Multimedia Systems
https://doi.org/10.1007/s00530-021-00786-6

REGULAR PAPER

Music mood and human emotion recognition based on physiological signals: a systematic review

Vybhav Chaturvedi[1] · Arman Beer Kaur[1] · Vedansh Varshney[1] · Anupam Garg[1] · Gurpal Singh Chhabra[1] · Munish Kumar[2]

Received: 20 September 2020 / Accepted: 24 March 2021

Abstract
Scientists and researchers have tried to establish a bond between the emotions conveyed and the subsequent mood perceived in a person. Emotions play a major role in terms of our choices, preferences, and decision-making. Emotions appear whenever a person perceives a change in their surroundings or within their body. Since early times, a considerable amount of effort has been made in the field of emotion detection and mood estimation. Listening to music forms a major part of our daily life. The music we listen to, the emotions it induces, and the resulting mood are all interrelated in ways we are unbeknownst to, and our survey is entirely based on these two areas of research. Differing viewpoints on this issue have led to the proposal of different ways of emotion annotation, model training, and result visualization. This paper provides a detailed review of the methods proposed in music mood recognition. It also discusses the different sensors that have been utilized to acquire various physiological signals. This paper will focus upon the datasets created and reused, different classifiers employed to obtain results with higher accuracy, features extracted from the acquired signals, and music along with an attempt to determine the exact features and parameters that will help in improving the classification process. It will also investigate several techniques to detect emotions and the different music models used to assess the music mood. This review intends to answer the questions and research issues in identifying human emotions and music mood to provide a greater insight into this field of interest and develop a better understanding to comprehend and answer the perplexing problems that surround us.

Keywords Emotion recognition · Music mood classification · Physiological signals · EEG feature extraction · K-NN · SVM · Random forest

1 Introduction

1.1 What are emotions

It has been a great necessity for human beings to understand each other to adapt and survive in their respective surroundings, with cooperation and in harmony with each other. 'Emotions' have always played a key role in the process of character development of individuals which in turn has aided in propelling the development and growth of their own communities and even the society when considered in totality. As a result, the topic of thoughts, feelings, and emotions experienced by people, aroused a great deal of interest and curiosity which led these subjects to attract a great amount of attention from many philosophers, psychiatrists, doctors, and researchers in the associated fields. They became interested in exploring how humans understood and responded to each other, to their circumstances and to their own bodies.

Over the centuries, with the extensive research work and studies carried out in this field, various definitions have been associated with the word 'emotion'. Some people describe it as a strong feeling that arises from the circumstances pertaining to an individual's experiences whether considered as a positive or a negative experience. Many characterize them to be a form of sensation or intuition or the response of an individual towards various things which hold personal

Communicated by I. Bartolini.

✉ Munish Kumar
munishcse@gmail.com

1 Computer Science and Engineering Department, Thapar Institute of Engineering and Technology, Patiala, Punjab, India

2 Department of Computational Sciences, Maharaja Ranjit Singh Punjab Technical University, Bathinda, Punjab, India

significance in their lives which can relate to both inanimate belongings and the people they hold dearly. Every person has a different perspective and understanding of this world and it is quite fair enough for them to understand these feelings in their own desirable manner. However, there lies a common underlying theme in these varying interpretations and it provides us a method to comprehend these emotions in a broader manner.

1.2 Impact of emotions on human beings

What makes this field of work and research more appealing is the complexity that it beholds. Talking about the subject, emotions are quite complex to understand which makes people's behavior unpredictable and difficult to apprehend on various occasions.

Emotions exert a great influence on several aspects of our daily lifestyle and the activities that we carry out throughout the day. The human mind is like a machine constantly busy producing thoughts, ideas, and thinking upon different course of actions. Upon consideration, these factors are very likely to trigger happiness, sadness, anxiety, fear, surprise, anger, and numerous other emotions in an individual. It can adversely affect our behavior, communication, decision-making skills, selection of hobbies and interests, and our well-being.

1.3 How are emotion and music related to each other

Emotions appear whenever a person perceives a change in their surroundings or within their body. It can be considered as a psychological state or a process that is designed to maintain a balance between the information processed by the brain and goals that it must achieve primarily. Many researchers have explored various Models of Emotions and the factors that give rise to the perception of emotion in music. Listening to music is a major part of our daily life and is even considered as a hobby by many. People spend hours listening to music and more than millions purchasing it. According to India Music 360[1] [1] music outranks all other interests/hobbies, and in total, 94% of online consumers listen to music throughout the year. The Digital Music Study[2] by Indian Music Industry shows an average Indian listens to music for 21.5 h/week, which is more than the global average of 17.8 h/week. In 2011, Zwaag et al. [109] claimed that listening to music lowers the chance of rash driving and even accidents contrary to the popular belief that music impairs drivers.

Various studies indicate the other benefits of music and songs; North et al. [78] concluded that music can have a wide range of positive commercial benefits. Silverman [92] concluded that live music can be an inexpensive method to positively impact people in the oncology waiting rooms and mentioned that live music can be used to help patients and their families in lowering anxiety and stress. In one review, Lehmberg et al. [56] pointed out the benefits of music for senior citizens, which include prolonged lifespan, lesser stress, better health, and overall happiness.

Thaut [104] compiled various benefits of music in Rehabilitation and Therapy. In 1998, McCraty et al. [67] studied the effect of 4 kinds of music on our mood and emotions and determined that depending upon the mood music can change our thinking process and concluded that only Designer Music is beneficial, as it has no negative effects. The latter is a very significant study as it shows that different types of music have different effects on our mood and feelings.

1.4 Challenges in the system

As stated earlier, the study of human emotions is a subject of complexity due to its unpredictable nature and can produce perplexing results. Extensive research work has been carried out to accurately determine the emotions experienced by an individual by recording their physiological signals and brain activity. Many of them have been successful in classifying emotions with a certain degree of correctness but have also faced several challenges during their study. A variety of techniques like High Order Crossing, Adaptive Filtering, Autoregressive Modelling, Genetic Algorithms, and their combinations have been employed in different works, but more optimized algorithms and techniques for improved overall performance of the process which would enable it to acquire a relatively higher accuracy are still required. This also includes the development of a unified algorithm which incorporates various biological signals, to detect human emotions more precisely. Many problems are even specific in their nature and depend exclusively upon their application. Considering the example of emotion detection using EEG signals where ascertaining the optimal positions for the placement of electrodes is still a challenging task. Several models of investigation are person-dependent; they require at least some labeled data from each subject for training purposes.

Researchers have investigated the problem of automatic emotion recognition related to music. Even the music industry has often faced a problem of labeling songs or the type of music with certainty. Consequently, a need arises to find ways to extract more powerful acoustic features that better

[1] Available at https://www.nielsen.com/in/en/insights/report/2018/india-music-360-report, The Nielsen Company (US) (Last Checked—6-Jan-2020) .

[2] Available at https://indianmi.org/?id=12060&t=Digital%20Music%20Study,%202019 (Last Checked—6-Jan-2020).

represent music primitives in mood perception, such as mode and articulation which are expected to improve the mood detection accuracy. Several research works have based their results solely either on lyrical or audio feature extraction. The selection of only a set of parameters to classify music might result in the neglecting of certain important aspects and it poses a challenge in this field.

On average, only 4–5 types of emotions have been included in every model of emotion so far examined. Therefore, extending the mood taxonomy to cover more mood types which are related to human responses and incorporate more ambiguous moods in the training set is still required. The majority of the experiments have been conducted on a group of few people, thus increasing the number of test subjects to obtain more data from more test subjects and be able to provide more solid conclusions is essential. Moreover, the conversion of the implemented project into a commercialized product is an onerous task. Fulfilling the expectations of the consumer that the respective sensor module will be implemented as a wearable device like wristwatch-type device or any other form suitable for everyday use also needs to be brought into consideration.

2 Motivation

The study of human behavior has always been a subject of eccentricity due to the element of uncertainty that has to be dealt with it. Considerable amount of research work and analysis has been carried to understand the connection between music and Human Psychology. Researchers have tried to examine the emotions experienced by their subjects using various samples of audio, video, and pictures to trigger certain responses in them and record the various physiological signals generated through different sensors. Analysis and identification of the signals which provide us with the highest degree of accuracy in emotion detection is our area of interest. Substantial amount of study has been conducted to classify the mood of a song after employing various feature extraction techniques for both its audio and lyrical features. Determination of the exact features and parameters that will produce the best results for our purpose of classification is crucial. Moreover, enthusiasm to study and employ algorithms and techniques that will produce results with a relatively higher degree of accuracy plays a key role in our survey.

3 Survey protocol

Survey protocol defines how a survey is implemented and plays a major role in quality data collection. It acts as rules and a set of guidelines that must be followed during the activity. It describes a comprehensive work structure for the project and even helps novice researchers in the field with a suitable amount of data and information. During the activity, no questionnaire or any such sort of method was adopted, the focus was solely on the available studies and compiling them together to achieve a better understanding of the field of human emotion detection and its relationship with music. The survey aims to help all the researchers interested in the field of Emotion detection, BCI, HCI, Music Mood analysis, and physiological signal analysis, by providing them useful information backed by data. In the further sections, important sub-concepts like research questions, source of study, quality assessment, sources of collecting datasets, and inclusion and exclusion criteria have been extensively discussed.

As the study revolves around available literature, all the well-known research papers, review papers, reports, and other literature from various journals were selected. After organizing them in chronological order, they were extensively studied to extract all the important information and to maintain the quality of our work, and the renowned papers were prioritized. Official articles related to various Taxonomies and Datasets were also included, so that the survey talks about the useful Models and datasets examined. Benchmarked datasets along with the notable ones were prioritized over self-arranged datasets, as they are not freely available and are harder to obtain.

3.1 Planning the survey

The survey planning commences with identifying the reason and purpose for conducting this survey. Our research work aims to target normal healthy participants. The whole planning procedure unfolds the methods employed to approach our goal, collecting the required relevant information, analyzing and implementing certain algorithms and techniques in accordance with the parameters included and excluded from the study, and hence obtaining the required results.

3.2 Research questions

- What is the relationship between Human Psychology and Music?
- Which physiological signals provide us with the best accuracy in terms of emotion detection?
- Is it possible to determine the mood of the song using audio and lyrical feature extraction?
- Which feature extraction techniques and classifiers will be best suited for our purpose?
- Is it possible to develop algorithms or techniques that will improve the accuracy of the previously obtained results?

- What data sets have been used and how frequently have they appeared?
- Is the size of the data sets enough to provide precise results?
- Is it possible to develop a system that will unify emotion detection along with song mood analysis?
- What are the application areas where it will prove to be most useful?

3.3 Inclusion and exclusion criteria

3.3.1 Inclusion criteria

- Physiological signals like EEG, EMG, ECG, EDA, SC, BVP, RSP, SKT, GSR, and HR have been included.
- Certain biomedical papers have been included in our study as they consist of informative feature extraction methods and datasets.
- Research papers that have investigated various Models of Emotions and Human Emotion Taxonomy have been included.
- Research papers solely based on Datasets that would be utilized for the sake of our study have been included.
- Research papers that analyze music and classify them under different labels have been included.
- It has been assumed that the test cases have a proper understanding of the language of the audio input (if song or video).

3.3.2 Exclusion criteria

- Physiological Signals like EOG, PCG, PPG, Optoacoustic, and Speech have been excluded.
- Vocal Intonation Recognition and Natural Language Understanding have been excluded.
- Facial Expressions Recognition and Image Processing do not fall in our area of interest.
- Techniques like Deep Learning require a huge amount of data for training and due to the lack of availability of EEG corpus we prefer to employ traditional Machine Learning techniques for the purpose of classification. The subject of Deep Learning has been slightly touched upon, but its detailed study is beyond our scope of the study.
- The difference between Left-Handed and Right-Handed people has not been considered.

3.4 Quality assessment

- High-quality research papers from prominent publishing journals like Springer, IEEE, ACM, Elsevier, and Science Direct.
- Research papers with a high number of citations and recommendations.
- Research papers that have employed a variety of techniques to not limit ourselves to specific methods.
- Research papers with a large dataset enough for carrying out comparative analysis.
- Research papers which have investigated parameters and factors relevant and confined to our area of interest

4 State-of-art surveys

A literature survey was conducted to understand the novel techniques, methods, and approaches and how they stand when compared to the classical approaches that were used in the past.

To incorporate as many studies as possible, as a part of the literature survey, we studied multiple acclaimed state-of-art reviews along with some of the more recent publications to cover as much time span as we can. We went through 15 review papers which were predominantly related to music, emotions, physiological signals, and also music theory. Table 1 lists all the surveys along with their covered areas, scope, and/or limitations.

We specifically chose reviews, which covered studies over a large period of time; for example, 30 years (1990–2020) [126], a specific study of all the protocols followed in EEG analysis over a span of 6 years (2009–2016) [4], and a generic study which aims to cover every important research in music classification over 90 years (1930–2020) [128]. The survey paper gave us novel insights and a better understanding of music mood detection and its commercial usage [39, 72]. As music recommendation systems are gaining more popularity with advancements in machine learning and AI, Paul et al. [82] and Song et al. [98] in the separate study compare multiple music recommendation systems across commercial as well as research platforms. The base of MIR is genre classification and music mood identification, and hence, a special effort was made towards accumulating multiple reviews and each of them provided a unique insight in the field of MIR. Warrenburg [128] provides a whole lot of information about multiple available datasets and taxonomies and their different implications on the result, while Yang et al. [114] and Weihs et al. [112] have tried to address the major issues faced in music mood recognition. Fu et al. [26] extensively worked on audio classification and related feature extraction. Scaringella et al. [88] are one of the classic reviews in the field of MIR. Changes in human emotion can be detected accurately using one or more physiological signals. Shu et al. [90] are a detailed review of six physiological signals and their resultant output for human emotion detection. A relatively higher number of surveys on EEG were also observed during our literature survey, Hosseini et al. [126] provided a detailed study on features extraction techniques used in EEG and also compared regression

Table 1 Details of the state-of-art surveys along with their scope and limitations

Serial no.	Author	Year	Covered areas	Scope and limitations
1	[82]	2020	The article segregated the current music recommendation system based on 5 most popular techniques and further more discussed the commercial application and techniques used for high scale music recommendation	The study has covered only a few music recommendation frameworks. Using a similar approach, more frameworks should be classified and the novel commercial means has not been explored The study proposed its own hybrid recommendation system
2	[126]	2020	An extensive survey, solely based on EEG feature extraction and classification. It comprised various studies over a span of 30 years. Details about classifiers and their accuracy parameters have been discussed. It also recommended classifiers and approaches for regression and classification separately	The paper correctly pointed out the sub-par accuracy levels of individual classifiers but has not discussed the possible combination of individual classifiers for a better accuracy score
3	[128]	2020	The article covered more than 300 previous studies over a span of 90 years in the field of music classification. The dataset used is the sole motive, and this provides a benchmark to compare with	The article is only focused on datasets and corpus used in the field of Music Information Retrieval. Various findings like abundance of research based on short stimuli and perceived information are thought-provoking and important to understand the need for well-structured dataset
4	[4]	2019	An overview of theory, recent research work, and future possibilities comprising various studies from 2009 to 2016 with detailed information about EEG and how it is being used to measure and detect emotions is presented. The article in great detail has discussed all the key points and proposed protocols during the phases of emotion recognition and corresponding implications using EEG	The step-by-step phases have been studied separately and hence can be used as a tool for debugging and back tracking. The survey is beginner-friendly and provides concise recommendations as well. At the same time, it has taken a deeper insight into the aspects involved in the process of recognizing emotions from EEG
5	[72]	2018	Article majorly covers Content-Based Music Information Retrieval (CB-MIR) in relation to the music industry. The article focuses on eight MIR-related tasks. The fundamental concepts of Indian classical music are also mentioned	This article points to some issues in CB-MIR and probable approaches to improve the efficiency of the existing CB-MIR systems. The article on its own is not a take on human emotion detection using music
6	[62]	2018	The review is focused on EEG-based Brain–Computer Interfacing (BCI) and is specifically dedicated to the review of classification algorithms used for BCI, their properties, and their evaluation. Moreover, it provides guidelines to help the reader with choosing the most appropriate classification algorithm for a given BCI experiment	This is a well-structured and detailed take on BCI and HCI. Although the techniques and algorithms are a bit outdated, but it still provides respectable comparisons between the classical algorithms
7	[2]	2018	The review is majorly based on Epilepsy and its relationship with brain activity. The article is well-structured and has detailed information regarding EEG like comparing various nonlinear features and various proposed techniques for easy comparison	Although meant for clinicians, the data provided for EEG is useful. As the study is comparatively newer, the features and techniques are up to date and hence give the best comparisons
8	[90]	2018	The survey is focused on emotion recognition using physiological signals and their possible application. Details and comparisons about feature extraction, machine learning techniques, and emotion taxonomy, and most widely used datasets have been provided	The article incorporates 6 physiological signals alongside neuromusicology which is a step-up compared to other similar studies
9	[98]	2012	Discussed various approaches used in music recommendation systems. The study divides all the systems based on major algorithms, user-based approaches, and all the key components of a recommendation system. The study further discussed 6 popular systems and 4 more potential systems	The survey provides a strong groundwork on major music recommendation systems, but completely ignores the subjective nature of the music which makes the study less user-friendly
10	[114]	2012	Discussed various taxonomy, annotation methods, and at the same time states various machine learning techniques that are being used compares the resultant output and visualization techniques	Study has listed down the most important issues faced during music mood recognition
11	[26]	2011	Discussed all the aspects and tasks related to MIR. It differentiated all the major steps like feature extraction, classification, and model training on the basis of different tasks that are related to MIR. Furthermore, the audio classification of music has been also addressed in depth	The study in depth has addressed the 4 concerns related to audio classification along with possible solutions. The paper also addressed the gaps in human emotions and genre classification

Table 1 (continued)

Serial no.	Author	Year	Covered areas	Scope and limitations
12	[39]	2011	The purpose of the review is to examine the influence of music on consumption experience and explore the relationships between musical variables and consumer responses in the context of retailing	The study has ignored the fact that music variables can relatively affect each other as well leading to a completely different effect The effect of environmental factors has been ignored
13	[112]	2007	The review has provided the solutions and open problems for music classification and the authors have concentrated on music similarity, structure analysis, cognitive psychology, and transcription	The study has addressed music transcription and its classification The study has not considered the emotions conveyed by music and has limited itself to genre classification only
14	[88]	2006	Music genre classification is one of the basic and ever-present methods. The article highlights the availability of extraction methods and the present work that is being done in the field	While the focus has been on genre classification, ignoring other methods of classification surely hampers the usage of the article in this study
15	[44]	2004	The study attempts to bridge the gap between social context of music listening and subsequent change in emotions. The survey also argues that emotion is strongly related to most people's primary motives for listening to music	The results provide preliminary estimates of the occurrence of various emotions in listening to music, at the same time mentions how music is used by listeners in several different emotional ways in various life contexts. The questionnaire also provides a good insight into the relationship between human emotions and motives

and classification methods. Lotte et al. [62] provide insight on EEG and BCI, and the article comprises multiple studies related to Human–Computer Interface (HCI) and Brain–Computer Interface (BCI). Acharya et al. [2] have taken into account the major development in using EEG as a cure for epilepsy, and the details about feature extraction provide a greater understanding of EEG signals in response to brain activity.

5 Background

5.1 Psychological effect of music

Music is a very common phenomenon across the world and people are attracted to music, because it is considered as one of the strongest tools for the arousal of emotions and feelings [44]. It is vastly theorized that music not only affects the brain activity but also our emotions and feelings, which might lead to a change in our perception towards the world. In 2011, Zwaag et al. [109] proposed "Listening to negative music compared to no music while driving leads to decreased Respiratory rate and listening to positive music compared to no music leads to slower driving speed". The study focuses on considering such a framework as our foundation for optimizing the emotion detection ability of the system.

North et al. [78] concluded that music can have a wide range of positive commercial benefits. Music can influence the places that customers go to, the atmosphere of these commercial premises, the amount that customers are prepared to spend in them, the amount that they actually spend, the products they buy, and various other perceptions and beliefs concerning related products. These are very positive and encouraging signs which strengthen the belief that music plays a vast impact in the day-to-day life of everyone who prefers to listen to music in their free time.

Various studies have investigated the usage of psychological and brain signals separately. Evaluating results from a prior investigation is a difficult task as various means have been used to accumulate, varying data numbers of test cases. Chanel [17], Hosseini [33], Kim [48], and Takahashi [103] have used various means and platforms for the detection of induced emotions from different stimuli and have achieved appreciable accuracies.

Li et al. [57] emphasized the fact that music does not only serve as a tool for entertainment, but its social and psychological effects are also essential as we experience a rapid growth in the field of music. Huron [36] pointed out the applications of music in many areas of our everyday life which includes playing songs in a restaurant that attract more customers and adds on to a better ambience, music selected by an aerobics instructor for exercises, tunes chosen by a film director for particular scenes of a movie to prompt certain emotions in the viewer or those used by physiotherapist for their patients. Automatic classification of music can serve useful in music database management systems. Tzanetakis et al. [108] highlighted the importance of automatic music analysis with respect to distribution of music content to the customers and the expansion of the digital market for music.

Several techniques for the classification of music based on music mood have been used. Kim et al. [50] introduced a game named MoodSwings which records dynamic mood ratings for music given by the people playing the game. Li et al. [57] used an approach that focuses upon the three parameters, namely timbre, intensity, and rhythm, to categorize music. Jiang et al. [43] utilized the spectral characteristics of a music clip through the Octave-based Spectral Contrast feature for music classification.

5.2 Human emotion based on physiological signals

The goal of this study is to investigate and understand electrical activity of the brain, skin conductance, respiratory rate, and heart rate with respect to their role in determining the moods and emotions experienced by a subject along with analyzing and identification of musical features and parameters that will aid in the former objective.

Physiological signals for a long time have been used to establish a relationship with changes in emotion and electrical activity of the brain, heart rate. In [90] studied six physiological signals and how to obtain appreciable accuracy in detecting human emotion using them. The most frequently used physiological signals are:

1. Electroencephalogram or EEG
2. Electrocardiogram or ECG
3. Heart rate
4. Galvanic skin response or GSR
5. Electromyogram or EMG
6. Respiratory rate or RSP.

Apart from these, skin temperature, eye twitching, blood vessel volume have also been incorporated some times.

EEG is the physiological method of choice to record the electrical activity generated by the brain via electrodes placed on the scalp surface.

ECG is a method of collecting electrical signals generated by the heart. This allows us to understand the level of physiological arousal that someone is experiencing.

Heart Rate is the number of times the heart beats within a certain time period, usually a minute. It is used to describe the frequency of the cardiac cycle.

GSR refers to changes in sweat gland activity that are reflective of the intensity of our emotional state, otherwise known as emotional arousal. It falls under Electrodermal Activity or EDA, both the terms are often used interchangeably.

EMG is used to measure and record the movement of our muscles. It measures the burst of electricity generated whenever a muscle contracts, the electrical activity can be measured through adjacent tissues and neighboring skin areas.

RSP is the number of breaths we take per minute.

EEG is one of the most used physiological signals for emotion recognition. The technique requires signals to be processed using electrodes over the scalp and forehead, usually in a form of wearable cap or headband. Multiple researchers have used different kinds of approaches, while using multiple electrodes like 32 or 64 is the mainstream, using lesser electrodes is also very common. Feature extraction and selection along with the classifier used for validation also varies a lot across the spectrum. Similarly, like EEG, ECG, EMG and GSR are two another important signals processed using a form of wearable equipment/machine with electrodes attached to it, though in case of ECG, different approach uses different number of electrodes, but in predetermined locations, GSR and EMG provide much more flexibility in this regard and multiple studies have pondered over different region where electrodes can be attached to and their resulting accuracies.

A change in human emotions has a resulting physiological effect and some theorists even argue that the same is true for vice versa. The Cannon–Bard theory [134] or the Thalamic theory of emotion argues that one feels emotions and experiences physiological reactions such as sweating, trembling, and muscle tension simultaneously. It even states that we react to a stimulus and experience the associated emotion at the same time. The physical reactions are not dependent upon the emotional reaction, or vice versa. Thus, proving that physiological signals and human emotion are actually interrelated and one can be the key to determining the other. Obviously, as human emotions are more related to feelings and signals can be quantized, it can be said that physiological signals can be used to determine human emotions in real time, and certain systems can be trained specifically for subjects as well [133].

6 Taxonomy

Many taxonomies can be considered while categorizing music according to different emotions. There exists no standard taxonomy to be universally followed, because a few basic emotions cannot represent the whole spectrum. Many researchers and theorists have classified emotions based on various theories and standards (Fig. 1). Psychologist

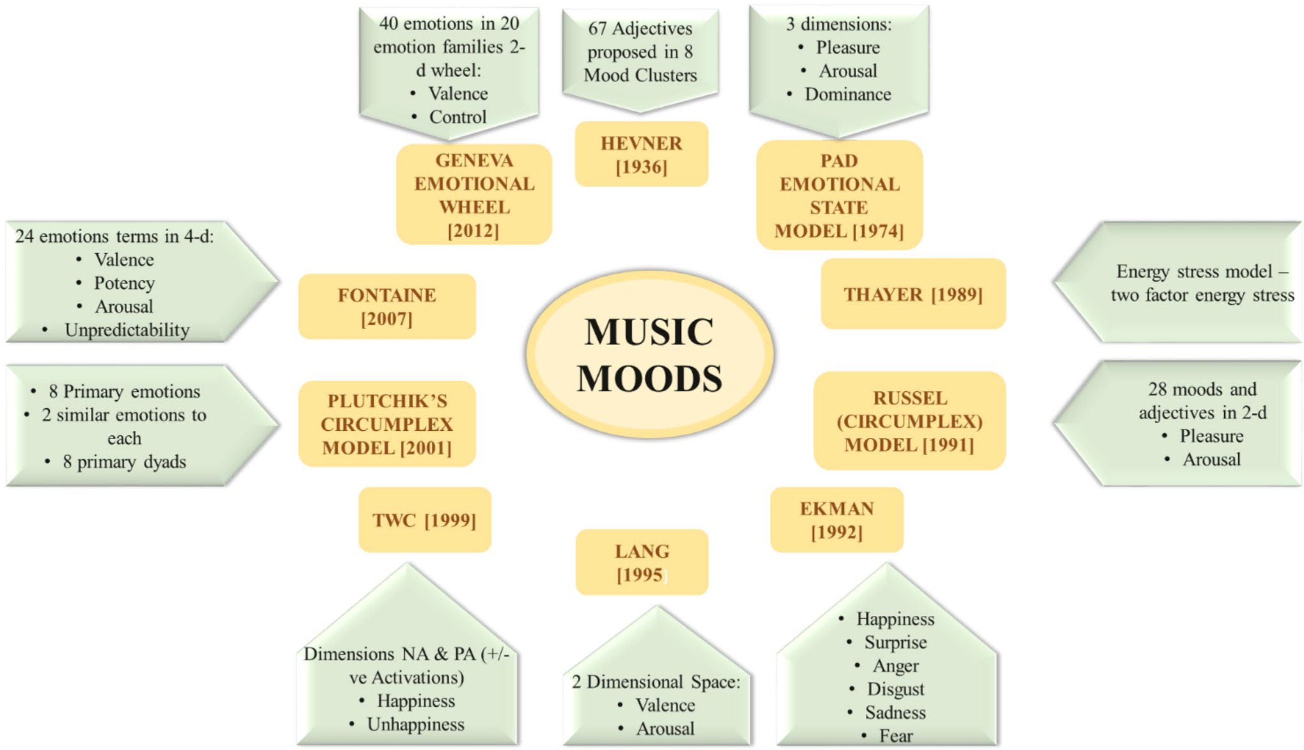

Fig. 1 All the major taxonomy taken into consideration during the study

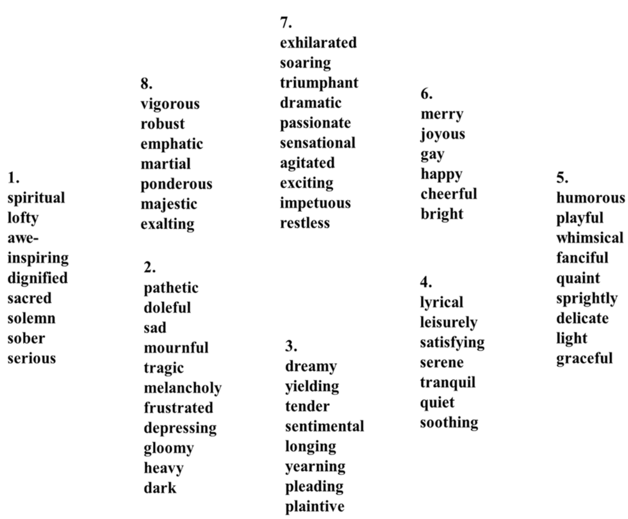

Fig. 2 Pictorial representation of Hevner's Eight Clustery

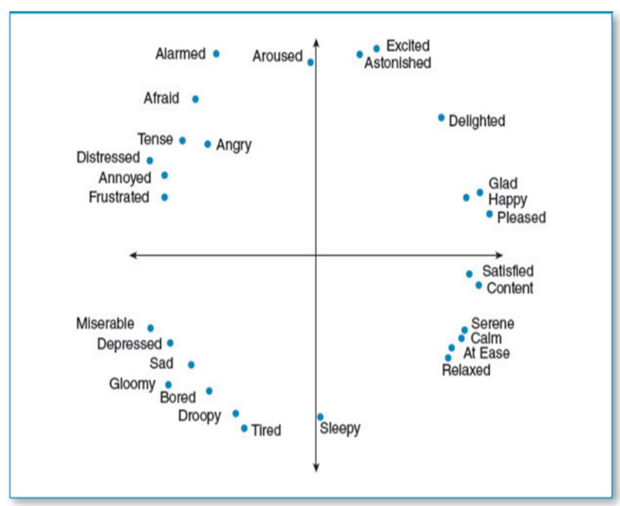

Fig. 3 Pictorial representation of Russell (Circumplex) model

Ekman [22] argued for basic emotions (Ekman proposed six emotions which are happiness, surprise, anger, disgust, sadness, and fear) and nine characteristics which can help to distinguish basic emotions from one another and other emotions such as smugness, hope, jealousy, grief, not being part of emotional families. There are two approaches followed when classifying according to emotions [50, 113], one approach is creating discrete emotions, i.e., forming clusters of adjectives that are said to express emotions in the English language, formed as a model by Hevner [31] (Fig. 2), who composed eight related groups of 67 adjectives, used to search similar musical files which achieved considerable accuracies in Jazz vocal tracks and Classical vocal tracks [57]. Other approach followed is the dimensional approach, conventionally the two most used are Circumplex emotional model of Russell [86] (Fig. 3) and Thayer's Model [122]

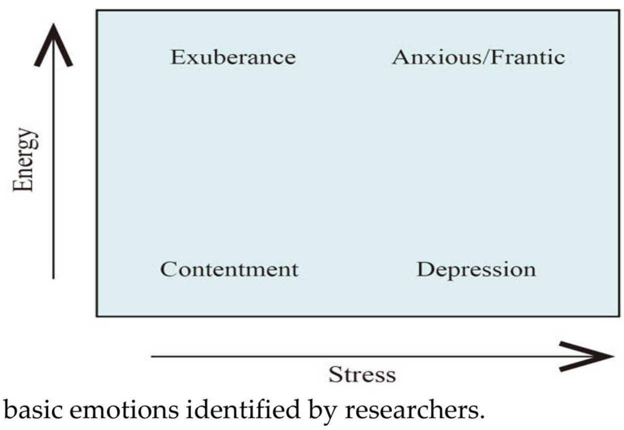

Fig. 4 Pictorial representation of Thayer's 2-D model

(Fig. 4). Russell scaled 28 affect (adjective) words on different scales, once directly, using multidimensional scaling, unidimensional scaling, with regression weights as a function of pleasure displeasure (horizontal axis) and degree of arousal (vertical axis). To measure approach/avoidance and to study the behavior of consumers and the influence of music on them [7, 7], PAD Emotional State Model developed by Mehrabian and Russell [69] which consists of three dimensions: pleasure, arousal (both from circumplex model), and dominance. Dominance is used as a measure for how much an individual feels dominated vs free to act and stay in relation to their environment. Thayer is an energy stress model that entails music from two factors: stress (happy/anxious) and energy (calm/energetic), and this model classifies music into four groups: exuberance/joy, contentment, anger/frantic, and depression. For classification of physiological signals according to emotions [42, 111], Ekman's and Lang's Model [53] have been used. Lang characterized emotions in two-dimensional space by their valence and arousal, images from the International Affective Picture System (IAPS) were organized by him based on the dimensions of valence and arousal. TWC model [105] is also used to address emotions with two dimensions, namely PA (Positive Activation) and NA (Negative Activation), independently accommodating a general bipolar happiness vs unhappiness dimension, they formed 120 item questionnaire sample and made it into 29 items with 3–4 items into eight categories: calm-ease, joy, interest, surprise, fear, anger/disgust/contempt, shame/guilt, and low energy. Like Ekman, Plutchik devised another set of eight basic emotions: joy, trust, fear, surprise, sadness, anticipation, anger, and disgust, which represent eight sectors designed and arranged as four pairs of opposites in Plutchik's three-dimensional circumplex model [85]. In this model, the cone's vertical dimension represents intensity, circle represents degrees of similarity among the emotions, and emotions in the blank are the primary dyads (emotions that are mixtures of the two of the primary emotions). This model was used as a method by Kim et al. [79] for automatic music mood classification using only lyrics from songs and to provide recommendations to users' mood. An Emotion Wheel with two dimensions was Geneva Emotion Wheel (GEW) by Scherer can wield feeling component, with two dimensions valence (negative to positive) and control (low to high), and divide mood into four dimensions, version 1 of the GEW was with 16 emotion terms, and then, version 2.0 of the GEW with 40 emotion terms arranged in 20 emotion families [91]. More dimensions were made into models like four-dimensional solution, which was created by Fontaine et al. [123] with 24 emotion terms: valence, potency, arousal, and unpredictability; it included all six of the components of basic emotions identified by researchers.

7 Data sets of music and physiological signals

A large number of datasets are available related to music, music mood detection, and physiological signals as described in Table 2 along with their reference and year of publication. The characteristics of the specified datasets are being discussed in the subsequent Sects. 7.1, 7.2, and 7.3.

7.1 Datasets related to physiological signals

DECAF—a multimodal dataset for interpreting physiological signals of the user in response to effective multimedia content. The dataset consists of brain signals acquired through magnetoencephalogram (MEG) sensor and the emotional responses elicited in 30 participants after watching movie clips. In addition, it comprises synchronously recorded near-infra-red (NIR) facial videos, horizontal Electrooculogram (hEOG), Electrocardiogram (ECG), and trapezius-Electromyogram (tEMG).

MIT-BIH Arrhythmia Database—The database consists of two-channel ambulatory ECG recordings recorded from 47 subjects studied by the BIH Arrhythmia Laboratory between 1975 and 1979.

NMED-H—The dataset contains EEG responses and behavioral ratings that were obtained after 48 adults listened to full-length Hindi/ Bollywood pop songs. It contains 29 h of auditory EEG responses, including over 7 h of responses to intact music.

NMED-T—The dataset contains EEG recordings and behavioral data like ratings of familiarity and enjoyment, which were collected during natural music listening (full-length songs with electronically produced beats at various tempos) conducted on 20 adult participants.

NMED-RP—The dataset contains EEG recordings collected, while five adult participants heard excerpts of natural

Table 2 List of various datasets along with reference and publishing year for music, music mood, and physiological signals

Serial number	Year	Dataset	References	Serial number	Year	Dataset	References
Dataset related to physiological signals				*Dataset related to music and music mood detection*			
1	2018	NMED-RP	[61]	15	2019	ElderReact	[65]
2	2017	NMED-T	[61]	16	2017	FMA†	[19]
3	2017	AMIGOS	[70]	17	2017	Audio Set†	[27]
4	2016	ASCERTAIN†	[102]	18	2015	AMG 1608	[18]
5	2016	SEED†	[118]	19	2015	Emotify	[6]
6	2014	NMED-H	[61]	20	2015	OpenMIIR†	[124]
7	2014	EEG Motor Movement	[89]	21	2015	MUSAN	[94]
8	2014	TUH EEG†	[30]	22	2014	EmoMusic	[96]
9	2012	Mahnob-HCI†	[97]	23	2014	MedleyDB†	[13]
10	2012	Bern-Barcelona†	[8]	24	2011	DEAP†	[51]
11	2007	DECAF	[1]	25	2011	Million Song Dataset†	[105]
12	2001	MIT-BIH†	[71]	26	2011	musiXmatch	[105]
Dataset that relates physiological signals with music				27	2002	GTZAN	[21]
13	2019	PMEmo†1	[127]	28	2007	SEMAINE†	[68]
14	2018	DEAM†	[5]	29	2005	MABD†	[32]
				30	2007	MIREX	[38]
				31	2003	Last.fm2	[130]

music or rhythms along with behavioral ratings of the stimuli like perceived rhythmic complexity, enjoyment rating, and ease in finding the beat.

EEG Motor movement/imagery dataset—The dataset consists of over 1500 EEG recordings obtained from 109 volunteers that performed different motor/imagery tasks such as opening/closing eyes and clenching/relaxing fists and feet in response to a stimulus.

TUH EEG—A corpus consisting of over 25,000 EEG studies along with a neurologist's interpretation of the test, a brief patient medical history, and demographic information about the patient such as their age and gender.

MAHNOB-HCI tagging database—The database contains synchronized, high-quality, multi-camera, multimodal recordings of spontaneous elicited emotions and responses to the correct and incorrect image or video tags. The recorded modalities are six camera views, sound from both a head-worn and a room microphone, eye gaze, and peripheral/central nervous system physiological signals.

Bern-Barcelona EEG database—The database consists of EEG signals belonging to two classes namely F (focal) and NF (non-focal) which were obtained from 5 epilepsy patients and both the classes contain 3750 pairs of signals.

AMIGOS—The database consists of physiological signals such as Electroencephalogram (EEG), Electrocardiogram (ECG), and Galvanic Skin Response (GSR) which were recorded using wearable sensors, while the 40 participants watched short- and long-length videos as an effective stimulus. Participants' emotions have been annotated with self-assessment of affective levels (valence, arousal, control, familiarity, like/dislike, and selection of basic emotion) felt by the participants and external assessment of participants' levels of valence and arousal during the whole experiment.

ASCERTAIN—The database consists of big-five personality scales and emotional self-ratings of 58 subjects along with synchronously recorded Electroencephalogram (EEG), Electrocardiogram (ECG), Galvanic Skin Response (GSR), and facial activity data elicited while watching affective movie clips.

SEED—The dataset contains EEG signals recorded when 15 subjects were watching carefully selected film clips intended to induce emotions like happiness, sadness, fear, and neutral reaction.

DEAP—The dataset consists of EEG and physiological signals recordings after 32 volunteers watched certain music videos and participants also rated the videos based on arousal, valence, and dominance. The frontal face video was also recorded for 22 participants.

OpenMIIR—The dataset contains EEG recordings of music perception and imagination for 12 short music stimuli which was collected from 10 subjects.

7.2 Datasets related to music mood recognition

ElderReact—A multimodal dataset for recognizing emotional response in aging adults. It contains 1323 video clips of 46 unique subjects with human annotations of six discrete emotions: anger, disgust, fear, happiness, sadness, and surprise along with valence.

FMA—The database consists of Creative Commons-licensed audio from 106,574 tracks from 16,341 artists and 14,854 albums, arranged in a hierarchical taxonomy of 161 genres. It provides full-length and high-quality audio, pre-computed features, together with track and user-level metadata, tags, and free-form text.

AudioSet—The dataset consists of an expanding ontology of 632 audio event classes and a collection of 2,084,320 human-labeled 10-second sound clips drawn from YouTube videos. The ontology is specified as a hierarchical graph of event categories, covering a wide range of human and animal sounds, musical instruments and genres, and common everyday environmental sounds.

AMG 1608—The dataset is composed of 1608 30-s music clips annotated by 665 subjects.

Emotify—The dataset includes 400 tracks and incorporates four music genres (classical, rock, pop, and electronic music), 100 tracks per genre, and 8407 annotations total. The annotations in the dataset verify the inducement of emotions like amazement, solemnity, tenderness, nostalgia, calmness, power, joyful activation, tension, and sadness.

MUSAN—A corpus of music, speech, and noise that is suitable for training models for voice activity detection (VAD) and music/speech discrimination. It consists of music from several genres, speech from 12 languages, and a wide assortment of technical and non-technical noises.

emoMusic—The dataset consists of arousal-valence for 744 songs which are split between the development set (619 songs) and the evaluation set (125 songs). Along with annotations, the dataset also provides the standard deviation of the annotations to get an idea about the margin of error.

MedleyDB—The dataset consists of 122 multitrack recordings along with specified annotations and genres like Singer/Songwriter, Classical, Rock, World/Folk, Fusion, Jazz, Pop, Musical Theatre, and Rap.

Million Song dataset—A freely available collection of audio features and metadata for a million contemporary popular music tracks. It contains 1 million songs/files, 44,745 unique artists, 7643 unique tags, and additional information relevant to MIR.

Last.fm—This dataset is a collection of song-level tags and pre-computed song-level similarity, where all of the data are associated with Million Song Dataset tracks (MSD). It contains 943,347 tracks matched with MSD, 505,216 tracks with at least one tag, 584,897 tracks with at least one similar track, 522,366 unique tags, and additional information relevant to MIR.

musiXmatch—The MXM dataset provides lyrics for many Million Song Dataset tracks. The lyrics come in bag-of-words format: each track is described as the word-counts for a dictionary of the top 5000 words across the set.

SEMAINE—An audio–visual database for building Sensitive Artificial Listener (SAL) agents that can engage a person in a sustained, emotionally colored conversation. 150 participants were involved in the experiment, for a total of 959 conversations with individual SAL characters, lasting approximately 5 min each. Solid SAL recordings are transcribed and extensively annotated: 6–8 raters per clip traced five affective dimensions and 27 associated categories.

GTZAN—This dataset consists of 1000 audio tracks that have been stored in folders that correspond to 10 different music genres like jazz, blues, rock, country, hip hop, classical, disco, reggae, metal, pop, and each folder contains 100 tracks.

MIREX—This dataset contains 903 audio clips, 764 lyrics, 193 MIDI files, and consists of folders named "clusters" which contain subfolders corresponding to different human emotions some of which are Aggressive, Boisterous, Humorous, Silly and Passionate. The number of songs in each folder varies.

Music Audio Benchmark Dataset—This dataset contains 1886 songs distributed across different genres, namely alternative, blues, electronic, folkcountry, funksoulrnb, jazz, pop, raphiphop, and rock with each genre containing varying number of samples.

7.3 Datasets that relate physiological signals with music mood

PMEmo—The dataset consists of emotion annotations of 794 songs as well as the simultaneous electrodermal activity (EDA) signals of the 457 subjects that were recruited for the experiment.

DEAM—The dataset contains 1802 excerpts and full songs annotated with valence and arousal values both continuously (per-second) and over the whole song. The emotional annotations are collected with the goal of detecting the emotions that are expressed by the music and musicians from the content.

Pros of the dataset used

- Few of the datasets used in the studies are benchmarked and certified by various authorities.
- DEAM, Million Song Dataset, Emotify provide a vast amount of data for computation.
- Last.fm has a very large database of songs and gives the most suitable tag according to the user's preference.
- DEAP, SEED, NMED, TUH EEG, Bern-Barcelona, and Mahnob-HCI provide high-quality EEG data generated from various subjects under various circumstances.

- As these datasets are regularly used by various researchers, it is easier to compare the finding, accuracy, and suitability of the technique.

Cons of the dataset used

- Most of the researchers prefer to develop their corpus, which is not publicly available; hence, it is not feasible to work with their datasets.
- Some datasets are too large and require a considerable amount of disk space.
- A few datasets do not provide the audio files, while this makes them corpus storage-friendly; a few researchers might find these unusable.

8 Music mood and human emotion recognition

8.1 Subjective annotations

Emotions are subjective, a song, for example, can induce different feelings in different people at the same time, and can also induce different feelings within the same person at different periods of time, and hence, collection of the ground truth data should be conducted carefully. Existing methods usually follow two lines, Expert-based suggestion or Subject-based selection. In the first method, an expert annotates emotion to a certain musical piece and simply abandons those pieces which cannot be annotated by him, whereas the second method is more about majority and consensus. Usually, the ground data are developed by averaging opinions of all subjects or the expert, whichever scheme is followed.

It is quite evident that the task of providing annotation is not only quite daunting but can also be time-consuming and cost-ineffective as well, and for this very reason, many researchers often try to use various techniques to make the task more user-friendly, time-friendly, and cost-friendly. Some popular examples are reducing the length of the music pieces [93, 115]; using exemplar songs to better articulate what each emotion class means [34]. Allowing the user to skip a song when none of the candidate emotion classes is appropriate to describe the effective content of the song [35]; designing a user-friendly annotation interface [116]. Traditional methods of data collection, such as the hiring of subjects or even experts, can be flawed, since labelling tasks are time-consuming, tedious, and expensive [114]. Recently, a significant amount of attention has been placed on the use of collaborative online games to collect such ground truth labels for difficult problems, so-called “Games with a Purpose”. Several such games have been proposed for the collection of music data, such as MajorMiner [66], Chorlody [59], Listen Game [107], and TagATune [55]. These implementations have primarily focused on the collection of descriptive labels for a relatively short audio clip [49].

Fernandez-sotos et al. [24] and Ito et al. [37] did a questionnaire for collecting data, questions were formed and presented to subjects, and they provided answers to their knowledge. Some researchers compiled their data by testing on subjects like Alakus et al. [125] collected phone data of 28 students from their university to predict future mood, health, and stress rate, and similarly, Andersson et al. [7] and Kim et al. [50] studied their 150 consumers and 100 users for their respective studies. Other music-based researches collected music data from sites such as last.fm[1], musiXmatch [105], to collect songs from different Genres or in researches like Besson et al. [11] collected 200 excerpts from operas from 16 participants. These annotations are mostly situation and experiment dependent to enhance the reliability of the emotion annotations, the subjective annotation is rarely longer than an hour. The number of songs a subject is asked to annotate is accordingly limited.

Even though various methods and techniques are being implemented, there are still various hurdles and doubts that have been ever-present like: Should the subjects be asked to deliberately ignore the lyrics? Are songs of a foreign language better for the subjects to eliminate the influence of lyrics? Which song to prefer, one with which the subject is familiar, or a completely new one? There seems to be no consensus on these issues so far. To deal with such difficulties, a recent trend is to obtain emotion tags from music websites that provide tags to songs according to the preferences of their users such as AMG and Last.fm. Typically, this can be done by a simple script-based URL lookup and is quite popular because of an abundant dataset. However, the quality of annotations might get severely hampered.

Annotations for physiological signals are a relatively tougher task compared to music, the power spectrum needs to be pre-processed, and providing annotations to it requires computational power and is quite daunting. As the signals cannot be effectively interpreted by humans, several attempts are being done to automate the process, multiple software and frameworks have been designed specifically for this purpose. EEG-Annotate [129] is a MATLAB coded tool for identifying particular responses and labelling them in unknown continuous EEG signals. The tool provides an easily accessible means to categorize data in a less-controlled environment. It incorporates EEGLAB [131] for pre-processing and results in visualization. MNE-Python in a python library currently under development, which has stable versions that can be used to label, visualize, and analyze EEG data obtained from humans. Similar tools are also available for ECG annotation; for example, WaveformECG is an open-source platform supporting interactive analysis, visualization, and annotation of ECGs. Using a Java backend and PostgreSQL server for data collection, an Apache-Hadoop powered server is used to collect, process, and visualize ECG data, and it follows the protocols laid by PhysioBank. Ghosh et al. [132] used wrist band for leveraging physiological signals and then made an annotation system for GSR, heart rate, and blood pressure.

8.2 Data collection

Researchers frequently used established labelling websites like AMG and last.fm to provide a label to any song or piece of music, considering that annual labor is user intensive, these websites have risen in popularity. Another method that is frequently used is using a database already developed during other studies, while this allows a person to compare accuracies and results to previous work, availability, and feasibility are the two major constraints. Another hurdle is that there is no universally accepted taxonomy for the purpose of human emotion detection, and some taxonomies are based on the basic emotions proposed by psychologists, while some are derived from clustering effective terms. One must also consider the fact that with the number of increasing digital tracks, it is becoming more and more difficult to sort them into different categories/genres/labels. While there are a few benchmarked datasets available, they often do not have a lot of annotated clips or are many a time recorded for a limited scope. Due to the absence of researcher-friendly datasets, it has been observed that researchers often tend to compile their own databases due to the lack of availability of a centralized database. Several reasons can be held accountable for hindering the development of a common database. First, the approach adopted by every research team is unique and there exists no common ground for agreement on the techniques that should be employed for emotion classification or the model of emotion that should be used. Second, violation of copyright issues poses an obstacle in gaining access and utilizing the audio, video clips, or even the medical databases important for the purpose of comparative analysis unless it has been made accessible for public use by the respective organization.

For physiological signals, there are abundant datasets, but some are not research-friendly, as most of the available datasets is of patients going through physiological signal scans [8], 30, 30], those datasets cannot be incorporated in every study; moreover, data obtained in controlled environment use multiple electrodes and signal filtering for obtaining EEG, ECG, SGR, and EMG outputs, data obtained from several electrodes cannot be replicated in every study because of the cost and hardware limits. Despite all the problems, the abundance of research in this track has resulted in multiple datasets specifically meant for studies. Moreover, cheaper hardware with lesser functionalities has been introduced commercially which can be used by researchers for an introductory study without the need of arranging a lot of funds. Examples are Nerosky Mndwave, Emotiv Insight, Intera Xon Muse (for EEG), Empatica E3 wristband, Shimmer3 GSR + (for GSR), Shimmer3 ECG, and Neurosky CardioChip (for ECG).

8.3 Data pre-processing and feature extraction

The raw data acquired while data collection is usually converted into a more useful and understandable form by data pre-processing. To compare the music clips fairly, music pieces are normally converted to a standard format (e.g., 22,050 Hz sampling frequency, 16-bit precision, and mono channel), although many times people convert them to other formats as per their needs. For example, each file can be converted to a 128-kbps format with a base rate of 44,100 Hz. [12, 92, 108] Moreover, since complete music pieces can contain sections with different emotions, a 20–60-s segment that is representative of the whole song is often selected to reduce the emotion variation within the segment and to lessen the burden of emotion annotation on the subjects. After normalizing basic features of songs such as Pitch, Tempo, Timber, parameters like zero crossing Rate, Flux, are extracted in accordance with the taxonomy followed. For the purpose of feature extraction, PyAudioAnalysis[3] and LibRosa[4] are mostly used. LibRosa is also used to determine the tonnetz of the musical files, which acts as another important feature. Pre-processing is one of the initial steps, and thereafter, various procedures are followed. It is quite evident that peak normalization and dynamic range compression normalization are most frequently used for removing distortions such as clipping, and noise variations can be removed. Furthermore, Root-Mean-Square normalization is also used by measuring negative and positive of a sinusoidal signal [47]. Regarding EEG and its feature extraction, as the output is full of noise and artifacts, although authors try to avoid artifacts (such as eye blinks and twitching of muscles) by paying attention to the posture, they may still occur, various techniques have been used and suggested to counter them.

Methods such as Blind Source Separation (BSS) (20%) and Independent Component Analysis (ICA) (10%) were applied to remove eye movements, blinking, twitching, loud heartbeats, and other distortions. Most of the works usually reference the electrodes being used, for example, Common Average Reference (CAR), or Average Mean Reference (AMR) (5.9%), or Laplacian (23.6%) [4]. Features essential for the purpose of our study are then extracted using various

[3] Available at https://github.com/tyiannak/pyAudioAnalysis (Last Checked—6-Jan-2020).

[4] Available at https://librosa.github.io/librosa (Last Checked—6-Jan-2020).

algorithms and techniques. Although, the criterion for feature selection may largely vary depending upon the objective of the research work. This is followed by the normalization of the input data which helps in filtering out the undesirable components like noise from it often which is carried out with the help of band-pass filters, band-stop filters, moving average filter (MAF), and high-pass FIR filter [47]. Bhatti et al. [12] extracted time, frequency, and wavelet domain features of the EEG signals using Short Time Fourier Transform (STFT). Petrantonakis et al. [83, 84] utilized the statistical and wavelet-based features of the EEG signal using High Order Crossing Analysis (HOC) and Hybrid Adaptive Filtering (HAF). Lan et al. [134] proposed a real-time system using the most stable features of EEG using Fractal Dimension (FD), five statistics features (standard deviation, mean of absolute values of the first differences, mean of absolute values of the first differences of normalized EEG, mean of absolute values of the second differences, mean of the absolute values of the second differences of the normalized EEG), 1st order Higher Order Crossings (HOC), and four band power features (alpha power, theta power, beta power, and theta/beta ratio), a subject dependent model was trained. Zheng et al. [119] worked on differential entropy and in [120] systematically evaluated the performance of six popular features: Power Spectral Density (PSD), Differential Entropy (DE), Differential Asymmetry (DASM), Rational Asymmetry (RASM), Asymmetry (ASM) and Differential Caudality (DCAU) features from EEG. Specifically, PSD was computed using Short Time Fourier Transform (STFT); DE was equivalent to the logarithmic power spectral density for a fixed length EEG sequence; DASM and RASM features were the differences and ratios between the DE features of hemispheric asymmetry electrodes; ASM features were the direct concatenation of DASM and RASM features. Sourina et al. [99, 100] introduced a system based on Fractal Dimensions but rather used arousal-valence model for emotion recognition. Using only three electrodes, Higuchi and box-counting algorithms were used for the EEG analysis and comparison. Wu et al.[133] used two channels of frontal EEG signals for emotion detection, using spatial, frequency, and asymmetry characteristics of EEG signals. Lu et al. [63] examined the intensity, timbre, rhythmic features for the purpose of analysis whereas Kim et al. [48] utilized features like mean, max, and standard deviation using frequency–time analysis.

8.4 Model selection

Our study involves the examination of several models of emotion that attempt to classify human emotions and mood of music accurately. To understand the effects of Tempo and Rhythmic effect in music on mood are presented by Fernandez-Sotos et al. [24]. They used the Circumplex model and depicted tempo and rhythmic units alongside the primary emotions of Russell's model. Russell's model is still being used in a wide range of studies due to its much easy dimensional approach, for the purpose of feature selection and comparing EEG signals [125]. Schatter et al. [91] applied the previously mentioned model and the Thayer's model independently for selecting mood for speech recorded musical collection to analyze emotions from the speech itself and when classifying music using both audio and music. Laurier et al. [54] used this model for classifying them in accordance with mood, providing better accuracies using the combination. Kim et al. [50] used Thayer's model for real-time mood and Liu et al. [60] made use of a model for automatic mood detection from acoustic music data. Li et al. [57] conducted similarity search and emotion detection with the help of Hevner's model, concluding that emotion detection from music is difficult than similarity search for music after getting lower accuracies for the same music database. Lang's model helped Wagner et al. [111] to classify into several moods the features extracted from physiological signals recorded from EMG, ECG, SC, and RSP beside three pattern recognition methods. As our study revolves around emotion detection as the result, these models can be used alongside physiological signals as well to obtain an emotional response.

8.5 Model training

A variety of classifiers have been employed along with machine learning algorithms and techniques for the purpose of classification or labelling and to train the system to work upon loads of unlabelled data. It was observed that models unique to subjects are also designed. Lan et al. [134] designed an emotion recognition model unique to subjects and hence is needed to be trained again for every new user. It was observed that for both music classification and human emotion classification using physiological signals, Support Vector Machine was the most used classifier in the majority of the cases and yields good accuracies over varying datasets and constraints. Aguilar et al. [23] used an advanced quadratic SVM classifier to obtain max accuracy of 87.4%, and Atkinson et al. [10], Naji et al. [73, 74] used SVM using RBF kernels. Sourina et al. [99] used SVM on IADS dataset for emotion detection using arousal-valence model, Kim et al. [48] used SVM for emotion detection using short-term monitoring of physiological signals, and Lin et al. [58] used SVM in his work and used EEG for emotion recognition. Zong et al. [121] used SVM along with HHT. Takahashi [103] also showed that SVM is most suitable for emotion detection. Li et al. [57] used SVM for similarity search and obtained accuracies up to 86%. Though SVM is still a useful technique, the newer versions of SVM are often preferred over the traditional algorithm, Jatupaiboon et al. [41]

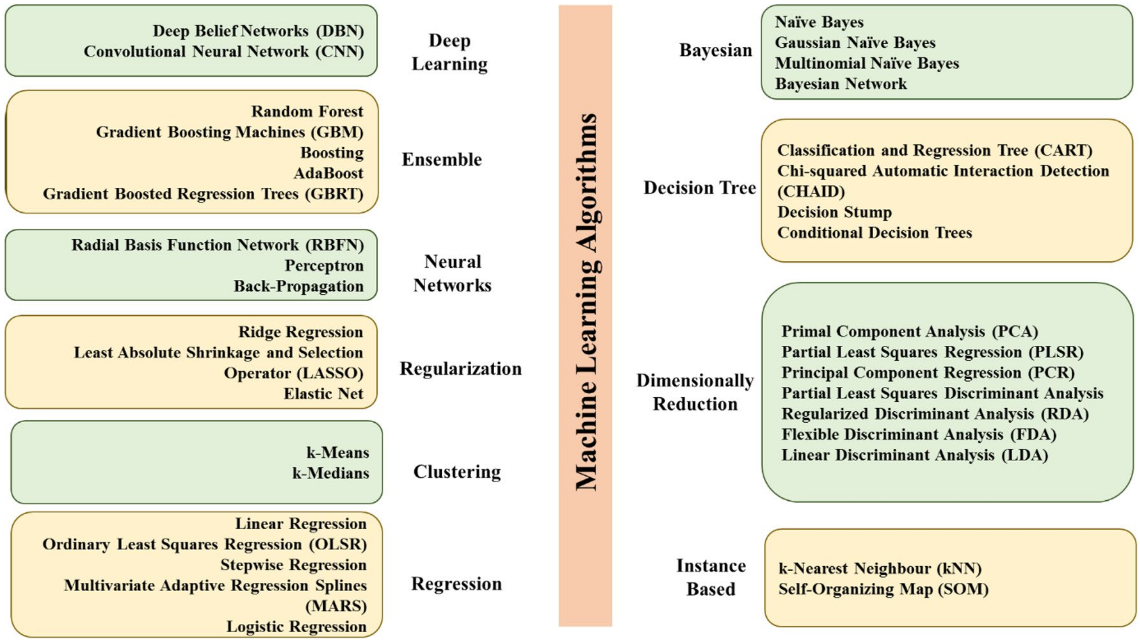

Fig. 5 Classification of most widely used machine learning algorithms in this study

used Gaussian SVM, Brown et al. [16] used Linear SVM, and Vijayan et al. [110] used multiclass SVM or MC-SVM. Soleymani et al. [95] combined PCA and Fisher Criterion with SVM and compared it with the traditional technique for facial features. Amin et al. [130] also used the SVM classifier with EEG signals over 48 practice patterns to classify the EEG outputs and obtained the highest accuracy of 98%.

Neural Network specifically k-NN was also used frequently by researchers, because it does not require any explicit specification and uses self-adaptive techniques to adjust the data. Lahane et al. [52] used k-NN for stress analysis using Teaser–Kaiser Energy operator. Alakus et al. [3] and Khosrowabadi et al. [45] used k-NN to classify eight different emotions using three different wavelet filters. Nasoz et al. [77] and Wagner et al. [111] used k-NN for emotion recognition using physiological signals. Hadjidimitriou et al. [28] used both SVM and k-NN for classifying vectors into two categories. Convolutional Neural Networks or CNN is also amongst the most used classifiers considering, CNNs build their own features from a raw signal and they rely on spatial features; for these very reasons, Yildirim et al. [117] preferred CNN over other classifiers for abnormal signal detection from EEG. Acharya et al. [2] used CNN for classification of ECG signals; similarly, Stober et al. [101] used CNN for feature learning from EEG recordings. Other variants of Neural Networks like DNN (Pandey et al. [80]) and RNN were also used when working with EEG. Along with SVM, Artificial Neural Networks, Naïve Bayes, k-means clustering, Decision Trees, Linear Discriminant analysis, Fuzzy Logic, and Ensemble learning are the most popular techinques [126]. The art of machine learning has led to the development and application of different techniques, which has made it possible for the computer to analyze and learn information from a given set of data, and make the desired prediction accordingly. By providing a sufficient amount of data, along with the help of precision coding and algorithms, it has been possible to train a computer in determining the exact output as required. Although it has been observed that to remove the possible discrepancies and drawbacks of the algorithms employed, researchers often selected a group of algorithms and clubbed them together as per the scope of the study to obtain better results. Classifiers like QDA, Mahalanobis distance have been employed by Petrantonakis et al. [83, 84], Hadjidimitriou et al. [29], and Jatupaiboon et al. [41]. Moreover, combinations of existing classifiers like MC-SVM, GA-SVM, QSVM, and LSTM-RNN have also been employed in different studies. SVR, SMO, CCRF, and MLP were also used, but are quite rare and hence probably are not the best option. Interestingly, Luo et al. [64] used combinations of relatively latest attention-based models such as Serial Linear Attention (SLA), Parallelized Linear Attention (PLA), and Parallelized CNN Attention (PCNNA) alongside Bidirectional RNN and obtained significant results, an important take away being that multilayer BRNN mostly outperforms single-layer BRNN.

Figure 5 contains a list of most widely used machine learning algorithms that were encountered during the study.

Table 3 Details of feature extracted and classifiers used by different researchers along with the accounted result in terms of accuracy

Citation	Features	Classifier	Number of emotions	Accuracy
Studies solely focused on music mood and its classification				
[25]	MFCC, tree quantizer	–	–	Cosine distance gave the best result
[43]	Octave-based spectral contrast feature	GMM with 16 components with expectation minimization	–	Average accuracy: 82.3% on 10-s clips 90.8% on whole clips
[54]	Timbral, rhythmic, tonal	–	–	–
[57]	MFCC, timbral features, DWCHs	SVM	6	70–86%
[63]	Intensity, timbre, and rhythm features, K–L transform	–	4	81.50%
[64]	Intensity, timbre, intensity	Simple Bayesian criteria	4	Overall classification accuracy: 86.3%
[79]	–	SVM	8	33.3–73.3%
[81]	Temperature, humidity, noise, illumination	Bayesian network, Sheffe's paired comparison	–	Used comparison test to determine, recommendation system better than subjective preferences
[108]	Timbral, rhythmic, pitch	Simple Gaussian classifier, GMM, k-NN	–	Music classification 86%
Studies solely focused on human emotions				
[9]	Pitch frequency, intensity, amplitude, voiced vs unvoiced	SOM, XY-fused	2 classes of emotion (arousal, valence)	SOM—68.39 ± 3.04% XY-fused—71.43 ± 2.23%
[91]	Temporal and spectral features	–	12	–
Studies primarily focused on human emotions in relation with music mood				
[24]	Tempo, rhythmic unit	–	5	–
[87]	Mean, min, max, standard deviation, quadrature pair using spatial temporal box filters (STBF)	Classifier constructed from a subset of all possible STBFs	Berlin Dataset: 7 emotions ORATOR Dataset: 7 emotions	Excellent performance
Studies solely focused on physiological signals				
[20]	ES, DE, DASM, RASM	k-NN, SVM	Stress	67–81%
[40]	Skin conductance, skin temperature, accelerometer data, GPS location data, phone data, surveys data, weather data, mood and wellbeing labels (daily)	Traditional-GP, NN; Personalized DA-GP, MTL-NN	2	Lower average error: 13–22%
[101]	Discriminative features	CNN	–	Determined best model using Hyper-parameter optimization
[117]	–	k-NN, Random Forest, CNN	–	79.34%
Studies primarily focused on human emotions and related physiological signals				
[3]	3 wavelet coefficients (Symlet, Daubechies, Morlet)	k-NN	3	79%
[10]	Statistical, band power, Hjorth parameter, fractal dimension	nRMR, GA-SVM	2	60.7–73.14%

Table 3 (continued)

Citation	Features	Classifier	Number of emotions	Accuracy
[15]	Fast Fourier Transform (FFT), binary linear FDA	–	Arousal valence model	Classification Rate – Modality: 82.1% Arousal: 92.3% Valence: 94.9%
[16]	Alpha power ratio, band power	QDC, SVM, k-NN	2 classes of emotion (arousal, valence) + neutral	82%
[23]	17 features including 5 statistical, 8 morphological and 4 temporals	Advance quadratic SVM	7	68.9%-87.4% for different emotions
[41]	Asymmetry, common spatial pattern, high order crossing, self-organized maps, higher order spectral	SVM, NB, QDA, k-NN, LDA, MLP	2	65.12–75.62%
[45]	Magnitude squared coherence estimate	k-NN	4 classes of emotion	84.50%
[48]	Mean, max, amplitude, standard deviation using frequency–time analysis	SVM	4	Classification ratio: 3 emotions: 78.43% 4 emotions: 61.76%
[52]	Energy, amplitude, frequency	SVM, k-NN, ANN, classification tree	2	83–93%
[60]	Time–frequency analysis using Fourier	LDA, SVM	7	86.43%
[75]	Gabor, wavelet coefficients	NDB, MLB, PNN and PSO, MEDO, IPSO (For Optimization)	4	Avg. Classification Accuracy: 64.78%
[77]	–	KNN, DFA	6	KNN—performed better for frustration, surprise; DFA—performed better for fear, sad, anger, amusement
[80]	Wavelet coefficients	DNN	4	57.67–64.25%
[83]	Statistical, wavelet-based features using high order crossing analysis	QDA, KNN, Mahalanobis distance, SVM	6	QDA: 62.30% SVM: 83.33%
[84]	Statistical, wavelet-based features using high order crossing analysis and hybrid adaptive filtering	QDA, KNN, Mahalanobis distance, SVM	6	QDA: 77.66% SVM: 85.17%
[100]	Fractal dimension	SVM	2 classes of emotion (arousal, valence)	73%
[110]	Mean, standard deviation using short time Fourier transform (STFT)	MC-SVM	4	94.10%
[111]	Statistical feature, amplitude, breathing rate, heartbeats	Linear discriminant function, k-NN	4	Recognition rates: 80% After Feature Reduction—92%
[119]	Differential entropy	DBN, SVM, LR, k-NN	2 classes of emotion (arousal, valence) + Neutral	72–86%
[120]	Power spectral density, differential entropy, differential asymmetry, rational asymmetry, asymmetry, differential Caudality	Fourier Transform, Hanning Window for feature extraction k-NN, SVM, LR, Graph regularized Extreme Learning Machine	3 Labels—positive, negative, and neutral	Max up to 90%

Table 3 (continued)

Citation	Features	Classifier	Number of emotions	Accuracy
[121]	HHT	SVM	–	Baseline: 71% Fission based: 76% Fusion based: 62%
[125]	DWD, DFA, H, Avg Eng, ShanEn, MSCEn, Std Dev, V, ZC using sequential forward selection	SAM, Bayesian Classification, QSVM	–	Average Accuracy Without Feature Selection—55 With Feature Selection—65
Studies primarily focused on music mood and observed physiological signals				
[28]	Spectrogram, HHT, ZAMT	k-NN, SVM	2 categories (based on liking)	78.90–91.02%
[29]	Frequency bands, reference states, time intervals, hemispheric asymmetries using time–frequency analysis	KNN, QDA, Mahalanobis distance, SVM	2 categories (based on liking)	86.52 ± 0.76% (using K-NN)
[37]	Fractal dimension, time and frequency feature	NN, FA, BP	–	–
[38]	Prefrontal cortex activity using genetic algorithm and fast Fourier transform (FFT)	–	–	Greater than 70%
Studies that relate human emotions with psychological signals and music as stimuli				
[12]	Time, frequency and wavelet domain features using short time Fourier transform	MLP, K-NN, SVM	4	MLP is most accurate
[46]	Prefrontal, frontal, central, parietal features using kernel smoothing density estimation (KSDE), Gaussian mixture model	BN, MLP, One-rule, random tree, radial basis function	6	leave-one-out accuracy: 90% intra-subject accuracy: 69.69%
[58]	DASM12, RASM12, PSD24, PSD30 using short time Fourier transform	SVM	4	82.29 ± 3.06%
[73]	Relative power, mean, Ang. nonlinear energy, high order crossing, poincare geometry, entropy, space mapping	SVM	2 classes of emotion (arousal, valence)	Avg. Accuracy: 88.78% Avg. Valence Accuracy: 94.91% Avg. Arousal Accuracy: 93.63%
[74]	Relative power, mean, Avg. nonlinear energy, high order crossing, entropy	SVM, k-NN, CFNN	2 classes of emotion (arousal, valence)	Avg. Accuracy: 87.05% Avg. Valence Accuracy: 93.66% Avg. Arousal Accuracy: 93.29%
[76]	Dual-tree complex wavelet packet transform (DT-CWPT)	SVD, QR factorization with column pivoting (QRcp), F-Ratio + SVM	2 classes of emotion (arousal, valence), dominance and liking	Valence: 64.3%, Arousal:66.2%, Dominance: 68.9% Like: 70.2%
[106]	Prefrontal cortex (PFC)	SVM	2 classes of emotion (arousal, valence)	> 80%

9 Results and discussion

To classify the different music moods and human emotions, different features are extracted upon which different classifiers are employed and their performance is evaluated based on accuracy metrics whose brief is specified in Table 3, and the discussion on the experimental results is done in Sects. 9.1–9.7.

9.1 Music mood and its classification

Considering the studies that primarily focused on studying Music Mood and classifying it using different existing taxonomies, it is observed that timbre and rhythm features that are extracted using Fast Fourier Transform (FFT) were a predominant feature along with Mel-frequency cepstral coefficients. Most studies focused on similar feature extraction techniques and utilized Fast Fourier and Discrete Cosine Transform for the purpose of pre-processing. K–L Transform is also used for exploiting the existing relativity among the different dimensions of raw features of music [63]. In terms of music classification, different approaches resulted in different accuracies, while an Octave-based spectral contrast [43] representation led to a high accuracy of 90.8% using the GMM model with 16 components. Other algorithms like SVM and Bayesian networks are more popular relatively but result in sub-par accuracies [57, 64, 79, 81].

9.2 Human emotions

In the studies that exclusively focused upon human emotions, Temporal and Spectral features and other features like Pitch, Frequency, Intensity, and Amplitude are taken into consideration. One of the studies utilized Self-Organizing Maps (SOM) and X–Y Fused networks (XYFs) to classify emotions into two classes and achieved accuracies of $68.39 \pm 3.04\%$ and $71.43 \pm 2.23\%$ respectively [9].

9.3 Physiological signals

Some of the studies are solely dedicated to exploring the different kinds of physiological responses and they focused upon EEG-related features like Differential Entropy (DE), DASM, and RASM, other physiological signals like Skin Conductance and skin temperature along with certain data that involved Accelerometer data, GPS location data, phone, and weather data [20, 40]. The most commonly used classifiers observed are k-NN, MTL-NN, and CNN. The highest accuracy of 79.34% is obtained when a combination of k-NN, CNN, and Random Forest is employed which is followed by an accuracy of 67–81% achieved through a combination of k-NN & SVM while identifying stress in the subjects [20, 117].

9.4 Human emotions in relation with music mood

For the studies that examined human emotions in relation to music mood, statistical features like mean, max, min, and standard deviation, and other features such as Tempo and rhythmic unit are explored [24, 87]. In particular, one of the studies utilized classifiers that are constructed from subsets of Spatial Temporal Box Filters (STBFs) produced excellent results in terms of accuracy while classifying human emotions into seven categories [87].

9.5 Human emotions and related physiological signals

For studies that focus upon human emotions and related physiological signals, it is observed that statistical, wavelet-based features, and features related to energy, power, and entropy are most commonly taken into consideration for different studies along with certain physiological signals like heartbeats and breathing rate in some of them. A variety of techniques like frequency–time analysis, Fourier Transforms, HOC analysis, HAV, and SFS are employed to extract the aforementioned features. Classifiers like k-NN, SVM, and QDA are predominantly utilized in these studies for the purpose of classifying emotions. The highest accuracy of 94.10% is obtained after using the MC-SVM classifier to classify emotions into four classes using STFT [110].

9.6 Music mood and observed physiological signals

For the studies focusing upon linking physiological signals observed in subject and the mood of the song used as stimuli, we observe that the features which were predominantly utilized include time and frequency features and physiological data like Hemispheric Asymmetries and Prefrontal Cortex Activity which are obtained using different methods like time–frequency analysis and FFT [29, 37, 38]. The most commonly used classifiers are k-NN and SVM and studies that involved using a combination of the two classifiers resulted in high rates of accuracy ranging from 78.9 to 91.0% while classifying into two classes of emotions [28, 29].

9.7 Human emotions with physiological signals and music as stimuli

Many studies had employed music as a stimulus to generate certain physiological responses in a subject and then record their emotions accordingly. Their analysis primarily has been based upon time, frequency, wavelet domain features, and other features related to energy, power, and entropy.

These studies also focused upon physiological data like EEG-related features which included DASM12. RASM12, PSD24, PSD30, and other features like prefrontal, frontal, central, and parietal features [46, 58]. The techniques that are utilized to extract these features involved STFT, GMM, and Kernel Smoothing Density Estimation (KSDE). It is observed that SVM and k-NN are the most commonly used classifiers in most of these studies. Most of these studies have reported high rates of accuracy while classifying emotions into different classes varying in number from two to six. The highest average accuracy achieved is 88.78% alongside an average valence accuracy of 94.91% and an average arousal accuracy of 93.63%, which are obtained through the sole use of SVM.

10 Limitations and future scope

The present study has tried to identify gaps in the existing research and correlate them with the practices related to musical variables. The study suggests that more research is required to correctly examine these relationships to magnify the outreach and feasibility of any such future work. The study also touches on various upcoming possibilities for future researches and some of the most important points are listed below,

- Studies on musical construct have been largely on structural elements. There is a need to explore the relative contribution of different compositional variables on consumption experience.
- The present study is confined only to the reported studies. Other researches based on physiological signals like Electrooculography (EOG), Phonocardiogram (PCG), Photoplethysmogram (PPG), Optoacoustic, and Speech recognition can be included in future work.
- The present study has reported the effects of music on customers, particularly in retail settings. The effect of music on employee behavior and its interrelationships with shopping experience can be explored with a combination of moderating variables.
- The studies on the effect of presence versus absence of music in emotion recognition have not been thought upon. Experimental studies are required to explore this dimension in greater detail.
- Various studies have either used one or two features in their works; with availability of refined and pre-processed datasets, the possibilities to use more features have certainly increased.
- Effect of music on behavioral responses and personality traits have not been thoroughly studied; these traits can induce a long-term emotional change in a person and needs to be examined in greater detail. Experimental studies can measure the effect of various musical variables in a controlled manner on a person's behavior in different environment resulting in change of preferences and daily lifestyle (food habits, dressing sense).
- The studies have not used the most recent classifiers and result boosting techniques, with availability of AdaBoost, LP Boost, and Gradient Boosting with the help of which better results can surely be expected in future works.
- Although a few studies touch up the field of mental health and its relationship with music, the nature, and intensity of emotions aroused using different musical variables and their effect on mental well-being such as treatment of stress, depression or migraine can be further explored.
- Interactive effects of music with other ambient factors such as language of song, surroundings, peripherals used, and visual appeal can be explored regarding different types of emotions that can be induced in terms of above-mentioned factors.
- Music and songs can revolutionize the field of surgery and treatment if used accordingly. Music therapy can help in reducing anxiety cases during surgeries [14]; although recently people are working on this field, it is still a new field with various possibilities.

11 Conclusion

Despite all the major advancements in the field of MER and real-time emotion detection, the field is still quite new, with many yet to be solved issues.

Be it lack of universally accepted taxonomy, problems regarding datasets, or simply the vastness and unlimited possibilities, there are still major milestones yet to be achieved. It is quite evident from our present study that the induction of certain emotions like happiness, sadness, anger, and calmness has been more frequently recorded in comparison to others. As stated earlier, substantial amounts of research work have been carried out to devise several taxonomies or models for classification of human emotions, but the majority of them have only succeeded in the identification of nearly 4–5 emotions on an average accurately. Adding to the fact that there exists no universally accepted taxonomy for our purpose only leads to every researcher having a different way of approaching our subject of concern. Perhaps, researchers should extend their experiments to capture these lesser identified emotions to broaden this narrow range of classification. This will gradually allow us to obtain a greater insight into a person's emotional behavior and hence will help us establish certain connections between music and human psychology.

References

1. Abadi, M.K., Subramanian, R., Kia, S.M., Avesani, P., Patras, I., Sebe, N.: DECAF: MEG-based multimodal database for decoding affective physiological responses. IEEE Trans. Affect. Comput. **6**(3), 209–222 (2015). https://doi.org/10.1109/TAFFC.2015.2392932
2. Acharya, U.R., Hagiwara, Y., Deshpande, S.N., Suren, S., Koh, J.E., Oh, S.L., Arunkumar, N., Ciaccio, E.J., Lim, C.M.: Characterization of focal EEG signals: a review. Future Gener. Comput. Syst. **91**(2019), 290–299 (2018). https://doi.org/10.1016/j.future.2018.08.044
3. Alakuş T.B., Turkoglu I. (2018). Determination of accuracies from different wavelet methods in emotion estimation based on EEG signals by applying KNN classifier. In: 3rd International Conference on Engineering Technology and Applied Science, pp 250–254 (2018)
4. Alarcão, S.M., Fonseca, M.J.: Emotions recognition using EEG signals: a survey. IEEE Trans. Affective Comput. **10**(3), 374–393 (2019). https://doi.org/10.1109/TAFFC.2017.2714671
5. Aljanaki, A., Yang, Y.H., Soleymani, M.: Developing a benchmark for emotional analysis of music. PLoS ONE **12**(3), 2017 (2017). https://doi.org/10.1371/journal.pone.0173392
6. Aljanaki, F.W., Veltkamp, R.C.: (2015) Studying emotion induced by music through a crowdsourcing game. Inf. Process. Manag. **52**(1), 115–128 (2015)
7. Andersson, P.K., Kristensson, P., Wastlund, E., Gustafsson, A.: Let the music play or not: the influence of background music on consumer behaviour. J. Retail. Consumer Behavior **19**(6), 553–560 (2012)
8. Andrzejak, R.G., Schindler, K., Rummel, C.: Non-Randomness, nonlinear dependence, and non-stationarity of electroencephalographic recordings from epilepsy patients. Phys. Rev. Am. Phys. Soc. E **86**, 2012 (2012). https://doi.org/10.1103/PhysRevE.86.046206
9. Arnrich, B., Marca, R.L., Ehlert, U.: Self-Organizing Maps for Affective State Detection. **2010**, 1–8 (2010)
10. Atkinson, J., Campos, D.: Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers. Expert Syst. Appl. **47**(C), 35–41 (2016). https://doi.org/10.1016/j.eswa.2015.10.049
11. Besson, M., Faïta, F., Peretz, I., Bonnel, A., Requin, J.: Singing in the brain: independence of lyrics and tunes. Psychol. Sci. **9**(6), 494–498 (1998)
12. Bhatti, A.M., Majid, M., Anwar, S.M., Khan, B.: Human emotion recognition and analysis in response to audio music using brain signals. Comput. Hum. Behav. **65**(2), 267–275 (2016). https://doi.org/10.1016/j.chb.2016.08.029
13. Bittner R., Salamon J, Tierney M., Mauch M., Cannam C., Bello J.P.: MedleyDB: a multitrack dataset for annotation-intensive MIR research. In: 15th International Society for Music Information Retrieval Conference, ISMIR, pp 155–160 (2014)
14. Bojorquez, G.R., Jackson, K.E., Andrews, A.K.: Music therapy for surgical patients. Crit. Care Nurs. Q. **43**(1), 81–85 (2020). https://doi.org/10.1097/cnq.0000000000000294
15. Bos D.O.: EEG-based Emotion Recognition. The Influence of Visual and Auditory Stimuli (2006)
16. Brown L., Grundlehner B., Penders J.: Towards wireless emotional valence detection from EEG. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2188–2191 (2011). https://doi.org/10.1109/IEMBS.2011.6090412
17. Chanel G., Ansari-Asl K., Pun T.: Valence-arousal evaluation using physiological signals in an emotion recall paradigm. In: IEEE International Conference on Systems, Man and Cybernetics, Montreal, pp. 2662–2667 (2007)
18. Chen, Y.-A., Yang Y-H., Wang J-C., Chen H.: The AMG1608 dataset for music emotion recognition. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). QLD, pp. 693–697 (2015). https://doi.org/10.1109/icassp.2015.7178058
19. Defferrard M., Benzi K., Vandergheynst P., Bresson X.: FMA: a dataset for music analysis. In: 18th International Society for Music Information Retrieval Conference 2017 (2017). arxiv.org/abs/1612.01840
20. Duan R., Zhu J., Lu B. Differential entropy feature for EEG-based emotion classification. In: 6th International IEEE/EMBS Conference on Neural Engineering (NER), pp. 81–84 (2013). https://doi.org/10.1109/NER.2013.6695876
21. Eerola, T., Vuoskoski, J.K.: A comparison of the discrete and dimensional models of emotion in music. Psychol. Music **39**(1), 18–49 (2010). https://doi.org/10.1177/0305735610362821
22. Ekman, P.: An Argument for Basic Emotions. Cognition Emotion **6**(3–4), 169–200 (1992). https://doi.org/10.1080/02699939208411068
23. Fernández-Aguilar, L., Martínez-Rodrigo, A., Moncho-Bogani, J., Fernández-Caballero, A., Latorre, J.M.: Emotion detection in aging adults through continuous monitoring of electro-dermal activity and heart-rate variability. IWINAC (2019). https://doi.org/10.1007/978-3-030-19591-5_26
24. Fernández-Sotos, A., Fernández-Caballero, A., Latorre, J.M.: Influence of tempo and rhythmic unit in musical emotion regulation. Front. Comput. Neurosci. **10**(80), 1–13 (2016)
25. Foote, J.: Content-based retrieval of music and audio. Multimed. Storage Arch. Syst. II **3229**, 1997 (1997). https://doi.org/10.1117/12.290336
26. Fu, Z., Lu, G., Ting, K.M., Zhang, D.: A survey of audio-based music classification and annotation. IEEE Trans. Multimed. **13**(2), 303–319 (2011)
27. Gemmeke J.F., Ellis D., Freedman D., Jansen A., Lawrence W., Moore R.C., Plakal M., Ritter M.: Audio Set: an ontology and human-labeled dataset for audio events. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp 776–780 (2017). https://doi.org/10.1109/ICASSP.2017.7952261.
28. Hadjidimitriou, S., Hadjileontiadis, L.J.: EEG-based classification of music appraisal responses using time-frequency analysis and familiarity ratings. IEEE Trans. Affect. Comput. **4**(2), 161–172 (2013). https://doi.org/10.1109/T-AFFC.2013.6
29. Hadjidimitriou, S.K., Hadjileontiadis, L.J.: Towards an EEG-based recognition of music liking using time-frequency analysis. IEEE Trans. Biomed. Eng. **59**(2012), 3498–3510 (2012). https://doi.org/10.1109/TBME.2012.2217495
30. Harati A., Lopez S., Obeid I., Picone J., Jacobson M., Tobochnik S.: The TUH EEG CORPUS: a big data resource for automated EEG interpretation. In: IEEE Signal Processing in Medicine and Biology Symposium, IEEE SPMB (2014). https://doi.org/10.1109/SPMB.2014.7002953
31. Hevner, K.: Experimental studies of the elements of expression in music. Am. J. Psychol. **48**(2), 246–268 (1936)
32. Homburg, H., Mierswa, I., Moller, B., Morik, K., Wurst, M.: A benchmark dataset for audio classification and clustering. Int. Symp. Music Inf. Retrieval **2005**, 528–531 (2005)
33. Hosseini, S.A., Naghibi-Sistani, M.B., Rahati-quchani, S.: Dissection and analysis of psychophysiological and EEG signals for emotional stress evaluation. J. Biol. Syst. **18**(spec01), 101–114 (2010)
34. Hu X., Downie J. S., Laurier C., Bay M., Ehmann A.F.: The 2007 Mirex audio mood classification task: Lessons learned.

In: ISMIR 9th International Conference on Music Information Retrieval, pp. 462–467 (2008). https://doi.org/10.5281/zenodo.1416380
35. Hu X., Sanghvi V.R., Vong B., On P.J., Leong C.N., Angelica J.: Moody: a web-based music mood classification and recommendation system. In: Proceedings of the International Conference on Music Information Retrieval (2008)
36. Huron D.: Perceptual and cognitive applications in music information retrieval. In: ISMIR, 1st International Symposium on Music Information Retrieval (2000)
37. Ito, S., Mitsukura, Y., Fukumi, M., Akamatsu, N.: A feature extraction of the EEG during listening to the music using the factor analysis and neural networks. Proc. Int. Jt. Conf. Neural Netw. **3**, 2263–2267 (2003). https://doi.org/10.1109/IJCNN.2003.1223763
38. Ito S., Mitsukura Y., Fukumi M., Cao J. (2007). Detecting method of music to match the user's mood in prefrontal cortex EEG activity using the GA. In: International Conference on Control, Automation and Systems, Seoul, South Korea, (2007). https://doi.org/10.1109/ICCAS.2007.4406685
39. Jain, R., Bagdare, S.: Music and consumption experience: a review. Int. J. Retail Distrib. Manag. **39**(4), 289–302 (2011). https://doi.org/10.1108/09590551111117554
40. Jaques N., Rudovic O., Taylor S., Sano A., Picard R.: Predicting tomorrow's mood, health, and stress level using personalized multitask learning and domain adaptation. In: IJCAI 2017 Workshop on Artificial Intelligence in Affective Computing, Journal of Machine Learning Research Vol. 66, pp. 17–33 (2017)
41. Jatupaiboon, N., Pan-Ngum, S., Israsena, P.: Real-time EEG-based happiness detection system. Sci. World J. (2013). https://doi.org/10.1155/2013/618649
42. Jerritta S., Murugappan M., Nagarajan R., Wan K.: Physiological signals based human emotion recognition: a review. In: IEEE 7th International Colloquium on Signal Processing and its Applications, pp. 410–415 (2011)
43. Jiang, D., Lu, L., Zhang, H., Tao, J., Cai, L.: Music type classification by spectral contrast feature. IEEE Int. Conf. Multimed. Expo **1**(2002), 113–116 (2002)
44. Juslin, P.N., Laukka, P.: Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening. J. New Music Res. **33**(2004), 217–238 (2004). https://doi.org/10.1080/0929821042000317813
45. Khosrowabadi R., Quek H.C., Wahab A., Ang K.K.: EEG-based emotion recognition using self-organizing map for boundary detection. In: 20th International Conference on Pattern Recognition, ICPR'10. IEEE Computer Society, pp. 4242–4245 (2010). https://doi.org/10.1109/ICPR.2010.1031
46. Khosrowabadi R., Wahab A., Ang K.K., Baniasad M.H.: Affective computation on EEG correlates of emotion from musical and vocal stimuli. In: International Joint Conference on Neural Networks, Atlanta, GA, USA (2009). https://doi.org/10.1109/IJCNN.2009.5178748
47. Kim, D., Koo, D.: Analysis of pre-processing methods for music information retrieval in noisy environments using mobile devices. Int. J. Contents **8**(2), 1–6 (2012). https://doi.org/10.5392/IJOC.2012.8.2.001
48. Kim, K.H., Bang, S.W., Kim, S.R.: Emotion recognition system using short-term monitoring of physiological signals. Med. Biol. Eng. Comput. **42**(2004), 419–427 (2004). https://doi.org/10.1007/BF02344719
49. Kim Y.E., Schmidt E.M., Migneco R., Morton B.G., Richardson P., Scott J.J., Speck J.A., Turnbull D.: State of the art report: music emotion recognition: a state-of-the-art review. In: International Society for Music Information Retrieval Conference ISMIR (2010)
50. Kim Y.E., Schmidt E.M., Emelle L.: MoodSwings: a collaborative game for music mood label collection. In: ISMIR, pp. 231–236 (2008)
51. Koelstra, S., Mühl, C., Soleymani, M., Jong-Seok, L., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., Patras, I.: DEAP: a database for emotion analysis using physiological signals. IEEE Trans. Affective Comput. **3**(1), 18–31 (2011). https://doi.org/10.1109/T-AFFC.2011.15
52. Lahane, P.U., Thirugnanam, M.: Human emotion detection and stress analysis using EEG signal. Int. J. Innov. Technol. Explor. Eng. (IJITEE) **8**(42), 96–100 (2019)
53. Lang, P.J.: The emotion probe: studies of motivation and attention. Am. Psychol. Assoc. **50**(5), 372–385 (1995)
54. Laurier C., Grivolla, N., Herrera P.: Multimodal music mood classification using audio and lyrics. In: 7th International Conference on Machine Learning and Applications, pp. 688–693 (2008). https://doi.org/10.1109/ICMLA.2008.96
55. Law E., Ahn L.V., Dannenberg R.B., Crawford M.J.: TagATune: a game for music and sound annotation. In: 8th International Conference on Music Information Retrieval ISMIR (2007)
56. Lehmberg, L.J., Fung, C.V.: Benefits of music participation for senior citizens: a review of the literature. Music Educ. Res. Int. **4**(2010), 19–30 (2010)
57. Li T., Ogihara M.: Content-based music similarity search and emotion detection. In: ICASSP, pp. 705–708 (1988). https://doi.org/10.1109/ICASSP.2004.1327208
58. Lin, Y.P., Wang, C.H., Jung, T.P., Wu, T.L., Jeng, S.K., Duann, J.R., Chen, J.H.: EEG-Based Emotion Recognition in Music Listening. IEEE Trans. Biomed. Eng. **57**(2010), 1798–1806 (2010). https://doi.org/10.1109/TBME.2010.2048568
59. Liu Y., Yan N., Hu D.: Chorlody: a music learning game. In: CHI Extended Abstracts on Human Factors in Computing Systems, pp. 277–280 (2014). https://doi.org/10.1145/2559206.2580098
60. Liu, Y.J., Yu, M., Zhao, G., Song, J., Ge, Y., Shi, Y.: Real-time movie-induced discrete emotion recognition from EEG signals. IEEE Trans. Affect. Comput. **9**(4), 550–562 (2018). https://doi.org/10.1109/TAFFC.2017.2660485
61. Losorelli S., Nguyen D.C., Dmochowski J.P., Kaneshiro B.: NMED-T: a tempo-focused dataset of cortical and behavioral responses to naturalistic music. In: 18th International Society for Music Information Retrieval Conference, pp. 339–346 (2017). https://doi.org/10.5281/zenodo.1417917
62. Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., Yger, F.: A review of classification algorithms for EEG-based brain–computer interfaces: a 10-year update. J. Neural Eng. **15**(3), 2018 (2018). https://doi.org/10.1088/1741-2552/aab2f2
63. Lu, L., Liu, D., Zhang, H.J.: Automatic Mood Detection and Tracking of Music Audio Signals. IEEE Trans. Audio Speech Lang. Process. **14**(2006), 5–18 (2005). https://doi.org/10.1109/TSA.2005.860344
64. Luo, S., Yu, Y., Liu, S., Qiao, H., Liu, Y., Feng, L.: Deep attention-based music genre classification. Neurocomputing **372**(2020), 84–91 (2020). https://doi.org/10.1016/j.neucom.2019.09.054
65. Ma K., Wang X., Yang X., Zhang M., Girard J.M., Morency L.P.: ElderReact: a multimodal dataset for recognizing emotional response in aging adults. In: 2019 International Conference on Multimodal Interaction, ICMI, pp. 349–357 (2019). https://doi.org/10.1145/3340555.3353747
66. Mandel, M.I., Ellis, D.P.: A web-based game for collecting music metadata. J. New Music Res. **37**, 365–366 (2007). https://doi.org/10.1080/09298210802479300

67. McCraty, R., Barrios-Choplin, B., Atkinson, M., Tomasino, D.: The effects of different types of music on mood, tension, and mental clarity. Altern. Ther. Health Med. **4**(1), 75–84 (1998)
68. McKeown, G., Valstar, M., Cowie, R., Pantic, M., Schroder, M.: The SEMAINE database: annotated multimodal records of emotionally colored conversations between a person and a limited agent. IEEE Trans. Affect. Comput. **3**(1), 5–17 (2012). https://doi.org/10.1109/T-AFFC.2011.20
69. Mehrabian, A., Russell, J.A.: An Approach to Environmental Psychology. MIT Press, Cambridge (1974)
70. Miranda, J., Abadi, M.K., Nicu, S., Patras, I.: AMIGOS: a dataset for mood, personality and affect research on individuals and groups. IEEE Trans. Affect. Comput. **2017**, 1–14 (2017). https://doi.org/10.1109/TAFFC.2018.2884461
71. Moody, G.B., Mark, R.G.: The impact of the MIT-BIH arrhythmia database. IEEE Eng. Med. Biomed. **20**(3), 45–50 (2001)
72. Murthy, Y.V.S., Koolagudi, S.G.: Content-based music information retrieval (CB-MIR) and its applications towards the music industry: a review. ACM Comput. Surv. **51**(3), 1–46 (2018)
73. Naji, M., Firoozabadi, M., Azadfallah, P.: Classification of music-induced emotions based on information fusion of forehead biosignals and electrocardiogram. Cognit. Comput. **6**(2), 241–252 (2013). https://doi.org/10.1007/s12559-013-9239-7
74. Naji, M., Firoozabadi, M., Azadfallah, P.: Emotion classification during music listening from forehead biosignals. Signal Image Video Process. **9**(6), 1365–1375 (2015). https://doi.org/10.1007/s11760-013-0591-6
75. Nasehi, S., Pourghassem, H.: An optimal EEG-based emotion recognition algorithm using Gabor features. WSEAS Trans. Signal Process. **8**, 87–99 (2012)
76. Naser, D.S., Saha, G.: Recognition of emotions induced by music videos using DT-CWPT. Indian Conf. Med. Inf. Telemed. (ICMIT) (2013). https://doi.org/10.1109/IndianCMIT.2013.6529408
77. Nasoz, F., Lisetti, C.L., Alvarez, K., Finkelstein, N.: Emotion recognition from physiological signals for user modelling of affect. Cognit. Technol. Work **6**(1), 4–14 (2003). https://doi.org/10.1007/s10111-003-0143-x
78. North, A.C., Hargreaves, D.J.: (2009) Music and consumer behavior. In: Hallam, S., Cross, I., Thaut, M. (eds.) The Oxford Handbook of Music Psychology, pp. 481–490. Oxford University Press, Oxford (2009)
79. Oh S., Hahn M., Kim J.: Music mood classification using intro and refrain parts of lyrics. In: 2013 International Conference on Information Science and Applications (ICISA), pp. 1–3 (2013). https://doi.org/10.1109/ICISA.2013.6579495
80. Pandey P., Seeja K.R.: Subject-independent emotion detection from EEG signals using deep neural network. In: Proceedings of ICICC 2018, Vol. 2, pp. 41–46 (2018). https://doi.org/10.1007/978-981-13-2354-6_5
81. Park, H.S., Yoo, J.O., Cho, S.B.: A Context-aware music recommendation system using fuzzy Bayesian networks with utility theory. In: Fuzzy Systems and Knowledge Discovery. FSKD 2006. Lecture Notes in Computer Science, vol. 4223. Springer, New York (2006). https://doi.org/10.1007/11881599_121
82. Paul, D.J., Kundu, S.: A survey of music recommendation systems with a proposed music recommendation system. In: Emerging Technology in Modelling and Graphics. Advances in Intelligent Systems and Computing, vol. 937. Springer, New York (2020). https://doi.org/10.1007/978-981-13-7403-6_26
83. Petrantonakis, P.C., Hadjileontiadis, L.J.: Emotion recognition from EEG using higher order crossing. IEEE Trans. Inf Technol. Biomed. **14**(2010), 186–197 (2009). https://doi.org/10.1109/TITB.2009.2034649
84. Petrantonakis, P.C., Hadjileontiadis, L.J.: Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossing analysis. IEEE Trans. Affect. Comput. **1**(2010), 81–97 (2010). https://doi.org/10.1109/T-AFFC.2010.7
85. Plutchik, R.: The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. Am. Sci. **89**(4), 344–350 (2001)
86. Russell, J.A.: Culture and the categorization of emotions. Psychol. Bull. **110**(3), 425–450 (1991)
87. Ruvolo P., Fasel I., Movellan J.: Auditory mood detection for social and educational robots. In: 2008 IEEE International Conference on Robotics and Automation, Pasadena, CA, USA (2008). https://doi.org/10.1109/ROBOT.2008.4543754
88. Scaringella, N., Zoia, G., Mlynek, D.: Automatic genre classification of music content: a survey. IEEE Signal Process. Mag. **23**(2), 133–141 (2006)
89. Schalk, G., McFarland, D.J., Hinterberger, T., Birbaumer, N., Wolpaw, J.R.: BCI2000: a general-purpose brain-computer interface (BCI) system. IEEE Trans. Biomed. Eng. **51**(6), 1034–1043 (2004)
90. Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., Xu, X., Yang, X.: A review of emotion recognition using physiological signals. Sensors **18**, 2074 (2018). https://doi.org/10.3390/s18072074
91. Shuman, V., Schlegel, K., Scherer, K.: Geneva Emotion Wheel rating study (Report). Swiss Centre for Affective Sciences, Geneva (2015)
92. Silverman, M.J.: Effects of live music in oncology waiting rooms: two mixed methods pilot studies. Int. J. Music Perform. Arts **3**(1), 1–15 (2015)
93. Skowronek J., McKinney M.F., Par S.V.: Ground truth for automatic music mood classification. In: Proceedings of the International Conference on Music Information Retrieval, pp. 395–396 (2006)
94. Snyder D., Chen G., & Povey D.: MUSAN: A Music, Speech, and Noise Corpus. ArXiv, abs/1510.08484 (2015)
95. Soleymani M., Asghari-Esfeden S., Pantic M., Fu Y. Continuous emotion detection using EEG signals and facial expressions. In: IEEE International Conference on Multimedia and Expo (ICME), pp. 1–6 (2014). https://doi.org/10.1109/ICME.2014.6890301
96. Soleymani M., Caro M.N., Schmidt E.M., Sha C.Y., Yang YH.: 1000 songs for emotional analysis of music. In: 2nd ACM international workshop on Crowdsourcing for multimedia (CrowdMM'13). ACM, pp. 1–6 (2013). https://doi.org/10.1145/2506364.2506365
97. Soleymani, M., Lichtenauer, J., Pun, T., Pantic, M.: A multimodal database for affect recognition and implicit tagging. IEEE Trans. Affect. Comput. **3**(1), 42–55 (2012). https://doi.org/10.1109/T-AFFC.2011.25
98. Song Y., Dixon S., Pearce M.T.: A survey of music recommendation systems and future perspectives. In: 9th International Symposium on Computer Music Modeling and Retrieval (CMMR'12) (2012)
99. Sourina O., Liu Y.: A fractal-based algorithm of emotion recognition from EEG using arousal-valence model. BIOSIGNALS. In: Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing, pp. 209–214 (2011)
100. Sourina, O., Wang, Q., Liu, Y., Nguyen, M.K.: A real-time fractal-based brain state recognition from EEG and its applications. In: International Joint Conference on Biomedical Engineering Systems and Technologies BIOSTEC, pp. 258–272 (2011)
101. Stober S., Sternin A., Owen A.M., Grahn J.A.: Deep Feature Learning for EEG Recordings. ArXiv, abs/1511.04306 (2015)
102. Subramanian, R., Wache, J., Abadi, M.K., Vieriu, R.L., Winkler, S., Sebe, N.: ASCERTAIN: emotion and personality recognition

using commercial sensors. IEEE Trans. Affect. Comput. **9**(2), 147–160 (2018). https://doi.org/10.1109/TAFFC.2016.2625250
103. Takahashi K.: Remarks on emotion recognition from bio-potential signals. In: 2nd International Conference on Autonomous Robots and Agents, pp. 186–191 (2004)
104. Thaut, M.H.: Rhythm, Music, and the Brain: Scientific Foundations and Clinical Applications. Routledge, New York (2005).. (**ISBN 0415973708**)
105. Bertin-Mahieux, T., Ellis, D.P.W., Whitman, B., Lamere, P. The million song dataset. In: 12th International Society for Music Information Retrieval Conference ISMIR, pp. 591–596 (2011)
106. Tseng K.C., Lin B.S., Han C.M., Wang P.S. Emotion recognition of EEG underlying favourite music by support vector machine. In: 1st International Conference on Orange Technologies (ICOT), pp. 155–158 (2013). https://doi.org/10.1109/ICOT.2013.6521181
107. Turnbull D., Liu R., Barrington L., Lanckriet, G.R. A game-based approach for collecting semantic annotations of music. In: International Society on Music Information Retrieval ISMIR, pp. 535–538 (2007)
108. Tzanetakis, G., Cook, P.: Musical genre classification of audio signals. IEEE Trans. Speech Audio Process. **10**(5), 293–302 (2002)
109. Van Der Zwaag, M.D., Dijksterhuis, C., De Waard, D., Mulder, B.L.J.M., Westerink, J.H.D.M., Brookhuis, K.A.: The influence of music on mood and performance while driving. Ergonomics **55**(1), 12–22 (2012). https://doi.org/10.1080/00140139.2011.638403
110. Vijayan A.E., Sen D., Sudheer A.P.: EEG-based emotion recognition using statistical measures and auto-regressive modelling. In: IEEE International Conference on Computational Intelligence & Communication Technology, pp. 587–591 (2015). https://doi.org/10.1109/CICT.2015.24
111. Wagner J., Kim J., Andre E.: From physiological signals to emotions: implementing and comparing selected methods for feature extraction and classification. In: 2005 IEEE International Conference on Multimedia and Expo, pp. 940–943 (2005)
112. Weihs C., Ligges U., Mörchen F., Müllensiefen D.: Classification in music research. Adv. Data Anal. Classification, 1(3): 255–291 (2007). 1111 S. (2013). Using Physiological Signals for Emotion Recognition. In: 6th International Conference on Human System Interactions. 2013: 556–561. DOI: https://doi.org/10.1109/HSI.2013.6577880
113. Yang, Y., Chen, H.H.: Machine recognition of music emotion: a review. ACM Trans. Intell. Syst. Technol. (TIST) **3**(3), 40:1-40:30 (2012). https://doi.org/10.1145/2168752.2168754
114. Yang, Y., Lin, Y., Su, Y., Chen, H.H.: A regression approach to music emotion recognition. IEEE Trans. Audio Speech Lang. Process. **16**(2), 448–457 (2008). https://doi.org/10.1109/TASL.2007.911513
115. Yang Y., Liu C.C., Chen H.H.: Music emotion classification: a fuzzy approach. In: Proceedings of the ACM International Conference on Multimedia, pp. 81–84 (2006). https://doi.org/10.1145/1180639.1180665
116. Yıldırım, Ö., Baloglu, U.B., Acharya, U.R.: A deep convolutional neural network model for automated identification of abnormal EEG signals. Neural Comput. Appl. (2018). https://doi.org/10.1007/s00521-018-3889-z
117. Zheng, W., Lu, B.: A Multimodal approach to estimating vigilance using EEG and forehead EOG. J. Neural Eng. **14**(2), 026017 (2016). https://doi.org/10.1088/1741-2552/aa5a98
118. Zheng, W., Lu, B.: Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. IEEE Trans. Auton. Mental Dev. **7**(3), 162–175 (2015). https://doi.org/10.1109/TAMD.2015.2431497
119. Zheng, W., Zhu, J., Lu, B.: Identifying stable patterns over time for emotion recognition from EEG. IEEE Trans. Affect. Comput. **10**(3), 417–429 (2016). https://doi.org/10.1109/TAFFC.2017.2712143
120. Zong C., Chetouani M.: Hilbert-Huang transform based physiological signals analysis for emotion recognition. In: IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 334–339 (2009). https://doi.org/10.1109/ISSPIT.2009.5407547
121. Thayer, R.E.: The Biopsychology of Mood and Arousal. Oxford University Press, New York (1989).. (**ISBN0-19-506827-0**)
122. Fontaine, J.R.J., Scherer, K.R., Roesch, E.B., Ellsworth, P.C.: The world of emotions is not two-dimensional. Psychol. Sci. **18**(12), 1050–1057 (2007). https://doi.org/10.1111/j.1467-9280.2007.02024.x
123. Stober S., Sternin A., Owen A.M., & Grahn, J.A.: Towards Music Imagery Information Retrieval: Introducing the OpenMIIR Dataset of EEG Recordings from Music Perception and Imagination. In: 16th International Society for Music Information Retrieval Conference, pp. 763–769 (2015). https://doi.org/10.5281/zenodo.1416270
124. Alakuş, T.B., Türkoglu, I.: Feature selection with sequential forward selection algorithm from emotion estimation based on EEG signals. Sakarya Univ. J. Sci. **23**(6), 1096–1105 (2019). https://doi.org/10.16984/saufenbilder.501799
125. Hosseini, M., Hosseini, A., Ahi, K.: A review on machine learning for EEG signal processing in bioengineering. IEEE Rev. Biomed. Eng. (2020). https://doi.org/10.1109/RBME.2020.2969915
126. Zhang, K., Zhang, H.B., Li, S., Yang, C., Sun, L.: The PMEmo dataset for music emotion recognition. Int. Conf. Multimed. Retrieval. (2018). https://doi.org/10.1145/3206025.3206037
127. Warrenburg, L.A.: Choosing the right tune: a review of music stimuli used in emotion research. Music Perception Interdiscip. J. **37**(3), 240–258 (2020). https://doi.org/10.1525/mp.2020.37.3.240
128. Su, K., Hairston, W., Robbins, K.: EEG-annotate: automated identification and labeling of events in continuous signals with applications to EEG. J. Neurosci. Methods **293**, 359–374 (2018)
129. Amin, H.U., Mumtaz, W., Subhani, A., Saad, M., & Malik, A.: Classification of EEG signals based on pattern recognition approach. Front. Comput. Neurosci. **11**, 103 (2017). https://doi.org/10.3389/fncom.2017.00103. eCollection 2017
130. Delorme, A., Makeig, S.: EEGLAB: an open-source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods **134**, 9–21 (2004)
131. Ghosh, A., Danieli, M., & Riccardi, G.: Annotation and prediction of stress and workload from physiological and inertial signals. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1621–1624 (2015)
132. Wu, S., Xu, X., Shu, L., Hu, B.: Estimation of valence of emotion using two frontal EEG channels. In: Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine, Kansas City, MO, USA, 13–16 November 2017; pp. 1127–1130 (2017)
133. Lan, Z., Sourina, O., Wang, L., Liu, Y.: Real-time EEG-based emotion monitoring using stable features. Vis. Comput. **32**, 347–358 (2016)
134. Dror, O.E.: The cannon-bard thalamic theory of emotions: a brief genealogy and reappraisal. Emot. Rev. **6**(1), 13–20 (2014). https://doi.org/10.1177/1754073913494898