracy of the recommendation results is expected to be low. To provide recommendation that better suit the user, it is essential to express the diverse characteristics and intensity of the user's emotions.

The emotion-based tourist destination recommendation method [7] combines content-based and collaborative filtering techniques to recommend tourist destinations based on the user's emotions. This approach involves collecting data on tourist attractions and quantifying emotions by calculating term frequencies based on words from eight emotion groups. Furthermore, research has been

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2025 doi:10.20944/preprints202503.1504.v1

conducted on recommending artwork by considering emotions [8], recommending fonts by analyzing emotions through text entered by users [9], and recommending emoticons [10].

A common limitation of these emotion-based content recommendation studies is that they only make recommendations that consider the user's current emotions. An individual's emotions change over time, and even the same emotion can be felt differently depending on how it changes. An individual's current emotional state may transition to a different emotional state over time or depending on the situation. Additionally, the speed at which emotions change may vary depending on an individual's emotional highs and lows. Research has shown that individuals experience regular change in their emotions over time [11,12], suggesting that certain emotions can change into other emotions with some regularity, making it possible to predict emotional changes in individual users. Additionally, individuals experience a variety of positive, negative, or mixed emotions in their daily lives [13,14]; therefore, there is a need to express their emotions as complex emotions. Thus, the user's emotional state can be predicted from the current point in time based on the user's emotional history. Predicting the user's emotional state allows us to anticipate their satisfaction with recommended items, thereby enabling the development of a recommendation system that aligns closely with the user's preferences and needs.

Therefore, in this study, we focus on music, which can significantly affect the user's emotions among various types of content. We propose a continuous music recommendation method that considers emotional changes over time.

2.2. Emotion Model

Psychological research related to emotions has been widely applied by computer scientists and artificial intelligence researchers aiming to develop computational systems capable of expressing human emotions similarly [23]. Emotional models are primarily categorized into individual emotion, dimensional, and evaluation models. The dimensional theory of emotion uses the concept of dimensions rather than individual categories to represent the structure of emotions [24]. It argues that emotions are not composed of independent categories but exist along an axis with two-dimensional polarities: valence and arousal [25]. The valence axis represents the degree to which emotions are positive or negative, and the arousal axis represents the degree of emotional intensity. Valence and arousal are correlated elements that emotions can be effectively expressed emotions [22,26,27]. Two representative figures in the research on emotion models are Russell [28] and Thayer [29].

Russell [28] argued that emotion is a linguistic interpretation of a comprehensive expression involving various cognitive processes, suggesting that specific theories can explain what appears to be emotion. Figure 1 illustrates Russell’s emotional model. The horizontal axis represents valence, ranging from pleasant to unpleasant intensities, with increasing pleasantness towards the right. The vertical axis represents arousal, ranging from mild to intense, with increasing intensity upwards. According to Russell, internal sensibility suggests that individuals can experience specific sensibility through internal actions alone, independent of external factors. He argued that these experiences can be understood at a simpler level than commonly recognized emotions like happiness, fear, and anger. However, Russell’s emotion model has been criticized for its lack of compatibility with modern emotions due to overlapping meanings and ambiguous adjective expressions inherent in its adjective-based approach [1]. Thayer's model has been proposed to address these limitations and compensate for these shortcomings.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2025 doi:10.20944/preprints202503.1504.v1

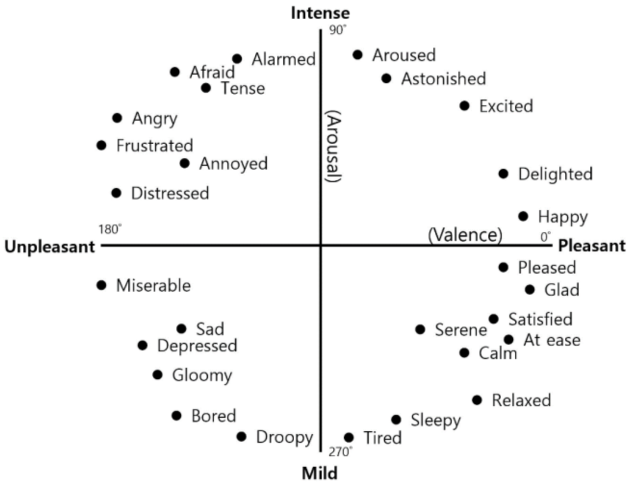


Figure 1. Russell's emotion model [28].

The emotion model presented by Thayer [29] is illustrated in Figure 2. This model compensates for the shortcomings of Russell's emotion model by providing a simplified and structured emotion distribution. It organizes the boundaries between emotions, making it easy to quantify and clearly express the location of emotions [30]. Thayer's emotion model employs 12 emotion words. In this model, the horizontal axis represents valence, indicating the intensity of positive and negative emotions at the extremes. The vertical axis represents arousal, reflecting the intensity of emotional response from low to high. Positivity increases towards the right on the horizontal axis, while emotional intensity increases upwards on the vertical axis. Each axis value is expressed as a real number between -1 and 1. Thayer's model is primarily used to express emotions in emotion-based studies.

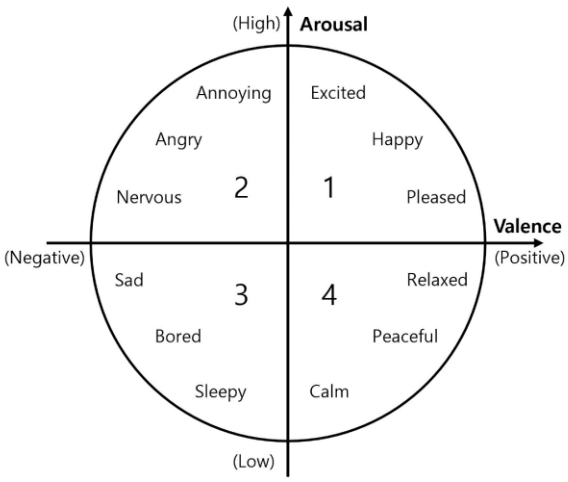


Figure 2. Thayer's emotion model [29].

Based on Thayer's emotion model, emotions are classified into quadrants based on ranges of valence and arousal values [23]. Additionally, a model has been proposed that classifies emotions into eight types, considering the characteristics of the content [30]. These studies interpret valence on the horizontal axis and arousal on the vertical axis similarly to Thayer's stablished emotion model.

Moon et al. [32] conducted an experiment where users evaluated their emotions toward music using Thayer's emotion model. To express emotions as a single point, a five-point scale was used for each emotional term, and user could select a maximum of three emotional terms. If multiple terms were selected, the total score could not exceed five points. Additionally, extreme emotional terms could not be evaluated simultaneously. Emotions can be felt in complex ways, and their intensity can vary among individuals; therefore, this evaluation method has limitations and is difficult to judge accurately.

In this study, we sought to generate music and user emotion vectors by dividing a quadrant into four emotion areas based on Thayer's emotion model. Music emotion vectors were created using valence and arousal values, and user emotion vectors were generated from user input on the level of emotional terms in each quadrant.

2.3. Emotion Analysis

Various studies have analyzed emotions in content. Laurier et al. [33]classified the music emotions using audio information and lyrics, selecting four levels of emotions: angry, happy, sad, and relaxed. Better performance in emotion analysis was shown when both audio and lyrics were used compared to lyrics alone. Hu et al. [34] also classified music emotions using audio and lyrics. They created emotional categories using tag information and conducted experiments comparing three methods: audio, lyrics, and audio plus lyrics. Performance varied by category. In categories such as happy and calm, the best performance was achieved when audio was used. On the other hand, in the romantic and angry categories, the best performance was achieved when lyrics were used.

In a previous study [20], focusing on visualizing music based on the emotional analysis of lyric text, the lyric text was linked to video visualization rules using Google's natural language processing API. The emotional analysis extracted score and magnitude values, which were mapped to valence and arousal to determine the location of each emotion. The score represents the emotion and is interpreted as positive or negative, and the magnitude represents its intensity. Based on this, six emotions (happy, calm, sad, angry, anticipated, and surprised) were classified into ten emotions according to their intensity levels. Consequently, we propose a rule for setting ranges based on values obtained from Google's natural language processing Emotion Analysis API.

In a previous study by Moon et al. [32], analyzing the relationship between mood and color according to music preference, multiple users were asked to select colors associated with mood and mood words for music. Emotional coordinates were calculated using Thayer's model based on the input values of the emotional intensity experienced by several users while listening to music. The sum of the intensity of each emotion for a specific piece of music among multiple users was calculated using Equation (1).

$$ed_i^s = \sum_{u=1}^{n} data_{u,i}^s \begin{cases} i = 1, 2, 3, \dots, 12 \\ n : number\ of\ volunteers \end{cases} \tag{1}$$

Here, ed_i^s represents the sum of emotions related to music, and $data_{u,i}^s$ denotes the level of emotions input by the user regarding the music. Using the results from Equation (1), the coordinate values for each emotion in the arousal and valence (AV) space were computed, as shown in Equation (2).

$$x_i = ed_i^s \times \cos(f\theta_i)\,, y_i = ed_i^s \times \sin(f\theta_i) \tag{2}$$

Here, x_i represents the valence value of emotion i in the AV space, y_i represents the arousal value of emotion in the AV space, and i represents the angle of emotion in the AV space. $f\theta_i$ is defined recursively as $f\theta_{i-1} + 30(2 \le i \le 12)$, and $f\theta_1$ is 15°. The average of the computed coordinate values for each emotion was then determined using Equation (3) to identify the representative emotion of the music.

$$\bar{x} = \frac{1}{12}\sum_{i=1}^{12} x_i\,, \bar{y} = \frac{1}{12}\sum_{i=1}^{12} y_i \tag{3}$$

In a previous study by Kim et al. [30], they designed a method to predict emotions by analyzing various elements of music. This method involved deriving an emotion formula that incorporates

 doi:10.20944/preprints202503.1504.v1

weights for five music elements: tempo, dynamics, amplitude change, brightness, and noise. These elements were mapped onto a two-dimensional emotion coordinate system. represented by X and Y coordinates. Next, the study involved plotting these X and Y values into on a two-dimensional plane and drawing a circle to determine the minimum and maximum radii that could encompass all eight emotions. The probability of each emotion was then calculated based on the proportion of the emotion's area within the circle. These probabilities were compared with emotions experienced by actual users through a survey. However, owing to conceptual differences in individual emotions, there are limitations in accurately quantifying emotions with this method.

When emotions are analyzed using music lyrics or audio information, there are limitations to evaluating the accuracy. Lyrics sometimes express emotions, but because they are generally lyrical and contain deep meanings, it is difficult to accurately predict emotions, and modern music, depending on the genre, has limitations in expressing emotions through audio information alone. Because emotions experienced when listening to music may vary based on individual emotional fluctuations and contexts, it is crucial to analyze how users experience emotions while listening to music.

Moon et al. [36] aimed to enhance music search performance by employing folksonomy-based mood tags and music AV tags. A folksonomy is a classification system that uses tags. Initially, they developed an AV prediction model for music based on Thayer's emotion model to predict the AV values and assign internal tags to the music. They stablished a mapping relationship between the constructed music folksonomy tags and the AV values of the music. This approach enables music search using tags, AV values, and their corresponding mapping information.

Cowen et al. [22] classified the emotional categories evoked by music into 13 emotions: joy (fun), irritation (displeasure), anxiety (worry), beauty, peace (relaxation), dreaminess, vitality, sensuality, rebellion (anger), joy, sadness (depression), fear, and victory (excitement). American and Chinese participants were recruited and asked to listen to various music genres, select the emotions they felt, and score their intensity. After analyzing the answers, the emotions experienced while listening to music were summarized into 13 categories. The study revealed a correlation between emotions and the elements of valence and arousal.

In a past study [37] on the emotions evoked by paintings, users were asked to indicate the ratio of positive and negative emotions they felt towards the artworks, which are artistic content in addition to music. Additionally, in a study [38] on emotions in movies, emotions were derived from the analysis of movie review data.

Various methods have been employed across various fields to analyze individual emotions. Psychology, dedicated to the study of emotions, actively conducts research on human emotions. Emotions are inferred through social factors, facial expressions, and voices recognized by humans [39-41]. Research has advanced from inferring emotions to predicting behavior based on other people's emotions and behavior [42,43]. This confirms that humans can predict emotions of others and that emotions can also predict future emotional states [26].

Thornton et al. [26] investigated the accuracy of mental models predicting emotional transitions. Participants rated the probability of transitioning between emotions based on 18 emotion words used in the study by Trampe et al. [13], assigning probabilities ranging from 0 to 100. They developed a mental model using average transition probabilities, depicted in Figure 3. Darker colors indicate higher probabilities of emotional transition, while lighter colors indicate lower probabilities.

To assess the accuracy of the mental model, they employed a Markov chain to measure how well the participants could predict real emotional transitions. The experiment utilized actual emotional history data, revealing significant prediction capability for emotions up to two levels deep. However, a limitation of the mental model presented in this study is its potential for errors due to the generalization of participants' individual emotions. Furthermore, because specific emotional transitions were evaluated using fragmented probabilities through one-to-one matching, it lacks consideration of the user's complete emotional history or temporal aspects. More accurate predictions could be achieved by predicting emotions based on an individual's personalized emotional transition

 doi:10.20944/preprints202503.1504.v1

probabilities. Additionally, brainwave data, facial expressions, and psychological tests were used to analyze user emotions.

In this study, we aimed to select music that is commonly employed in emotion analysis research, known to affect emotions. To analyze music emotions, we used music emotion data sourced from Cowen et al. [22]. Using a multiple regression model, we analyzed the relationship between valence and arousal within each emotional domain. To analyze user emotions, we utilized data as detailed by Trampe et al. [13], with considerations for the limitations identified in the mental model proposed by Thornton et al.[26].

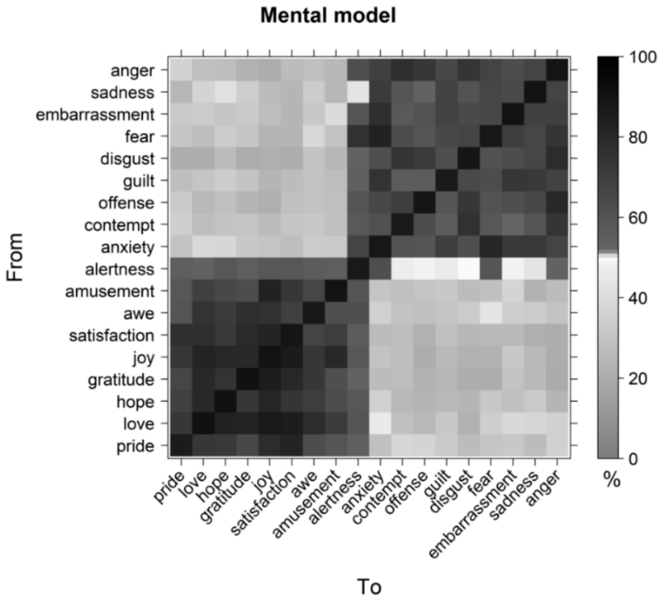


Figure 3. A mental model for emotional transitions [26].

3. Generation of Music Emotion Vectors

This section presents a method for generating emotion vectors through music data analysis for music recommendations. To recommend music that aligns with the user's emotions, we created an emotion vector based on the characteristics of the music. In this study, Thayer's emotion model was used to generate music emotion vectors.

Figure 4 illustrates the process of generating music emotion vectors.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2025 doi:10.20944/preprints202503.1504.v1

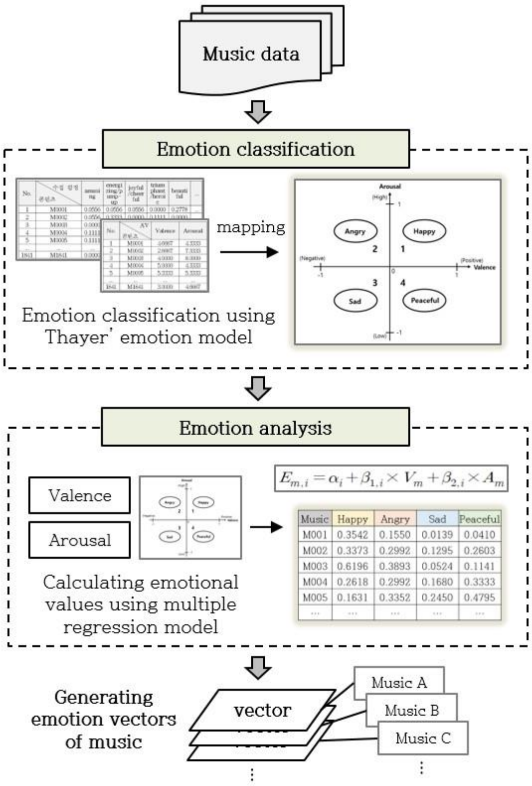


Figure 4. Process of generating music emotion vectors.

Emotions evoked by the music were analyzed using numerical values, valence, and arousal ratings provided by users. Following Thayer's emotion model, these emotions were mapped to represent each quadrant and classified based on the music's valence and arousal values. The relationships between valence, arousal, and emotions across the four emotional domains were examined using a multiple regression model. Multiple regression equations were used to calculate each emotional state. Emotion vectors for the music were then computed by applying multiple regression equations to calculate emotional values corresponding to the valence and arousal levels within each emotional area.

3.1. Music Data

This section describes the data analyzed in this study to generate music emotion vectors. Data provided by Cowen et al. [22] were utilized, including emotional, valence, and arousal values of the music. These numerical data integrate results from a survey on emotions experienced while listening to music by a large number of users. The emotional values for the music were classified into 13 emotions, and these values were normalized to range between 0 and 1. The valence and arousal values of the music ranged from 1 to 9.

For the emotional classification of the music, the valence and arousal values were normalized according to Thayer's emotion model. Valence and arousal values were normalized to range between -1 and 1 using Equation (4).

$$V_m = \left(\frac{Valance_m - \min(Valance)}{\max(Valance) - \min(Valance)}\right) \times 2 - 1,$$
$$A_m = \left(\frac{Arousal_m - \min(Arousal)}{\max(Arousal) - \min(Arousal)}\right) \times 2 - 1 \quad (4)$$

Here, V_m represents the normalized valence value of music m, A_m denotes the normalized arousal value of m, $Valence_m$ signifies the valence value of m, and $Arousal_m$ indicates the arousal value of m. We classified the emotional areas based on the normalized valence and arousal values of the music.

 doi:10.20944/preprints202503.1504.v1

Additionally, other forms of content, such as movies, art, and books, can also express emotions, valence, and arousal. Therefore, the values of emotion, valence, and arousal can be utilized depending on the type of collected content.

3.2. Emotion Classification of Music

In this section, we classify music emotions using Thayer's emotion model. According to Thayer's model, emotions in each quadrant were classified into Happy, Angry, Sad, and Peaceful, as illustrated in Figure 5. The representative emotions for each quadrant can be adjusted and defined based on the situation.

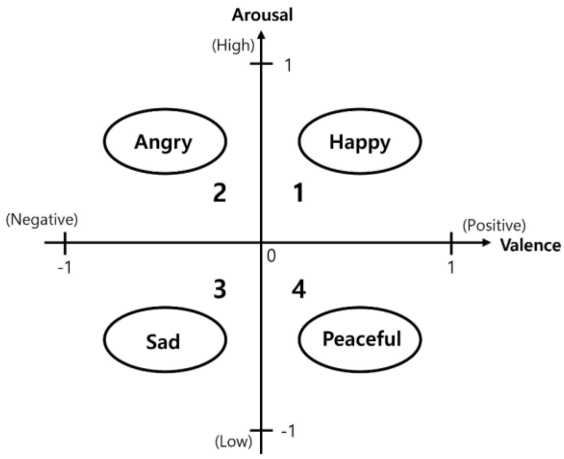


Figure 5. Classification of emotions into our types based on Thayer’s emotion model.

Based on the emotion model and the standard classification of emotion words [50], the 13 emotions (collected emotions) of music from the study by Cowen et al. [22] were categorized into four emotions. Table 1 presents the results of this emotion mapping.

Table 1. Results of emotion mapping.

Emotion	Collected emotions
Happy	Amusing, Beautiful, Energizing/pump-up, Joyful/cheerful, Triumphant/heroic
Angry	Annoying, Anxious/tense, Indignant/defiant, Scary/fearful
Sad	Sad/depressing
Peaceful	Calm/relaxing/serene, Dreamy, Erotic/desirous

Based on Table 1, music emotions were categorized into four primary emotions. Equation (5) illustrates how emotional value $E_{m,i}$ for each emotion i of music m is determined. Based on the calculated emotional value of music, the emotion with the highest value was classified as the representative emotion.

$$E_{m,i} = \frac{\sum_{k=1}^{N_i} degree_{i,k}}{N_i} \ (\mathrm{i} = 1, 2, \dots, \mathrm{n}) \tag{5}$$

Here $E_{m,i}$ represents the emotional value of music m for emotion i, $degree_{i,k}$ denotes the degree of the collected emotion k that belongs to emotion i, and N_i is the total number of collected emotions categorized under emotion i.

Based on Thayer's emotion model, music emotions are classified according to the range of valence and arousal values. Equation (6) was used for emotion classification based on the ranges of valence and arousal values. To analyze the relationship between valence, arousal, and each emotion, cases in which the valence or arousal value was zero were excluded.

 doi:10.20944/preprints202503.1504.v1

$$Emotion_m \begin{cases} Happy & (V_m > 0 \; and \; A_m > 0) \\ Angry & (V_m < 0 \; and \; A_m > 0) \\ Sad & (V_m < 0 \; and \; A_m < 0) \\ Peaceful & (V_m > 0 \; and \; A_m < 0) \\ N/A & (V_m = 0 \; and \; A_m = 0) \end{cases} \tag{6}$$

$Emotion_m$ is the emotion according to the valence and arousal values of the music m, V_m is the valence value of the music m, and A_m is the arousal value of the music m.

3.3. Emotion Analysis of Music and Vector Generation

In this section, we explore the relationship between the valence and arousal of music and each emotion, detailing the process of generating emotion vectors for music. We employed a multiple regression model to analyze the relationship between valence and arousal within the four emotional domains. Initially, we refined the data by comparing the representative emotions of the music, as classified in Section 3.2 with emotions categorized based on their valence and arousal values. Subsequently, for the multiple regression analysis, we further refined the data to include only those instances where emotions matched the valence, arousal, and representative categories within each emotional area.

Based on the refined data, we analyzed the relationship between valence, arousal, and each emotion using the Pearson correlation equation, shown in Equation (7). The correlation coefficient r measures the strength and direction of the liner relationship between two variables:

$$\mathrm{r} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{7}$$

Here, r represents the correlation coefficient, x and y denote the variables under analysis. As a result of the correlation analysis between valence, arousal, and each emotion, significant correlations were observed. Subsequently, we computed the emotional value of the music based on its valence and arousal values across the four emotional areas.

Multiple regression analysis was used to calculate the emotional value of music. During the regression analysis, valence and arousal were used as independent variables. Equation (8) creates a multiple-regression model using valence and arousal.

$$E_{m,i} = \alpha_i + \beta_{1,i} \times V_m + \beta_{2,i} \times A_m \tag{8}$$

In this equation, $E_{m,i}$ represents the emotion i value of music m, V_m denotes the valence value of music m, A_m denotes the arousal value of music m, α_i signifies the regression coefficient of emotion i, $\beta_{1,i}$ represents the regression coefficient of valence for emotion i, and $\beta_{2,i}$ represents the regression coefficient of arousal for emotion i.

Each emotional value of the music, calculated through multiple regression analysis, was represented as an emotion vector. Equation (9) defines the emotion vector for content.

$$\overrightarrow{V_m} = (E_{m,happy}, E_{m,angry}, E_{m,sad}, E_{m,peaceful}) \tag{9}$$

In this equation, $\overrightarrow{V_m}$ denotes the emotion vector for music m, where $E_{m,happy}, E_{m,angry}, E_{m,sad}$ and $E_{m,peaceful}$ represent the calculated emotional values for happiness, anger, sadness, and peacefulness, respectively.

4. Generating User Emotion Vectors Based on Emotional Changes

This section describes the process of generating emotion vectors based on changes in a user's emotions. User characteristics can be defined in various ways to meet specific needs, and similarities can only be identified when they align with previously defined music characteristics. Therefore, to continually recommend content that considers a user's emotional changes, an emotion vector was generated for each user's perspective based on the user's emotional information. The user emotion

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2025 doi:10.20944/preprints202503.1504.v1

vector is expressed by four emotion values derived from Thayer's emotion model, similar to how the music emotion vector is generated. Additionally, the user's emotion vector is generated by predicting future emotions based on changes in the user's emotions.

To predict user emotions continuously over time, we propose a probabilistic model based on a Markov chain. The Markov chain is a probability model where transitions between states depend on k previous states. Here, k refers to the number of states that can influence the determination of the next state, and "transition" denotes a change in state. Considering the characteristics of emotional changes, transitions between states depend on previous states. Therefore, the Markov chain process was utilized. Figure 6 shows a flowchart of generating user emotion vectors based on changes in user emotion.

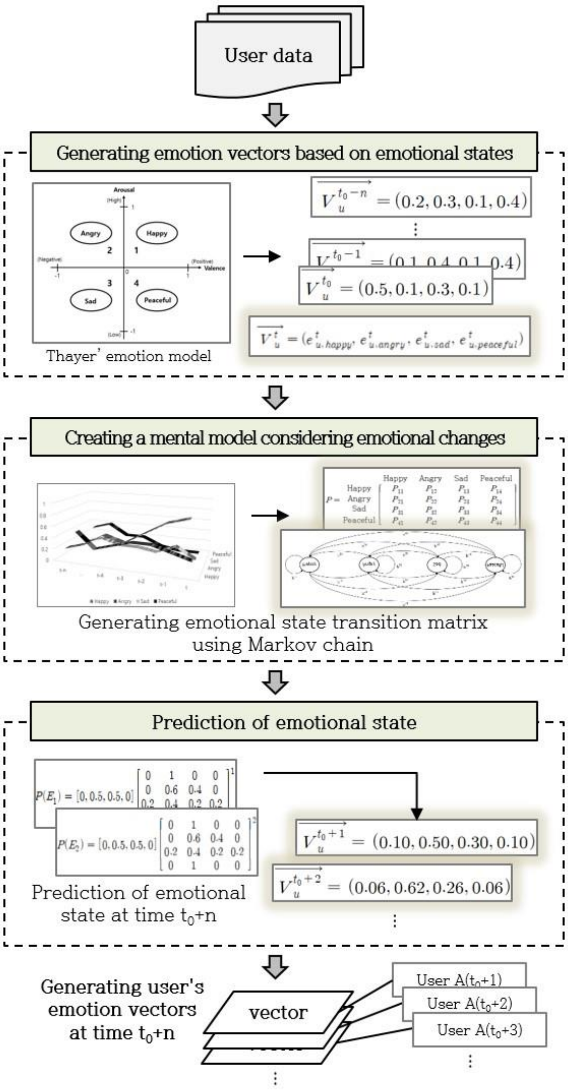


Figure 6. Flowchart for generating user emotion vectors based on changes in user emotion.

In this study, emotions were analyzed based on the user's emotional history. In the absence of user history, the initial emotional state and transition probabilities are input by the user. Using historical data, the emotional state was identified at regular intervals, and an emotion vector from the past to the present was generated. If there was no emotional data at given interval, an interpolation method was used to estimate the intermediate emotional state. If there were no user history data, the current emotion vector was generated by considering the input emotional state. Each time, emotional changes were identified through the generated emotion vectors, and a personalized mental model was created with an emotional state transition matrix using a Markov chain. If there

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2025 doi:10.20944/preprints202503.1504.v1

were no emotional history data, an emotional state transition matrix was created using the input emotional transition probability, defining it as a mental model. The user's emotional state at time t_0+n was predicted using the generated current-state emotion vector and emotional-state transition matrix, thereby generating emotion vectors for each user's time according to their emotional state.

4.1. User Data

This section describes the data analyzed to generate user emotion vectors. To analyze a user's emotions, various types of media can be utilized to record daily life, such as the emotional state value input by the user, the user's content viewing history, and social network services (SNS). Additionally, all data that can extract the user's emotions can be used, including wearable sensor data that records the user's voice, blood pressure, electrocardiogram, and brain wave data. Psychological testing methods, such as profile of mood states and the positive and negative affect schedule, can also be used to calculate emotion vectors. Emotional history data are required to identify changes in user emotions. These data represent the emotional state values at regular intervals over time.

In this study, based on Thayer's emotion model, user emotions were expressed as four emotion vectors. Table 2 shows an example of the user emotional history data over time. Based on this, changes in the user's emotions can be identified, and emotion vectors are generated at each time point. t_0 represents the current time.

Table 2. Example of user emotional history data over time.

Time	Happy	Angry	Sad	Peaceful
t_0	0	1	0	0
t_0-1	0	0.5	0.5	0
t_0-2	0	0.5	0.5	0
t_0-3	0.25	0.25	0.25	0.25
...	...	...	...	...
t_0-n	0	0	1	0

In the initial state, where there is no emotional history for the user, the emotional state is input by the user, and an initial value is assigned. In this study, the user's emotional state was used as input to calculate the user's emotion vector. The method for receiving the emotional state input involves obtaining the degree value for each emotion at each point in time based on Thayer's emotional model. To input the emotional state, information on each of the 12 emotions in each quadrant was presented, and each emotion was scored on a scale of 0 to 5. The scale of emotions varied depending on the situation. As users can experience multiple emotions simultaneously, they were asked to enter the degree of each emotion. Thus, the user's score for each emotion was determined, and the values of the four emotions were calculated based on the user’s emotional state.

In this study, we identify emotional changes based on a user's emotional history and create a personalized mental model. The personalized mental model is expressed as an emotional-state transition matrix using a Markov chain. Thus, we aim to predict user emotions. If a user has an emotional history, an emotional state transition matrix can be generated by calculating the emotional transition probability based on the emotional history data for each time at certain intervals. In the initial state, in which the user has no emotional history, the emotional transition probability is input by the user, and an emotional state transition matrix is generated. In this study, the probability of an emotional transition after a certain time was input to predict a user's emotion. To input the probability of emotion conversion, four emotions were used according to the emotion classification of the content, based on Thayer's emotion model. Figure 7 shows an example of the emotion conversion probability input. The probability of switching between emotions was set to a value between 0 and 100.

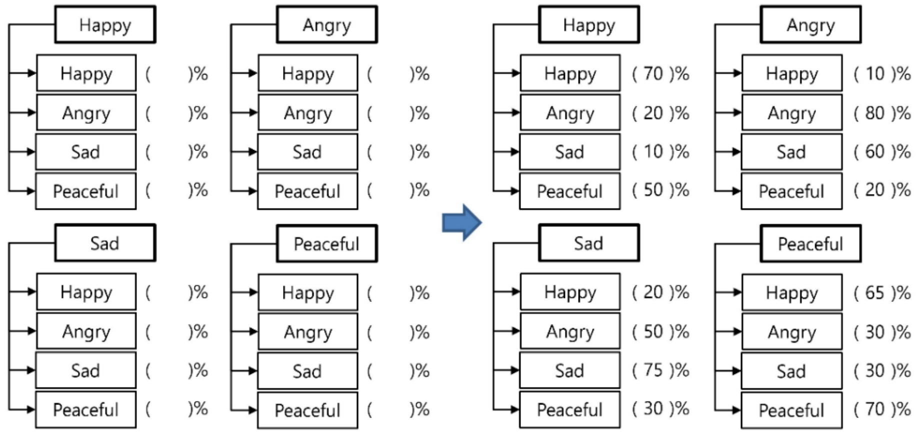


Figure 7. Examples of user emotional transition probability input.

4.2. Generation of Emotion Vectors for Each Time Point Considering Emotional State

In this section, the emotion vectors for each time were generated using the user's emotional history data or emotional state input. Based on Thayer's emotion model, four representative emotion values for each quadrant were calculated using the input user emotion level values. The emotions in each quadrant were classified based on Thayer's emotion model (Figure 5), as shown in Table 3.

Table 2. Results of emotion mapping.

Quadrant	Emotion	Thayer's Emotions
1	Happy	Pleased, Happy, Excited
2	Angry	Annoyed, Angry, Nervous
3	Sad	Sad, Bored, Sleepy
4	Peaceful	Calm, Peaceful, Relaxed

The value of each emotion was calculated using Equation (10). The ratio of the emotional degree values corresponding to each emotional quadrant was calculated. Values were calculated at regular intervals and stored as historical data.

$$E_{u,i} = \frac{\sum_{k=1}^{N_i} degree_{i,k}}{\sum_{k=1}^{N} degree_k} \ (\mathrm{i} = 1, 2, \dots, \mathrm{n}) \quad (10)$$

where $E_{u,i}$ is the emotion i value of the user u, $degree_{i,k}$ is the degree of Thayer's emotion axis k included in emotion i, $degree_k$ is the degree of Thayer's emotion axis k, and i is the degree of Thayer's emotion axis k. The index i, k refers to Thayer emotions included in emotion i, where N_i represents the number of Thayer emotions included in emotion i, and N represents the total number of Thayer emotions.

If there is no emotional history at a certain interval, the emotional state at an intermediate point in time without a history is estimated using polynomial interpolation to generate an emotion vector. The emotional state value at an intermediate point in time was calculated using polynomial interpolation based on existing emotional history data. Polynomial interpolation involves fitting a polynomial that passes through n points and can be expressed as a polynomial of order n-1. Common methods for finding such polynomials include the method of indeterminate coefficients and Newton's interpolation method, which mitigates the disadvantage of lengthy computation. Equation (11) represents Newton's interpolation formula used to calculate the emotional state value $E_i(x)$ of emotion i at a specific time x, based on emotional state values at n time points when there is no history.

$$E_i(x) = a_1 + a_2(x - x_1) + a_3(x - x_1)(x - x_2) + \cdots + a_n(x - x_1)(x - x_2) \dots (x - x_{n-1}) \quad (11)$$

 doi:10.20944/preprints202503.1504.v1

The coefficients are determined using an nth-order finite-difference approximation method. Consequently, each user's emotional value was represented as an emotion vector. Equation (12) represents the emotion vector from each user perspective.

$$\overrightarrow{V_u^t} = (E_{u,happy}^t, E_{u,angry}^t, E_{u,sad}^t, E_{u,peaceful}^t) \tag{12}$$

4.3. Generating a Personalized Mental Model Considering Emotional Changes

This section describes the method used to develop a personalized mental model that considers changes in user emotions. The personalized mental model derived by constructing an emotional-state transition matrix using a Markov chain process, which is a probabilistic model. A Markov chain process is a mathematical technique for modeling systems that identifies their dynamic characteristics based on past changes and predicts future changes. In this study, we aim to predict the emotional state at time t_0+n by employing a Markov chain process to analyze a user's emotional changes. The emotional state transition matrix is generated using either the user's visual emotion vector or input values for emotional transition probabilities.

If emotional history is available, matrix elements can be calculated by analyzing patterns of emotional change. The probability of emotional transitions is determined using emotion vectors collected at regular intervals over time. However, if the emotion value is zero, the sum of the P_{ij} values is zero. Equation (13) calculates the transition probability based on emotion vectors. Thus, an emotional-state transition matrix was generated for each time point.

$$tp_{ij} = \sum_{n=k}^{t_0} (E_{u,i}^{t_0-n} \times E_{u,i}^{t_0-n+1}) \tag{13}$$

where tp_{ij} represents the probability of switching between emotions, t_0 denotes the current time, k signifies the time difference between the current time t_0 and the start of historical data, $E_{u,i}$ is the value of emotion i of user u, and i and j are the emotion indices.

An emotional-state transition matrix was generated by calculating the probability of emotional change. Equation (14) calculates the element P_{ij} of the emotional state transition matrix using the emotional transition probabilities. If no emotional history is available, the value entered in Figure 6 was used.

$$p_{ij} = \frac{tp_{ij}}{\sum_{j=1}^{n} tp_{ij}} \ (i, j = 1, 2, \dots, n) \tag{14}$$

where P_{ij} represents the element value of the emotional state transition matrix, tp_{ij} represents the transition probability between emotions, and i and j are indices representing different emotions. The user's emotional state transition matrix for the four emotions classified in this study is expressed by Equation (15).

$$P_{ij} = \begin{array}{r} \\ Happy \\ Angry \\ Sad \\ Peaceful \end{array} \begin{array}{c} \begin{array}{cccc} Happy & Angry & Sad & Peaceful \end{array} \\ \begin{bmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \\ P_{41} & P_{42} & P_{43} & P_{44} \end{bmatrix} \end{array} \tag{15}$$

Figure 8 shows a state diagram created using Equation 15. The result is expressed as the user's personalized mental model.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2025 doi:10.20944/preprints202503.1504.v1

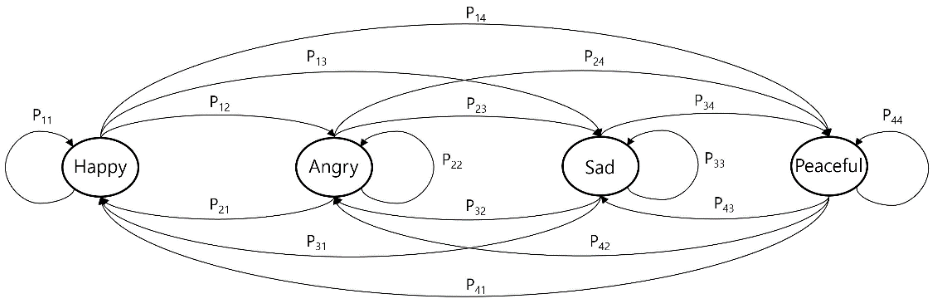


Figure 8. User's personalized mental model.

4.4. Emotion State Prediction and Vector Generation

In this section, the emotional state at time t_0+n is predicted using the user's current emotional state and emotional state transition matrix, and an emotion vector is generated. Equation (16) predicts the emotional state of user u at time t_0+n.

$$P_u(E_n) = V_u^{t_0} \times (P_u)^n \tag{16}$$

Here, $P_u(E_n)$ is the emotional state of user u at time t_0+n, $V_u^{t_0}$ is the emotional state of user u at the current time t_0, and P_u is the emotional state transition matrix of user u. Based on the user's emotional history data, if the regular interval time is 1 h, the emotional state at time t_0+n is predicted at 1-h intervals. Additionally, in the initial state, if the time standard for the emotional transition probability is 30 min, the emotional state at time t_0+n is predicted at 30-min intervals. By predicting the emotional state of the user, emotion vectors were generated at each time point. Equation (17) generates an emotion vector for each user perspective.

$$\overrightarrow{V_u^t} = (E_{u,happy}, E_{u,angry}, E_{u,sad}, E_{u,peaceful}) \tag{17}$$

5. Continuous Recommendation of Music Reflecting Emotional Changes

This section describes a continuous music recommendation method that reflects changes in user emotions. Figure 9 shows a flow chart of continuous music recommendation.

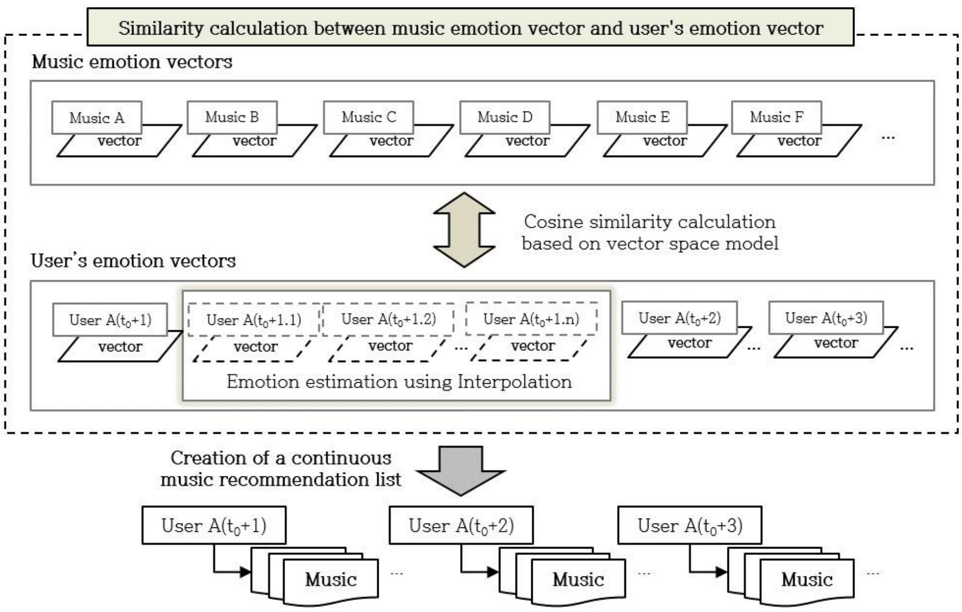


Figure 9. Flow of continuous music recommendations.

 doi:10.20944/preprints202503.1504.v1

A recommendation list is created by calculating the similarity between the emotion vector of the music and the emotion vector from each user's perspective. The similarity between the two vectors was calculated using a cosine similarity calculation formula based on the vector space model. In addition, to continuously recommend music, an interpolation method was used to estimate the emotional value between emotion vectors each time and generate a vector. A recommendation list is continuously generated according to the emotion vector created by the user.

5.1. Calculate the Similarity Between Music Emotion Vector and User Emotion Vector

This section describes the method for calculating the similarity between music and user emotion vectors. By calculating the similarity between the user emotion vector and the music emotion vector each time, a content recommendation list was created. Various similarity calculation methods can be used to calculate the similarity between two vectors. In this study, we used the cosine similarity calculation formula based on the vector space model, which calculates the similarity between two vectors. The similarity value is between 0 and 1; the closer the similarity value is to 1, the more similar it is judged to be between the user's emotional state and the music's emotions. To calculate the similarity between the music emotion vector and the user emotion vector each time, Equation (18) is used.

$$\cos(\overrightarrow{V_m}, \overrightarrow{V_u^t}) = \frac{\sum_{i=1}^{N}(V_{m,i} \times V_{u,i}^t)}{\sqrt{\sum_{i=1}^{N}(V_{m,i})^2} \times \sqrt{\sum_{i=1}^{N}(V_{u,i}^t)^2}} \tag{18}$$

$\overrightarrow{V_m}$ is the emotion vector of music m, $\overrightarrow{V_u^t}$ is the emotion vector of the user u at each time, $V_{m,i}$ is the value of the emotion vector of music m, $V_{u,i}^t$ represents the value of the emotion vector of the user u at each time, and N represents the number of emotion vector elements.

5.2. Emotion Estimation Between Visual Emotion Vectors

To continuously recommend music, the emotion between the emotion vectors for each user's perspective was estimated, and an emotion vector was created. If emotion vectors are generated at 1-h intervals and music is recommended, the emotional state between time t_0+1 and time t_0+2 is estimated, and a recommendation list is continuously generated. For this purpose, a linear interpolation formula, which is generally used to estimate the value between two points, was employed. Equation (19) is a linear interpolation formula that calculates the state value of each emotion at a specific interval between two times.

$$E_i(x) = d_2 \times E_i(x_1) + d_1 \times E_i(x_2) \tag{19}$$

where $E_i(x)$ is the state value of emotion i at a specific time x, $E_i(x_1)$ and $E_i(x_2)$ are the state values of emotion i at two times (x_1, x_2), and d_1 and d_2 are the time ratios from time x to times x_1 and x_2, respectively. The sum of d_1 and d_2 is 1: Thus, each user's emotional value was created as an emotion vector. Equation (20) represents the emotion vector for each specific time of the day for the user.

$$\overrightarrow{V_u^t} = (E_{u,happy}^t, E_{u,angry}^t, E_{u,sad}^t, E_{u,peaceful}^t) \tag{20}$$

Based on the emotion vector generated each time, the similarity with the music vector was calculated using Equation (18).

5.3. Generating a Music Continuous Recommendation List

This section describes a method for continuously generating a music-recommendation list based on changes in user emotions. The similarity between the user emotion vector and the music emotion vector at each time point was calculated, and a music list was created in the order of highest similarity. A recommendation list was created according to the emotional state at time t_0+n, and a recommendation list was created for a limited time. Equation (21) generates a continuous recommendation list based on the emotional state at each time point.

 doi:10.20944/preprints202503.1504.v1

$$Recommendation\ list_m = \{(m_1, m_2, m_3, \dots, m_{N_r})\} = \{(m_k)\}_k^{N_r} (\sum_{k=1}^{N_r} time_k \leq \mathrm{T}) \tag{21}$$

Recommendation $list_m$ is a music recommendation list according to emotional state at each time, m is music, k is the music index, N_r is the number of music recommendations, *time* is music playback time, and T is the time limit. N_r is the number of recommended music lists according to the emotional state at each time and is the number of pieces of music that can be played during a limited time. If the time limit is set to T, a recommendation list is created at each interval of T based on the user's emotional state.

6. Experiments and Evaluation

6.1. Experimental Method

To prove the effectiveness of the proposed continuous music recommendation method that reflects emotional changes, we conducted emotional analysis of music, user emotion prediction considering emotional changes, and evaluation of music recommendation results according to user emotions. The data analysis programming language R was used to evaluate the experiments.

The results of the emotional analysis of the music were evaluated based on the accuracy of the calculation of the emotional value of the music. The user-emotion prediction results considering emotional changes were evaluated based on the accuracy of the predicted emotions by analyzing the emotional history of the user. The recommendation results based on the user emotions were assessed based on the accuracy of recommending music that matched the user's emotional state.

Evaluation metrics used to determine the accuracy of data analysis and prediction results include the mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and normalized root mean squared error (NRMSE). These metrics measure the difference between actual and predicted values, with lower values indicating better model performance.

Precision, recall, F1-score, balanced accuracy, and similarity were measured to evaluate the recommendation results based on user emotions. A value closer to 1 signifies higher performance, and the results are often expressed as a percentages, with values closer to 100 indicating better performance. To evaluate the recommendation results, we selected the top n items from the list and conducted an experiment. Accuracy measures the proportion of correct predictions where the actual and predicted values align across all data. Balanced accuracy is a method used to mitigate exaggerated performance estimates in imbalanced datasets. In this study, balanced accuracy was calculated to account for the experimental data characteristics. Similarity assesses how closely recommended music matches the user's actual emotions. It quantifies the resemblance between the user's actual emotion vector and that of the recommended music.

6.2. Experimental Data

In this section, we describe the data used to demonstrate the effectiveness of the proposed method. The music data used in this study included emotional values, valence, and arousal data sourced from Cowen et al.[22]. The emotional numerical data of the music were classified into 13 types of emotions and stored as values ranging between 0 and 1. The valence and arousal data for the music ranged from one to nine.

Based on the collected data, the procedures outlined in Sections 3.1 and 3.2 were conducted. Table 4 displays a subset of the refined data used in the experiments. The data were categorized by emotional area in the following order: Happy, Peaceful, Angry, and Sad.

Table 4. Part of the preprocessed data for emotion analysis of music.

No	Emotion	Valance	Arousal	Happy	Angry	Sad	Peaceful
1	Angry	-0.5833	0.5833	0.3396	0.6604	0.0000	0.0000

 doi:10.20944/preprints202503.1504.v1

2	Happy	0.0833	0.0833	0.4504	0.1206	0.0000	0.4290
3	Happy	0.5000	0.4167	0.5038	0.1145	0.0000	0.3817
4	Peaceful	0.0833	-0.0833	0.0678	0.0282	0.4520	0.4520
5	Peaceful	0.2500	-0.4167	0.1212	0.0303	0.2424	0.6061
…	...	...	...	...	...	...	...
570	Peaceful	0.7500	-0.2500	0.1109	0.1039	0.2771	0.5081

The user data used in this study were emotional history data sourced from Trampe et al. [13]. These data represent a survey of the emotions experienced by 12,212 users on specific days and times of the week. The user's emotional history data were classified into 18 emotions and stored as values of 1 or 0 by checking the emotions felt by the user. ID is the number of users, Day is a value between 1 and 7, where 1 starts from Monday and 7 represents Sunday, and Hours is a value based on 24 h.

For experimental data, the 18 collected emotions were mapped onto the four emotions used in this study. Table 5 presents the results of mapping these 18 emotions to four emotions based on the standard word classification for emotion [50].

Table 5. Results of emotion mapping.

Emotion	Collected emotions
Happy	Pride, Love, Gratitude, Joy, Awe, Amusement, Alertness
Angry	Anxiety, Disdain, Offense, Guilt, Disgust, Fear, Embarrassment, Anger
Sad	Sadness
Peaceful	Hope, Satisfaction

Equation (22) was used to calculate the user's emotional level based on the classification results into the four primary emotions.

$$E_{u,i} = \frac{\sum_{k=1}^{N_i} select_{i,k}}{\sum_{k=1}^{N} select_k} \ (\mathrm{i} = 1, 2, \dots, \mathrm{n}) \tag{22}$$

where $E_{u,i}$ is the emotion i value of the user u, $select_{i,k}$ indicates whether the collected emotion axis k included in emotion i is selected, $select_k$ is whether the collected emotion axis k is selected, i is the index of emotion i, k is the index of collected emotions included in emotion i, N_i represents the number of collected emotions included in emotion i, and N represents the total number of emotions.

Table 6 shows an example of the numerical data for each user emotion, classified into four categories. To analyze the user's emotions by time of day, data points without selected emotions at the given time were excluded. Additionally, user data with a history of regular intervals of days and times were selected for the experiment. Although there were limitations in selecting data from users who had emotional histories recorded at equal intervals, the experiment was conducted by setting the selection criteria according to the data distribution.

Table 6. Example of the data for each user's emotional value.

ID	Day	Time	Happy	Angry	Sad	Peaceful
1	1	19	0.0000	0.5000	0.5000	0.0000
1	2	14	0.0000	1.0000	0.0000	0.0000
1	3	15	0.0000	1.0000	0.0000	0.0000
...	...	...	...	...	...	...
64	6	12	0.6667	0.0000	0.0000	0.3333
64	6	16	0.5000	0.0000	0.0000	0.5000
64	6	20	0.5000	0.0000	0.0000	0.5000

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2025 doi:10.20944/preprints202503.1504.v1

...	...	...	...	...	...	...
1,925	1	9	0.6667	0.0000	0.0000	0.3333
1,925	2	13	0.6000	0.0000	0.0000	0.4000
1,925	3	15	0.3333	0.3333	0.0000	0.3333
...	...	...	...	...	...	...
12,212	1	10	0.7500	0.0000	0.0000	0.2500
12,212	1	19	0.5000	0.0000	0.0000	0.5000
12,212	1	22	1.0000	0.0000	0.0000	0.0000

6.3. Experimental Results and Evaluation

6.3.1. Evaluation of Music Emotion Analysis Results

To classify emotions in music, we first analyzed the correlation between valence and arousal in each emotional area and emotion and confirmed that the absolute value of the correlation coefficient $|r|$ (Equation 7) was higher than 0.3, indicating a moderate correlation between the variables. A total of 570 songs were used, and the emotional value of the music was calculated based on the valence and arousal values for each emotional area using a multiple regression model. Table 7 presents the regression coefficients for each emotional area used as a factor in the multiple regression equations.

Table 7. Regression coefficient by emotion.

Emotion	*Quadrant*	α_i	$\beta_{1,i}$	$\beta_{2,i}$
Happy	*1*	*0.22368*	*0.10001*	*0.05696*
	2	*0.53479*	*0.10941*	*0.00000*
	3	*0.12853*	*0.00000*	*-0.10995*
	4	*0.11300*	*0.00000*	*-0.05642*
Angry	*1*	*0.62323*	*-0.01513*	*0.13190*
	2	*0.22235*	*0.05899*	*0.00000*
	3	*0.04441*	*0.00000*	*-0.06108*
	4	*0.11002*	*0.00000*	*-0.12981*
Sad	*1*	*0.09973*	*0.01176*	*-0.00089*
	2	*0.09031*	*0.12525*	*0.00000*
	3	*0.63701*	*0.00000*	*-0.09162*
	4	*0.17295*	*0.00000*	*-0.03274*
Peaceful	*1*	*0.17360*	*-0.05470*	*-0.02478*
	2	*0.07192*	*0.01919*	*0.00000*
	3	*0.15074*	*0.00000*	*-0.00715*
	4	*0.60374*	*0.00000*	*0.01274*

Table 8 presents selected results from the calculation of the emotional values of music. These results were obtained using Equation (8) employing the regression coefficients specific to each emotional area detailed in Table 7, and based on the experimental data.

Table 8. Example of the results from the calculation of the emotional values of music.

No	Music	Happy	Angry	Sad	Peaceful
1	M001	0.1986	0.4710	0.0644	0.0801
2	M002	0.6330	0.2273	0.0393	0.0992
3	M003	0.6706	0.2518	0.0190	0.0559

 doi:10.20944/preprints202503.1504.v1

4	M004	0.1711	0.0735	0.1513	0.6027
5	M005	0.1702	0.0767	0.1537	0.5984
6	M006	0.2010	0.4892	0.0919	0.0942
7	M007	0.6549	0.2273	0.0291	0.0776
8	M008	0.2010	0.4892	0.0919	0.0942
9	M009	0.7108	0.2666	0.0014	0.0127
10	M010	0.1727	0.0751	0.1531	0.5995
…	…	…	…	…	…
570	M570	0.1388	0.0863	0.1525	0.6006

In this study, the emotional values of music were determined using multiple regression analysis. The accuracy of these results was evaluated using the MSE, RMSE, and MAE. Figure 10 illustrates the average errors in the calculation results for each emotion.

Experimental evaluation indicated that the MSE, RMSE, and MAE values were small. The average values across all emotions were 0.0146, 0.1190, and 0.0785, respectively. Additionally, NRMSE was computed to further assess the accuracy of the results. Figure 11 depicts the error rates in the calculation results for each emotion. On average, the error rate across all emotions was 12.74%, indicating an accuracy of 87.26%. The error rate represents the difference between the emotional value of the music as evaluated by users and the calculated values. Increasing the number of user evaluations or comparing them with exact values can improve the accuracy of these results.

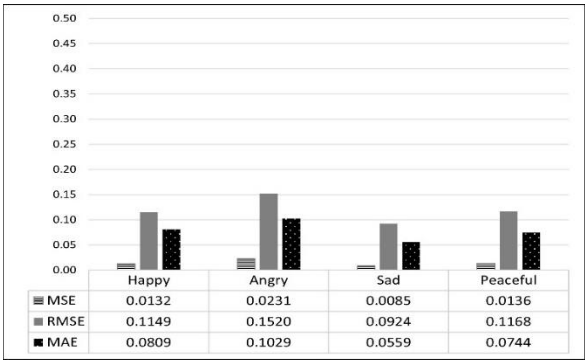


Figure 10. Average error of the calculation results for each emotion.

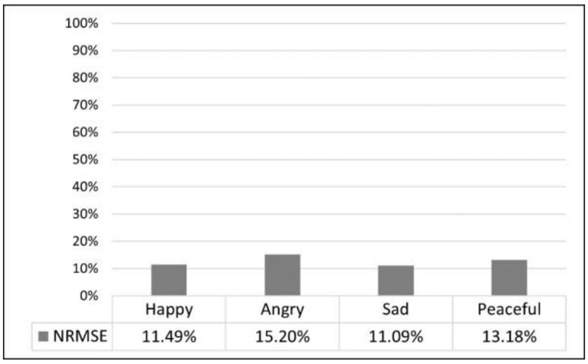


Figure 11. Error rate of calculation results for each emotion.

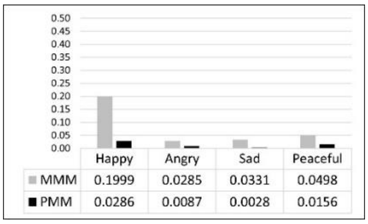
6.3.2. Evaluation of user emotion prediction results considering emotional changes

To demonstrate the effectiveness of the proposed method, a comparative experiment was conducted between the emotion prediction results using the mental model proposed in a previously reported study [26] and the emotion prediction results by applying the PMM proposed in this paper.

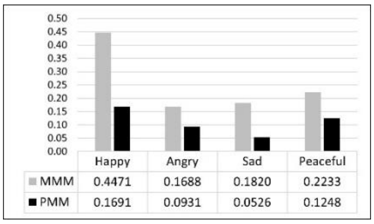
Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2025 doi:10.20944/preprints202503.1504.v1

In this comparison, the mental model from the previous study was adapted by mapping its emotions to four categories. This modified mental model is referred to as the MMM.

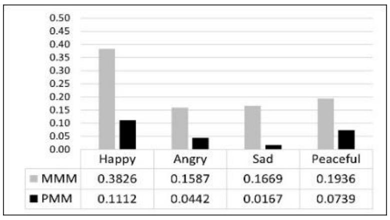
An experiment was conducted using the emotional history data from 50 users. Users were selected based on having at least six emotional experiences, each separated by at least one hour. A mental model was developed using historical data to predict emotions at times t_0+1 and t_0+2. For comparison with actual emotional data, prediction were made based on emotional history data. MSE, RMSE, and MAE were calculated to assess prediction accuracy. Figures 12 and 13 display graphs comparing the average prediction errors for each emotion at times t_0+1 and t_0+2 using PMM and MMM.



(a) MSE

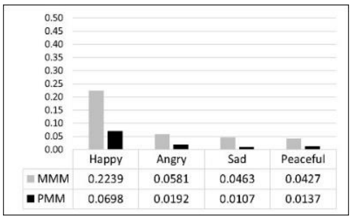


(b) RMSE

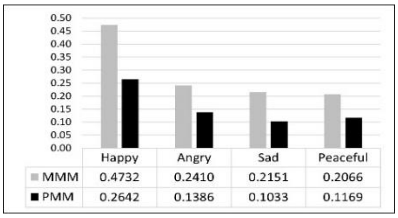


(c) MAE

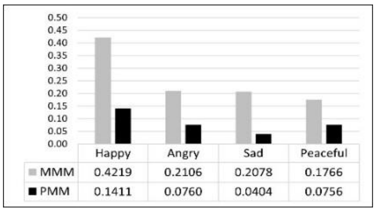
Figure 12. Comparison of average prediction errors for each emotion at time t_0+1.



(a) MSE



(b) RMSE



(c) MAE

Figure 13. Comparison of average prediction errors for each emotion at time t_0+2.

Comparing the results with the actual emotional data confirmed that the prediction error using the PMM was smaller than that using the MMM. Additionally, the average MSE, RMSE, and MAE of the total error values were calculated. Figures 14 and 15 present the results comparing the overall average prediction errors of emotions at times t_0+1 and t_0+2 using the MMM and PMM.

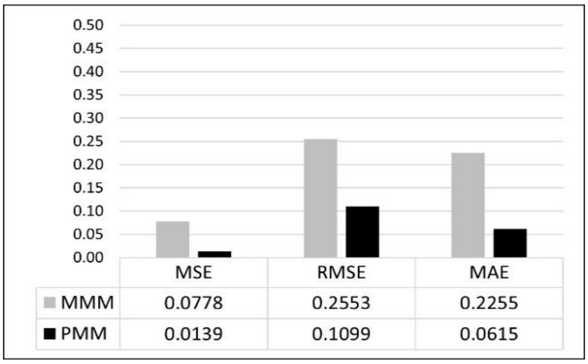


Figure 14. Comparison of the overall average prediction errors at time t_0+1.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2025 doi:10.20944/preprints202503.1504.v1

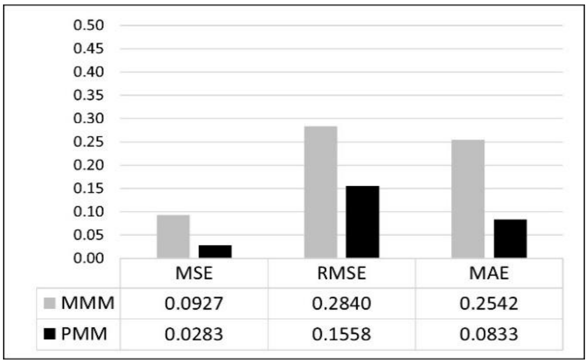


Figure 15. Comparison of the overall average prediction errors at time t_0+2.

Overall, it was confirmed that the error using the PMM was smaller than that using the MMM, indicating higher accuracy in the predicted values. Additionally, NRMSE was calculated to further evaluate the accuracy of the predictions. Figures 16 and 17 compare the average error rates of predicted values for each emotion at times t_0+*1* and t_0+2.

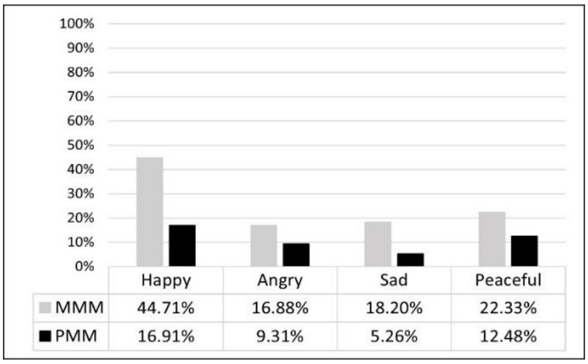


Figure 16. Comparison of average error rates of predicted values for each emotion at time t_0+*1*.

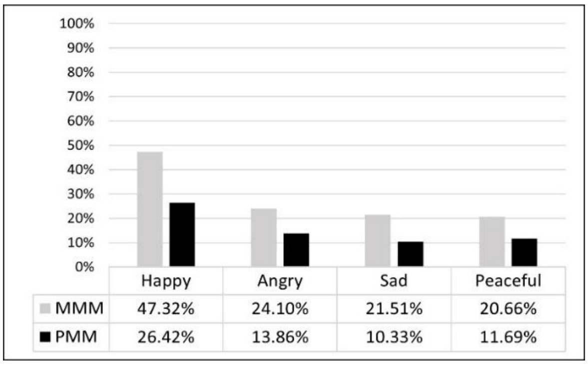


Figure 17. Comparison of average error rates of predicted values for each emotion at time t_0+2.

Figure 18 presents the comparison of overall average error rates. For predicting emotions at time t_0+*1*, the error rates were 25.53% for MMM and 10.99% for PMM. For time t_0+2, the error rates were 28.40% for MMM and 15.58% for PMM, indicating higher error rates for MMM compared to PMM. These results show a 13.68% improvement in error rate for PMM over MMM when compared with actual data, demonstrating increased accuracy.

 doi:10.20944/preprints202503.1504.v1

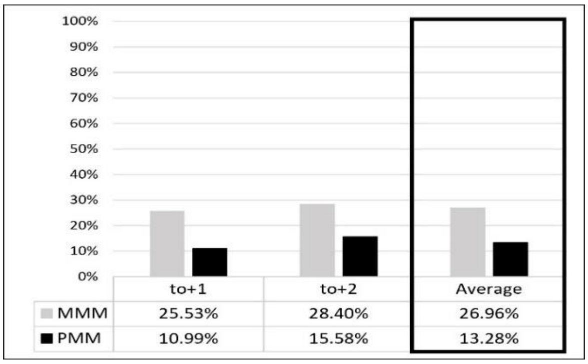


Figure 18. Comparison of overall average error rates of emotion prediction.

6.3.3. Evaluation of Recommendation Results Based on User Emotions

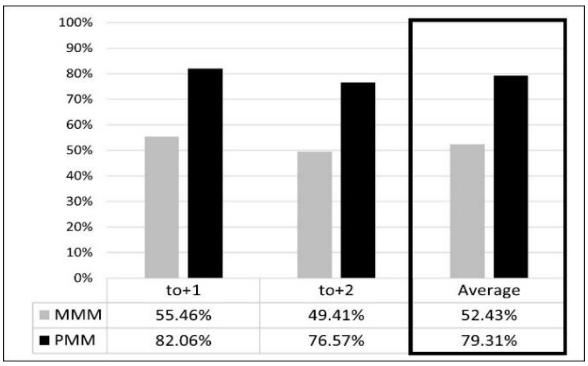
To assess the recommendation results based on user emotions, a recommendation list was created following the procedures outlined in Section 6.3.2. The similarity between the music emotion vector and user emotion vector was calculated for each time of the day. Considering that the emotional history data were collected at 1-h intervals, a set of 10 songs was recommended by sorting them in descending order of similarity. For experimental evaluation, the similarities between the actual emotion vector, emotion vector using MMM, emotion vector using PMM, and music vector were computed. Precision, recall, F1-score, and balanced accuracy were measured to evaluate the recommendation results based on user emotions. The recommendation list based on the actual emotion vector was compared with that based on the predicted emotion vector to calculate these metrics. Table 9 presents the precision, recall, and F1-score as evaluation results for the recommendation outcomes.

Table 9. Comparison of average precision, recall, and F1-score for the recommendation results.

Method	t_0+1			t_0+2		
	Precision	**Recall**	**F1-score**	**Precision**	**Recall**	**F1-score**
MMM	0.2777	0.2777	0.2777	0.2733	0.2733	0.2733
PMM	0.6722	0.6766	0.6744	0.6255	0.6311	0.6288

As observed from the comparison with actual emotional data, PMM demonstrates higher precision, recall, and F1-score values compared to MMM, indicating its effectiveness in recommending music suitable for users.

Figure 19 illustrates the overall average balanced accuracy of the recommendation results using MMM and PMM. The top ten recommendation lists were extracted and compared based on the actual emotion vector, emotion vector using MMM, and emotion vector using PMM. The comparison shows that PMM achieves a higher balance accuracy of 26.88% compared to MMM.



Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 March 2025 doi:10.20944/preprints202503.1504.v1

Figure 19. Comparison of overall average balanced accuracy for recommendation results.

Additionally, to assess the accuracy of the recommendation list, the similarity between the recommended music vector and the actual value was calculated. Figure 20 depicts a graph comparing the average similarity of the recommendation results obtained using PMM and MMM. The comparison confirms that PMM can recommend music better suited to the user's actual emotions compared to MMM.

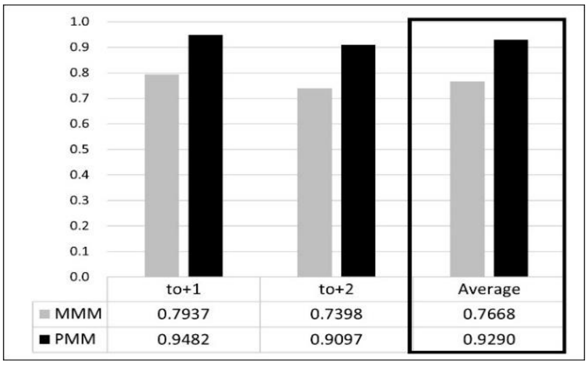


Figure 20. Comparison of average similarity of recommendation results.

7. Conclusion and Future Research

With ongoing advancements in emotion-based recommendation research, services that recommend various types of content to users are being developed. In such emotion-based recommendation system, predicting the user's emotional state can enhance satisfaction with the recommended content. However, current research on content recommendation often focus solely on the user's current emotions without considering changes in individual emotion states.

In this study, we propose a method for continuous music-recommendation that reflects emotional changes. First, we construct emotion vectors for music using valence and arousal values, which are elements of Thayer's emotional model, which were categorized in four emotional areas. Using a multiple regression model, we explored the relationships between valence, arousal, and these four emotions to generate music emotion vectors. Second, we developed a user emotion vector based on emotional changes using the user's historical emotional data. These vectors were updated over time, and an individual's emotional-state transition matrix was created using a Markov chain process. A PMM was defined, and the user's emotion vector was generated by predicting the emotional state at time t_0+n. Third, we implemented a method to continuously recommend content reflecting emotional changes. To create a recommendation list, the similarity between the content emotion vector and the user emotion vector was calculated using cosine similarity. In addition, we used an interpolation method to estimate the emotional state between times t_0+n, generating a recommendation list according to the estimated emotional state.

To validate the effectiveness of the proposed method, we assessed the results of music emotion analysis, user emotion prediction considering emotional changes, and music recommendation based on user emotions. The music emotion analysis indicated an average error rate across all emotions of 12.74%, achieving an accuracy of 87.26%. To enhance the accuracy of our evaluations, it would be beneficial to utilize more precise and reliable evaluation metrics. In a comparative experiment, we evaluated user emotion prediction using an MMM and our PMM, which adapts the mental model from existing research. Predicting emotions at times t_0+1 and t_0+2 using the MMM and PMM revealed that PMM improved the error rate by 13.68% over MMM, increasing accuracy. Additionally, in the evaluation of music recommendation experiments, PMM demonstrated a higher balanced accuracy of 26.88% compared to MMM. The similarity between the recommended music vector and the actual emotion vector was higher with PMM. These findings confirm our ability to predict user emotions accurately and recommend music that aligns closely with user preferences.

It is anticipated that satisfaction with the music provided can be enhanced by continuously recommending music according to the user's emotions and situations. Beyond music, this approach is expected to extend its utility to aiding in product recommendations and psychological treatments by reflecting user's emotional changes.

To enhance the accuracy of emotion prediction in the future, it is crucial to research continuous recommendation methods that detect changes in users' emotions based on recommendation outcomes and incorporate feedback derived from these recommendations. Furthermore, research is needed to address the limitations of experimental data and extend recommendations beyond music to include other content types such as movies and books.

Author Contributions: Conceptualization, SI.B. and YK.L.; methodology, SI.B. and YK.L.; software, SI.B.; validation, SI.B. and YK.L.; formal analysis, SI.B.; investigation, SI.B. and YK.L.; resources, SI.B.; data curation, SI.B.; writing—original draft preparation, SI.B.; writing—review and editing, SI.B. and YK.L.; visualization, SI.B.; supervision, YK.L.; project administration, YK.L.; funding acquisition, SI.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development (IITP-2025-RS-2023-00254592) grant funded by the Korea government (MSIT).

References

1. Kim, T. Y.; Song, B. H.; Bae, S. H. A design and implementation of Music & image retrieval recommendation system based on emotion. *Journal of the Institute of Electronics and Information Engineers*, **2010**, 47(1), 73–79.
2. Wang, Y.; Song, W.; Tao, W.; Liotta, A.; Yang, D.; Li, X.; ... Zhang, W. A systematic review on affective computing: Emotion models, databases, and recent advances. *Information Fusion*, **2022**, 83–84, 19–52. doi:10.1016/j.inffus.2022.03.009
3. Liu, Z.; Xu, W.; Zhang, W.; Jiang, Q. An emotion-based personalized music recommendation framework for emotion improvement. *Information Processing and Management*, **2023**, 60(3, May). doi:10.1016/j.ipm.2022.103256
4. Baek, S. I. A method for continuously recommending contents reflecting emotional changes. Seoul, Republic of Korea: Department of Computer Science Engineering, Dongguk University. **2024**.
5. Shirwadkar, A.; Shinde, P.; Desai, S.; Jacob, S. Emotion based music recommendation system. International Journal for Research in Applied Science and Engineering Technology, **2022**, 10(12, December), 690–694. doi:10.22214/ijraset.2022.47996
6. Wakil, K.; Bakhtyar, R.; Ali, K.; Alaadin, K. Improving web movie recommender system based on emotions. International Journal of Advanced Computer Science and Applications, **2015**, 6(2), 218–226. doi:10.14569/IJACSA.2015.060232
7. Piumi Ishanka, U. A.; Yukawa, T. The Prefiltering techniques in emotion based place recommendation derived by user reviews. *Applied Computational Intelligence and Soft Computing*, **2017**, 5680398.
8. Messina, P.; Dominguez, V.; Parra, D.; Trattner, C.; Soto, A. Content-based artwork recommendation: Integrating painting metadata with neural and manually engineered visual features. *User Modeling and User-Adapted Interaction*, **2019**, 29(2), 251–290. doi:10.1007/s11257-018-9206-9
9. Kim, H.-Y.; Lim, S. B. Emotion-based hangul font recommendation system using crowdsourcing. *Cognitive Systems Research*, **2018**, 47, 214–225. doi:10.1016/j.cogsys.2017.10.004
10. Kim, J. H.; Kim, Y.-R.; Byun, H. W. SNS context-based emoji recommendation using hierarchical KoBERT. *Journal of Digital Contents Society*, **2023**, 24(6), 1361–1371. doi:10.9728/dcs.2023.24.6.1361
11. Heller, A. S.; Casey, B. J. The neurodynamics of emotion: Delineating typical and atypical emotional processes during adolescence. *Developmental Science*, **2016**, 19(1), 3–18. doi:10.1111/desc.12373
12. Cunningham, W. A.; Dunfield, K. A.; Stillman, P. E. Emotional states from affective dynamics. Emotion Review, 2013, 5(4), 344–355. doi:10.1177/1754073913489749

13. Trampe, D.; Quoidbach, J.; Taquet, M. Emotions in everyday life. *PLOS ONE*, **2015**, 10(12), e0145450. doi:10.1371/journal.pone.0145450
14. Larsen, J. T.; McGraw, A. P.; Cacioppo, J. T. Can people feel happy and sad at the same time? *Journal of Personality and Social Psychology*, **2001**, 81(4), 684–696. doi:10.1037/0022-3514.81.4.684
15. Jeong, Y. E.; Kim, J. H.; Son, J. Differences in consumers' emotional response to the same Music: Focusing on the possibility of emotions that advertising music can induce. *Productivity Research*, **2021**, 39(1), 141–149.
16. Sammler, D.; Grigutsch, M.; Fritz, T.; Koelsch, S. Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant Music. *Psychophysiology*, **2007**, 44(2), 293–304. doi:10.1111/j.1469-8986.2007.00497.x
17. Balasubramanian, G.; Kanagasabai, A.; Mohan, J.; Seshadri, N. P. Music induced emotion using wavelet packet decomposition—an EEG study. *Biomedical Signal Processing and Control*, **2018**, 42, 115–128. doi:10.1016/j.bspc.2018.01.015
18. North, A. C.; Hargreaves, D. J. Lifestyle correlates of musical preference: 1. Relationships, living arrangements, beliefs, and crime. *Psychology of Music*, **2007a**, 35(1), 58–87. doi:10.1177/0305735607068888
19. North, A. C.; Hargreaves, D. J. Lifestyle correlates of musical preference: 2. Media, leisure time and music. *Psychology of Music*, **2007b**, 35(2), 179–200. doi:10.1177/0305735607070302
20. North, A. C.; Hargreaves, D. J. Lifestyle correlates of musical preference: 3. Travel, money, education, employment and health. *Psychology of Music*, **2007c**, 35(3), 473–497. doi:10.1177/0305735607072656
21. Sievers, B.; Polansky, L.; Casey, M.; Wheatley, T. Music and movement share a dynamic structure that supports universal expressions of emotion. *Proceedings of the National Academy of Sciences of the United States of America*, **2013**, 110(1), 70–75. doi:10.1073/pnas.1209023110
22. Cowen, A. S.; Fang, X.; Sauter, D.; Keltner, D. What music makes us feel: At least 13 dimensions organize subjective experiences associated with music across different cultures. *Proceedings of the National Academy of Sciences of the United States of America*, 2020, 117(4), 1924–1934. doi:10.1073/pnas.1910704117
23. Kim, H. -R. Development of the Artwork using Music Visualization based on Sentiment Analysis of Lyrics. *Journal of the Korea Contents Association*, **2020**, 20(10), 89–99.
24. Sohn, J. H. Review on discrete, appraisal, and dimensional models of emotion. *Journal of the Ergonomics Society of Korea*, **2011**, 30(1), 179–186. doi:10.5143/JESK.2011.30.1.179
25. Hwang, M.-C.; Kim, J.-H.; Moon, S.-C.; Sang-in, P. Emotion modeling and recognition technology. *Journal of Robotics*, **2011**, 8(4), 34–44.
26. Thornton, M. A.; Tamir, D. I. Mental models accurately predict emotion transitions. *Proceedings of the National Academy of Sciences of the United States of America*, **2017**, 114(23), 5982–5987. doi:10.1073/pnas.1616056114
27. Tamir, D. I.; Thornton, M. A.; Contreras, J. M.; Mitchell, J. P. Neural evidence that three dimensions organize mental state representation: Rationality, social impact, and valence. *Proceedings of the National Academy of Sciences of the United States of America*, **2016**, 113(1), 194–199. doi:10.1073/pnas.1511905112
28. Russell, J. A. A Circumplex model of affect. *Journal of Personality and Social Psychology*, **1980**, 39(6, December), 1161–1178. doi:10.1037/h0077714
29. Thayer, R. E. The biopsychology of Mood and arousal. *Oxford: Oxford University Press*, **1989**.
30. Kim, D.; Lim, B.; Lim, Y. Music Emotion Control Algorithm based on Sound Emotion Tree. *Journal of the Korea Contents Association*, **2015**, 15(3), 21–31. doi:10.5392/JKCA.2015.15.03.021
31. Lee, H.; Shin, D.; Shin, D. A research on the emotion classification and precision improvement of EEG(electroencephalogram) data using machine learning algorithm. *Journal of Internet Computing and Services*, **2019**, 20(5), 27–36.
32. Moon, C. B.; Kim, H.S.; Lee, H. A.; Kim, B. M. Analysis of relationships between Mood and color for different musical preferences. *Color Research and Application*, **2014**, 39(4), 413–423. doi:10.1002/col.21806
33. Laurier, C.; Grivolla, J.; Herrera, P. Multimodal music mood classification using audio and lyrics. *Proceedings of the IEEE Seventh International Conference on Machine Learning and Applications*, **2008**, 688-693. doi:10.1109/ICMLA.2008.96
34. Hu, X.; Stephen, J. Downie, and Andreas F. Ehmann, Lyric text mining in music mood classification. *International Society for Music Information Retrieval Conference*, **2009**, 183, 411–416.

 doi:10.20944/preprints202503.1504.v1

35. Lee, J.; Lim, H.; Kim, H.-J. Similarity evaluation of popular music based on emotion and structure of lyrics. *KIISE Transactions on Computing Practices*, **2016**, 22(10), 479–487. doi:10.5626/KTCP.2016.22.10.479
36. Moon, C. B.; Kim, H. S.; Kim, B. M. Music Retrieval Method using Mood Tag and Music AV Tag based on Folksonomy. *Journal of Korean Institute of Information Scientists and Engineers*, **2013**, 40(9), 526–543.
37. Lee, H.; Lee, T. A study on the classification of emotions and the generation of custom paintings based on EEG. *Journal of Digital Contents Society*, **2021**, 22(10), 1577–1586. doi:10.9728/dcs.2021.22.10.1577
38. Kim, Y.; Song, M. A study on analyzing sentiments on movie reviews by multi-level sentiment classifier. *Journal of Intelligence and Information Systems*, **2016**, 22(3), 71–89. doi:10.13088/jiis.2016.22.3.071
39. Atkinson, A. P.; Dittrich, W. H.; Gemmell, A. J.; Young, A. W. Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, **2004**, 33(6), 717–746.
40. Barrett, L. F.; Mesquita, B.; Gendron, M. Context in emotion perception. *Current Directions in Psychological Science*, **2011**, 20(5), 286–290. doi:10.1177/0963721411422522
41. Zaki, J.; Bolger, N.; Ochsner, K. Unpacking the informational bases of empathic accuracy. *Emotion*, **2009**, 9(4), 478–487. doi:10.1037/a0016551
42. Frijda, N. H. ‘Emotions and Action,’ feelings and emotions: The Amsterdam symposium. *Cambridge: Cambridge University Press*. **2004**, 158-173.
43. Tomkins, S. S. Affect imagery consciousness, I. The Positive Affects, *Springer Publishing Company*. **1962**.
44. Sharma, V. P.; Gaded, A. S.; Chaudhary, D.; Kumar, S.; Sharma, S. Emotion-based music recommendation system. *Proceedings of the IEEE 9th International Conference on Reliability, Infocom Technologies and Optimization*, **2021**. doi:10.1109/ICRITO51393.2021.9596276
45. Joseph, M. M.; Treessa Varghese, D. T.; Sadath, L.; Mishra, V. P. Emotion based music recommendation system. *In Proceedings of the Of 2023 International Conference on Computational Intelligence and Knowledge Economy*, **2023**, 505–510. doi:10.1109/ICCIKE58312.2023.10131874
46. Mahadik, A.; Milgir, S.; Patel, J.; Jagan, V. B.; Kavathekar, V. Mood based music recommendation system. *International Journal of Engineering Research and Technology*, **2021**, 10(6), 553–559.
47. Abdul, A.; Chen, J.; Liao, H. Y.; Chang, S. H. An emotion-aware personalized music recommendation system using a convolutional neural networks approach. *Applied Sciences*, 2018, 8(7), 1103. doi:10.3390/app8071103
48. Lee, S. J.; Seo, B.-G.; Park, D.-H. Development of Music Recommendation System Based on Customer Sentiment Analysis. *Journal of Intelligence and Information Systems*, **2018**, 24(4), 197–217. doi:10.13088/jiis.2018.24.4.197
49. Jazi, S. Y.; Kaedi, M.; Fatemi, A. An emotion-aware music recommender system: Bridging the User’s interaction and music recommendation. *Multimedia Tools and Applications*, **2021**, 80(1), 13559–13574.
50. David, S. ‘3 Ways to Better Understand Your Emotions,’ emotional intelligence. Harvard Business Review. **2016**.