

# A Detailed Review of Music Mood and Human Emotion Recognition Based on Physiological Signals

Vybhav Chaturvedi\*, Arman Beer Kaur\*, Vedansh Varshney\*

Anupam Garg†, Gurpal Singh Chabbra†

**Abstract** - Scientists and numerous researchers have tried to establish a bond between the emotions conveyed and the subsequent mood perceived in a person. Emotions form an important part of our daily life and 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. 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. When we are happy, or in the mood of celebrating, we dance to songs and when sad, then either we try to improve our mood or find songs with lyrics that relate to our mood. Thus, listening to music forms a major part of our daily life. Since early times, a considerable amount of effort has been made in the field of emotion detection and mood estimation. 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 aims to provide a detailed review of the methods proposed in music mood recognition. It also discusses about various sensors that have been utilized in order to acquire various physiological signals like EEG, ECG, EMG, GSR etc. 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 with the objective of providing a greater insight into this field of interest and develop a better understanding in order 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 in order 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 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 its quite fair enough for them to understand these feelings in their own desirable manner. But 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 at 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 <sup>1</sup> music outranks all other interests/hobbies and in total, 94% of online consumers listen to music throughout the year. The Digital Music Study <sup>2</sup> by Indian Music Industry shows, an average Indian listen to music for 21.5 Hours per week, which is more than the global average of 17.8 hrs./week. In 2011 Zwaag et al. [116], claimed that listening to music lowers the chance of rash driving and even accidents contrary to the popular belief that music impairs drivers.

\*Author Vybhav Chaturvedi, Arman Beer Kaur, Vedansh Varshney are with Thapar University, Punjab, India, and are student in Computer Science and Engineering Department.

†Mentor and Instructor Anupam Garg and Gurpal Singh Chabbra are with Thapar University, Punjab, India, and are professor in Computer Science and Engineering Department

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)

Various researches indicate the other benefits of music and songs, North et al. [82] concluded that music can have a wide range of positive commercial benefits. Silverman [98] 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, Lehmborg et al. [60] pointed out the benefits of music for senior citizens, which include prolonged lifespan, lesser stress, better health, and overall happiness. Thaut [111] compiled various benefits of music in Rehabilitation and Therapy. In 1998, McCraty et al. [71] 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 in order 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 in the course of their study. A variety of techniques like High Order Crossing, Adaptive Filtering, Autoregressive Modelling, Genetic Algorithms etc. and their combinations have been employed in different works, but more optimized algorithms and techniques for an 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 such as EEG, EMG, SCR, respiration rate etc., in order 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 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. Selection of only a set of parameters in order to classify music might result in the neglect of certain important aspects and it poses as a challenge in this field.

On an 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 training set is still required. Majority of the experiments have been conducted on a group of few people, thus increasing the number of test subjects in order 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

Study of the human behavior has always been a subject of eccentricity due to the element of uncertainty that has to be dealt along with it. Considerable amount of research work and analysis has been carried in order 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 in order 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 in order to classify the mood of a song after employing various feature extraction techniques for both of 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. BACKGROUND

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 [48]. 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. [116] 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. [82] 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 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, etc. The latter being quite important if the aim is to develop a user-based system. Chanel [18], Hosseini [37], Kim [52], and Takahashi [109] have used various means and platforms for the detection of induced emotions from different stimuli and have achieved appreciable accuracies.

Li et al. [61] emphasized on the fact that music doesn't 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[40] 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. [115] 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 classification of music based on music mood have been used. Kim et al. [54] introduced a game named MoodSwings which records dynamic mood ratings for music given by the people playing the game. Li et al. [61] used an approach that focusses upon the three parameters namely timbre, intensity, rhythm in order to categorize music. Jiang et al. [47] utilized the spectral

characteristics of a music clip through the Octave-based Spectral Contrast feature for music classification.

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

## 4. EXISTING WORK

Table 1: Already Available Survey Articles, with Remarks (Scope and/or Limitations)

| Serial No. | Author | Covered Areas   | Scope and Limitations   |
|------------|--------|---|---|
| 1.         | [76]   | Article majorly covers Content-Based Music Information Retrieval (CB-MIR) in relation with 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. |
| 2          | [28]   | This article also points to some general research issues in CB-MIR and probable approaches toward their solutions to improve the efficiency of the music being provided to consumer on the basis of their own taste.  | While it is mainly concentrated on similarity approaches with specific features, it also explores the field of Music Retrieval.   |
| 3.         | [104]  | The review approaches the development of music recommendation systems that are first-time reviewed. It had given information about the types of methods and user modeling techniques.   | The survey is well structured and detailed. However, the techniques available for definitions have been provided by this article instead of existing research and scope.                                    |
| 4.         | [119]  | 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.  | Although it is a take on musical classification using similarity and transcription, with availability of better techniques and methods, better results can be obtained.                                     |
| 5.         | [93]   | Music Genre Classification is one of the basic and ever-present method. 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.   |
| 6.         | [66]   | The review is focused on EEG based Brain Computer interfacing, and is specifically dedicated to the review of classification algorithms used for BCI, their properties and their evaluation. Moreover, it provides guidelines in order 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.  | [43]  | 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.  | Although not directly related to this study, the article provides an insight on the further possibilities regarding this field, as music induces behavioral changes it can be exploited in various fields.   |
| 8.  | [48]  | An overview of theory and research concerning expression, perception, and induction of emotion in music. Unique approach by also considering Social Context, the survey also argues that emotion is strongly related to most people's primary motives for listening to music. | The results are novel and unique as they provide preliminary estimates of the occurrence of various emotions in listening to music, at the 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 relation between human emotions and motives. |
| 9.  | [121] | Emotion retrieval and music organization. Discuss about various taxonomy and at the same time states various Machine Learning techniques that are being used.   | Provides some basic solutions addressing some basic issues faced in music emotion recognition.   |
| 10. | [2]   | 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 gives the best comparisons.  |
| 11. | [86]  | The survey is revolving around the ever-increasing numbers of music recommendation systems. The article is concise and suggests its own recommendation system after going through the pros and cons of existing one.  | The comparisons prove to be really helpful and can be used in any research as per the need.  |
| 12. | [4]   | An overview of theory, recent research work and future possibilities with detailed information about EEG and how it is being used to measure and detect emotions.   | Provides important insight in complex working of EEG and compares various studies, which provides strong basis for comparison of results obtained. The article is well structured and carries important details.   |
| 13. | [96]  | A one of a kind reviews on emotion recognition and physiological signal. Details about feature extraction, Machine Learning techniques and emotion taxonomy.  | The article is well structured and gives detailed knowledge about the recent works being undertaken in the field of neuromusicology in accordance with neurophysiology. It never deviates from the point and also provides important datasets.   |

## 5. TAXONOMY

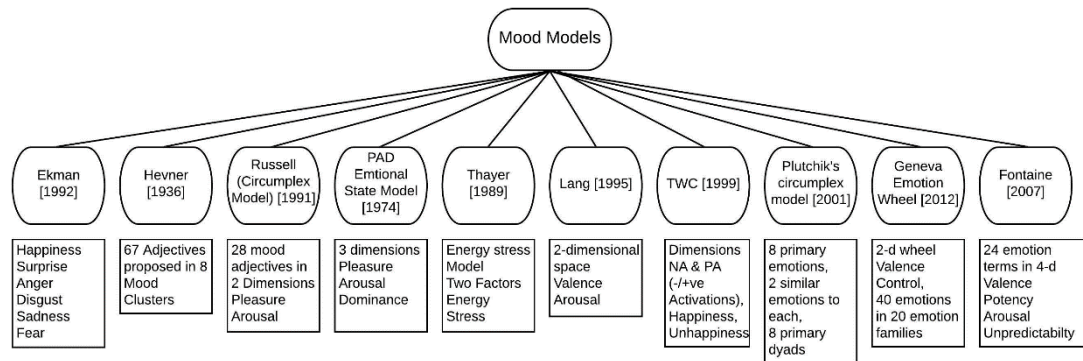


Figure 1: All the major Taxonomy taken into consideration during the Study

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 (Figure 1). Psychologist Ekman [24] argued for basic emotions (Ekman proposed six emotions which are happiness, surprise, anger, disgust, sadness and fear) and nine characteristics which can help distinguish basic emotions from one another and other emotions like Smugness, Hope, Jealousy, Grief, etc, not being part of emotional families. There are two approaches followed when classifying according to emotions [54], [120], one approach is creating discrete emotions i.e. forming clusters of adjectives that are said to express emotions in English Language, formed as a model by Hevner [35] (Figure 3), 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 [61]. Other approach followed is the dimensional approach, conventionally the two most used are Circumplex Emotional Model of Russell [90] (Figure-4) and Thayer's Model [130.] (Figure-5). 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 behaviour of consumers and the influence of music on them [92] [8], PAD Emotional State Model developed by Mehrabian and Russell [73] which consists of three dimensions: Pleasure, Arousal (both from circumplex model) and Dominance. Dominance is used as a measure for how much an individual feel 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), this model classifies music into four groups: Exuberance/Joy, Contentment,

Anger/Frantic, and Depression. For classification of physiological signals according to emotions [46] [118], Ekman's and Lang's Model [57] have been used. Lang characterized emotions on 2-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 [112] 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 item 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 pair of opposites in Plutchik's three dimensional circumplex model [89]. 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 [83] for automatic music mood classification using only lyrics from songs and to provide recommendations to users' mood. An Emotion Wheel with 2-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, then version 2.0 of the GEW with 40 emotion terms arranged in 20 emotion families[97]. More dimensions were made into models like 4-dimensional solution, was created by Fontaine et. al. [131] with 24 emotion terms: Valence, Potency, Arousal and Unpredictability, it included all six of the components of basic emotions identified by researchers. .

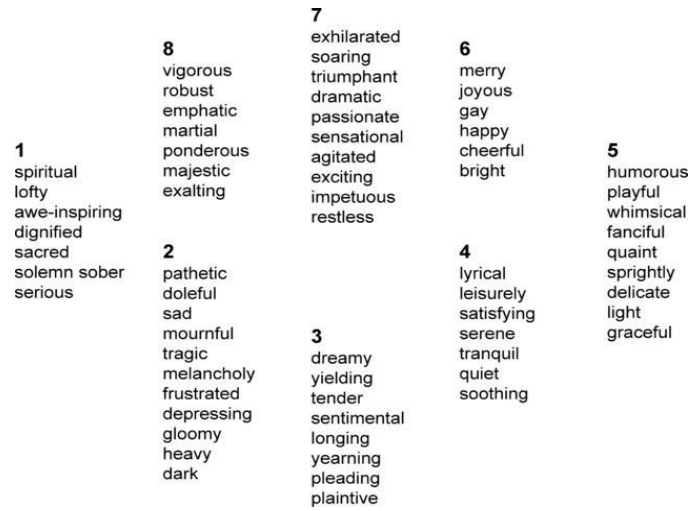


Figure 2: Hevner's Eight cluster

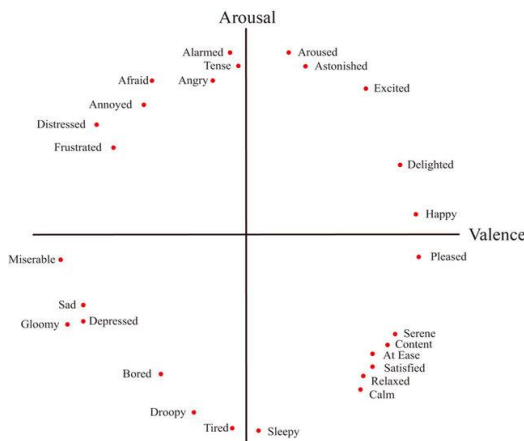


Figure 4: Russell's Circumplex Model

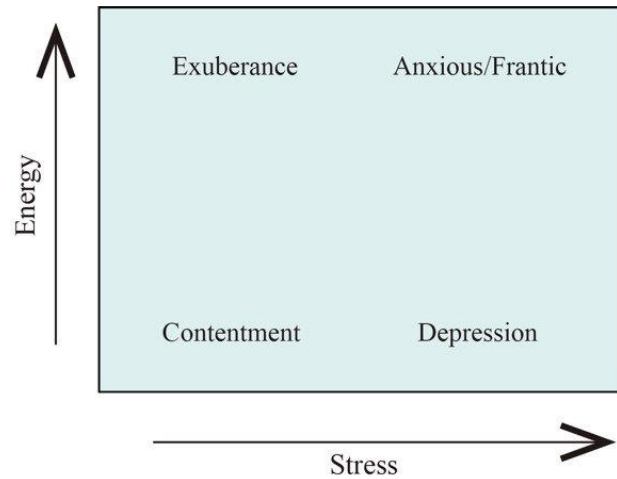


Figure 3: Thayer's 2-D Model

## 6. 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 compile them together in order 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, etc., 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, inclusion and exclusion criteria have been extensively discussed.

As the study revolves around available literature, all the well-known research papers, review papers, reports, etc. from various journals were selected. After organizing them in chronological order, they were extensively studied to extract all the important information and in order 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.

### 6.1 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 in order 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.

## 6.2 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?

## 6.3 Data Sets

In the next section, the datasets used in various studies are compiled, while many researchers prefer to use publicly available datasets, it has been observed that most prefer to develop their databases during the course of the study, this is most predominant where people are working on an EEG or are trying to develop a user-friendly HCI. All the datasets mentioned are publicly available unless stated otherwise and are easily available at the same time, although it is important to mention that some are only available on a request basis.

Table: 2, lists down the datasets in groups based on their similarities and usage, while Table: 3, lists down the dataset individually along with their instances and information provided in them.

Table 2: Existing Datasets w.r.t specific tasks

| Serial Number | Task                               | Dataset Available   | Scope and Usage   |
|---------------|------------------------------------|---|---|
| 1.            | Emotion Recognition using Audio    | DEAP, AMG-1608, DEAM, EmoMusic, Emotify, MedleyDB, OpenMIIR, SoundTracks, Million Song Dataset, musicXmatch, Last.fm, Music Audio Benchmark Dataset, FMA, Audio Set | DEAP and FMA dataset are some of the most frequently used datasets. Emotify provides 9 categories of moods. Last.fm builds a detailed profile for each user's musical taste and provides respective labels. |
| 2.            | EEG                                | NMED-RP, NMED-H, NMED-T, EEG Motor Movement, ASCERTAIN, MIT-BIH, ElderReact, Manhab HCI, TUH EEG, Bern Barcelona, SEMAINE   | Manhab HCI, TUH EEG, SEMAINE are most widely used datasets and provide the most accurate results.   |
| 3.            | Other Inputs (ECG, GSR, Respirary) | DECAF, Physio bank MIT-BIH, ASCERTAIN, DEAP, AMIGOS   | These datasets contain various information of different Inputs, apart from the required input.  |

Table 3: Details of Datasets used in the Study

| Serial Number | Year | Dataset               | Reference | Instances               |
|---------------|------|-----------------------|-----------|-------------------------|
| 1.            | 2011 | DEAP†                 | [55]      | 32                      |
| 2.            | 2015 | AMG 1608              | [19]      | 1608                    |
| 3.            | 2018 | DEAM†                 | [5]       | 1802                    |
| 4.            | 2014 | EmoMusic              | [102]     | 744                     |
| 5.            | 2015 | Emotify               | [6]       | 400                     |
| 6.            | 2014 | MedleyDB†             | [14]      | 122                     |
| 7.            | 2015 | OpenMIIR†             | [132]     | 10 subject 12 fragments |
| 8.            | 2010 | SoundTracks           | [23]      | 110                     |
| 9.            | 2011 | Million Song Dataset† | [112]     | 1000000                 |

|     |      |                             |       |   |
|-----|------|-----------------------------|-------|---|
| 10. | 2011 | musiXmatch                  | [112] | 237662                                    |
| 11. | 2003 | Last.fm <sup>†</sup>        |       | -   |
| 12. | 2005 | MABD <sup>†</sup>           | [36]  | 1886                                      |
| 13. | 2017 | FMA <sup>†</sup>            | [20]  | 106574                                    |
| 14. | 2017 | Audio Set <sup>†</sup>      | [29]  | 2084320                                   |
| 15. | 2018 | NMED-RP                     | [65]  | 15  |
| 16. | 2014 | NMED-H                      | [65]  | 62  |
| 17. | 2017 | NMED-T                      | [65]  | 10  |
| 18. | 204  | EEG Motor Movement          | [94]  | 10  |
| 19. | 2016 | ASCERTAIN <sup>†</sup>      | [108] | 58  |
| 20. | 2001 | MIT-BIH <sup>†</sup>        | [75]  | 25  |
| 21. | 2019 | ElderReact                  | [69]  | 1323                                      |
| 22. | 2012 | Manhab-HCI <sup>†</sup>     | [103] | 20  |
| 23. | 2014 | TUH EEG <sup>†</sup>        | [33]  | 25000                                     |
| 24. | 2012 | Bern Barcelona <sup>†</sup> | [9]   | 5 Subject                                 |
| 25. | 2007 | SEMAINE <sup>†</sup>        | [72]  | 959                                       |
| 26. | 2007 | DECAF                       | [1]   | 30  |
| 27. | 2017 | AMIGOS                      | [74]  | 40  |
| 28. | 2015 | MUSAN                       | [100] | 109 hours of Music, 12 different language |
| 29. | 2016 | SEED <sup>†</sup>           | [126] | 15  |

<sup>†</sup> Benchmarked Database

1-Available at <https://www.last.fm/>, The Nielsen Company (US) (Last Checked - 6-Jan-2020)

### 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, Manhab 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 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.
- Few datasets don't provide the audio files, while this makes them corpus storage-friendly, a few researchers might find these unusable.

## 6.4 Inclusion and Exclusion criteria

### 6.4.1 Inclusion Criteria:

- Physiological signals like EEG, EMG, ECG, EDA, SC, BVP, RSP, SKT, GSR, 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).

#### 6.4.2 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 differences between Left-Handed and Right-Handed people has not been considered.

#### 6.5 Quality Assessment

- High quality research papers from prominent publishing journals like Springer, IEEE, ACM, Elsevier, Science Direct etc.
- Research papers with 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

#### 6.6 Feature Extracted

Table 4: Details of Feature Extracted, Classifiers used and respective Accuracies

| Citation | Features   | Classifier                          | Number of Emotions                                | Accuracy                           |
|----------|--|-------------------------------------|---|------------------------------------|
| [25]     | 17 features including 5 statistical, 8 morphological and 4 temporals | Advance Quadratic SVM               | 7   | 68.9%-87.4% for different emotions |
| [56]     | Energy, Amplitude, Frequency   | SVM, k-NN, ANN, Classification Tree | 2   | 83-93%                             |
| [125]    | -  | k-NN, Random Forest, CNN            | -   | 79.34%                             |
| [84]     | Wavelet Coefficients   | DNN                                 | 4   | 57.67-64.25%                       |
| [3]      | 3 Wavelet Coefficients (Symlet, Daubechies, Morlet)                  | k-NN                                | 3   | 79%                                |
| [64]     | Time – Frequency Analysisusing Fourier                               | LDA, SVM                            | 7   | 86.43%                             |
| [107]    | -  | CNN                                 | -   | -                                  |
| [11]     | Statistical, Band Power, Hjorth parameter, Fractal Dimension         | nRMR, GA-SVM                        | 2   | 60.7-73.14%                        |
| [127]    | Differential Entropy   | DBN, SVM, LR, k-NN                  | 2 Classes of Emotion (Arousal, Valence) + Neutral | 72-86%                             |

|       |   |  |   |   |
|-------|---|--|---|---|
| [77]  | 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%<br>Avg. Valence Accuracy – 94.91%<br>Avg Arousal Accuracy – 93.63% |
| [78]  | Relative Power, Mean, Avg. Nonlinear energy, High Order Crossing, Entropy                                   | SVM, k-NN, CFNN                                    | 2 Classes of Emotion (Arousal, Valence)           | Avg Accuracy – 87.05%<br>Avg. Valence Accuracy – 93.66%<br>Avg Arousal Accuracy – 93.29%  |
| [22]  | ES, DE, DASM, RASM  | k-NN, SVM  | Stress  | 67-81%  |
| [45]  | ASM, CSP, HOC, SOM, HOS, FD, ASP  | SVM, NB, QDA, k-NN, LDA, MLP                       | 2   | 65.12-75.62%  |
| [31]  | Spectrogram, HHT, ZAMT  | k-NN, SVM  | 2 Categories (Based on liking)                    | 78.90-91.02%  |
| [79]  | Gabor, Wavelet Coefficients   | NDB, MLB, PNN & PSO, MEDO, IPSO (For Optimization) | 4   | Avg Classification Accuracy – 64.78%  |
| [17]  | Alpha Power Ratio, Band Power   | QDC, SVM, k-NN                                     | 2 Classes of Emotion (Arousal, Valence) + Neutral | 82%   |
| [106] | Fractal Dimension   | SVM  | 2 Classes of Emotion (Arousal, Valence)           | 73%   |
| [49]  | Magnitude Squared Coherence Estimate  | k-NN   | 4 Classes of Emotion                              | 84.5%   |
| [129] | HHT   | SVM  | -   | Baseline - 71%<br>Fission based – 76%<br>Fusion based – 62%                               |
| [10]  | Pitch frequency, Intensity, Amplitude, Voiced vs Unvoiced   | SOM, XY-fused                                      | 2 Classes of Emotion (Arousal, Valence)           | SOM - 68.39±3.04%<br>XY-fused - 71.43±2.23%   |
| [13]  | Time, Frequency and Wavelet domain features using Short Time Fourier Transform (STFT)                       | MLP, K-NN, SVM                                     | 4   | MLP is most accurate  |
| [117] | Mean, Standard Deviation using Short Time Fourier Transform (STFT)  | MC-SVM   | 4   | 94.10%  |
| [101] | Power Spectral  | MLR, SVR, CCRF, LSTM-RNN                           | Valence   | –   |
| [113] | Prefrontal Cortex (PFC)   | SVM  | 2 Classes of Emotion (Arousal, Valence)           | >80%  |

|       |  |  |   |   |
|-------|--|--|---|---|
| [80]  | 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 & Liking | Valence - 64.3%, Arousal - 66.2%, Dominance- 68.9% Like- 70.2%      |
| [32]  | 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)   |
| [87]  | Statistical, Wavelet-based features using High Order Crossing (HOC) Analysis   | QDA, KNN, Mahalanobis distance, SVM                              | 6   | QDA - 62.30%<br>SVM - 83.33%  |
| [88]  | Statistical, Wavelet-based features using High Order Crossing (HOC) Analysis and Hybrid Adaptive Filtering (HAF)               | QDA, KNN, Mahalanobis distance, SVM                              | 6   | QDA- 77.66%; SVM- 85.17%  |
| [62]  | DASM12, RASM12, PSD24, PSD30 using Short Time Fourier Transform (STFT)   | SVM  | 4   | 82.29±3.06 %  |
| [50]  | Prefrontal, Frontal, Central, Parietal features using Kernel Smoothing Density Estimation (KSDE), Gaussian Mixture Model (GMM) | BN, MLP, One-Rule, Random Tree, Radial Basis Function            | 6   | 1. leave-one-out accuracy- 90%<br>2. intra-subject accuracy- 69.69% |
| [91]  | Mean, Min, Max, Standard deviation, Quadrature pair using Spatial Temporal Box Filters (STBF)                                  | Classifier constructed from subset of all possible STBFs.        | Berlin Dataset: 7 emotions<br>ORATOR Dataset: 7 emotions    | excellent performance   |
| [42]  | Prefrontal Cortex Activity using Genetic Algorithm and Fast Fourier Transform (FFT)  | –  | –   | greater than 70%  |
| [67]. | Intensity, Timbre and Rhythm features using Fast Fourier Transform (FFT), K-L Transform  | –  | 4   | 81.50%  |
| [16]  | Fast Fourier Transform (FFT), binary linear FDA  | –  |   | > 80%   |
| [52]  | Mean, Max, Amplitude, Standard Deviation using Frequency-Time analysis   | SVM  | 4   | Classification ratio:<br>3 emotions- 78.43%<br>4 emotions- 61.76%   |

|       |  |  |    |   |
|-------|--|--|----|---|
| [81]  | –  | KNN, DFA   | 6  | 1. KNN- performed better for frustration, surprise;<br>2. DFA- performed better for fear, sad, anger, amusement |
| [44]  | 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%  |
| [26]  | Tempo, Rhythmic unit   | -  | 5  | -   |
| [83]  | -  | SVM using LibSVM library                             | 8  | 33.3 – 73.3%  |
| [133] | DWD, DFA, H, AvgEng, ShanEn, MSCEn, StdDev, V, ZC using SFS (Sequential Forward Selection)   | SAM, Bayesian Classification, QSVM                   | -  | Average Accuracy Without Feature Selection - 55<br>With Feature Selection - 65                                  |
| [97]  | Temporal and Spectral Features   | -  | 12 |   |
| [118] | Statistical Feature, Amplitude, Breathing rate, Heartbeats   | Linear Discriminant Function, k-NN                   | 4  | Recognition rates - 80%<br>After Feature Reduction - 92%  |
| [61]  | MFCC, Timbral Features, DWCHs  | SVM  | 6  | 70-86%  |
| [41]  | Fractal Dimension, Time and Frequency Feature  | NN, FA, BP   | -  | -   |
| [68]  | Intensity, Timbre, Intensity   | Simple Bayesian Criteria                             | 4  | Overall Classification accuracy - 86.3%   |
| [47]  | Octave-based Spectral Contrast Feature   | GMM with 16 components with Expectation Minimization | -  | Average Accuracy - 82.3%<br>on 10-sec clips<br>90.8% on whole clips   |
| [115] | Timbral, Rhythmic, Pitch   | Simple Gaussian Classifier, GMM, k-NN                | -  | Music classification -86%   |
| [27]  | MFCC, Tree Quantizer   | -  | -  | Cosine Distance gave best result  |

|       |   |   |   |   |
|-------|---|---|---|---|
| [58]  | Timbral, Rhythmic, Tonal  | -   | -   | -   |
| [85]  | Temperature, Humidity, Noise, Illumination  | Bayesian Network, Sheffe's Paired Comparison  | -   | Used comparison test to determine, recommendation system better than subjective preferences |
| [128] | 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 upto 90%  |

## 7. RELATED WORK

### 7.1 Subjective Annotation

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, hence collection of the ground truth data should be conducted carefully. Existing methods usually follow 2 lines, Expert-based suggestion or Subject-based selection. In the first method an expert annotate emotion to a certain musical piece and simply abandons those pieces which can't be annotated by him whereas the second method is more about majority and consensus. Usually, the ground data is 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, for this very reason many researchers often try to use various techniques in order to make the task more user-friendly, time-friendly, cost-friendly, etc. Some popular examples are - Reducing the length of the music pieces [99; 122]; using exemplar songs to better articulate what each emotion class means [38]. allowing the user to skip a song when none of the candidate emotion classes is appropriate to describe the effective content of the song [39]; designing a user-friendly annotation interface [124]. 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 [121]. 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 [70], Chorlody [63] Listen Game [114], and TagATune [59]. These implementations have primarily focused on the collection of descriptive labels for a relatively short audio clip [53].

Fernandez-sotos et al. [26] and Ito et al. [41] 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. [133] collected phone data of 28 students from their University to predict future mood, health and stress rate and similarly Andersson et al. [8] and Kim et al. [54] studied their 150 consumers and 100 users for their respective Researches. Other Music Based Researches collected Music Data from sites like last.fm<sup>1</sup>, musiXmatch [112], etc to collect songs from different Genres or in Researches like Besson et al. collected 200 excerpts from operas from 16 participants. These annotations are mostly situation and

experiment dependent in order 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 to 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.

### 7.2 Data Collection

Researchers frequently used established labelling websites like AMG and last.fm in order to provide a label to any song or piece of music, considering that annual labour 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, 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 in different categories/genres/labels, etc. While there are a few benchmarked datasets available, they often don't 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. Firstly, 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. Secondly, violation of copyright issues poses as 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.

Even if there are various problems associated with data collection, Benchmarked datasets like DEAP, Audio Set, DEAM, TUH-EEG, MANHAB-HCI, Bern Barcelona, SEMAINE, etc provide good enough corpus for any project.

### 7.3 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-bits 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 128kbps format with a base rate of 44100 Hz. [13, 98, 115] Moreover, since complete music pieces can contain sections with different emotions, a 20–60 second 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 normalising basic features of songs such as Pitch, Tempo, Timber, parameters like zero crossing Rate, Flux, etc., are extracted in accordance with the taxonomy followed. For the purpose of extraction PyAudioAnalysis<sup>1</sup> and LibRosa<sup>2</sup> were mostly used. LibRosa was also used to determine the tonnetz of the musical files, which acts as another important feature. Pre-processing is one of the milestone step and various procedures are followed. It's 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. [51] 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, twitching of muscles etc) 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 percent) and Independent Component Analysis (ICA) (10 percent) 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 percent), or Laplacian (23.6 percent) [4]. Features essential for the purpose of our study are then extracted using various 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 pas FIR filter [51]. Bhatti et al. [13] extracted time, frequency and wavelet domain features of the EEG signals using Short Time Fourier Transform (STFT). Petrantonakis et al. [87, 88] utilized the statistical and wavelet-based features of the EEG signal using High Order Crossing Analysis (HOC) and Hybrid Adaptive Filtering (HAF). Zheng et al. [127] worked on differential entropy and Sourina et al. [105, 106] used fractal dimension for their study. Lu et al. [67] examined the intensity, timbre, rhythmic features for the purpose of analysis whereas Kim et al. [52] utilized features like mean, max, standard deviation using Frequency-Time Analysis.

### 7.4 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 effects of Tempo and Rhythmic effect in Music on Mood Fernandez-Sotos et al [26]. Used the Circumplex Model and depicted tempo and rhythmic units alongside primary emotions of Russell's Model. Russell's Model is still being used in wide range of researches due to its much easy dimensional approach, for the purpose of feature selection and comparing EEG signals Alakus et al. [133] used it. Schatter et al. [97] applied the previously mentioned model and the Thayer's Model independently for selecting mood for speech recorded musical collection to analyse emotions from the speech itself and when classifying music using both Audio and Music. Laurier et al. [58] used this model for classifying them in accordance with mood, providing with better accuracies by using the combination. Kim et al. [54] used Thayer's Model for real-time mood and Liu et al. [64] made use of a model for automatic mood detection from Acoustic Music Data. Li et al. [61] conducted similarity search and emotion detection with the help of Hevner's Model, concluding that Emotion Detection from Music is difficult then similarity search for Music after getting lower accuracies for the same Music Database. Much Modified Dimensional Model, Lang's Model helped Wagner et al. [118] to classify into several moods the features extracted from physiological signals recorded from electromyogram (EMG), electrocardiogram (ECG), skin conductivity (SC) and Respiration Change (RSP) beside three pattern recognition methods tested.

### 7.5 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. Support Vector Machine was the most used classifier in majority of the cases and yields good accuracies over varying datasets and constraints. Aguilar et al. [25] used an advanced quadratic SVM classifier to obtain max accuracy of 87.4%, Atkinson et al. [11], Naji et al. [77, 78] used SVM using RBF Classifiers. Sourina et al. [105] used SVM on IADS dataset for emotion detection using Arousal-Valence Model, Kim et al. [52] used SVM for emotion detection using short-term monitoring of physiological signals, Lin et al. [62] used SVM in his work and used EEG for emotion recognition. Zong et al. [129] used SVM along with HHT. Takahashi [109] also showed that SVM is most suitable for emotion detection. Li et al. [61] used SVM for similarity search and obtained accuracies up to 86%. Though SVM is still a useful technique, the newer versions of SVM is often preferred over the traditional algorithm, Jatupaiboon et al. [45] used Gaussian SVM, Brown, et al. [17] used Linear SVM, Vijayan et al. [117] used multiclass SVM or MC-SVM. Soleymani et al. [101] combined PCA and Fisher Criterion with SVM and compared it with the traditional technique for facial features. Neural Network specifically k-NN was also used frequently by researchers because it doesn't require any explicit specification and uses self-adaptive techniques to adjust the data. Lahane et al [56] used k-NN for stress analysis using Teaser – Kaiser Energy operator. Lopez et al used k-NN on a single channel to detect Abnormal EEG signals. Alakus et al. [3] and Khosrowabadi et al. [49] used k-NN to classify 8 different emotions using 3 different wavelet filters. Nasoz et al. [81] and Wagner et al. [118] used k-NN for emotion recognition using physiological signals. Hadjidimitriou et al. [31] used both SVM and k-NN for classifying vectors into 2 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. [125] preferred CNN over other classifiers for abnormal signal detection from EEG. Acharya et al. [2]

1-Available at <https://github.com/tjannak/pyAudioAnalysis> (Last Checked - 6-Jan-2020)

2-Available at <https://librosa.github.io/librosa/> (Last Checked - 6-Jan-2020)

used CNN for classification of ECG signals, similarly Stober et al. [107] used CNN for feature learning from EEG recordings. Other variants of

Neural Networks like DNN (Pandey et al. [84]) and RNN were also used when working with EEG.



Figure 5: Most widely used Machine Learning Model

It has been observed that in order to remove the discrepancies of the algorithm employed, researchers often selected a group of algorithms and clubbed them together as per the scope of the study in order to obtain better results. Classifiers like QDA, Mahalanobis distance have been employed by Petrantonakis et al. [87, 88], Hadjidimitriou et al. [32] and Jatupaiboon et al. [45]. Moreover, combinations of existing classifiers like MC-SVM, GA-SVM, QSVM, and LSTM-RNN have also been employed in different studies. SVR, SMO, CCRF, MLP were also used, but are quite rare and hence

probably aren't the best option. Interestingly, Luo et al. [68] 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.

Table 5 has a combination of properties of all the major classifiers that have been observed in the study.

Table 5: Properties of the classifiers used in various studies

| Classifier           | Linear | Non-Linear | Generative | Discriminant | Dynamic | Static | Regular | Stable | Unstable | High Dimension Robust |
|----------------------|--------|------------|------------|--------------|---------|--------|---------|--------|----------|-----------------------|
| Linear SVM           | •      |            |            | •            |         | •      |         | •      |          | •                     |
| RBF SVM              |        | •          |            | •            |         | •      | •       | •      |          | •                     |
| MLP                  |        | •          |            | •            |         | •      | •       |        | •        |                       |
| BLR NN               |        | •          |            | •            |         | •      |         |        | •        |                       |
| ALN NN               |        | •          |            | •            |         | •      |         |        | •        |                       |
| Gaussian NN          |        | •          |            | •            |         | •      |         |        | •        |                       |
| Perception           | •      |            |            | •            |         | •      |         | •      | •        |                       |
| RBF NN               |        | •          |            | •            |         | •      |         |        | •        |                       |
| Fuzzy Logic          |        | •          |            | •            |         | •      | •       |        | •        |                       |
| NN                   |        |            |            |              |         | •      |         |        |          |                       |
| HMM                  |        | •          |            |              | •       |        |         |        | •        |                       |
| Bayes Quadratic      |        | •          | •          |              |         | •      |         |        | •        |                       |
| Bayes Graphical      |        |            | •          |              |         | •      |         |        | •        |                       |
| k-NN                 |        | •          |            | •            |         | •      |         |        | •        |                       |
| FLDA                 | •      |            |            | •            |         | •      |         | •      |          |                       |
| BLDA                 |        |            |            |              |         |        |         |        |          |                       |
| Random Forest        |        | •          |            | •            |         |        |         |        |          |                       |
| Decision tree        |        | •          |            | •            |         |        |         |        |          |                       |
| Mahalanobis Distance |        | •          |            | •            |         | •      |         |        | •        |                       |



## 9. DRAWBACKS & FUTURE SCOPE

The present study has tried to identify gaps in the existing research and correlate them with the practices related to musical variable. The study suggests that more research is required to correctly examine these relationships in order 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 was confined only to the reported studies. Other researches based on physiological signals like EOG, PCG, 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. Effect of music on employee behaviour 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 has not been thought upon. Experimental studies are required to explore this dimension in greater details.
- Various researches have either used one or two features in their works, with availability of refined and pre-processed datasets the possibilities to use more features has certainly increased.
- Effect of music on behavioural responses, 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 details. Experimental studies can measure the effect of various musical variables, in a controlled manner, person's behaviours in different environment resulting into change in preferences and daily lifestyle (food habits, dressing sense etc).
- The studies have not used the most recent classifiers and result boosting techniques, with availability of AdaBoost, LP Boost and Gradient Boosting, 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 wellbeing like treatment of Stress, Depression, Migraine etc can be further explored.
- Interactive effect of music with other ambient factors such as language of song, surroundings, peripheral 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 revolutionise the field of surgery and treatment if used accordingly. Music Therapy, can help in reducing anxiety cases during surgeries [15], although recently people are working on this field, it's still a new field with various possibilities.

## 10. 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, calmness has been more frequently recorded in comparison to others. As stated earlier, substantial amount of research work has been carried out in order to devise several taxonomies or models for classification of human emotions but 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 in order to capture these lesser identified emotions in order to broaden this narrow range of classification. This will gradually allow us to obtain a greater insight into a person's emotional behaviour and hence will help us establish certain connections between music and the Human Psychology.

## REFERENCES

- [1] Abadi M. K., Subramanian R., Kia S. M., Avesani P., Patras I., Sebe N., (2015) DECAF: MEG-Based Multimodal Database for Decoding Affective Physiological Responses. *IEEE Transactions on Affective Computing*, Volume 6(3), 2015: 209-222. DOI: 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. (2018). Characterization of focal EEG signals: A review. *Future Generation Computer System*, Volume: 91, 2019:290-299. DOI: 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. *3<sup>rd</sup> International Conference on Engineering Technology and Applied Science*. 2018: 250-254
- [4] Alarcão S.M., & Fonseca M.J. (2019). Emotions Recognition Using EEG Signals: A Survey. *IEEE Transactions on Affective Computing*, Volume 10 (3), 2019: 374-393. DOI: 10.1109/TAFFC.2017.2714671
- [5] Aljanaki A., Yang Y.H., Soleymani M. (2017) Developing a benchmark for emotional analysis of music. *PLOS ONE*. Volume 12(3), 2017. DOI: 10.1371/journal.pone.0173392
- [6] Aljanaki, F. Wiering, R. C. Veltkamp. (2015) Studying emotion induced by music through a crowdsourcing game. *Information Processing & Management*, Volume 52(1), 2015: 115-128
- [7] Al-Nafjan A., Hosny M., Al-Ohali Y., & Al-Wabil A. (2017). Review and Classification of Emotion Recognition Based on EEG Brain-Computer Interface System Research: A Systematic Review. *International Journal of Recent Technology and Engineering* Volume-8(1C), 2019: 207-211 DOI: 10.3390/app7121239
- [8] Andersson P. K., Kristensson P., Wastlund E. and Gustafsson A. (2012). Let the music play or not: The influence of background music on consumer behaviour. *Journal of Retailing and Consumer Behavior*, Elsevier, Volume 19(6) 2012: 553-560.
- [9] Andrzejak R.G., Schindler K., Rummel C. (2012) Non-Randomness, nonlinear dependence, and non-stationarity of electroencephalographic recordings from epilepsy patients. *Physical Review American Physical Society*. E-86, 2012. DOI: 10.1103/PhysRevE.86.046206
- [10] Arnrich B., Marca R.L., & Ehlert U. (2010). Self-Organizing Maps for Affective State Detection. 2010: 1-8.
- [11] Atkinson J., Campos D. (2016). Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers. *Expert System with Application*, Volume 47 (C), 2016:35-41. 10.1016/j.eswa.2015.10.049
- [12] Besson M., Faïta F., Peretz I., Bonnel A., and Requin, J. (1998). Singing in the Brain: Independence of Lyrics and Tunes. *Psychological Science*, Volume 9(6), 1998: 494-498.
- [13] Bhatti A.M., Majid M., Anwar S.M. and Khan B. (2016). Human emotion recognition and analysis in response to audio music using brain signals. *Computers in Human Behaviour*, Elsevier, Volume 65 (C), 2016: 267-275. DOI: 10.1016/j.chb.2016.08.029
- [14] Bittner R., Salamon J, Tierney M., Mauch M., Cannam C., Bello J. P. (2014) MedleyDB: A Multitrack Dataset for Annotation-Intensive MIR Research. *15th International Society for Music Information Retrieval Conference, ISMIR* 2014: 155-160
- [15] Bojorquez G. R., Jackson K. E., & Andrews A. K. (2020). Music Therapy for Surgical Patients. *Critical Care Nursing Quarterly*, Volume 43(1) 2020: 81-85. DOI:10.1097/cnq.0000000000000294
- [16] Bos D.O. (2006). EEG-based Emotion Recognition. *The Influence of Visual and Auditory Stimuli*.
- [17] Brown L., Grundlehner B., Penders J. (2011). Towards wireless emotional valence detection from EEG. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011:2188-2191. DOI: 10.1109/IEMBS.2011.6090412
- [18] Chanel G., Ansari-Asl K. and Pun T. (2007). Valence-arousal evaluation using physiological signals in an emotion recall paradigm. *IEEE International Conference on Systems, Man and Cybernetics*, Montreal, 2007:2662-2667.
- [19] Chen, Y-A., Yang Y-H., Wang J-C., Chen H. (2015). The AMG1608 dataset for music emotion recognition. *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. QLD, 2015: 693-697 DOI: 10.1109/icassp.2015.7178058
- [20] Defferrard M., Benzi K., Vandergheynst P., Bresson X., (2017). FMA: A Dataset for Music Analysis. *18th International Society for Music Information Retrieval Conference* 2017. DOI: arxiv.org/abs/1612.01840
- [21] Devillers L., Vidrascu L., Lamel L. (2005). Challenges in real-life emotion annotation and machine learning based detection. *Neural Networks*, Elsevier, Volume 18, 2005: 407-422. DOI: 10.1016/j.neunet.2005.03.007
- [22] Duan R., Zhu J., Lu B. (2013). Differential entropy feature for EEG-based emotion classification. *6th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2013: 81-84. DOI: 10.1109/NER.2013.6695876
- [23] Eerola, T., Vuoskoski, J. K. (2010). A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, Volume 39(1), 2010: 18-49. DOI: 10.1177/0305735610362821
- [24] Ekman P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, Volume 6 (3-4), 1992: 169 -200. DOI: 10.1080/02699939208411068
- [25] Fernández-Aguilar L., Martínez-Rodrigo A., Moncho-Bogani J., Fernández-Caballero A., Latorre, J.M. (2019). Emotion Detection in Aging Adults Through Continuous Monitoring of Electro-Dermal Activity and Heart-Rate Variability. *IWINAC* 2019: 252-261. DOI: 10.1007/978-3-030-19591-5\_26
- [26] Fernández-Sotos, A., Fernández-Caballero, A., and Latorre, J. M. (2016). Influence of Tempo and Rhythmic Unit in Musical Emotion Regulation. *Frontiers in Computational Neuroscience*, Volume 10 (80) 2016: 1-13.
- [27] Foote, J. (1997). Content-based retrieval of music and audio. *Multimedia Storage and Archiving Systems II*, Volume 3229, 1997. DOI: 10.1117/12.290336
- [28] Fu Z., Lu G., Ting K.M., Zhang D. (2011). A survey of audio-based music classification and annotation. *IEEE Transaction on Multimedia*. Volume 13(2). 2011:303-319.
- [29] Gemmeke J.F., Ellis D., Freedman D., Jansen A., Lawrence W., Moore R.C., Plakal M., Ritter M. (2017). Audio Set: An ontology and human-labeled dataset for audio events. *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* 2017: 776-780. DOI: 10.1109/ICASSP.2017.7952261.

- [30] Griffiths, D., Cunningham, S., and Weinel, J. (2013). Automatic Music Playlist Generation using Affective Computing Technologies.
- [31] Hadjidimitriou S., Hadjileontiadis L.J. (2013) EEG-Based Classification of Music Appraisal Responses Using Time-Frequency Analysis and Familiarity Ratings. *IEEE Transactions on Affective Computing*, Volume 4(2), 2013: 161-172. DOI: 10.1109/T-AFFC.2013.6
- [32] Hadjidimitriou S.K. and Hadjileontiadis L.J. (2012). Towards an EEG-Based Recognition of Music Liking Using Time-Frequency Analysis. *IEEE Transactions on Biomedical Engineering*, Volume 59, 2012: 3498-3510. DOI: 10.1109/TBME.2012.2217495
- [33] Harati A., Lopez S., Obeid I., Picone J., Jacobson M., Tobochnik S. (2014) The TUH EEG CORPUS: A big data resource for automated EEG interpretation. *IEEE Signal Processing in Medicine and Biology Symposium, IEEE SPMB 2014*. DOI: 10.1109/SPMB.2014.7002953.
- [34] Hatamikia, S., Maghooli, K., and Nasrabadi, A. M. (2014). The emotion recognition system based on autoregressive model and sequential forward feature selection of electroencephalogram signals. *Journal of medical signals and sensors*, Volume 4(3), 2014: 194–201.
- [35] Hevner K. (1936). Experimental studies of the elements of expression in music. *American Journal of Psychology*, Volume 48 (2), 1936: 246–268.
- [36] Homburg H., Mierswa I., Moller B., Morik K., Wurst M. (2005) A Benchmark Dataset for Audio Classification and Clustering. *International Symposium on Music Information Retrieval 2005*: 528-531
- [37] Hosseini S.A., Naghibi-Sistani M.B. and Rahati-quchani S. (2010). Dissection and analysis of psychophysiological and EEG signals for emotional stress evaluation. *Journal of Biological Systems*, 18(spec01):101-114.
- [38] Hu X., Downie J. S., Laurier C., Bay M., & Ehmann A. F. (2008). The 2007 Mirex audio mood classification task: Lessons learned. *ISMIR 9th International Conference on Music Information Retrieval*. 2008: 462-467. DOI: 10.5281/zenodo.1416380
- [39] Hu X., Sanghvi V.R., Vong B., On P.J., Leong C.N., & Angelica J. (2008). Moody: A web-based music mood classification and recommendation system. In *Proceedings of the International Conference on Music Information Retrieval*. 2008
- [40] Huron D. (2000). Perceptual and Cognitive Applications in Music Information Retrieval. *ISMIR, 1st International Symposium on Music Information Retrieval*, 2000.
- [41] Ito S., Mitsukura Y., Fukumi M. and Akamatsu N. (2003). A feature extraction of the EEG during listening to the music using the factor analysis and neural networks. *Proceedings of the International Joint Conference on Neural Networks*, Volume 3, 2003: 2263-2267. DOI: 10.1109/IJCNN.2003.1223763.
- [42] Ito S., Mitsukura Y., Fukumi M. and Cao J. (2007). Detecting Method of Music to Match the User's Mood in Prefrontal Cortex EEG Activity Using the GA. *International Conference on Control, Automation and Systems*, Seoul, South Korea, 2007. DOI: 10.1109/ICCAS.2007.4406685
- [43] Jain R., Bagdare S. (2011). Music and consumption experience: A Review. *International Journal of Retail & Distribution Management*, Volume 39(4):289-302. DOI: 10.1108/09590551111117554
- [44] Jaques N., Rudovic O., Taylor S., Sano A. and Picard R. (2017). Predicting Tomorrow's Mood, Health, and Stress Level using Personalized Multitask Learning and Domain Adaptation. *IJCAI 2017 Workshop on Artificial Intelligence in Affective Computing*, *Journal of Machine Learning Research* 66, 2017: 17-33.
- [45] Jatupaiboon N., Pan-Ngum S., & Israsena P. (2013) Real-Time EEG-Based Happiness Detection System. *The Scientific World Journal*. 618649. DOI: 10.1155/2013/618649.
- [46] Jerritta S., Murugappan M., Nagarajan R. and Wan K. (2011). Physiological signals based Human Emotion Recognition: A Review. *IEEE 7<sup>th</sup> International Colloquium on Signal Processing and its Applications*, 2011: 410-415.
- [47] Jiang D., Lu L., Zhang H., Tao J. and Cai L. (2002). Music type classification by spectral contrast feature. *IEEE International Conference on Multimedia and Expo*, Volume 1, 2002: 113-116.
- [48] Juslin P.N., Laukka P. (2004). Expression, Perception, and Induction of Musical Emotions: A Review and a Questionnaire Study of Everyday Listening. *Journal of New Music Research*, Volume 33, 2004: 217-238. DOI: 10.1080/0929821042000317813
- [49] Khosrowabadi R., Quek H.C., Wahab A., & Ang K.K. (2010). EEG-based Emotion Recognition Using Self-Organizing Map for Boundary Detection. *20th International Conference on Pattern Recognition, ICPR '10*. IEEE Computer Society, 2010: 4242-4245. DOI: 10.1109/ICPR.2010.1031
- [50] Khosrowabadi R., Wahab A., Ang K.K. and Baniasad M.H. (2009). Affective computation on EEG correlates of emotion from musical and vocal stimuli. *International Joint Conference on Neural Networks*, Atlanta, GA, USA, 2009. DOI: 10.1109/IJCNN.2009.5178748
- [51] Kim D., & Koo D. (2012). Analysis of Pre-Processing Methods for Music Information Retrieval in Noisy Environments using Mobile Devices. *International Journal of Contents*, Volume 8(2), 2012: 1–6. DOI: doi.org/10.5392/IJOC.2012.8.2.001
- [52] Kim K.H., Bang S.W. and Kim S.R. (2004). Emotion recognition system using short-term monitoring of physiological signals. *Medical and Biological Engineering and Computing*, Springer, Volume 42, 2004: 419-427. DOI: 10.1007/BF02344719
- [53] Kim Y.E., Schmidt E.M., Migneco R., Morton B.G., Richardson P., Scott J.J., Speck J.A., Turnbull D. (2010). State of the Art Report: Music Emotion Recognition: A State-of-the-Art Review. *International Society for Music Information Retrieval Conference ISMIR*, 2010
- [54] Kim Y.E., Schmidt E.M., & Emelle L. (2008). MoodSwings: A Collaborative Game for Music Mood Label Collection. *ISMIR*, 2008: 231-236.
- [55] Koelstra S., Mühl C., Soleymani, M., Lee Jong-Seok, Yazdani A., Ebrahimi T., Pun T., Nijholt A., Patras I. (2011). DEAP: A Database for Emotion Analysis Using Physiological Signals. *IEEE Transactions on Affective Computing*. Volume 3(1). 2012: 18-31. DOI - 10.1109/T-AFFC.2011.15.
- [56] Lahane P.U., Thirugnanam M. Human Emotion Detection and Stress Analysis using EEG Signal. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)* ISSN: 2278-3075, Volume-8 Issue-4S2, 2019: 96- 100.
- [57] Lang P. J. (1995). The Emotion Probe: Studies of Motivation and Attention. *The American Psychologist Association*, Volume 50 (5), 1995: 372-385.
- [58] Laurier C., Grivolla and Herrera P. (2008). Multimodal Music Mood Classification Using Audio and Lyrics. *7<sup>th</sup> International*

- Conference on Machine Learning and Applications, 2008: 688-693. DOI: 10.1109/ICMLA.2008.96.
- [59] Law E., Ahn L.V., Dannenberg R.B., & Crawford M.J. (2007). TagATune: A Game for Music and Sound Annotation. 8th International Conference on Music Information Retrieval ISMIR, 2007
- [60] Lehmberg L.J. and Fung C.V. (2010). Benefits of Music Participation for Senior Citizens: A Review of the Literature. Music Education Research International, Volume 4, 2010: 19-30
- [61] Li T. and Ogihara M. (2004). Content-based music similarity search and emotion detection. ICASSP, 1988: 705-708. DOI: 10.1109/ICASSP.2004.1327208.
- [62] Lin Y.P., Wang C.H., Jung T.P., Wu T.L., Jeng S.K., Duann J.R. and Chen J.H. (2010). EEG-Based Emotion Recognition in Music Listening. IEEE Transactions on Biomedical Engineering, Volume 57, 2010: 1798-1806. DOI:10.1109/TBME.2010.2048568
- [63] Liu Y., Yan N., & Hu D. (2014). Chorlody: a music learning game. CHI Extended Abstracts on Human Factors in Computing Systems. 2014: 277-280 DOI: doi.org/10.1145/2559206.2580098
- [64] Liu Y.J., Yu M., Zhao G., Song J., Ge Y., Shi Y. (2018). Real-Time Movie-Induced Discrete Emotion Recognition from EEG Signals. IEEE Transactions on Affective Computing, Volume 9(4), 2018:550-562. DOI: 10.1109/TAFFC.2017.2660485
- [65] Losorelli S., Nguyen D.C., Dmochowski J.P., Kaneshiro B. (2017) NMED-T: A TEMPO-FOCUSED DATASET OF CORTICAL AND BEHAVIORAL RESPONSES TO NATURALISTIC MUSIC. 18th International Society for Music Information Retrieval Conference. 2017: 339-346. DOI: 10.5281/zenodo.1417917
- [66] Lotte F., Bougrain L., Cichocki A., Clerc M., Congedo M., Rakotomamonjy A., Yger F. (2018). A review of classification algorithms for EEG-based brain-computer interfaces: A 10-year update. Journal of Neural Engineering, Volume 15(3). 2018. DOI: 10.1088/1741-2552/aab2f2
- [67] Lu L., Liu D. and Zhang H.J. (2005). Automatic Mood Detection and Tracking of Music Audio Signals. IEEE Transactions on Audio, Speech, and Language Processing, Volume 14, 2006: 5-18. DOI: 10.1109/TSA.2005.860344
- [68] Luo S., Yu Y., Liu S., Qiao H., Liu Y., & Feng, L. (2020). Deep attention-based music genre classification. Neurocomputing, Volume 372, 2020: 84-91. DOI: doi.org/10.1016/j.neucom.2019.09.054.
- [69] Ma K., Wang X., Yang X., Zhang M., Girard J.M., Morency L.P. (2019) ElderReact: A Multimodal Dataset for Recognizing Emotional Response in Aging Adults. In 2019 International Conference on Multimodal Interaction, ICMI 2019:349-357. DOI: 10.1145/3340555.3353747
- [70] Mandel M.I., & Ellis D.P. (2007). A Web-Based Game for Collecting Music Metadata. Journal of New Music Research. 2007. Volume 37. 2007: 365-366. DOI: 10.1080/09298210802479300.
- [71] McCraty R., Barrios-Choplin B., Atkinson M. and Tomasino D. (1998). The effects of different types of music on mood, tension, and mental clarity. Alternative therapies in health and medicine, 4(1):75-84.
- [72] McKeown G., Valstar M., Cowie R., Pantic M., Schroder M. (2012) The SEMAINE Database: Annotated Multimodal Records of Emotionally Colored Conversations between a Person and a Limited Agent. IEEE Transactions on Affective Computing. Volume 3(1). 2012:5-17. DOI: 10.1109/TAFFC.2011.20
- [73] Mehrabian, A. and Russell J. A. (1974). An Approach to Environmental Psychology. Cambridge, MA: MIT Press.
- [74] Miranda J., Abadi M.K., Nicu S., Patras I. (2017) AMIGOS: A dataset for Mood, personality and affect research on Individuals and Groups. IEEE Transactions on Affective Computing. 2017: 1-14. DOI:10.1109/TAFFC.2018.2884461.
- [75] Moody G.B., Mark R.G. (2001) The impact of the MIT-BIH Arrhythmia Database. IEEE Engineering in Medical and Biomedical. Volume 20(3), 2001:45-50.
- [76] Murthy Y. V. S. and Koolagudi S. G. (2018). Content-Based Music Information Retrieval (CB-MIR) and Its Applications towards the Music Industry: A Review. ACM Computing Surveys, Volume 51 (3), 2018: 1-46
- [77] Naji M., Firoozabadi M., Azadfallah P. (2013). Classification of Music-Induced Emotions Based on Information Fusion of Forehead Biosignals and Electrocardiogram. Cognitive Computation, Volume 6(2), 2014: 241-252. DOI: 10.1007/s12559-013-9239-7
- [78] Naji M., Firoozabadi M., Azadfallah P. (2015). Emotion classification during music listening from forehead biosignals. Signal, Image and Video Processing, Volume 9(6), 2015:1365-1375. DOI: 10.1007/s11760-013-0591-6
- [79] Nasehi S., Pourghassem H. (2012) An Optimal EEG-based Emotion Recognition Algorithm Using Gabor Features. Volume 8. 2012: 87-99
- [80] Naser D.S. and Saha G. (2013). Recognition of Emotions Induced by Music Videos Using DT-CWPT. Indian Conference on Medical Informatics and Telemedicine (ICMIT), 2013. DOI: 10.1109/IndianCMIT.2013.6529408
- [81] Nasoz F., Lisetti C.L., Alvarez K. and Finkelstein N. (2003). Emotion Recognition from Physiological Signals for User Modelling of Affect. Cognition, Technology & Work, Volume 6 (1), 2004: 4-14. DOI: 10.1007/s10111-003-0143-x
- [82] North A.C. and Hargreaves, D.J. (2009) Music and consumer behavior, in S. Hallam, I. Cross and M. Thaut (ed). The Oxford Handbook of Music Psychology, Oxford: Oxford University Press, 2009: 481-490.
- [83] Oh S., Hahn M. and Kim J. (2013). Music mood classification using intro and refrain parts of lyrics. 2013 International Conference on Information Science and Applications (ICISA), 2013: 1-3. DOI:10.1109/ICISA.2013.6579495.
- [84] Pandey P., Seeja K.R. (2018). Subject-Independent Emotion Detection from EEG Signals Using Deep Neural Network. Proceedings of ICICC 2018, Volume 2. 2019: 41-46 DOI:10.1007/978-981-13-2354-6\_5.
- [85] Park H.S., Yoo J.O., & Cho S.B. (2006). A Context-Aware Music Recommendation System Using Fuzzy Bayesian Networks with Utility Theory. Fuzzy Systems and Knowledge Discovery. FSKD 2006. Lecture Notes in Computer Science, vol 4223. Springer, 2006. DOI: doi.org/10.1007/11881599\_121
- [86] Paul D.J., & Kundu S. (2020). A Survey of Music Recommendation Systems with a Proposed Music Recommendation System. Emerging Technology in Modelling and Graphics. Advances in Intelligent Systems and Computing, vol 937. Springer, 2020. DOI: doi.org/10.1007/978-981-13-7403-6\_26
- [87] Petrantonakis P.C. and Hadjileontiadis L.J. (2009). Emotion Recognition From EEG Using Higher Order Crossing. IEEE Transactions on Information Technology in Biomedicine,

Volume 14, 2010: 186-197. DOI: 10.1109/TITB.2009.2034649

- [88] Petrantonakis P.C. and Hadjileontiadis L.J. (2010). Emotion Recognition from Brain Signals Using Hybrid Adaptive Filtering and Higher Order Crossing Analysis. *IEEE Transactions on Affective Computing*, Volume 1, 2010: 81-97. DOI: 10.1109/T-AFFC.2010.7
- [89] Plutchik R. (2001). The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American Scientist*, Volume 89(4), 2001: 344 – 350.
- [90] Russell J. A. (1991). Culture and the Categorization of Emotions. *Psychological Bulletin*, Volume 110 (3), 1991: 425 – 450.
- [91] Ruvolo P., Fasel I. and Movellan J. (2008). Auditory Mood Detection for Social and Educational Robots. 2008 IEEE International Conference on Robotics and Automation, Pasadena, CA, USA, 2008. DOI: 10.1109/ROBOT.2008.4543754
- [92] Santos E. B. A. and Freire O. B. D. L. (2013). The Influence of Music on Consumer Behaviour. *Independent Journal of Management & Production*, Volume 4 (2) 2013: 537-548.
- [93] Scaringella N., Zoia G., Mlynek D. (2006). Automatic genre classification of music content: a survey. *IEEE Signal Processing Magazine*, Volume 23(2), 2006:133-141
- [94] Schalk G., McFarland D.J., Hinterberger T., Birbaumer N., Wolpaw J.R. BCI2000: A General-Purpose Brain-Computer Interface (BCI) System. *IEEE Transactions on Biomedical Engineering* 51(6):1034-1043, 2004.
- [95] Schatter G. and Kramer P. (2011). Mood-based selection of music collections by speech. 4th International Conference on Human System Interactions, HSI 2011, 2011: 165-168. DOI: 10.1109/HSI.2011.5937360.
- [96] Shu L., Xie J., Yang M., Li Z., Li Z., Liao D., Xu X., & Yang X. (2018). A Review of Emotion Recognition Using Physiological Signals. *Sensors*. DOI: 10.3390/s18072074Shu, L.; Xie, J.; Yang, M.; Li, Z.; Li, Z.; Liao, D.; Xu, X.; Yang, X. A Review of Emotion Recognition Using Physiological Signals. *Sensors* **2018**, *18*, 2074.
- [97] Shuman V., Schlegel K. and Scherer K. (2015). Geneva Emotion Wheel rating study (Report). Geneva, Swiss Centre for Affective Sciences.
- [98] Silverman M. J. (2015). Effects of Live Music in Oncology Waiting Rooms: Two Mixed Methods Pilot Studies. *International Journal of Music and Performing Arts*, 3(1):1-15.
- [99] Skowronek J., McKinney M.F., & Par S.V. (2006). Ground truth for automatic music mood classification. In *Proceedings of the International Conference on Music Information Retrieval*. 2006 395–396.
- [100] Snyder D., Chen G., & Povey D. (2015). MUSAN: A Music, Speech, and Noise Corpus. *ArXiv*, *abs/1510.08484*.
- [101] Soleymani M., Asghari-Esfeden S., Pantic M. and Fu Y. (2014). Continuous emotion detection using EEG signals and facial expressions. *IEEE International Conference on Multimedia and Expo (ICME)*, 2014: 1-6. DOI: 10.1109/ICME.2014.6890301
- [102] Soleymani M., Caro M.N., Schmidt E.M., Sha C.Y., Yang YH. (2013). 1000 songs for emotional analysis of music. 2nd ACM international workshop on Crowdsourcing for multimedia (CrowdMM '13). *ACM*, 2013:1-6. DOI: 10.1145/2506364.2506365
- [103] Soleymani M., Lichtenauer J., Pun T., Pantic M. (2012) A Multimodal Database for Affect Recognition and Implicit Tagging. *IEEE Transactions on Affective Computing*, Volume 3(1). 2012:42-55. DOI: 10.1109/T-AFFC.2011.25
- [104] Song Y., Dixon S., & Pearce M.T. (2012). A Survey of Music Recommendation Systems and Future Perspectives. In *Proceedings of the 9th International Symposium on Computer Music Modeling and Retrieval (CMMR'12)*.
- [105] Sourina O., Liu Y. (2011). A Fractal-based Algorithm of Emotion Recognition from EEG using Arousal-Valence Model. *BIO SIGNALS*. *Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing*. 2011: 209-214.
- [106] Sourina, O., Wang, Q., Liu, Y., and Nguyen, M.K. (2011). A Real-time Fractal-based Brain State Recognition from EEG and its Applications. *International Joint Conference on Biomedical Engineering Systems and Technologies BIOSTEC*, 2011: 258-272.
- [107] Stober S., Sternin A., Owen A.M., Grahn J.A. (2015). Deep Feature Learning for EEG Recordings. *ArXiv*, *abs/1511.04306*.
- [108] Subramanian R., Wache J., Abadi M. K., Vieriu R. L., Winkler S., Sebe N. (2018) ASCERTAIN: Emotion and Personality Recognition Using Commercial Sensors. *IEEE Transactions on Affective Computing*. Volume 9(2), 2018: 147-160. DOI: 10.1109/TAFFC.2016.2625250
- [109] Takahashi K. (2004). Remarks on Emotion Recognition from Bio-Potential Signals. 2nd International Conference on Autonomous Robots and Agents, 2004:186-191
- [110] Tellegen A., Watson D. and Clark L. A. (1999). On the Dimensional and Hierarchical Structure of Affect. *Psychological Science*, Volume 10 (4), 1999: 297-303.
- [111] Thaut M.H. (2005). *Rhythm, music, and the brain: Scientific foundations and clinical applications*. Routledge, New York. ISBN 0415973708.
- [112] Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere. The Million Song Dataset. 12th International Society for Music Information Retrieval Conference ISMIR 2011: 591-596
- [113] Tseng K.C., Lin B.S., Han C.M. and Wang P.S. (2013). Emotion Recognition of EEG Underlying Favourite Music by Support Vector Machine. 1st International Conference on Orange Technologies (ICOT), 2013: 155-158. DOI: 10.1109/ICOT.2013.6521181
- [114] Turnbull D., Liu R., Barrington L., & Lanckriet, G.R. (2007). A Game-Based Approach for Collecting Semantic Annotations of Music. *International Society on Music Information Retrieval ISMIR*, 2007: 535-538
- [115] Tzanetakis G. and Cook P. (2002). Musical Genre Classification of Audio Signals. *IEEE Transactions on Speech and Audio Processing*, Volume 10(5), 2002: 293-302.
- [116] Van Der Zwaag M.D., Dijksterhuis C., De Waard D., Mulder B.L.J.M., Westerink J.H.D.M. and Brookhuis K.A. (2012). The influence of music on mood and performance while driving. *Ergonomics*, Volume 55(1), 2012: 12-22. DOI: 10.1080/00140139.2011.638403
- [117] Vijayan A.E., Sen D. and Sudheer A.P. (2015). EEG-based Emotion Recognition using Statistical measures and Auto-regressive modelling. *IEEE International Conference on Computational Intelligence & Communication Technology*, 2015: 587-591. DOI: 10.1109/CICT.2015.24
- [118] Wagner J., Kim J. and Andre E. (2005). From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extraction and Classification. 2005 IEEE

International Conference on Multimedia and Expo, 2005: 940-943.

- [119] Weihs C., Ligges U., Mörchen F., Müllensiefen D. (2007). Classification in music research. *Advances in Data Analysis and Classification*, Volume 1(3), 2007: 255-291.
- [120] 1111 S. (2013). Using Physiological Signals for Emotion Recognition. 6th International Conference on Human System Interactions, 2013: 556-561. DOI: 10.1109/HSI.2013.6577880
- [121] Yang Y., Chen H.H. (2012). Machine Recognition of Music Emotion: A Review. *ACM Transactions on Intelligent Systems and Technology (TIST)*, Volume 3(3), 2012. 40:1-40:30. DOI: 10.1145/2168752.2168754.
- [122] Yang Y., Lin Y., Su Y., & Chen H.H. (2008). A Regression Approach to Music Emotion Recognition. *IEEE Transactions on Audio, Speech, and Language Processing*, Volume 16(2), 2008: 448-457. DOI: 10.1109/TASL.2007.911513
- [123] Yang Y., Liu C.C., & Chen H.H. (2006). Music emotion classification: a fuzzy approach. In *Proceedings of the ACM International Conference on Multimedia*, 2006: 81–84. DOI: 10.1145/1180639.1180665
- [124] Yang Y., Su Y., Lin Y., & Chen H.H. (2007). Music emotion recognition: the role of individuality. In *Proceedings of the international workshop on Human-centered multimedia*. 2007: 13-22. DOI: <https://doi.org/10.1145/1290128.1290132>
- [125] Yıldırım Ö., Baloglu U.B., Acharya U.R. (2018). A deep convolutional neural network model for automated identification of abnormal EEG signals. *Neural Computing and Applications*, 2018:1-12. DOI: 10.1007/s00521-018-3889-z
- [126] Zheng W., & Lu B. (2016). A Multimodal Approach to Estimating Vigilance Using EEG and Forehead EOG. *Journal of neural engineering*, Volume 14 (2), 2016: 026017. DOI:10.1088/1741-2552/aa5a98
- [127] Zheng W., Lu B. (2015). Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks. *IEEE Transactions on Autonomous Mental Development*, Volume 7(3), 2015:162-175. DOI: 10.1109/TAMD.2015.2431497
- [128] Zheng W., Zhu J., & Lu B. (2016). Identifying Stable Patterns over Time for Emotion Recognition from EEG. *IEEE Transactions on Affective Computing*, Volume 10(3), 2019: 417-429. DOI: 10.1109/TAFFC.2017.2712143
- [129] Zong C., & Chetouani M. (2009). Hilbert-Huang transform based physiological signals analysis for emotion recognition. *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, 2009: 334-339. DOI: 10.1109/ISSPIT.2009.5407547
- [130] Thayer R. E. (1989). *The Biopsychology of Mood and Arousal*. Oxford University Press, New York. ISBN0-19-506827-0
- [131] Fontaine J. R. J., Scherer K. R., Roesch E. B., & Ellsworth P. C. (2007). The World of Emotions is not Two-Dimensional. *Psychological Science*, Volume 18(12), 2007: 1050–1057. DOI: 10.1111/j.1467-9280.2007.02024.x
- [132] Stober S., Sternin A., Owen A.M., & Grahn, J.A. (2015). Towards Music Imagery Information Retrieval: Introducing the OpenMIIR Dataset of EEG Recordings from Music Perception and Imagination. 16th International Society for Music Information Retrieval Conference, 2015: 763–769. DOI: [doi.org/10.5281/zenodo.1416270](https://doi.org/10.5281/zenodo.1416270)
- [133] Alakuş T.B., & Türkoglu I. (2019). Feature Selection with Sequential Forward Selection Algorithm from Emotion Estimation based on EEG Signals. *Sakarya University Journal of Science*, Volume 23 (6). 2019: 1096-1105. DOI: 10.16984/saufenbilder.501799