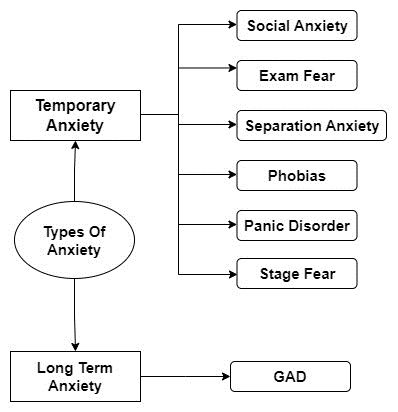
Predicting Anxiety among Young Adults Using Machine Learning Algorithms

Pralhad R Yadawad  
*School electronics and communication*  
*KLE Technological University*Karnataka, India  
ORCID: 0009-0002-0627-0661  
pralhadyadawad@gmail.com

Kaushik Mallibhat  
School of Electronics and Communication  
*KLE Technological University*Karnataka, India

*Abstract*— Anxiety is a common mental health concern among young adults, and early identification of temporary or long-term anxiety episodes can significantly enhance overall well-being and quality of life. This research paper presents a comprehensive approach to predicting both temporary and long-term anxiety among young adults by integrating electroencephalogram (EEG) signals and validated psychological questionnaires like GAD-7 and DASS-21. The study uses a structured experimental methodology involving participants in a range of activities specifically designed to trigger different types of Anxieties, including Phobias, Panic Disorder, Separation Anxiety, Social Anxiety, Stage Fear, and Exam Fear. EEG signals are recorded during these activities, enabling the capture of changes in beta wave activity, a known indicator of anxiety. To assess long-term anxiety, participants complete questionnaires such as GAD-7 and DASS-21, which provide self-reported data regarding chronic anxiety symptoms.

EEG data undergoes thorough preprocessing, and machine learning algorithms, including Support Vector Machines (SVM) and K-nearest neighbours (K-nearest), are employed for predictive modelling. The analysis achieves high accuracy in identifying various types of anxiety.

Keywords—EEG, GAD, Anxiety, Machine Learning, SVM (keywords)

# Introduction

Anxiety is a multifaceted emotional state and psychological process, typically triggered by perceived threats. Research indicates that while a certain level of anxiety can enhance performance and help tackle challenges, excessive anxiety can negatively affect sleep, attention bias, and the ability to think abstractly. At present, identifying the anxiety disorders is primarily dependent on clinical manifestations or a variety of surveys including the Beck Anxiety Inventory, State-Trait Anxiety Inventory, Manifest Anxiety Scale, Self-Rating Anxiety Scale, and the Hamilton Anxiety Scale. Nonetheless, this diagnostic approach may not be fully unbiased and presents a potential for subjective prejudice, which could affect the accuracy of the assessment. Neuroimaging technologies like Magnetic Resonance Imaging (MRI) and Electroencephalographs (EEG), known for their high accuracy and relative objectivity, are extensively used in emotion assessment. Specifically, due to its high temporal resolution and immediate nature of emotion generation, EEG has clear lead over other techniques. Presently, numerous anxiety studies are based on EEG. These studies suggest that anxiety primarily manifests as alterations in neurobehavioral function, which in turn affect EEG patterns.

Anxiety disorders pose a considerable mental health obstacle, impacting a reasonable segment of the young adult demographic. Effective counseling and treatment depend on the early detection and classification of diverse anxiety types, which include long-term illnesses like Generalized Anxiety Disorder (GAD) as well as transitory episodes brought on by specific circumstances[6].

This study embarks on a multifaceted journey, leveraging utilizing province tools for psychological testing and prediction of anxiety kinds in young individuals. Specifically, we integrate electroencephalogram (EEG) signals, widely recognized for their potential to capture neural responses, and validated psychological questionnaires, including the Generalized Anxiety Disorder 7-item scale (GAD-7) and the Depression, Anxiety, and Stress Scale (DASS-21). By combining these objective and subjective measures, our research endeavors to achieve a comprehensive understanding of the anxiety spectrum.

Accordingly, we organize a range of activities carefully designed to induce distinct temporary anxiety types, from Specific Phobias to Panic Disorder, Separation Anxiety, Social Anxiety, and performance anxieties like Stage Fear and Exam Fear. During these activities, EEG data is recorded and analyzed to capture dynamic changes in beta wave activity, a well-established marker of anxiety. To determine long-term anxiety, subjects complete GAD-7 and DASS-21 questionnaires, which provide self-reported data regarding chronic anxiety symptoms[11].

The machine learning framework, encompassing of various machine learning methods such as random forest (RF), XGBoost (XGB), is employed to process EEG data, extract important features, and predict the presence and intensity of anxiety. The amalgamation of EEG signals, questionnaire responses, and activity-induced beta wave analysis forms the cornerstone of our research.

The significance of these endeavors extends to the potential identification of distinctive neural and behavioral markers for various anxiety types, informing the development of tailored interventions and support strategies. By harnessing the power of machine learning and cutting-edge neuroscience, we aim to improve the overall well-being of young adults grappling with the multifaceted challenges of anxiety.

The techniques, findings, and discussions supporting our effort to identify and categorise different types of anxiety in young people are presented in the following parts, which will contribute significantly to the larger field of mental health research and treatment.

## Objectives:

1. Develop a machine learning algorithm to detect and classify the type of anxiety using EEG signals.
2. Perform a literature review to comprehend the state of art towards detection and measurement of anxiety detection.
3. A recently acquired EEG dataset was gathered with the express purpose of categorizing varying levels of anxiety intensity.

# literature survey

A crucial component of mental health research is understanding and predicting anxiety among young adults, as anxiety disorders constitute a significant mental health challenge that affect an important percentage of the young adult population. The initial focus of the paper was on predicting anxiety among students, but as we carried out additional research, we realized that the age group of students is not limited to those who have completed their education in schools or colleges. As an outcome, we concluded that we could also predict anxiety among young adults, or those who are between the ages of 18 and 30. Numerous studies on anxiety monitoring in young adults have already been published, but they mostly concentrate on identifying whether a student has anxiety permanently. Different research papers identified anxiety in individuals using different techniques. Researchers employed skin temperature (ST), photoplethysmogram (PPG), electroencephalography (EEG), electrocardiogram (ECG), heart rate (HR), oximetry, respiration (RSP), and electrodermal activity (EDA), alongside a questionnaire, in a comprehensive study to collect and analyze various physiological parameters.

Various papers have used different parameters and combinations of these parameters to get maximum accuracy in predicting anxiety [6]. The most accuracy reached in predicting anxiety is through the support vector machine learning algorithm which was about 78% the particular paper used EDA, ST, Ox, RSP, and HR but a common gap that we found in most of the papers was that these people were oriented in identifying whether the given test subject (human being) had permanent anxiety or not, during our literature survey were able to find only a few papers which identifies with the given test person has temporary anxiety or not and if he or she has temporary anxiety then what kind of temporary anxiety whether it is stage fear, social anxiety, exam fear, specific phobias, panic disorder, separation anxiety and also the intensity of that particular anxiety. These studies typically capture EEG signals while the subjects are engaged in activities, pre-process the frequency bands, and then use machine learning algorithms. Here our paper acts as the gap filler, and identifies as well as classify the level of anxiety whether the given person has temporary anxiety or not, we have used EEG signals and question air based on DASS-21 and GAD-7[1] conventions and we are also using different triggering activities to trigger the specific kind of anxiety.

EEG signals have been widely utilised in the study of anxiety. The presence of specific patterns, particularly increased beta wave activity, is indicative of heightened anxiety levels [9]. EEG signals have been employed to investigate real-time anxiety responses during anxiety-inducing activities [9]. Psychological questionnaires, such as the Generalized Anxiety Disorder 7 (GAD-7) and Depression, Anxiety, and Stress Scale (DASS-21), provide valuable self-reported data regarding chronic anxiety symptoms [1]. These questionnaires help assess long-term anxiety conditions and offer a complementary perspective to the neural data obtained from EEG signals [1]. This research has embraced a holistic approach that integrates EEG signals, questionnaire responses, and the utilization of activity-induced beta wave changes [10]. By engaging participants in anxiety-inducing activities specific to certain anxiety types, such as specific phobias, panic disorder, or social anxiety, researchers can effectively trigger and identify patterns in EEG data that correspond to these conditions.

The existing literature underscores the potential of EEG signals, machine learning, and psychological questionnaires in advancing our ability to predict and address anxiety among young adults. As the field continues to evolve, it holds great promise for more accurate and personalised mental health interventions [9]

|  |  |  |  |
| --- | --- | --- | --- |
| Method | Description | Advantages | Disadvantages |
| Hamilton Anxiety Rating Scale  (HAM-A) | The Hamilton Anxiety Rating Scale (HAM-A) is a psychological questionnaire used to quantify the severity of anxiety symptomatology, often used in psychotropic drug evaluation, with 14 items rated 0-4 and a total score range of 0-56. | It provides a comprehensive assessment of anxiety symptoms. It is widely used in both clinical and research settings. | It requires a trained clinician to administer and interpret. |
| State-Trait Anxiety Inventory  (STAI) | The State-Trait Anxiety Inventory (STAI) is a self-assessment tool that measures both temporary ‘state’ and long-standing ‘trait’ anxiety through various statements, rated on a 4-point scale, with higher scores indicating higher anxiety levels. | It distinguishes between the temporary condition of “state anxiety” and the more general and long-standing quality of “trait anxiety”. It helps in understanding how an individual’s personality and coping mechanisms can affect their experience of anxiety. | It is self-reported, so it may be influenced by the individual’s current state or willingness to accurately report their feelings. |
| Beck Anxiety Inventory  (BAI) | The BAI contains 21 questions, each answer being scored on a scale value of 0 (not at all) to 3 (severely). Higher total scores indicate more severe anxiety symptoms. | It is quick and easy to administer. It focuses on somatic symptoms of anxiety, which can be useful in differentiating anxiety from depression. | It may not be as effective in individuals who do not experience somatic symptoms. It may also be influenced by the individual’s willingness to accurately report their feelings. |
| DASS-21  (Depression Anxiety  Stress Scale-21) | The DASS-21 is a short version of a 42-item self-report instrument designed to measure three related negative emotional states: depression, anxiety and tension/stress. | The DASS-21 had good internal reliability (Cronbach’s alpha), and its ordinal alpha demonstrated good internal reliability for all its sub-scales4. Only the DASS-D and EPDS demonstrated a satisfactory ability to discriminate cases from non-cases4. | Information not found. |
| General anxiety disorder  (GAD-7) | The GAD-7 is a self-reported questionnaire for screening and severity measuring of generalized anxiety disorder (GAD). The GAD-7 is normally used in outpatient and primary care settings for referral to a psychiatrist pending outcome1. | It’s free to use, available in several languages, relatively brief, and easy to score, and correlations with other psychiatric instruments support the reliability and validity of the GAD-72. | The GAD-7 provides only probable diagnoses that should be confirmed by further evaluation2. |
| Electrodermal Activity  (EDA) | EDA measures changes in perspiration by detecting the changes in the electrical conductivity of the skin. Under stress conditions, there are notable changes in physiological signals such as perspiration. EDA can identify stress and, by extension, anxiety. | EDA reflects the output of integrated attentional, affective and motivational processes within the central nervous system acting on the body. It is a valuable tool in behavioural medicine as a biomarker of individual characteristics of emotional responsiveness. EDA biofeedback has been studied as a treatment tool for anxiety states and stress-sensitive physical disorders. | The main disadvantage of EDA recording is poor spatial resolution. The EEG signal is not useful for pinpointing the exact source of activity. |
| Electrocardiogram  (ECG) | ECG monitors heart rate and rhythm. Anxiety can cause changes in heart rate, and ECG can help detect these changes. It can also help differentiate between anxiety-related heart rate changes and those caused by other medical conditions. | ECG is helpful to measure three basic parameters of clinical interest viz. rhythm and heart rate, axis of the heart, and state of myocardial muscle. ECG represents data in the topographic form which provides higher diagnostical information. | ECG monitoring should document symptoms such as syncope and palpitations, but 24 hours is often too short a period, therefore other devices have been introduced. |
| Electroencephalography  (EEG) | EEG monitors brain activity. Anxiety disorders are associated with certain patterns of brain activity, which can be detected using EEG. This method can help identify anxiety by monitoring arousal-related brain activity. | EEG is a functionally fast, relatively cheap, and safe way of checking the functioning of different areas of the brain. High-precision time measurements can be achieved. | The main disadvantage of EEG recording is poor spatial resolution. |
| Respiration  (RSP) | Changes in respiration rates and patterns can be indicative of anxiety. Monitoring respiration can help detect these changes and provide data for anxiety detection. | RSP can provide continuous monitoring. | Information not found. |
| Skin Temperature  (ST) | Anxiety can cause changes in blood flow, which can affect skin temperature. Monitoring skin temperature can help detect these changes. | Information not found. | Information not found. |
| Photoplethysmogram  (PPG) | PPG monitors blood volume changes in the microvascular bed of tissue, which can be affected by anxiety. It’s a non-invasive method that involves neurocognitive training through a brain–computer interface. | PPG is simple, reliable, and inexpensive. It can easily be integrated into wearable healthcare devices for various health-related measurements such as pulse rate (or heart rate), blood flow, Heart Rate Variability (HRV), etc. | PPG sensors lack accuracy. They can’t show anomalies in the heart because these detectors are not fixed in that area and work on a different principle. |
| Oximetry  (Ox) | Oximetry measures oxygen saturation levels in the blood. Anxiety can cause shortness of breath or other respiratory issues that may affect oxygen saturation. | Oximetry is a safe, comfortable, and inexpensive method with no need for end-user calibration. It can usually be self-applied and does not require a medical specialist. | For people with health anxiety, or whose doctors have not helped them understand the role of an oximeter, this can cause unnecessary worry. |
| Heart Rate  (HR) | Heart rate monitoring is a common method for detecting anxiety. Increased heart rate is a common physiological response to anxiety, and continuous monitoring can provide real-time data about a person’s anxiety levels. | HR monitoring techniques that rely on PPG sensors have several advantages over traditional ECG-based systems. For instance, PPG sensors use simpler hardware implementation and have lower costs, and for operation, only a single sensor is required to be placed on the body. | Information not found. |

# Pipeline of Methodology

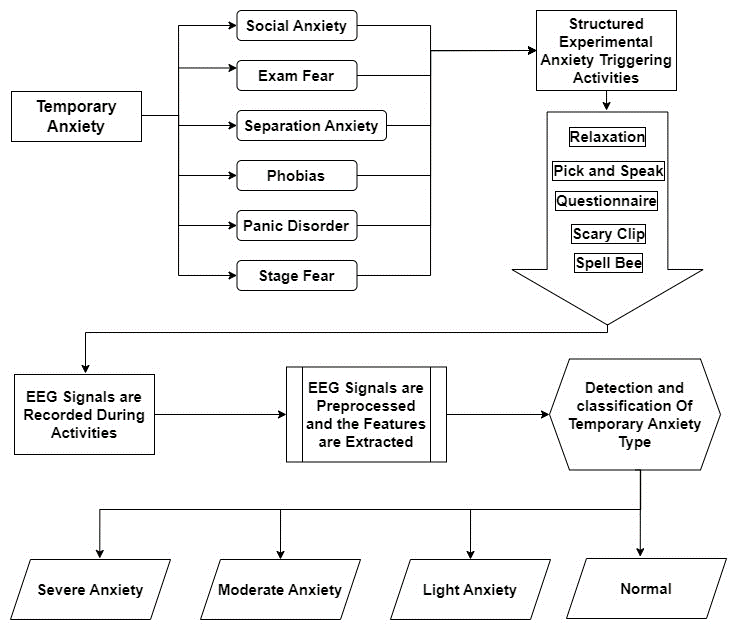


Figure 2

# Dataset

## Participants

The current data set includes 9 individuals (9 guys). All of the participants are between the ages of 18 and 21 and are pursuing a bachelor's degree. All of the participants have finished primary school at an English board school. The subjects declared that they had not ingested alcohol or any form of neural drugs before the dataset collection, which could have affected the study, and proper consent was obtained.

## Apparatus

For capturing the EEG signals we used EMOTIV EPOC+, which was a precise and easy-to-go device to record and process the EEG waves. The EMOTIV EPOC+ is a 14-channel portable EEG system with high resolution. It was created to be quick and simple to fit and measure in real-world research applications. The electrodes were positioned at AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8 and AF4. The electrodes were placed at the following locations: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4. The EMOTIV headband can record the EEG data at a sampling rate of 128 SPS / 256 SPS (2048 Hz internal). The EEG device has a resolution of 14 bits 1 LSB = 0.51V (16-bit ADC, 2 bits instrumental noise floor is discarded), but the settings can be changed to 16 bits and a bandwidth of 0.2 - 45Hz, with digital notch filters at 50Hz and 60Hz. To hydrate all of the electrodes and connect them to the sensors, a saline solution containing 0.7% - 4% w-w sodium chloride was used. The Emotiv Launcher, EmotivPRO, and EmotivLABS software were used to connect the device to the PC via a Bluetooth dongle (for a better connection) and store the captured EEG data during triggering activities.

## Experimental Procedure

Each participant completed five activities. There were 5 structured anxiety-triggering activities of 30 seconds each, and the EEG data was recorded simultaneously. The study comprised five distinct sessions designed to explore various facets of participants' experiences and responses. In the first session, individuals were instructed to put on earphones and listen to relaxing music, with the primary objective being to induce a state of relaxation. The second session, known as "Pick and Speak," involved participants selecting a topic from a spinning wheel and delivering a spontaneous 30-second speech on the chosen subject. The third session focused on a “questionnaire” tailored to each participant, aiming to uncover individual anxiety levels through specific inquiries. The fourth session, titled "Scary Scene", exposed participants to a brief, 30-second video clip from a horror film, serving as a stimulus to gauge fear responses. The fifth and final session, "Spell Bee," required participants to spell out seemingly simple yet intricately challenging words. Each session provided unique insights into participants' emotional and cognitive reactions across a spectrum of scenarios.

# Labeling Of Data.

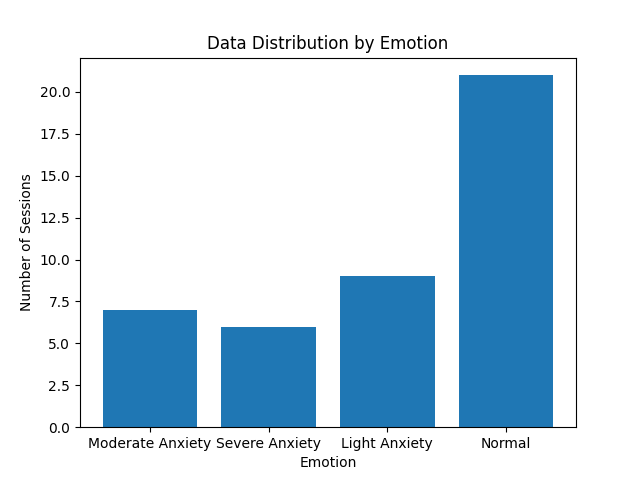
The labelling of the data acquired was according to the SAM method, for each session the valence and arousal of every test subject were recorded, and the labelling of each session was labelled on the CSV file itself as light anxiety, moderate anxiety, severe anxiety, and normal. The process of subject labelling is done Under the SAM rating-based labelling, any experimental trials with a valance score below 5 and an arousal score above 5 are deemed as Normal. Trials that have valance scores from 0 to 2 and arousal scores from 7 to 9 are designated as Severe. Subjects falling within the valance score range of 2 to 4 and arousal scores between 6 and 7 are recognized as having Moderate anxiety. Finally, subjects who have valance scores from 4 to 5 and arousal scores between 5 and 6 are classified as experiencing Light Anxiety. 

Figure 3

# Methodology

The EEG signals recorded from participants were acquired and subjected to feature selection in the conducted study. Specifically, the band power feature of the EEG data was chosen, focusing on the frequency information power within the theta (4-8Hz), alpha (8-12Hz), low beta (12-16Hz), high beta (16-25Hz), and gamma (25-45Hz) frequency bands across channels. In Electroencephalography data analysis, band power denotes the mean power or strength of the signal within a designated frequency range. Typically, the EEG signal is decomposed into functionally distinct frequency bands. In this study, band power is a crucial element for predicting and classifying anxiety states. The computation of band power provides a singular metric that encapsulates the influence of specific frequency bands on overall signal power, contributing significantly to our objective.

For example, during deep sleep, the prevalence of slow waves in the delta band (0.5 to 4 Hz) aligns with our approach, reflecting synchronized brain activity. Conversely, the reduced delta activity and increased higher-frequency activity during wakefulness support our focus on using band power to differentiate anxiety states. Our choice aligns with the inherent characteristics of EEG wave bands, where lower-frequency bands (delta and theta) exhibit significantly larger values than higher-frequency bands (alpha and beta). This selection enables effective tracking of temporal changes and facilitates comparing band strengths, enhancing the precision of our anxiety state predictions. Utilizing the EPOX PRO software from Emotive, band power values for each sensor and across all EEG frequency bands i.e. alpha, gamma, high beta, low beta and theta were directly obtained, resulting in a feature vector with a length of 70.

Data packets serve as discrete units transmitted between devices or systems across a network. In the EmotivPRO context, these packets encapsulate information acquired from the headset's sensors and transmitted to a computer. The recurring one-second cycle of headset data, operating at either 128 Hz or 256 Hz, is illustrated by grey sawtooth lines. The utilization of a USB extender to connect the headset's dongle to the computer ensured no data packet loss, preserving information integrity. All 70 features within the feature vector were utilized for classification using various models, including Random Forest, Decision Trees, Logistic Regression, Support Vector Machine, K-nearest Neighbour, Naïve Bayes, XGBoost, Voting Classifiers (combinations of RF, DT, SVC), Simple Neural Network, and Complex Neural Network. The highest accuracy, at 71.017%, was achieved with the Random Forest model.

## Algorithm

## Algorithm for EEG Feature Selection based on Band Power

**Input:**

1) Extracted feature group FV = Band Power.

2) Class Labels = Normal, Light anxiety, Moderate anxiety

3) Classifier =RF, DT, LR, SVM, KNN, NB, XGb, Voting Classifier, Neural networks

4) EEG frequency bands = α, β, δ, γ, θ.

5) Number of EEG bands = N.

**Output:** Process EEG Data for Emotion Classification

If the EEG data DataFrame is not empty:

* Set parameters for segmenting EEG signals.
* Define target frequency bands: Alpha, Theta, BetaL, BetaH, Gamma.
* Initialize an empty dictionary to store values for each frequency band.

For each frequency band in each segment:

For each target frequency band:

* Initialize an empty Series for the current frequency band.
* Iterate over EEG data segments:

For each segment:

* Process frequency values for the current segment and frequency band.
* Concatenate processed values to the overall frequency values.
* Calculate and print the average value for the current frequency band.

End

End

End

Prepare data for machine learning:

* Create a DataFrame 'X' with columns for each frequency band.
* Drop rows with missing values (NaN) from 'X'.

Assign emotion labels based on the file path:

* Use a function (e.g., determine\_emotion\_label) to determine the emotion label.
* Update the counter for emotion occurrences.

Print the assigned emotion label for verification.

Else,

if the EEG data DataFrame is empty:

Print a message: "The DataFrame is empty. Please check the file contents."

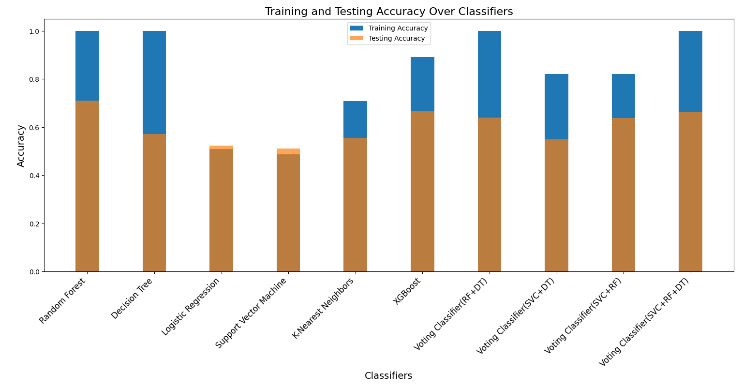
End

End

End

# Results

This paper presents an experimental study aiming to identify an optimal EEG recording phase for classifying perceived mental stress into multiple categories. Features extracted from various EEG bands were analyzed using band power during activity phases in an open-eye condition. Notably, features contributing to enhanced classification accuracy were selected. During this process, we also found out that including all band power spectra i.e. (α, β, δ, γ, θ) results in a better accuracy score, as the study using SAM which might also be reasons for low accuracy, the study employed fourteen classification algorithms, with the Random Forest classifier yielding the best results for two- and three-class stress classification, achieving the highest accuracy of 70.01% and 72.97% (10-fold cross-validation). Future work aims to extend this approach to classify acute mental stress based on EEG signals during such activities to classify various types of anxiety.



# Conclusion

This research presents a pilot study to predict and classify anxiety levels as well as provide a methodology for recorded EEG data. From the alpha, gamma, low beta, high beta, and theta bands of EEG data collected during the experimentally designed tasks which trigger multiple kinds of anxiety, one group of features—band power—was retrieved. activities include the following: 1. Pick and Speak 2. Questionnaire 3. Scary Scene 4. Relaxing Music Session 5. Spell Bee. The EEG frequency bands were used to select a single feature with a feature vector length of 70 to help determine the degree of anxiety.

##### References

1. Shaurya Bhatnagara, Jyoti Agarwalb\*, Ojasvi Rajeev Sharmaa, Detection and classification of anxiety in university students through the application of machine learning,2013, India
2. Jetli Chung and Jason Teo, Mental Health Prediction Using Machine Learning: Taxonomy, Applications, and Challenges, Hindawi Applied Computational Intelligence and So Computing, Volume 2022,2022
3. Hyeon-Sang Hwang 1 and Eui-Chul Lee 2, Non-Contact Respiration Measurement Method Based on RGB Camera Using 1D Convolutional Neural Networks, published 2021.
4. Ezekiel Victora Zahra M. Aghajan, Ph. D.b Amy R. Stewart, M.A., C.Philc Ray Christian., B.B.A, Detecting Depression Using a Framework Combining Deep Multimodal Neural Networks with a Purpose-Built Automated Evaluation,
5. Amogh Gudi 1,2, Marian Bittner 1,2, and Jan van Gemert 2, Real-Time Webcam Heart-Rate and Variability Estimation with Clean Ground Truth for Evaluation,2020
6. Florian Ritsert 1,2, Mohamed Elgendi 1, Valeria Galli 1 and Carlo Menon 1, Heart and Breathing Rate Variations as Biomarkers for Anxiety Detection,2022
7. Lou Ancillon 1,2, Mohamed Elgendi 1, and Carlo Menon 1, Machine Learning for Anxiety Detection Using Biosignals: A Review,2022
8. Ashigha, Rasha Shajahan, Shifana Abdul Azeez, Kashish S Shah, Veena K Viswam, Mental Health Monitoring Using Machine Learning: A Survey Maria, International Journal of Computer Science Trends and Technology (IJCST) – Volume 10 Issue 3, May-Jun 2022
9. Ziyu Li, Xia Wu, Xueyuan Xu, Hailing Wang, Zhenghao Guo, Student Member, IEEE, Zhichao Zhan, and Li Yao, The Recognition of Multiple Anxiety Levels Based on Electroencephalograph, IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, VOL. 13, NO. 1, JANUARY-MARCH 2022.
10. <https://www.neuroscienceinc.com/news/2020/stress-anxiousness-and-the-brain-wave-connection#:~:text=Anxiousness%20is%20associated%20with%20decreased,imbalanced%20and%20poorly%20regulated%20brain>
11. Aamir Arsalan, Muhammad Majid; Amna Rauf Butt; Syed Muhammad Anwar, “Classification of Perceived Mental Stress Using A Commercially Available EEG Headband”