Computers in Biology and Medicine 107 (2019) 182–196

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Human stress classification using EEG signals in response to music tracks

Anum Asif[a], Muhammad Majid[a,*], Syed Muhammad Anwar[b]

[a] Department of Computer Engineering, University of Engineering and Technology, Taxila, 47050, Pakistan
[b] Department of Software Engineering, University of Engineering and Technology, Taxila, 47050, Pakistan

ARTICLE INFO

Keywords:
Electroencephalography (EEG)
Human stress
Music
State trait anxiety inventory (STAI)
Feature extraction
Classification

ABSTRACT

Stress is inevitably experienced by almost every person at some stage of their life. A reliable and accurate measurement of stress can give an estimate of an individual's stress burden. It is necessary to take essential steps to relieve the burden and regain control for better health. Listening to music is a way that can help in breaking the hold of stress. This study examines the effect of music tracks in English and Urdu language on human stress level using brain signals. Twenty-seven subjects including 14 males and 13 females having Urdu as their first language, with ages ranging from 20 to 35 years, voluntarily participated in the study. The electroencephalograph (EEG) signals of the participants are recorded, while listening to different music tracks by using a four-channel MUSE headband. Participants are asked to subjectively report their stress level using the state and trait anxiety questionnaire. The English music tracks used in this study are categorized into four genres i.e., rock, metal, electronic, and rap. The Urdu music tracks consist of five genres i.e., famous, patriotic, melodious, qawali, and ghazal. Five groups of features including absolute power, relative power, coherence, phase lag, and amplitude asymmetry are extracted from the preprocessed EEG signals of four channels and five bands, which are used by the classifier for stress classification. Four classifier algorithms namely sequential minimal optimization, stochastic decent gradient, logistic regression (LR), and multilayer perceptron are used to classify the subject's stress level into two and three classes. It is observed that LR performs well in identifying stress with the highest reported accuracy of 98.76% and 95.06% for two- and three-level classification respectively. For understanding gender, language, and genre related discriminations in stress, a *t*-test and one-way analysis of variance is used. It is evident from results that English music tracks have more influence on stress level reduction as compared to Urdu music tracks. Among the genres of both languages, a noticeable difference is not found. Moreover, significant difference is found in the scores reported by females as compared to males. This indicates that the stress behavior of females is more sensitive to music as compared to males.

1. Introduction

Stress in everyday terms has been recognized as a state, where an individual feel burdened and struggles to manage with psychological or social demands. These demands can be related to work, relationships, finances, and other similar situations in which an individual must perform under sheer pressure. A rapid change in the surroundings that requires the human body to react and adjust is typically described as stress. It is the body's natural defense against predators and threats, where the body gears up to evade or confront the perceived danger. This is termed as a fight or flight mechanism. Stress impacts an individual's way of living and has the potential to be a major health threat facilitating illness and disease [1,2]. It can be detrimental for both physical and mental well-being [3]. Stress is a risk element for hypertension [4], stroke [5], coronary artery disease [6], cardiac arrest [7–9] and other physical disorders such as irritable bowel syndrome (IBS), back pain, and gastro esophageal reflux disease (GERD) [10]. Similarly, mental illnesses like depression and generalized anxiety disorder are caused by stress [11].

Psychological stress has been linked as a co-factor of various diseases. Therefore, it is essential to learn how to effectively measure stress. Several responses are used to quantify the stress level and its fluctuations. Clinically, an individual's stress level has been assessed using self-report questionnaires, which are subjective indicators. These questionnaires are based on rating stress levels by an individual on various scales such as relative stress scale, perceived stress scale (PSS) [12], and state-trait anxiety inventory (STAI) [13]. Alternatively, physiological and physical measures have also been employed as objective methods to quantify stress [14]. Some physical features sensitive to stress are facial expressions [15], blink rate [16], pupil dilation [17,18],

* Corresponding author.
E-mail address: m.majid@uettaxila.edu.pk (M. Majid).

https://doi.org/10.1016/j.compbiomed.2019.02.015
Received 6 August 2018; Received in revised form 19 February 2019; Accepted 19 February 2019

eye gaze [19] and voice [20–22]. On the other hand, stressful and emergency situations induce dynamic changes in autonomous nervous system (ANS) [23], which is organized into sympathetic and parasympathetic nervous system. They regulate heart rate variability (HRV) [24], blood pressure, skin conductance [25], respiration [26], and brain waves. However, latest neuroscience considers the human brain as the main target of mental stress, because of its ability to determine whether, a situation is stressful and threatening [27,28]. The functional changes in the brain activity are commonly analyzed using electroencephalography (EEG) [29], functional magnetic resonance imaging (fMRI) [30], and positron emission tomography (PET) [31]. EEG is considered as a suitable and non-invasive neuroimaging modality to obtain the cortical response to stress. It is preferred due to its low cost, less intrusive equipment, and high temporal resolution. Moreover, EEG has shown association with other response measures such as heart rate and HRV specifically in terms of stress [32].

Stress is an inevitable phenomenon and a prolonged stress related experience is generally linked with the poor well-being of an individual. Therefore, it is important to develop efficient stress prevention and management approaches, which help an individual to recover the lost stability. Different intermediations have been introduced in order to promote an efficient stress recovery. Music listening is one of the most frequently used approaches [33,34]. Music is an essential part of human life, which can be used to express as well as evoke emotions [35–37]. The controlled use of specific kind of music and its ability to influence behavioral, emotional, and psychological changes in human being during the treatment of a disability or illness is usually referred to as music therapy [38,39]. It has shown beneficial effects on stress related physiological and emotional processes [40,41]. Two brain structures i.e., hippocampus and amygdala are associated with the music induced emotions and are known to be involved in regulating the hypothalamic–pituitary–adrenal (HPA) axis [42], which is activated when a situation is interpreted as stressful. ANS is another prominent stress-sensitive system in the human body, which is involved in mediating physiological changes induced by music. Therefore, the effect and influence of music listening on stress and anxiety has been studied extensively. Numerous studies have shown that the music could improve an individual's emotional health by reducing the stress level. A significant reduction in self-reported anxiety under controlled laboratory conditions are highlighted in several studies [43,44]. An alteration in the brain activity has been reported due to music listening, which promotes the relaxation state [45]. Thus, the use of music has received special interest as a non-invasive, easily implementable, economic, and highly acknowledged tool in stress management and stress related health concerns.

Several models have been developed by researchers for stress assessment using EEG signals, which determine stress level based on the brain activity. The electrical signals recorded in EEG are generated by the underlying neural activity in the brain. These waveforms are recorded by electrodes placed on the human scalp and are usually characterized by shape, amplitude and frequency. The frequency and amplitude variations in EEG signals are measured in hertz (Hz) and microvolt range respectively. The frequency behavior of the EEG signals is generally classified into five different bands i.e., delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–32 Hz), and gamma ($>$ 32 Hz), where each band could signify a different physiological state of mind. A state of mind can be activated by using different stimuli such as audio and video [37]. During negative emotions, the activities in right hemisphere take over the activities in left hemisphere [46], which indicate the area for stress detection. Alpha and beta bands signify consciousness, while delta and theta bands represent unconscious state of mind [47]. Rapid beta wave frequencies are considered as the characteristic indicator of stress. For EEG based stress evaluation methods, researchers have employed various features and classification techniques. The features used for the purpose of stress classification are frequency band power, peak frequency in alpha band, cross-correlation between band powers and Hjorth parameters, which are time-based characteristics of EEG waveform [48]. Similarly, the ratio of power spectral densities of alpha and beta bands has been computed for the analysis of physical stress [49]. EEG signal has been analyzed in frequency, time, and spatial domain using signal processing techniques such as Fourier transform. The spatio-temporal EEG features are reported to successfully classify traumatic brain injury and controlled condition [50]. The changes in EEG absolute power and other connectivity measures such as coherence have been observed due to stress [51]. EEG alpha asymmetry is found to be influenced during stress therapy [52]. It has also been utilized in a virtual reality environment to reveal stress related issues [53]. Some other features such as EEG eigenvalue decomposition [54] and brain wave balancing index [55] has been discussed in studies to assess the stress level. In the context of stress identification and stress state classification using EEG based approaches, different computational methods are used, such as decision trees [56], artificial neural network (ANN) [57], support vector machines (SVM) [58,59], random forest [60], Bayesian classifiers [46], and K-nearest neighbors [54,61]. For an improved measure of stress, EEG has been used in fusion with other modalities such as electrocardiography (ECG) [62] and skin conductance [59].

Music listening is assumed to influence stress related cognitive and physiological responses as it can initiate multitude of cognitive processes in human brain [63]. In the context of music listening after a stressful experience, a decline in the perceived level of psychological stress and an increase in the stress coping ability on an individual is reported [64,65]. Music listening has been shown to effect anxiety, which is considered as an adaptive response to a stressful or threatening experience. As music listening is assumed to trigger brain activities related to the intense emotional experience [66–68], therefore it could modulate anxiety levels induced by a stressful experience. The most consistent findings in this line of research have reported a reduction in anxiety levels, after listening to music both in the laboratory settings [69,70] and in out-of-lab experiments [71]. However, not all findings reported a decrease in anxiety after music listening [72,73]. The selection of appropriate music is another significant consideration, since all kinds of music cannot successfully reduce stress. The effect of different music genres on reducing stress has been explored in several studies. In particular, classical music is found to be more stress relieving than non-classical music such as heavy metal and hard rock [70]. Similarly, sedative music or silence reduces tension and has a relaxing effect, as compared to stimulative music or noise [74,75]. In contrast, no significant difference in stress reduction has been found in the effects of silence, sedative, and stimulative music [76]. Therefore, due to an inconsistent pattern of reported results, a firm conclusion cannot be drawn regarding the influence of music on emotional and cognitive components of stress response.

Besides the lack of research in exploring and investigating the influence of music on stress response, the existing literature report discrepant findings, when physiological markers of stress are compared with self-reported stress. The methodological limitations, such as a small sample size and different music selection might be the reason for such discrepancies. Even though music listening seems to have an inherent ability to reduce the psychobiological stress response, but due to the inconsistent research pattern, a definitive conclusion about its effectiveness cannot be drawn. In view of these considerations, we set out to examine the effect of music listening on human stress response using brain signals for healthy participants in a laboratory setting. We hypothesize that music listening could affect the EEG rhythms, which can quantitatively identify the stress level if analyzed properly. Furthermore, we are interested in whether language, gender, and genre related differences influence stress reducing effects of music listening. Music tracks from two languages i.e., English and Urdu are considered for the purpose of analysis. Five groups of features are extracted from EEG signals recorded by a commercially available MUSE headband. These are used to classify human stress into multiple states using four different classifiers including sequential minimal optimization (SMO), stochastic gradient descent (SGD), logistic regression (LR), and multilayer perceptron (MLP). The major contributions of the paper in this line of research are,

1. A new dataset based on EEG signals while listening to music tracks of two different languages i.e., English and Urdu are acquired by using a commercially available four-channel brain sensing MUSE headband.
2. An objective framework of features based on EEG signals is proposed to enhance human stress classification accuracy.
3. The influence of different language music tracks, music genres, and gender related differences in stress reducing effect of music listening are analyzed.

The rest of the paper is structured as follows. The detailed methodology is explained in Section 2. Section 3 describes the experimental results of stress classification in response to music and the statistical significance of music with respect to gender and language is presented. The discussion of results and comparison is provided in Section 4, followed by conclusion in Section 5.

2. Proposed methodology

The proposed framework for stress classification from EEG signals in response to music stimuli is shown in Fig. 1, which consists of four steps i.e., EEG data acquisition, preprocessing, feature extraction, and classification. The brain activity of 27 subjects is recorded using a four-channel brain sensing MUSE EEG headband in response to different language music tracks. The noise in the recorded EEG signals is removed by using the built-in preprocessing mechanism of MUSE headband. After preprocessing, five groups of features are extracted from each band of each channel, which are arranged in the form of a row vector for each subject. The subjects are labeled based on the state anxiety score after listening to the music track. Four different classification algorithms are used to classify human stress into multiple stress states. To validate the classification algorithms, a 10-fold cross validation is used, which splits the data into ten equal segments. For each round, nine segments are treated as training data and one segment is treated as test data. The following subsections describe the details of each block.

2.1. EEG data acquisition

This stage includes recording of the brain signals of participants, before and after the music stimuli is presented. The details of participants involved, music stimuli, apparatus used for data acquisition, self-report questionnaires, and experimental protocol followed in this study is presented in the following subsections.

2.1.1. Participants

A total of 30 healthy subjects including 15 males and 15 females participated voluntarily in this study with a signed consent. The participants belonged to diverse cultural and educational backgrounds with Urdu being their first language and ages ranging from 20 to 35 years. All participants have normal hearing power and no reported history of neurological disorder. Out of 30 subjects, the data for 3 participants (1 male and 2 female) were excluded because of the corrupted EEG recordings. The participants were naïve in respect of this study. Before the start of the experiment, the subjects were informed about the protocol and scope of the study. The participation in the study was voluntary and the procedures were followed according to the principles of Helsinki declaration. The study was approved by the board of advanced studies and research at the University of Engineering and Technology, Taxila.

2.1.2. Music stimuli

As an external stimulus, nine audio music tracks in mp3 format sampled at 44.1 kHz were used for the stress classification experiment based on EEG signals. Out of these 9 tracks, 4 tracks are in English, while the remaining 5 tracks are in Urdu language. The English music tracks consisted of four genres i.e., rock (0–20.2 kHz), metal (0–16 kHz), electronic (0–16.1 kHz), and rap (0–15.1 kHz). The Urdu music tracks included five genres i.e., famous (0–15.9 kHz), patriotic (0–15.8 kHz), melodious (0–20.2 kHz), qawali (0–10.6 kHz), and ghazal (0–19 kHz). The music tracks for each genre were selected based on their YouTube rating. To avoid the likelihood of boredom during the experiment, an extract of 1-min from each music track is used as a stimulus.

2.1.3. Self-reports and questionnaires

The clinical and socio demographic data about psychological symptoms like depression, stress, and anxiety is collected through self-report questionnaires filled by the participants. Initially, a demographic questionnaire was used to collect information such as name, age, gender, and education etc. To analyze the effect of different situations on an individual's feelings and to assess the perception of stress, state-trait anxiety inventory (STAI) questionnaire was used. It is a psychological questionnaire typically used to assess the state and trait anxiety in a clinical or non-clinical setting [13]. STAI consists of two forms i.e., Y1 and Y2, both having a total of 20 questions, to assess state anxiety and trait anxiety respectively. State anxiety refers to a person's feeling at the time of perceived danger such as fear and discomfort. On the other hand, the trait anxiety is usually linked to the feelings across typical situations experienced daily such as worry and stress. Both forms are based on a four-point scale and have a score range between 20 and 80.

2.1.4. Apparatus

The electrical activity of the brain was recorded using a commercially available non-invasive four channel MUSE EEG headband at a sampling rate of 256 Hz. EEG data obtained by MUSE headband gives a real time insight into the human brain. The headband is easily adjustable, and the electrodes are placed on the scalp according to the

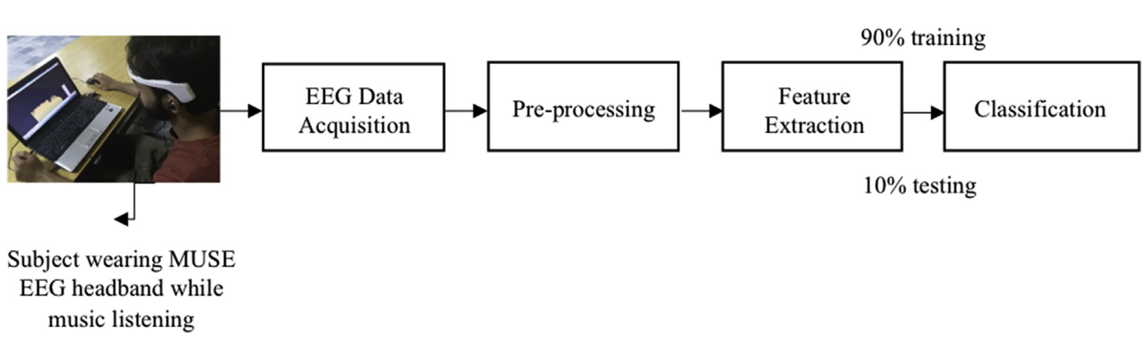

Fig. 1. Stress classification framework using EEG signals.

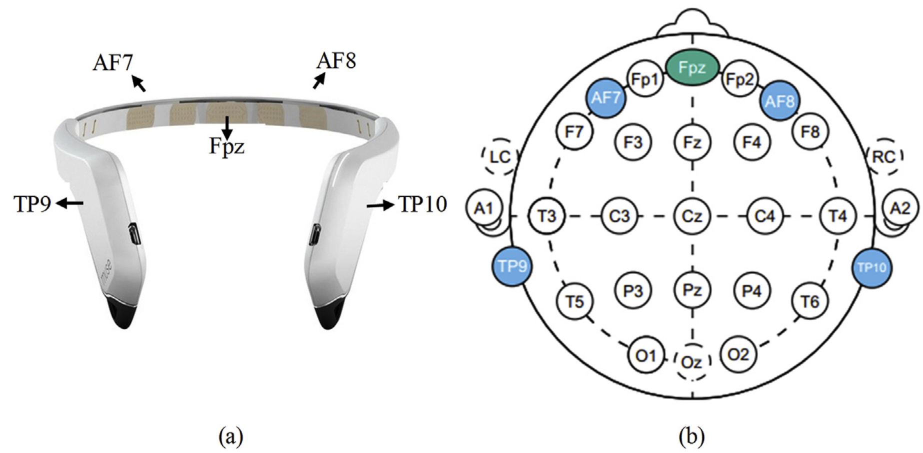

Fig. 2. (a) MUSE EEG headband (b) MUSE EEG headband electrode positioning according to 10–20 electrode positioning system.

10–20 electrode positioning system at locations AF7, AF8, TP9, and TP10, with a reference electrode at Fpz. The material used for AF7 and AF8 is made of silver, while conductive silicon rubber is used for TP9 and TP10. MUSE EEG headband along with the electrode positioning according to the 10–20 positioning system is shown in Fig. 2. It offers a wide range of benefits such as flexibility, extreme accessibility and lightweight. Moreover, it is wireless and can be paired with any smartphone or tablet device. The EEG data was recorded on a smart phone by using the MUSE monitor application, which is then transferred via Bluetooth for offline processing. All EEG recordings were carried out in a noise free and isolated environment.

2.1.5. Experimental procedure

Before the start of experiment, the entire experimental procedure was explained to each participant. On the day of recording, the experimental setup was organized such as to record brain signals from EEG MUSE headband, while listening to the music. Fig. 3 represents the experimental procedure followed for data acquisition in this study. The EEG data was acquired during the stages represented in blue color. First, the participants filled the consent, demographic, and trait anxiety (Y2) forms, followed by the baseline EEG recording of 3 min. After baseline recording, the state anxiety form (Y1(i)) was filled by the participant and then a music track belonging to a genre and language was presented. The subject was seated on a comfortable chair and headphones were used for music listening. The loudness of the music track was adjusted according to the participant's comfort level. The EEG signals of the participant were also recorded, while listening to the music track by using the MUSE EEG headband. After listening to the music track, the state anxiety form (Y1(ii)) was again filled by the subject, which was used for labeling the data for stress classification task. The same procedure was repeated three times for each subject by presenting music tracks of a different language and genre.

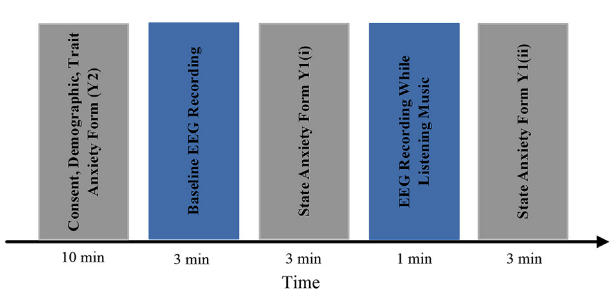

Fig. 3. The experimental procedure followed for data acquisition in the proposed study.

2.2. Signal preprocessing

The recorded EEG signals are preprocessed to remove noise, before feature extraction and classification. The MUSE headband provides both raw and preprocessed EEG signals. A notch filter is applied to the raw EEG data, whose value can be adjusted. Notch filter is usually referred to as a band rejection or band stop filter with a narrow stop band. The raw EEG signals are passed through the filter to remove frequencies between 45 and 64 Hz inclusive. A computationally efficient, noise-free preprocessed signal in MUSE is obtained by over-sampling followed by down-sampling to yield an output sampling rate of 256 Hz, with 2uV (rms) noise. Active noise cancellation is achieved by using an on-board driven right leg (DRL) circuit between the frontal electrodes and Fpz electrode. DRL is a feedback circuit designed to remove the noise from EEG signals and ensures that the EEG electrodes have proper skin contact. The characteristics of the incoming EEG signal i.e., variance, amplitude and kurtosis are used in a decision tree, where signals with low values for variance, amplitude, and kurtosis are regarded as clean. On-board digital signal processing module performs fast Fourier transform on the raw EEG signal using a 256-sample window size with an overlap of 90% to give the required frequency bands. The five EEG bands obtained from muse headband include delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–44 Hz) bands. These frequency ranges are specified in the MUSE hardware design specifications.

2.3. Feature extraction

EEG signals contain a large amount of information regarding the functional patterns of brain. The main objective of feature extraction from EEG signals is to extract meaningful information, which can be used for the classification task. In this study, five groups of features are extracted from preprocessed EEG signals from four channels (AF7, AF8, TP9, TP10) and five bands. The details of each feature group are presented in following subsections.

2.3.1. Absolute power

For the estimation of absolute power, EEG signal is first transformed into frequency domain using the fast Fourier transform (FFT). A 256-sample tapered cosine window with 75% overlapping is used to apply FFT. The cosine tapered window, which is a rectangular window with the initial and last sample equal to the parts of cosine, is computed as,

$$w(n) = \begin{cases} \frac{1}{2}\left\{1 + \cos\left(\frac{2}{u}\left[n - \frac{u}{2}\right]\right)\right\}, & 0 \le n \le \frac{u}{2} \\ 1, & \frac{u}{2} \le n \le -\frac{u}{2} \\ \frac{1}{2}\left\{1 + \cos\left(\frac{2}{u}\left[n - 1 + \frac{u}{2}\right]\right)\right\}, & 1 - \frac{u}{2} \le n \le 1 \end{cases} \tag{1}$$

The EEG absolute power is computed by taking the absolute of FFT of signal and dividing it by the length of the signal. The EEG absolute power is calculated for all four channels across five frequency bands comprising of 20 features i.e., f_1-f_{20}.

2.3.2. Relative power

Relative power is typically derived from absolute power and is used to determine the rhythmicity of EEG signals. It is computed by dividing the specific band power to the total power in all the bands and is calculated as,

$$Relative\ Power = \frac{Power\ in\ band}{Total\ power} \times 100\% \tag{2}$$

The relative power is also calculated for each of the four scalp locations and five frequency bands comprising of 20 features i.e., f_{21}-f_{40}.

2.3.3. Coherence

To determine the functional interaction between brain locations, a connectivity measure known as coherence is computed. It reports the degree of association of EEG recorded at two scalp locations. The idea behind using coherence as a feature is to indicate the functional connectivity between the stress and non-stress conditions. The mathematical representation of coherence is given as,

$$C_{xy} = \frac{|H_{xy}^2|}{|H_x||H_y|} \tag{3}$$

where H_{xy} is the cross spectral density between x and y, and H_x and H_y are the autospectral density of x and y respectively. It is the ratio between the cross spectrum of the two signals and auto spectra of the individual signals. Coherence is calculated for six pairs of channel locations across each frequency band. The corresponding six pairs of channel locations are (AF7, AF8), (AF7, TP9), (AF7, TP10), (AF8, TP9), (AF8, TP10), and (TP9, TP10) respectively. The number of features extracted in this group is 30, which are represented as f_{41}-f_{70}.

2.3.4. Phase lag

Phase lag is another functional connectivity measure, which is used to determine the lead or lag between two EEG signals from distinct brain locations. The phase is defined as the ratio of FFT's quadrature components and is generally described as,

$$\theta = \arctan\left(\frac{b}{a}\right) \tag{4}$$

where a represents the in-phase real component, while b represents the out of phase imaginary component of signal computed from the FFT. By subtracting the individual phases of the signals from two locations, phase difference is calculated as,

$$\theta - \theta_1 = \arctan\left(\frac{b2}{a2}\right) - \arctan\left(\frac{b1}{a1}\right) \tag{5}$$

The square root of the squared phase difference results in the absolute phase delay. The features of phase difference are computed for six electrode pairs for each frequency band comprising of 30 features i.e., f_{71}-f_{100}.

2.3.5. Amplitude asymmetry

Amplitude asymmetry is a measure of brain connectivity extracted from EEG signal, which indicates the relative excitation between two electrode locations. The difference between amplitudes of the signals and its normalization by the sum of these amplitudes is used to find the asymmetry and is calculated as,

$$Asymmetry = \frac{X - Y}{X + Y} \tag{6}$$

where X and Y indicates the amplitudes of signals at particular instances. Amplitude asymmetry is also computed at six electrode pairs across five frequency bands comprising of 30 features i.e., f_{101}-f_{130}. The features extracted from each group are concatenated to create a feature vector represented as,

$$FV = [f_1 \quad f_2 \quad f_3 \cdots f_{129} \quad f_{130}] \tag{7}$$

The features extracted from the EEG recordings and the corresponding class labels are arranged in the form of a data matrix. The columns of the data matrix contain feature values f_1-f_{130} for each EEG recording, when a subject is listening to each type of music. The rows of the data matrix indicate data points, which depends on the number of subjects and EEG recordings, while listening to the music track. As 27 subjects listens to 3 different music tracks, therefore the number of rows in the data matrix is 81 in this study. All 130 features are selected to train the classification model. Five different types of features are considered in this study, the length of the feature vector indicates how these feature values are made more representative by considering various EEG bands and electrode pairs. The class labels are obtained by the state anxiety form Y1(ii) filled by the subject after listening to each music track.

2.4. Classification

Different classification algorithms have been used to classify EEG recording in to stress categories based on the meaningful information learned from the extracted features [77]. To model the association between the set of features and corresponding target conditions, the following classifiers are used in this study.

2.4.1. Sequential minimal optimization (SMO)

Sequential minimal optimization is a classification algorithm extensively used to train support vector machines via efficient quadratic programming (QP) solvers [78]. SVM is a supervised machine learning model, which analyzes data for the purpose of classification [79]. The feature space is classified based on the hyperplane that separates the stress and non-stress conditions using class labels. However, an optimization problem arises during training process of SVM, which may refrain, from obtaining optimal hyperplane. Among several specialized approaches, SMO is one way to resolve the quadratic programming problem inherent in SVM. The iterative algorithm used by SMO, partitions the optimization problem into a series of smaller sub-problems. These QP based sub-problems are then solved analytically by entirely avoiding numerical QP optimization.

2.4.2. Stochastic gradient descent (SGD)

Stochastic gradient descent also known as incremental gradient descent is an algorithm for learning a wide range of linear models in machine learning such as binary class support vector machines, logistics regression, linear regression and other graphical models [80]. It uses iterative method and statistical estimation to analyze the problem of objective function minimization, which is usually expressed in the form of sum of differential functions. The algorithm probes through the training set until convergence and converts nominal attributes into target binary conditions i.e., stress and non-stress. As compared to ordinary gradient descent, SGD reduces the computational cost by economically evaluating the true gradient from just a few examples instead of the entire training set [81].

2.4.3. Logistic regression (LR)

Logistic regression is a machine learning technique related to the

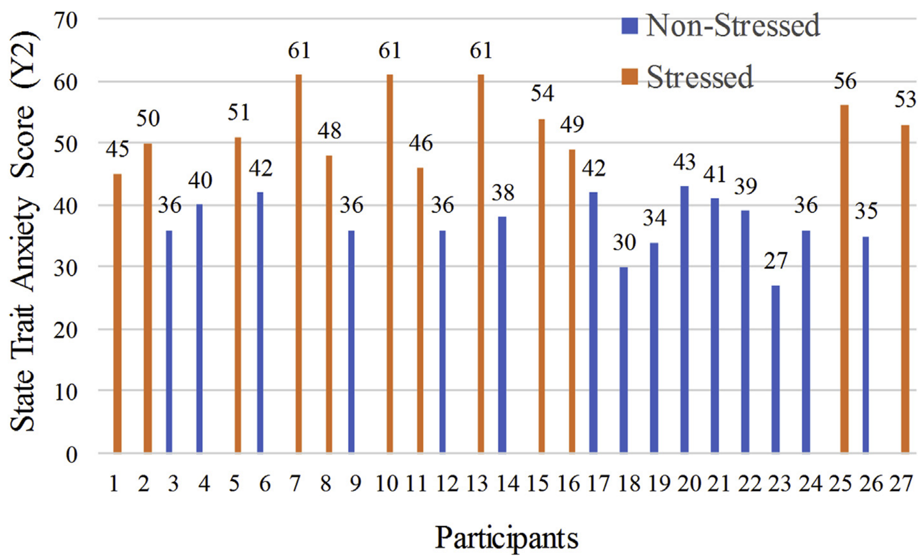

Fig. 4. Trait anxiety score of all the participants participated in this study.

Table 1
Language and Gender related differences among the subjects of stressed and non-stressed groups based on state anxiety scores before Y1(i) and after Y1(ii) listening music tracks using two sample *t*-test.

Stressed Group			Non-Stressed Group		
Language/ Gender	Null hypothesis (h)	p-value	Language/ Gender	Null hypothesis (h)	p-value
English	**1**	**0.0053**	English	0	0.5522
Urdu	0	0.7694	Urdu	0	0.6597
Male	0	0.7286	Male	0	0.8938
Female	**1**	**0.0099**	Female	**1**	**0.0126**

field of statistics, which is used to solve the binary classification problem. It is a linear probabilistic classifier that projects an input vector on to multiple hyperplanes, each corresponding to an individual class. LR uses the link function to show the relationship between the set of EEG features and the conditional outcome. A standard link function, also known as the logistic function is defined as,

$$F(t) = \frac{1}{1 + e^{-t}} \tag{8}$$

where t represents the linear combination of multiple variables such as EEG features and class labels. The logistic regression model is obtained from the logistic function by expressing t as,

$$t = \beta_0 + \beta_1 x_1 + \cdots + \beta_n x_n \tag{9}$$

The logistic function can now be expressed as,

$$F(t) = \frac{1}{1 + e^{-(\sum \beta_i x_i)}} \tag{10}$$

A likelihood value between the range of 0 and 1 is returned as a result of LR classifier, which indicates the association of the subject with stress and non-stress categories.

2.4.4. Multilayer perceptron (MLP)

Multilayer perceptron is the easiest and most commonly used feed forward neural network, which consists of multiple layers of computational units [82]. MLP uses supervised back propagation learning algorithm to classify instances [83]. It consists of an input layer, an output layer, and multiple hidden layers. The weighted inputs, to the output of each neuron are mapped by using transfer functions such as hyperbolic tangent, sigmoid, and rectified linear unit. In this case, a sigmoid function has been used to determine the state y_i, using a total weighted input given as,

$$y_i = \frac{1}{1 + e^{-xi}}, \tag{11}$$

The total weighted input X_i is calculated as,

$$X_i = \sum_j^n y_j W_{ij}, \tag{12}$$

where y_j represents the state level and the weight between the ith and jth connection is represented by W_{ij}. Furthermore, the classification task is a two-stage process i.e., training and testing. The basic objective of the training process is to determine the weight values that will match the neural network output with the specific target values as closely as possible.

2.4.5. Classification validation model

Once the classifier algorithm is selected, its performance needs to be assessed over a range of features until it converges. For this purpose, a 10-fold cross validation method is adopted in the proposed study to evaluate the performance of the classification algorithm. This provides better validation, since the number of instances is limited. In 10-fold cross validation, data of all subjects are divided into 10 equal portions. The classification model is trained using 9 portions, while one portion is utilized for testing and the process is repeated for different number of iterations. Every iteration gives a different value for the performance parameters, therefore average results of evaluation parameters are presented in result section.

3. Experimental results

A machine learning framework is proposed in this paper for stress classification using EEG signals in response to different music tracks. The experimental results are presented using, (i) statistical analysis of human stress behavior in response to different language and genres of music tracks, and (ii) performance evaluation of two- and three-class stress classification using EEG signals.

3.1. Statistical significance

For statistical analysis, participants involved in the study are categorized into stressed and non-stressed group based on their trait anxiety score i.e., Y2. The trait anxiety score for each subject participated in

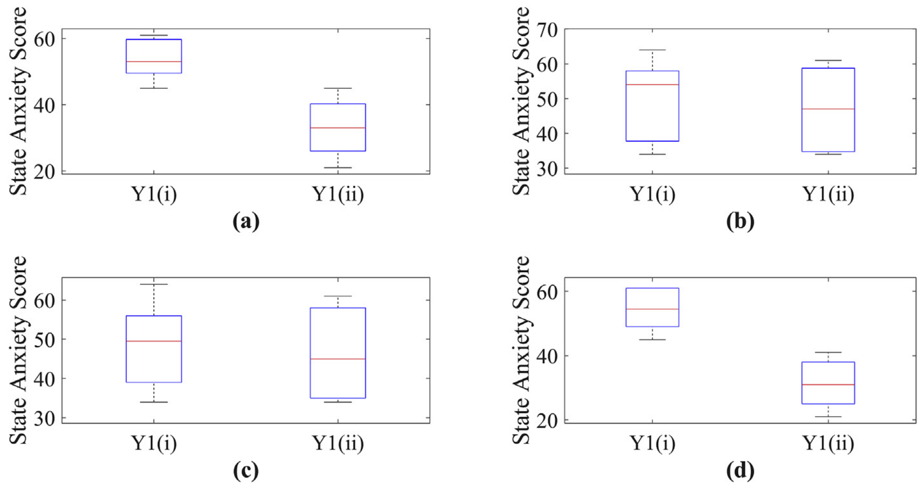

Fig. 5. Box plot representation of stressed group (a) participants listened to English Music (b) participants listened to Urdu music (c) Male participants (d) Female participants.

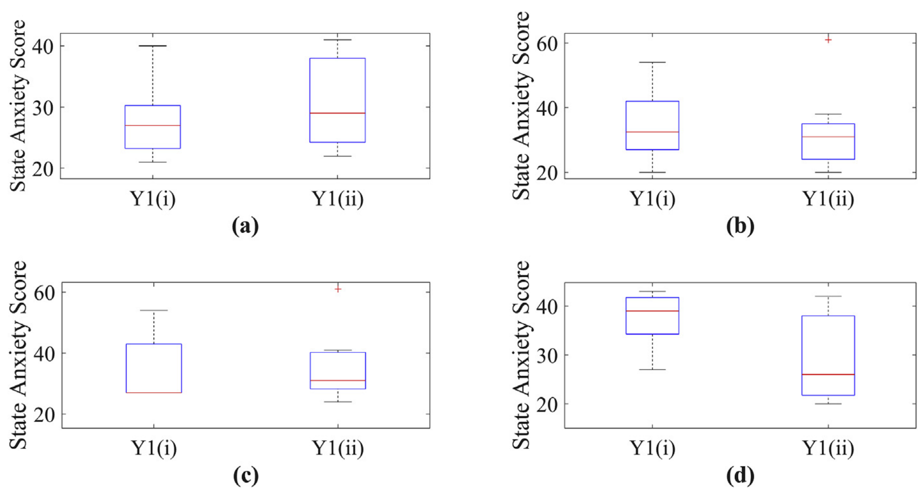

Fig. 6. Box plot representation of non-stressed group (a) participants listened to English Music (b) participants listened to Urdu music (c) Male participants (d) Female participants.

Table 2
Genre related differences among the subjects based on state anxiety scores before Y1(i) and after Y1(ii) listening music tracks of English music and Urdu music using one-way ANOVA.

Music Track Language (Genres)/State Anxiety Score	Null Hypothesis (h)	F-score	p-value
English (Rock, Electronic, Metal, Rap)/Y1(i)	0	0.27	0.4113
English (Rock, Electronic, Metal, Rap)/Y1(ii)	0	0.21	0.3695
Urdu (Famous, Patriotic, Melodious, Qawali, Ghazal)/Y1(i)	0	0.42	0.6430
Urdu(Famous, Patriotic, Melodious, Qawali, Ghazal)/Y1(ii)	0	0.65	0.7922

this study is shown in Fig. 4. A subject is in the stressed group, if their trait anxiety score is greater than the mean value of the trait anxiety score for all participants. The average trait anxiety score in this study is 44. The stressed group contains twelve subjects including six males and six females. Similarly, fifteen participants are part of the non-stressed group including eight males and seven females.

For a better understanding of gender and language related discrimination in stress reduction in response to music, both groups are analyzed by applying a two-sample *t*-test based on the state anxiety score before listening to the music track i.e., Y1(i) and after listening to the music track i.e., Y1(ii). The test decision is based on the value of null hypothesis (h) and p-value (p). If $p < 0.05$ and $h = 1$, then the differences are statistically significant in the groups. The *t*-test results based on language and gender for stressed and non-stressed groups are shown in the Table 1. As music tracks in two different languages are presented to the subjects, both groups contain participants who have listened to either English and Urdu music tracks. Subjects belonging to the stressed group shows a significant difference, when English music tracks are used a stimulus. On the other hand, no significant difference is found in stressed group, when Urdu music tracks are used. Moreover, no language related difference was found among the participants of non-stressed group. A similar kind of analysis is performed to find gender related differences for both groups. Females of both groups have shown statistical difference in their state anxiety scores. However, no significant difference is found in scores reported by males in both groups.

The distributional characteristics of the scores for stressed and non-stressed group are presented using box plots, which describe an overall

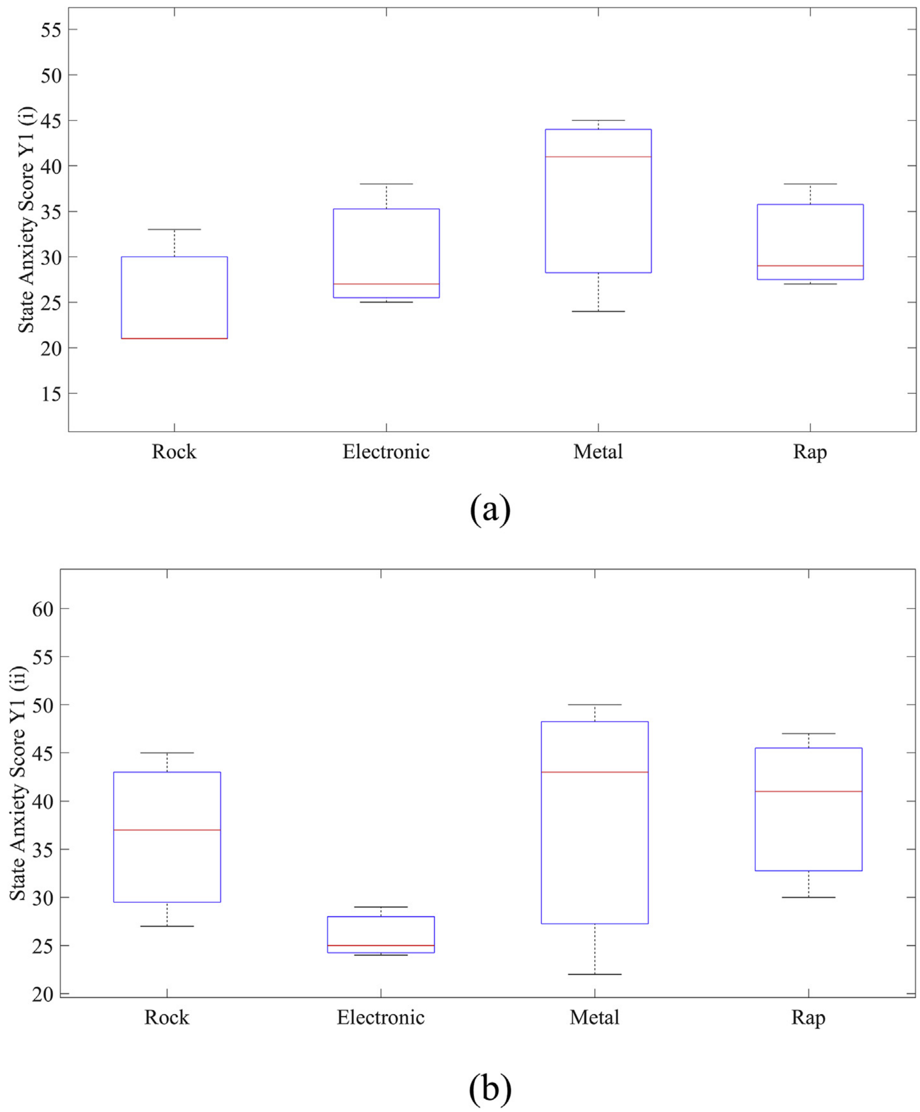

Fig. 7. Boxplot of English language music genres with respect to state trait anxiety scores (a) Y1(i) (b) Y1(ii).

response pattern for a group. The bottom and top 25% of the data values are represented by the whiskers, while the middle 50% scores for the group are represented by the interquartile range boxes. The centre of data represented by a line in the box is marked by the median. The box plot of scores for participants in response to listening to songs with English and Urdu lyrics, for both male and female subjects in a stressed group before listening to music, Y1(i) and after listening to music, Y1(ii) is presented in Fig. 5. The box plot of Y1(i) is much higher than the box plot of Y1(ii) in case of English music tracks. Moreover, an obvious variation can be seen between their median values. This suggests that the data values of Y1(ii) are significantly different from the data values of Y1(i). The distribution pattern differs between the box plots, when the scores of participants from stressed group who listened to music with Urdu lyrics are compared. However, no significant difference is found between the data values. Similarly, the variation between the middle quartile is not significant, which suggests that the difference is minimal. From the analysis of the distributions and median values of these box plots, it is evident that scores for male participants are not significantly different. However, the box plots for females show a significant difference in the data values. The characteristics of non-stressed group with respect to gender and language are also analyzed using box plots as shown in Fig. 6. The only notable difference is found in the scores of female participants that belonged to non-stressed group.

The difference between the level of stress based on music genres is also analyzed by using one-way analysis of variance (ANOVA), which is a statistical technique to assess the potential differences among several categories of data. In ANOVA, a larger F-statistic corresponds to a small p-value, which casts doubt on the null hypothesis and indicates the existence of significant difference among the groups. The significance level considered in this study is 0.05. The analysis is carried based on state anxiety scores of participants before and after listening English and Urdu music genres and the results are presented in Table 2. The p-value for English genres with respect to Y1(i) and Y1(ii) is 0.4113 and 0.3695 respectively. Similarly, the p-value of Urdu genres in both cases is 0.6430 and 0.7922, which indicates that no significant difference lies in the scores of participants who listened to the music of different genres. Fig. 7 and Fig. 8 shows the box plots for state anxiety scores of participants before and after listening to music tracks of different genres from English and Urdu language respectively. It is evident from the results that no significant difference is found among the state anxiety scores of participants in response to different genres of English and Urdu language music.

3.2. Classification results

In this study, each participant listened to three music tracks of 1 min

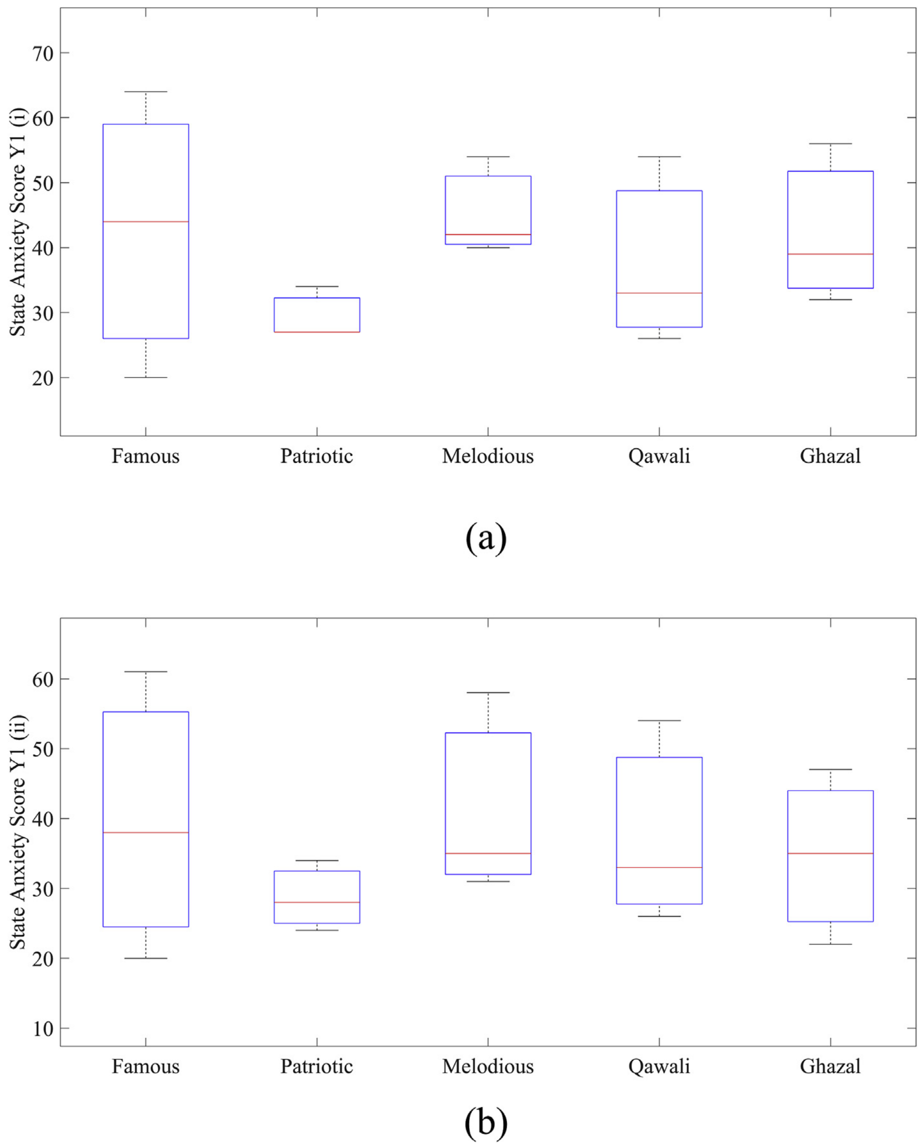

Fig. 8. Boxplot of Urdu language music genres with respect to state trait anxiety scores (a) Y1(i) (b) Y1(ii).

each from three different genres. Music is randomly presented to the participants, irrespective of their gender, age and preference. Each subject had to perform a subjective evaluation for scoring the anxiety level by filling the state anxiety form before and after listening to music. The state anxiety score obtained from subjective assessment after music listening i.e., Y1(ii) is used as class labels to train the classifiers. For the two-class problem, subjects are labeled into two groups i.e., stressed and non-stressed, whereas for the three-class problem subjects are labeled to be either in non-stressed, medium stressed and highly stressed. The mean and standard deviation (STD) of the state anxiety scores after listening to music for 27 subjects participated in this study is 36 and 11.8 respectively. For the two-class problem, participants having a Y1(ii) score less than the mean value i.e., 36 are considered as non-stressed and participants having a score greater or equal than 36 are considered as stressed. For the three-class problem, participants having Y1(ii) scores between 0 and 29, 30 to 42, and 44 to 80 are considered as non-stressed, medium stressed, and highly stressed respectively. These values are calculated using, $mean \pm \left(\frac{STD}{2}\right)$, and the results are shown in Fig. 9 (a) and Fig. 9 (b) for the two- and three-class scenarios respectively. Five groups of features i.e., absolute power, relative power, coherence, amplitude asymmetry, and phase lag are extracted from four channels i.e., AF7, AF8, TP9, and TP10 and five EEG bands i.e., delta, theta, alpha, beta, and gamma. The mean and standard deviation of five feature groups from all channels and bands are shown in Table 3 for the two- and three-class problems.

Four different classifiers including SMO, SGD, LR, and MLP are used to classify the subject's stress level into two and three states. To achieve consistent results, a 10-fold cross validation technique is used. The results of different classification algorithms are evaluated and compared based on accuracy, F-measure, kappa statistics, root mean square error (RMSE), and mean absolute error (MAE). The degree of closeness of classified samples to the true value of data samples is referred as accuracy. F-measure is defined as the fusion measure of precision and recall with a value ranging from 0 (worst value) to 1 (best value). Similarly, kappa statistics compares the observed and expected accuracy values. Kappa coefficient does not depend on the number of classes and number of samples per class. Its value lies within the range of 0 (chance-based classification) and 1 (perfect classification). MAE and RMSE parameters assess the error performance of the classifier.

Table 4 shows a comparison of performance metrics for different classifiers for the two-class stress classification problem based on features extracted from EEG signals recorded during music presentation. These results show that the overall performance of LR is better as compared to other classifiers with an accuracy of 98.76% and a kappa

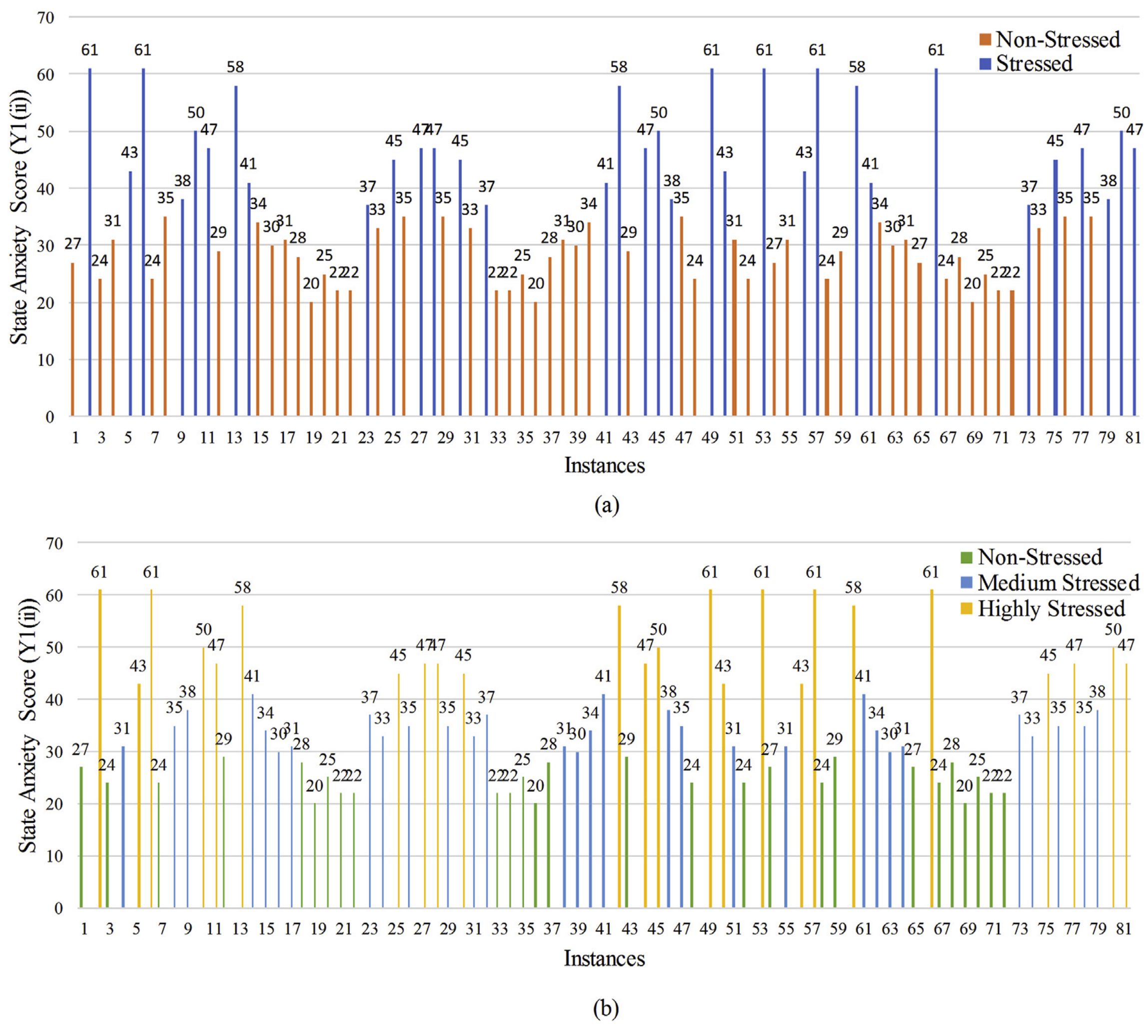

Fig. 9. Class labels based on state anxiety score after music listening for stress classification (a) two class (b) three class.

Table 3
Mean and Standard deviation of five feature groups extracted from EEG signals for two and three class stress classification problem.

Feature Name	Two Class				Three Class					
	Stressed		Non-Stressed		Low stress		Medium Stress		High Stress	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Absolute Power	410.1	128.1	413.3	86.3	417.6	656.7	423.0	671.46	357.8	480.3
Relative Power	0.2	0.002	0.2	0.002	0.2	0.002	0.2	0.0022	0.2	0.002
Coherence	0.298	0.058	0.301	0.040	0.285	0.036	0.284	0.0422	0.276	0.008
Amplitude Asymmetry	−0.069	0.038	−0.046	0.037	0.036	0.024	0.033	0.0329	0.038	0.030
Phase Lag	−0.001	0.004	−0.002	0.003	−0.001	0.004	−0.001	0.0048	−0.002	0.004

Table 4
Performance comparison of different classifiers for two-class stress classification based on EEG recordings while music listening.

Classifier	Accuracy (%)	Kappa	F-measure	MAE	RMSE
SMO	97.5309%	0.9493	0.975	0.0247	0.1571
SGD	96.2963%	0.9229	0.963	0.037	0.1925
LR	98.7654%	0.9743	0.988	0.0158	0.1126
MLP	92.5926%	0.8451	0.926	0.0959	0.2188

statistic value of 0.97. Similarly, LR gives F-measure, which is close to its best value, and smaller values for the error performance measures i.e., MAE and RMSE. The classification results of SMO and SGD are comparable with accuracies of 97.53% and 96.29%, respectively. The non-stress state is classified with a higher individual accuracy except for the case of SMO. However, significant difference is not seen in the individual accuracy values for the stress condition. The confusion matrices of SMO, SGD, LR, and MLP classifiers are shown in Table 5. It can be observed that non-stressed class is classified with a higher individual accuracy except for the case of SMO classifier.

Table 6 shows the comparison of performance metrics of different classifiers for two-class stress classification based on the features extracted from baseline EEG signals recorded before music presentation. It is evident from the results that LR outperforms all other classifiers with an accuracy of 94.44%. The confusion matrices of SMO, SGD, LR, and MLP classifiers are shown in Table 7. It is observed that the stressed

Table 5
Confusion matrices of different classifiers for two-class stress classification based on EEG recordings while listening music tracks.

SMO				SGD			
a	B	Classified as	Accuracy (%)	a	b	Classified as	Accuracy (%)
46	2	a = Non-stressed	95.8%	47	1	a = Non-stressed	97.9%
0	33	b = Stressed	100%	2	31	b = Stressed	93.9%
LR				**MLP**			
a	B	Classified as	Accuracy (%)	a	b	Classified as	Accuracy (%)
47	1	a = Non-Stressed	97.9%	46	2	a = Non-Stressed	95.8%
1	32	b = Stressed	97.0%	4	29	b = stressed	87.9%

Table 6
Performance comparison of different classifiers for two-class stress classification based on EEG recordings before music listening.

Classifier	Accuracy (%)	Kappa	F-measure	MAE	RMSE
SMO	85.1852%	0.7017	0.853	0.1481	0.3849
SGD	75.9259%	0.5185	0.761	0.2407	0.4907
Logistic	94.44%	0.8873	0.945	0.0654	0.2431
MLP	75.9259%	0.5118	0.761	0.258	0.4249

class has the highest individual classification accuracy, when compared to the non-stressed class.

A performance analysis of the three-class stress classification is also performed in this study. SMO, LR, and MLP classifiers are used to classify three classes namely non-stressed, medium stressed and highly stressed. SGD, being a binary classifier is not included in this case. Table 8 shows the comparison of performance metrics for different classifiers based on the features extracted from EEG signals recorded during music presentation. The highest reported accuracy is 95.06% using the LR classifier. The accuracies achieved by SMO and MLP are 85.185% and 88.89%, respectively. The confusion matrices of SMO, LR, and MLP classifiers for the three-class stress classification are shown in Table 9. It is observed that LR gives the highest individual accuracies. Moreover, medium stressed class is classified with the highest accuracy among other classes irrespective of the classifier used.

Table 10 shows the comparison of performance metrics for different classifiers for the three-class stress classification based on features extracted from baseline EEG signals recorded before music presentation. LR again outperforms other classifiers with an accuracy of 85.18% and kappa coefficient of 0.7769. However, the results of SMO and MLP classification are almost similar with an accuracy of 75.92% in both cases. The confusion matrices of SMO, LR, and MLP classifiers for the three-class stress classification problem are shown in Table 11. It is evident that LR gives the highest individual accuracies. Moreover, highly stressed class is classified with the highest accuracy among other classes irrespective of the classifier.

Table 8
Performance comparison of different classifiers for three-class stress classification based on EEG recordings while music listening.

Classifier	Accuracy (%)	Kappa	F-measure	MAE	RMSE
SMO	85.1852	0.777	0.852	0.2606	0.334
LR	95.0617	0.9255	0.951	0.0453	0.188
MLP	88.89	0.8329	0.889	0.1114	0.2478

Table 9
Confusion matrices of different classifiers for three-class stress classification based on EEG recordings while listening music tracks.

SMO				
a	b	c	Classified as	Accuracy (%)
25	4	1	a = Non-stressed	83.3%
3	20	1	b = Highly stressed	83.3%
2	1	24	b = Medium Stressed	88.9%
LR				
a	b	c	Classified as	Accuracy (%)
29	1	0	a = Non-stressed	96.7%
2	22	0	b = Highly stressed	91.7%
1	0	26	b = Medium Stressed	96.3%
MLP				
a	b	c	Classified as	Accuracy (%)
26	2	2	a = Non-stressed	86.7%
1	22	1	b = Highly stressed	91.7%
2	1	24	b = Medium Stressed	88.9%

Table 10
Performance comparison of different classifiers for three-class stress classification based on EEG recordings before music listening.

Classifier	Accuracy (%)	Kappa	F-measure	MAE	RMSE
SMO	68.518%	0.5248	0.683	0.3169	0.4091
LR	83.33%	0.7515	0.837	0.1166	0.3133
MLP	74.0741%	0.6103	0.740	0.1988	0.3547

Table 7
Confusion matrices of different classifiers for two-class stress classification based on EEG recordings before listening music tracks.

SMO				SGD			
a	B	Classified as	Accuracy (%)	a	b	Classified as	Accuracy (%)
26	6	a = Non-stressed	81.3%	23	9	a = Non-stressed	71.9%
2	20	b = Stressed	90.9%	4	18	b = Stressed	81.8%
LR				**MLP**			
a	B	Classified as	Accuracy (%)	a	b	Classified as	Accuracy (%)
29	3	a = Non-stressed	90.6%	24	8	a = Non-stressed	75%
0	22	b = Stressed	100%	5	17	b = Stressed	77.3%

Table 11
Confusion matrices of different classifiers for three-class stress classification based on EEG recordings before listening music tracks.

SMO				
a	b	c	Classified as	Accuracy (%)
12	4	4	a = Non-stressed	60%
5	10	1	b = Highly stressed	62.5%
3	0	15	b = Medium Stressed	83.3%
LR				
a	b	c	Classified as	Accuracy (%)
15	4	1	a = Non-stressed	75%
1	15	0	b = Highly stressed	93.8%
0	3	15	b = Medium Stressed	83.3%
MLP				
a	b	c	Classified as	Accuracy (%)
13	4	3	a = Non-stressed	65%
4	12	0	b = Highly stressed	75%
2	1	15	b = Medium Stressed	83.3%

4. Discussion

This paper presents a framework for the classification of stress in response to music tracks. Moreover, the influence of music belonging to different languages and genres on stressed and non-stressed groups is also determined. It is evident from results that English language music tracks help in reducing the stress level of participants in the stressed group. Similarly, female participants in the stressed group have shown a reduction in their stress level after listening to music. For the non-stressed group, music does not play any significant role for reducing stress level in male participants, but a reduction in stress level is observed in female participants. For two- and three-class stress classification, logistic regression is found to be the best classifier based on EEG recording of participants, while they are listening to music.

Some of the major studies that have analyzed human stress behavior in response to music are summarized and compared with the proposed method in Table 12. In Ref. [37], the effect of five different English music genres i.e., electronic, rap, metal, rock, and hip-hop on human emotions using brain signals is presented. The results indicated that rap and rock genres induced sad and happy emotions in the subjects, respectively. The effect of three different conditions i.e., relaxing music, sound of rippling water, and rest without acoustic stimulation prior to stress test is presented in Ref. [44]. Sixty female participants are assessed based on the salivary alpha amylase, cortisol, respiratory sinus, heart rate, subjective stress perception and anxiety level. The results reported lower cortisol concentration in the condition of rippling water before stress task. Electrocardiogram and respiration rate have been used to determine the effect of familiar and preferred music listening [74]. Stimulative music increased the heart and respiration rate and increased joyous and energetic feelings but have no effect on calm and relaxed emotions.

In Ref. [83], the impact of preferred music selected by a subject after stress exposure is presented. The study involves 58 participants, which are randomly divided into control and experimental groups. Self-reported emotional states and cardiovascular measures are monitored during the stress induction and recovery stages. Participants in experimental group have shown a higher level of cardiovascular recovery and a lower level of self-reported anxiety level. In gender related discrimination, females exhibit greater negative emotional states as compared to male participants. The impact of listening to music alone and in the presence of others is presented in Ref. [84]. Salivary cortisol and alpha amylase are used as markers for stress. Random music for seven consecutive days, five times per day is presented to 53 participants. The subjective stress level which has been observed using questionnaires is reduced, when music is played in the presence of others. In Ref. [85], 200 females listened to music after performing a mental arithmetic task. State-trait anxiety and valence and arousal scores are used for the analysis of music on human stress. Low arousal music has a greater effect in stress reduction as compared to music with high arousal. Moreover, preferred music contributes to relaxation as it induces positive emotions.

A comparison of the proposed scheme with recent EEG based human stress classification methods is presented in Table 13. A machine learning framework for stress classification using 128-channel EEG cap is presented in Ref. [28]. Multilevel stress is induced in participants by using Montreal imaging stress task. An accuracy of 94.6% and 83.4% is achieved for two-level and multi-level stress classification. A multi-

Table 12
Comparison of the proposed study with different studies available in literature related to music and stress.

Method, Year	Number of Subjects	Measures	Stimuli	Findings
[37] 2016	30	EEG	• English music tracks from five different genres i.e., electronic, rap, metal, rock, and hiphop	• Hiphop and Rock music genres induce sad and happy emotions respectively. • Metal and Rap music genres evoke angry and sad emotions
[44] 2013	60	Salivary alpha amylase, cortisol, respiratory sinus, heart rate, anxiety and mood.	• Relaxing music • Sound of rippling water • Rest without acoustic stimulation	Sound of water have stronger relaxation effect as compared to music.
[74] 2009	38	ECG and respiration rate.	• Stimulative and Sedative music	• Stimulative music increased heart and respiration rate and increased joyous and energetic feelings but didn't effect calm and relaxed emotions. • Both sedative and aroused emotions were induced by sedative music.
[83] 2017	58	ECG	• Preferred relaxing music	• Preferred music improved cardiovascular recovery. • Females exhibited greater negative emotional state than males.
[84] 2016	53	Salivary cortisol and alpha amylase.	• Random Music listening at specified intervals	Listening to music in the presence of others have beneficial effects.
[85] 2016	200	State anxiety level, valence and arousal score	• 32 Music excerpts selected by the experimenters	• Music low in arousal have greater effect in stress reduction • Preferred music contributes to relaxation as it induces positive emotions.
Proposed	27	EEG	• 4 English music tracks of metal, rock, electronic, and rap genres • 5 Urdu music tracks of famous, melodious, patriotic, qawali, and ghazal genres	• English music tracks have more influence on stress reduction as compared to Urdu music tracks. • Females seemed to be more effected by music listening than males.

Table 13
Comparison of the proposed study with different studies available in literature for stress classification using EEG.

Method, Year	Number of Subjects	Number of EEG Channels	Accuracy % (Classes)	Classifier
[28], 2017	22	128	94.6% (2) 83.4% (multiple)	Naïve Bayes
[58], 2016	22	7	91.70% (2)	Support Vector Machine
[62], 2017	9	1	83.3% (2)	Support Vector Machine
[86], 2017	28	1	71.4% (2)	Naïve Bayes
Proposed	27	4	98.76% (2) 95.06% (3)	Logistic Regression

modal method for stress classification is presented in Ref. [58], where EEG and functional near infrared spectroscopy (fNIRS) based features are used. The acute stress is classified with an accuracy of 91.7% and 96.6%, when only EEG, and EEG and fNIRS based hybrid features are used, respectively. In Ref. [62], an EEG based study is proposed to classify between relaxed and stressful states by using a single channel headset. An accuracy of 83.3% is achieved for 9 participants in case of the two-class problem. The perceived stress classification method based on the single channel EEG signal recorded in closed eye condition is proposed in [89]. An accuracy of 71.4% is achieved by using the Naïve Bayes classifier. The methods are compared in terms of number of subjects participated in the study, EEG channels used for feature extraction, and accuracy. Apart from Ref. [62] and [89], all other methods have used a minimum of 7 electrodes to classify stress. The maximum accuracy achieved by the single channel EEG method for two-class classification is 83.3%. On the other hand, in the proposed scheme an accuracy of 98.76% and 95.06% is achieved in case of two-class and three-class classification with an addition of three electrodes. The method presented in Ref. [28], achieves comparable classification accuracy, but uses data from 128 EEG channels.

5. Conclusion

In this paper, the classification of human stress in response to music is analyzed using the brain signals. EEG signals of 27 participants were acquired using a commercially available brain sensing MUSE headband. The signals were recorded before and during the presentation of music stimuli. The recorded data was analyzed based on five features, which includes absolute and relative power, coherence, phase lag, and amplitude asymmetry. These features were classified into stressed and non-stressed states using four different classifiers. The classification problem was also extended for three classes i.e., non-stressed, medium stressed and highly stressed. The results indicate that the LR classifier performed better than SMO, SGD, and MLP. The highest accuracy achieved by LR for the two- and three-level classification is 98.76% and 95.06%, respectively. Furthermore, a statistical analysis of the subjective scores was conducted to analyze the genres, gender, and language related differences. It is evident from results that English music tracks have more influence on stress behavior as compared to Urdu music tracks, since a significant difference is reported in the results. Moreover, females are more affected by listening to music as compared to males, while no significant genre related differences were found.

References

[1] Bruce S. McEwen, Protective and damaging effects of stress mediators, N. Engl. J. Med. 338 (3) (1998) 171–179.
[2] Bruce S. McEwen, Central effects of stress hormones in health and disease: understanding the protective and damaging effects of stress and stress mediators, Eur. J. Pharmacol. 583 (2–3) (2008) 174–185.
[3] Alison Abbott, Stress and the city: urban decay, Nature News 490 (7419) (2012) 162.
[4] Thomas G. Pickering, Mental stress as a causal factor in the development of hypertension and cardiovascular disease, Curr. Hypertens. Rep. 3 (3) (2001) 249–254.
[5] Claudia Espinosa-Garcia, Sayeed Iqbal, Seema Yousuf, Fahim Atif, Elena G. Sergeeva, Gretchen N. Neigh, Donald G. Stein, Abstract TP83: Stress Exacerbates Global Ischemia-Induced Inflammatory Response: Intervention by Progesterone, (2017) ATP83-ATP83.
[6] Sari D. Holmes, David S. Krantz, Heather Rogers, John Gottdiener, Richard J. Contrada, Mental stress and coronary artery disease: a multidisciplinary guide, Prog. Cardiovasc. Dis. 49 (2) (2006) 106–122.
[7] Florian Lederbogen, Peter Kirsch, Leila Haddad, Fabian Streit, Heike Tost, Philipp Schuch, Stefan Wüst, et al., City living and urban upbringing affect neural social stress processing in humans, Nature 474 (7352) (2011) 498.
[8] Andrew P. Allen, Paul J. Kennedy, John F. Cryan, Timothy G. Dinan, Gerard Clarke, Biological and psychological markers of stress in humans: focus on the Trier Social Stress Test, Neurosci. Biobehav. Rev. 38 (2014) 94–124.
[9] Holger Ursin, Hege R. Eriksen, The cognitive activation theory of stress, Psychoneuroendocrinology 29 (5) (2004) 567–592.
[10] H. Mönnikes, J.J. Tebbe, M. Hildebrandt, P. Arck, E. Osmanoglou, M. Rose, B. Klapp, B. Wiedenmann, I. Heymann-Mönnikes, Role of stress in functional gastrointestinal disorders, Dig. Dis. 19 (3) (2001) 201–211.
[11] Ronald S. Duman, Neurobiology of stress, depression, and rapid acting antidepressants: remodeling synaptic connections, Depress. Anxiety 31 (4) (2014) 291–296.
[12] Sheldon Cohen, Tom Kamarck, Robin Mermelstein, A global measure of perceived stress, JHSB (J. Health Soc. Behav.) (1983) 385–396.
[13] Charles D. Spielberger, Fernando Gonzalez-Reigosa, Angel Martinez-Urrutia, Luiz FS. Natalicio, Diana S. Natalicio, The state-trait anxiety inventory, Revista Interamericana de Psicologia/Interamerican Journal of Psychology 5 (3 & 4) (2017).
[14] Nandita Sharma, Tom Gedeon, Objective measures, sensors and computational techniques for stress recognition and classification: a survey, Comput. Methods Progr. Biomed. 108 (3) (2012) 1287–1301.
[15] Andréa Deschênes, Hélène Forget, Camille Daudelin-Peltier, Daniel Fiset, Caroline Blais, Facial expression recognition impairment following acute social stress, J. Vis. 15 (12) (2015) 1383-1383.
[16] Sowjanya Gowrisankaran, Niru K. Nahar, John R. Hayes, James E. Sheedy, Asthenopia and blink rate under visual and cognitive loads, Optom. Vis. Sci. 89 (1) (2012) 97–104.
[17] J.T.-Y. Wang, Pupil dilation and eye tracking, in: A.K.M. Schulte-Mecklenbeck, J.G. Johnson (Eds.), A Handbook of Process Tracing Methods for Decision Research: A Critical Review and User's Guide, Psychology Press, New York, NY, USA, 2011, pp. 185–204.
[18] Armando Barreto, Zhai Jing, Malek Adjouadi, Non-intrusive physiological monitoring for automated stress detection in human-computer interaction, International Workshop on Human-Computer Interaction, Springer, Berlin, Heidelberg, 2007, pp. 29–38.
[19] Liao, Wenhui, Weihong Zhang, Zhiwei Zhu, and Qiang Ji. "A real-time human stress monitoring system using dynamic Bayesian network." In Computer Vision and Pattern Recognition-Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on, pp. 70-70. IEEE, 2005.
[20] Tin Lay Nwe, Say Wei Foo, Liyanage C. De Silva, Speech emotion recognition using hidden Markov models, Speech Commun. 41 (4) (2003) 603–623.
[21] Dimitrios Ververidis, Constantine Kotropoulos, Emotional speech recognition: resources, features, and methods, Speech Commun. 48 (9) (2006) 1162–1181.
[22] Jeremy Ang, Rajdip Dhillon, Ashley Krupski, Elizabeth Shriberg, Andreas Stolcke, Prosody-based automatic detection of annoyance and frustration in human-computer dialog, Seventh International Conference on Spoken Language Processing, 2002.
[23] Gwendolyn C. Dieleman, Anja C. Huizink, Joke HM. Tulen, Elisabeth MWJ. Utens, Hanneke E. Creemers, Jan van der Ende, Frank C. Verhulst, Alterations in HPA-axis and autonomic nervous system functioning in childhood anxiety disorders point to a chronic stress hypothesis, Psychoneuroendocrinology 51 (2015) 135–150.
[24] Subahni, Ahmad Rauf, Likun Xia, and Aamir Saeed Malik. "Association of mental stress with video games." In Intelligent and Advanced Systems (ICIAS), 2012 4th International Conference on, Vol. 1, pp. 82-85. IEEE, 2012.
[25] Alexandros Liapis, Christos Katsanos, Dimitris Sotiropoulos, Michalis Xenos, Nikos Karousos, Recognizing emotions in human computer interaction: studying stress using skin conductance, Human-Computer Interaction, Springer, Cham, 2015, pp. 255–262.
[26] Joseph Wielgosz, Brianna S. Schuyler, Lutz Antoine, Richard J. Davidson, Long-term mindfulness training is associated with reliable differences in resting respiration rate, Sci. Rep. 6 (2016) 27533.
[27] Katarina Dedovic, Robert Renwick, Najmeh Khalili Mahani, Veronika Engert, Sonia J. Lupien, Jens C. Pruessner, The Montreal Imaging Stress Task: using functional imaging to investigate the effects of perceiving and processing psychosocial stress in the human brain, J. Psychiatr. Neurosci. 30 (5) (2005) 319.

[28] Subhani, Rauf Ahmad, et al., Machine learning framework for the detection of mental stress at multiple levels, IEEE Access 5 (2017) 13545–13556.
[29] Daniel Novák, Lenka Lhotská, Vladimír Eck, Milan Sorf, "EEG and VEP Signal processing." Cybernetics, Faculty of Electrical Eng, 2004, pp. 50–53.
[30] Julien Dubois, Ralph Adolphs, Building a science of individual differences from fMRI, Trends Cognit. Sci. 20 (6) (2016) 425–443.
[31] N. Vanitha, Positron emission tomography in neuroscience research, Ann. Neurosci. 18 (2011) 2.
[32] Subhani, Rauf Ahmad, Likun Xia, Aamir Saeed Malik, Zahiruddin Othman, Quantification of physiological disparities and task performance in stress and control conditions, Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE, IEEE, 2013, pp. 2060–2063.
[33] Nicola Dibben, Victoria J. Williamson, An exploratory survey of in-vehicle music listening, Psychol. Music 35 (4) (2007) 571–589.
[34] Anneli B. Haake, Individual music listening in workplace settings: an exploratory survey of offices in the UK, Music. Sci. 15 (1) (2011) 107–129.
[35] Sandrine Vieillard, Isabelle Peretz, Nathalie Gosselin, Stéphanie Khalfa, Lise Gagnon, Bernard Bouchard, Happy, sad, scary and peaceful musical excerpts for research on emotions, Cognit. Emot. 22 (4) (2008) 720–752.
[36] Carlos Silva Pereira, João Teixeira, Patrícia Figueiredo, João Xavier, São Luís Castro, Elvira Brattico, Music and emotions in the brain: familiarity matters, PLoS One 6 (11) (2011) e27241.
[37] Adnan Mehmood Bhatti, Muhammad Majid, Syed Muhammad Anwar, Bilal Khan, Human emotion recognition and analysis in response to audio music using brain signals, Comput. Hum. Behav. 65 (2016) 267–275.
[38] Kelly Johnston, Jacqueline Rohaly-Davis, An introduction to music therapy: helping the oncology patient in the ICU, Crit. Care Nurs. Q. 18 (4) (1996) 54–60.
[39] Gwendolyn R. Watkins, Music therapy: proposed physiological mechanisms and clinical implications, Clin. Nurse Spec. 11 (no. 2) (1997) 43–50.
[40] Urs M. Nater, Elvira Abbruzzese, Monika Krebs, Ulrike Ehlert, Sex differences in emotional and psychophysiological responses to musical stimuli, Int. J. Psychophysiol. 62 (2) (2006) 300–308.
[41] Patrik N. Juslin, John A. Sloboda, Music and emotion, The Psychology of Music, third ed., 2013, pp. 583–645.
[42] Stefan Koelsch, Brain correlates of music-evoked emotions, Nat. Rev. Neurosci. 15 (3) (2014) 170.
[43] Hui-Ling Lai, Yin-Ming Li, The effect of music on biochemical markers and self--perceived stress among first-line nurses: a randomized controlled crossover trial, J. Adv. Nurs. 67 (11) (2011) 2414–2424.
[44] Myriam V. Thoma, Roberto La Marca, Rebecca Brönnimann, Linda Finkel, Ulrike Ehlert, Urs M. Nater, The effect of music on the human stress response, PLoS One 8 (8) (2013) e70156.
[45] Nechama Yehuda, Music and stress, J. Adult Dev. 18 (2) (2011) 85–94.
[46] Robert Horlings, Dragos Datcu, Leon JM. Rothkrantz, Emotion recognition using brain activity, Proceedings of the 9th International Conference on Computer Systems and Technologies and Workshop for PhD Students in Computing, ACM, 2008, p. 6.
[47] Erik Hoffmann, Brain Training against Stress, (2005), pp. 3–7.
[48] Bo Hjorth, EEG analysis based on time domain properties, Electroencephalogr. Clin. Neurophysiol. 29 (3) (1970) 306–310.
[49] Noor Hayatee Abdul Hamid, Norizam Sulaiman, Siti Armiza Mohd Aris, Zunairah Hj Murat, Mohd Nasir Taib, Evaluation of human stress using EEG power spectrum, Signal Processing and its Applications (CSPA), 2010 6th International Colloquium on, IEEE, 2010, pp. 1–4.
[50] Pavel Chernyavskiy, Caitlin M. Hudac, Dennis L. Molfese, David B. Marx, Using the spatio-temporal variogram for the classification ofelectroencephalographic (EEG) assessment, METMA VII and GRASPA14, 2014, pp. 1–4.
[51] J.F. Alonso, S. Romero, M.R. Ballester, R.M. Antonijoan, M.A. Mañanas, Stress assessment based on EEG univariate features and functional connectivity measures, Physiol. Meas. 36 (7) (2015) 1351.
[52] Inga Dziembowska, Paweł Izdebski, Anna Rasmus, Janina Brudny, Marta Grzelczak, Piotr Cysewski, Effects of heart rate variability biofeedback on EEG alpha asymmetry and anxiety symptoms in male athletes: a pilot study, Appl. Psychophysiol. Biofeedback 41 (2) (2016) 141–150.
[53] Pietro Cipresso, Andrea Gaggioli, Silvia Serino, Federica Pallavicini, Simona Raspelli, Alessandra Grassi, Luigi Sellitti, Giuseppe Riva, EEG alpha asymmetry in virtual environments for the assessment of stress-related disorders, Stud. Health Technol. Inf. 173 (2012) 102–104.
[54] Paul Murugesa Pandiyan, Sazali Yaacob, Mental stress level classification using eigenvector features and principal component analysis, Communications in Information Science and Management Engineering 3 (5) (2013) 254.
[55] Adnan, Nadia, Zunairah Hj Murat, Ros Shilawani S. Abdul Kadir, and Noradibah Hj Mohamad Yunos. "University students stress level and brainwave balancing index: comparison between early and end of study semester." In Research and Development (SCOReD), 2012 IEEE Student Conference on, pp. 42-47. IEEE, 2012.
[56] Z. Dharmawan, Analysis of computer games player stress level using EEG data, Master of Science Thesis Report, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Netherlands, 2007.
[57] Andrea Gaggioli, Pietro Cipresso, Silvia Serino, Giovanni Pioggia, Gennaro Tartarisco, Giovanni Baldus, Daniele Corda, et al., A decision support system for real-time stress detection during virtual reality exposure, MMVR, 2014, pp. 114–120.
[58] Fares Al-Shargie, Masashi Kiguchi, Nasreen Badruddin, Sarat C. Dass, Ahmad Fadzil Mohammad Hani, Tong Boon Tang, Mental stress assessment using simultaneous measurement of EEG and fNIRS, Biomed. Opt. Express (10) (2016) 3882–3898.
[59] Yuchae Jung, Yong Ik Yoon, Multi-level assessment model for wellness service based on human mental stress level, Multimed. Tool. Appl. 76 (9) (2017) 11305–11317.
[60] Kyriaki Kalimeri, Charalampos Saitis, Exploring multimodal biosignal features for stress detection during indoor mobility, Proceedings of the 18th ACM International Conference on Multimodal Interaction, ACM, 2016, pp. 53–60.
[61] Khosrowabadi, Reza, Chai Quek, Kai Keng Ang, Sau Wai Tung, and Michel Heijnen. "A Brain-Computer Interface for classifying EEG correlates of chronic mental stress." In Neural Networks (IJCNN), the 2011 International Joint Conference on, pp. 757-762. IEEE, 2011.
[62] A. Secerbegovic, S. Ibric, J. Nisic, N. Suljanovic, A. Mujcic, Mental workload vs. stress differentiation using single-channel EEG, CMBEBIH 2017, Springer, Singapore, 2017, pp. 511–515.
[63] Isabelle Peretz, Robert J. Zatorre, Brain organization for music processing, Annu. Rev. Psychol. 56 (2005) 89–114.
[64] Jason Burns, Elise Labbé, Kathryn Williams, Jennifer McCall, Perceived and physiological indicators of relaxation: as different as Mozart and Alice in chains, Appl. Psychophysiol. Biofeedback 24 (3) (1999) 197–202.
[65] Karen Allen, Lawrence H. Golden, Joseph L. Izzo Jr., Marilou I. Ching, Alan Forrest, Charles R. Niles, Philip R. Niswander, Jared C. Barlow, Normalization of hypertensive responses during ambulatory surgical stress by perioperative music, Psychosom. Med. 63 (3) (2001) 487–492.
[66] Anne J. Blood, Robert J. Zatorre, Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion, Proc. Natl. Acad. Sci. Unit. States Am. 98 (20) (2001) 11818–11823.
[67] V.J. Menon, W.R. Butt, J.F. Thiercelin, R. Gomeni, P.L. Morselli, D.R. London, Effects in man of progabide on prolactin release induced by haloperidol or domperidone, Psychoneuroendocrinology 9 (2) (1984) 141–146.
[68] Stefan Koelsch, Towards a neural basis of music-evoked emotions, Trends Cognit. Sci. 14 (3) (2010) 131–137.
[69] Wendy EJ. Knight, Nikki S. Rickard, Relaxing music prevents stress-induced increases in subjective anxiety, systolic blood pressure, and heart rate in healthy males and females, J. Music Ther. 38 (4) (2001) 254–272.
[70] Elise Labbé, Nicholas Schmidt, Jonathan Babin, Martha Pharr, Coping with stress: the effectiveness of different types of music, Appl. Psychophysiol. Biofeedback 32 (3–4) (2007) 163–168.
[71] T. Ventura, M.C. Gomes, T. Carreira, Cortisol and anxiety response to a relaxing intervention on pregnant women awaiting amniocentesis, Psychoneuroendocrinology (1) (2012) 148–156.
[72] Ulrica Nilsson, The anxiety-and pain-reducing effects of music interventions: a systematic review, AORN J. 87 (4) (2008) 780–807.
[73] Terry Richards, Jennifer Johnson, Amy Sparks, Howard Emerson, The effect of music therapy on patients' perception and manifestation of pain, anxiety, and patient satisfaction, Medsurg Nurs. 16 (1) (2007) 7.
[74] Joseph Lingham, Tores Theorell, Self-selected “favourite” stimulative and sedative music listening–how does familiar and preferred music listening affect the body? Nord. J. Music Ther. 18 (2) (2009) 150–166.
[75] Gillian M. Sandstrom, Frank A. Russo, Music hath charms: the effects of valence and arousal on recovery following an acute stressor, Music and Medicine 2 (3) (2010) 137–143.
[76] Valerie N. Stratton, Annette H. Zalanowski, The relationship between music, degree of liking, and self-reported relaxation, J. Music Ther. 21 (4) (1984) 184–192.
[77] Ethem Alpaydin, Introduction to Machine Learning, MIT press, 2014.
[78] John Platt, Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines, (1998).
[79] Christopher JC. Burges, A tutorial on support vector machines for pattern recognition, Data Min. Knowl. Discov. 2 (2) (1998) 121–167.
[80] Jenny Rose Finkel, Alex Kleeman, Christopher D. Manning, Efficient, feature-based, conditional random field parsing, Proceedings of ACL-08: HLT, 2008, pp. 959–967.
[81] Olivier Bousquet, Léon Bottou, The tradeoffs of large scale learning, Advances in Neural Information processing Systems, 2008, pp. 161–168.
[82] Jamal M. Nazzal, Ibrahim M. El-Emary, Salam A. Najim, Multilayer Perceptron Neural Network (MLPs) for Analyzing the Properties of Jordan Oil Shale 1, (2008).
[83] Alejandro de la Torre-Luque, Carolina Díaz-Piedra, Gualberto Buela-Casal, Effects of preferred relaxing music after acute stress exposure: a randomized controlled trial, Psychol. Music 45 (6) (2017) 795–813.
[84] Linnemann, Strahler Jana, M. Urs, Nater, “The Stress-Reducing Effect of Music Listening Varies Depending on the Social Context”, Psychoneuroendocrinology, 2016.
[85] Jiang, Jun, Daphne Rickson, and Cunmei Jiang. " The mechanism of music for reducing psychological stress: music preference as a mediator." Arts Psychotherapy, pp. 62-68, 2016.
[86] Saeed, Umar Sanay Muhammad, Anwar Syed Muhammad, Muhammad Majid, Quantification of human stress using commercially available single channel EEG headset, IEICE Trans. Info Syst. 100 (9) (2017) 2241–2244.

Anum Asif received his BSc. in Computer Engineering from COMSATS Islamabad. Currently she is pursuing his MSc. in Computer Engineering from UET Taxila. His research interest includes human stress assessment and brain signal processing.

Muhammad Majid received BSc. in Computer Engineering with honors from University of Engineering and Technology (UET) Taxila, Pakistan in 2005, MSc. in Data Communications with distinction and PhD in Electronic and Electrical Engineering from the University of Sheffield, UK in 2007 and 2011 respectively. He is currently an Associate Professor at Department of Computer Engineering, UET Taxila and heading Signal, Image, Multimedia Processing and LEarning (SIMPLE) research group since January 2012. He was a recipient of Faculty Development Program Scholarship of Higher Education Commission (HEC) Pakistan for his MSc and PhD Studies. His research interests include image and video coding, multimedia signal processing, biomedical signal processing and emotion classification in response to multimedia content.

Syed Muhammad Anwar received BSc. in Computer Engineering with honors from University of Engineering and Technology (UET) Taxila, Pakistan in 2005, MSc. in Data Communications with distinction and PhD in Electronic and Electrical Engineering from the University of Sheffield, UK in 2007 and 2012 respectively. He is currently an Associate Professor at Department of Software Engineering, UET Taxila. His research interests include medical imaging, data communication, and human computer interaction.