Anxiety detection using EEG data

Submitted By

Harsh Arora

22MCEC02

INSTITUTE OF TECHNOLOGY

NAAC ACCREDITED 'A+' GRADE

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2024

Anxiety detection using EEG data

Major Project - II

Submitted in partial fulfillment of the requirements

for the degree of

Master of Technology in Computer Science and Engineering

Submitted By

Harsh Arora

(22MCEC02)

Guided By

Dr. Swati Jain

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

INSTITUTE OF TECHNOLOGY

NIRMA UNIVERSITY

AHMEDABAD-382481

May 2024

Certificate

This is to certify that the major project entitled **"Anxiety detection using EEG data"** submitted by **Harsh Arora (Roll No: 22MCEC02)**, towards the partial fulfillment of the requirements for the award of degree of Master of Technology in Computer Science and Engineering of Nirma University, Ahmedabad, is the record of work carried out by him under my supervision and guidance. In my opinion, the submitted work has reached a level required for being accepted for examination. The results embodied in this major project part-II, to the best of my knowledge, haven't been submitted to any other university or institution for award of any degree or diploma.

21/5/2024

Dr. Swati Jain
Guide & Associate Professor,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

21/05/24

Dr. Sudeep Tanwar
Professor,
Coordinator M.Tech - CSE
Institute of Technology,
Nirma University, Ahmedabad

Dr. Madhuri Bhavsar
Professor and Head,
CSE Department,
Institute of Technology,
Nirma University, Ahmedabad.

Himanshu Soni

Dr
Director,
School of Technology, Nirma
University, Ahmedabad

Statement of Originality

I, **Harsh Arora**, Roll. No. **22MCEC02**, give undertaking that the Major Project entitled "**Anxiety detection using EEG data**" submitted by me, towards the partial fulfillment of the requirements for the degree of Master of Technology in **Computer Science & Engineering** of Institute of Technology, Nirma University, Ahmedabad, contains no material that has been awarded for any degree or diploma in any university or school in any territory to the best of my knowledge. It is the original work carried out by me and I give assurance that no attempt of plagiarism has been made.It contains no material that is previously published or written, except where reference has been made. I understand that in the event of any similarity found subsequently with any published work or any dissertation work elsewhere; it will result in severe disciplinary action.

H.A.Arora

Signature of Student
Date: 21/5/2024
Place:Ahmedabad

21/5/2024

Endorsed by
Dr. Swati Jain
(Signature of Guide)

Acknowledgements

I am heartily thankful to my supervisor **Dr. Swati Jain**, whose encouragement, supervision and support from the preliminary to the concluding level enabled me to develop an understanding of the subject. At the end, I offer my regards and blessings to all of those who supported us in any respect during the completion of the project and to our college for providing resources and materials.

It gives me an immense pleasure to thank **Dr. Madhuri Bhavsar**, Head of Computer Science and Engineering Department, Institute of Technology, Nirma University, Ahmedabad for his kind support and providing basic infrastructure and healthy research environment.

A special thank you is expressed wholeheartedly to **Dr. R. N. Patel**, Hon'ble Director, Institute of Technology, Nirma University, Ahmedabad for the unmentionable motivation he has extended throughout course of this work.

I would also thank the Institution, all faculty members of Computer Engineering Department, Nirma University, Ahmedabad for their special attention and suggestions towards the project work.

Harsh Arora
22MCEC02

Abstract

These days, new methods of analysis and prediction are required due to the increasing incidence of brain-related diseases such as depression, anxiety, and stress. This study provides an innovative approach to creating datasets by using videos that feature anxiety-inducing sequences of animal hunting intercut with humorous footage. 30 subjects (14 boys +15 girls) pursuing MTech degrees are part of the data collection using a single-channel Narosky Mind Wave Starter kit.

A binary classification into two categories—anxiety and non-anxiety—is made simpler by the dataset. The main goal is to identify the brain signals that are released when worried information is seen. This discovery has the potential to help physicians accurately diagnose patients and identify those who suffer from serious brain problems.

The research also uses explainable Artificial Intelligence (XAI) algorithms for insightful analysis, such SHAP (SHapley Additive exPlanations) and LAME (Local Additive Model Explanation). The project's goal is to find dominating frequency bands connected to anxiety within the generated dataset by using these XAI techniques. This all-encompassing method not only advances our knowledge of how the brain reacts to anxiety-inducing events, but it also provides invaluable resources for medical professionals in their pursuit of diagnosis.

Abbreviations

XAI	Explainable artificial intelligence
EEG	Electroencephalography
SHAP	Shapley additive explanations.

Contents

List of Figures

Chapter 1

Introduction

1.1 Aim of Project :-

This research paper delves into the exploration of brainwave patterns linked to anxiety using Electroencephalogra- phy (EEG) data. The study focuses on understanding the neural underpinnings of emotions, particularly anxiety, by leveraging EEG's non-invasive and real-time data capture capabilities. The literature review encompasses an analysis of key EEG datasets of 29 subjects(14 Males + 15 Females) and multiple machine learning algorithms mainly random forest applied to emotional classification using EEG data. The experimental setup involves data collection with the NeuroSky Brainwave starter kit, data preprocessing, and the application of machine learning algorithms, leading to the identification of the Random Forest model as the most effective. Additionally, Explainable AI (XAI) techniques, specifically SHAP, are utilized to unveil the critical EEG frequency bands contributing to anxiety. The findings underscore the significance of beta(13Hz - 30Hz) and lower gamma(30Hz - 60Hz) frequency bands in the EEG signals of individuals experiencing anxiety, thereby providing valuable insights into emotional processing and affective neuroscience.

1.2 Scope of project :-

This research unfolds in two phases: first, the creation of a tailored dataset featuring anxiety-inducing videos, and second, the application of advanced XAI algorithms for in-depth analysis. Beyond contributing to the understanding of anxiety at a neurological level, this interdisciplinary effort aspires to bridge the gap between behavioral stimuli and analytical insights. The project's significance lies not only in its potential to redefine diagnostic practices in mental health but also in its broader impact on the intersection of neuroscience and data science, fostering collaborative research that transcends disciplinary boundaries.

This study is conducted in two stages: first, a customized dataset of movies that elicit anxiety is created, and then XAI algorithms are used for in-depth analysis. This interdisciplinary endeavor aims to close the gap between behavioral stimuli and analytical insights, in addition to adding to our understanding of anxiety from a neurobiological perspective. The initiative is significant not only because it has the potential to revolutionize mental health diagnosis procedures but also because it will have a wider influence on the nexus between data science and neuroscience, encouraging cross-disciplinary study.

1.3 Electroecephlography :-

Electroencephalography (EEG) is playing an important role in study of Emotions. It is a non-invasive technique. In recent years the development of neuromaging methods gives us essential new understandings of human emotions, cognitive functions, and mental health disorders. EEG is one such method that uses electrodes attached to the scalp to measure the electrical activity produced by the brain. EEG has shown to be especially useful in the investigation of emotional responses. The majority of communication between brain cells occurs through electrical impulses, which are constantly active even when you're asleep[1]. EEG recording, are organized into distinct frequency bands, each associated with different states of brain activity and cognitive processes.

1.4 Emotions :-

Emotion is an intricate process of psycho-physiology which consisting of electrical pulses due to brain activity that represents various states of feelings, action and psychological shifts. Emotions has different forms, which are joy, sadness, anger, fear, satisfaction, amazed, love, and lessy.[2]. There are numerous methods for determining a person's emotion, which includs body language [3], facial expressions [4], and verbal exchanges.

But one can fake in all of these. Because it is so difficult to spoof, brain signals it become more popular. This aids researchers in analyzing emotional behavior [5].

Majorly emotions can be classified in 3 classes 1) Positive, 2) Negative and 3) Nutral. Positive emotions, such as happiness and contentment, are characterized by increased energy and a sense of well-being. Conversely, negative emotions, like sadness and anxiety, lead to decreased energy and distress[6]. Negative emotions are affecting more to brain. Anxiety disorders are common and severe, affecting people in many different groups which is basically negative emotion. It is essential to understand the brain foundations of anxiety in order to create networks of assistance and interventions which are effective. This work uses EEG to explore the subtle aspects of anxiety reactions, emphasizing the roles that various frequency bands have in the whole experience. In the field of neuroscience and mental health studies, it is critical to understand the complex relationship between emotional experiences and brain activity. As EEG is non-invasive and can record real-time electrical impulses produced by the brain, it is an exceptionally useful tool for deciphering the secrets of human emotions[7].

1.5 Dataset :-

A variety of datasets are available that are intended to investigate different emotions in diverse settings. These datasets cover a wide range of experimental methods and stimuli designed to elicit a wide range of emotional reactions. For example, movies, photos, or audio clips that portray situations linked to various emotional states[8], such as happiness, sorrow, anger, fear, or surprise, are included in some datasets[9]. Some datasets aim to elicit particular emotions like empathy, amusement, or wonder through interactive challenges, virtual surroundings, or real-world events[10].Through the utilisation of these varied datasets, scientists can acquire a more all-encompassing comprehension of the intricate relationship among emotions, cognition, and neural activity[11]. This will ultimately propel the domain of affective neuroscience forward and aid in the creation of more refined emotional processing models. There is often a less focused area to specificity in the stimuli designed to elicit anxiety compare to other emotions.

1.6 Machine learning and XAI Algorithm :-

Machine learning models have been used more and more in the past several years, especially in the fields of psychology and healthcare. A critical component in guaranteeing the transparency and interpret ability of these models is XAI. This work investigates the usage of popular XAI methods for EEG data analysis: SHAP (SHapley Additive exPlanations). The principal aim is to elucidate the roles played by distinct EEG frequency patterns affecting the recorded anxiety levels, so providing insight into the inner workings of the machine learning models utilized for the research. This investigation not only makes the brain correlates of anxiety more understood, but it also emphasizes how crucial XAI techniques are to closing the gap between sophisticated model outputs and actual psychological experiences[12].

The use of SHAP for the analysis of result given by machine learning algorithm on EEG data is very beneficial in the context of our study. The global interpretability of SHAP provides a more comprehensive viewpoint on the cumulative significance of various EEG parameters throughout the full dataset, assisting in the discovery of recurring patterns associated with anxiety.

In this paper, section II is of literature review. Literature review contains 2 types of review which includes literature review of dataset acquisition and another is for algorithms applied on EEG data. Section III contains experimental setup, the flow of whole process which contains 1) Dataset collection, 2) Pre-Processing steps and 3) Classification and XAI on top of applied algorithm to find dominant frequency band affecting when any subject feeling anxious. Section IV representing results of this experiment. Section V gives conclusion derived from this project.

Chapter 2

Literature Survey

2.1 Understanding of terminologies :-

2.1.1 EEG Waves :-

An electroencephalogram is a test that measures brain electrical activity using microscopic metal discs or electrodes attached to the scalp. The majority of communication between brain cells occurs through electrical impulses, which are constantly active even when you're asleep. On an EEG recording, waves can be visible during this action.

EEG Signals are classified into some range of frequencies and all frequencies having different meaning or brain activity associated with it shown in table below.

Techniques shown in table:

Sr. NO	Frequency with Range	Brain state
1	Delta Waves (0.5-4 Hz)	deep sleep and unconsciousness.
2	Theta Waves (4-8 Hz)	drowsiness, daydreaming, and REM (rapid eye movement) sleep.
3	Alpha Waves (8-12 Hz)	when the eyes are closed and the individual is in a quiet, meditative state.
4	Beta Waves (12-30 Hz)	active wakefulness, attention, and concentration.
5	Gamma Waves (30-100 Hz)	associated with complex cognitive processing, sensory integration, and attention

Table 2.1: Waveform Classification of EEG signals with activity

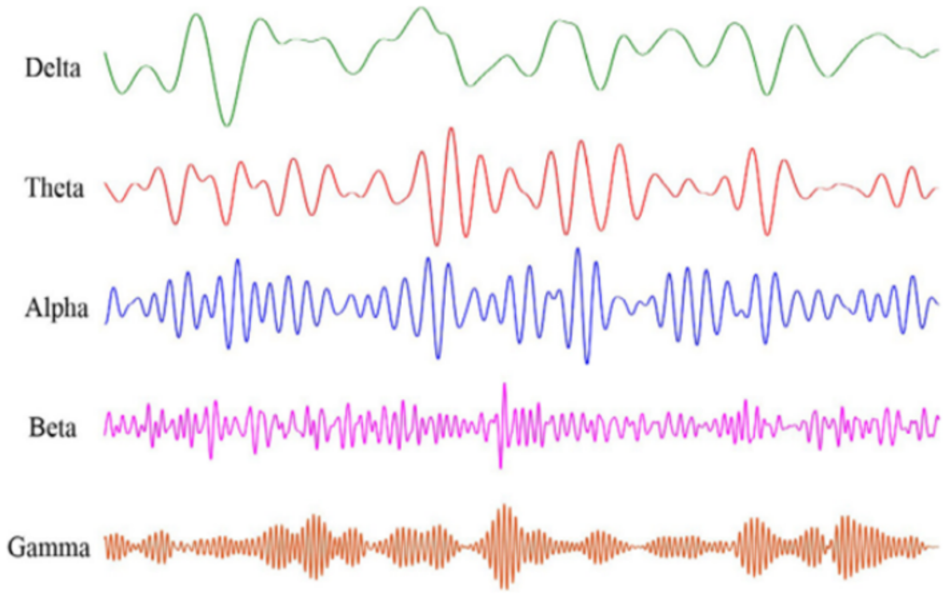


Figure 2.1: Waveform Classification of EEG
[13]

Some studies have suggested that increased activity in the beta frequency range (12-30 Hz) in the frontal cortex may be associated with anxiety. Other studies have suggested that increased theta activity (4-8 Hz) and decreased alpha activity (8-12 Hz) may also be associated with anxiety.

Also, Additionally, some studies have suggested that changes in the ratio of different types of brain waves, such as an increased theta/beta ratio, may be indicative of anxiety. However, it's important to note that the specific patterns of brain waves associated with anxiety may vary depending on the individual and the type of anxiety being studied.

2.1.2 Explainable AI :-

The field of Explainable Artificial Intelligence (XAI) is leading the way in resolving the issues of interpretability and transparency that arise from sophisticated machine learning models.Two well-known algorithms in the field of XAI, SHapley Additive exPlanations (SHAP) and Local Additive Model Explanation (LIME), have been highly effective at explaining how black-box models make decisions. The development and deployment of these techniques, which provide insights into how particular variables contribute to model predictions, have been driven by the necessity to comprehend and trust the results of ma-

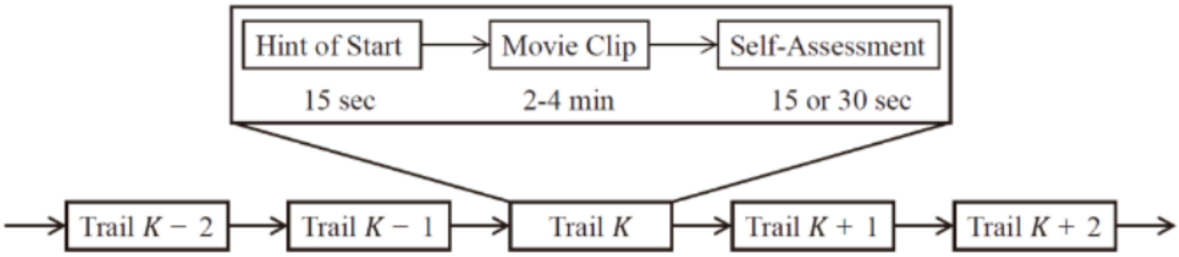


Figure 2.2: Flow chart of videos while creating dataset [10]

chine learning models. As its name implies, LIME focuses on building local, interpretable models around specific predictions, illuminating the variables affecting the result of a certain instance. Conversely, SHAP, which is based on cooperative game theory, gives each feature a value that quantifies its influence on model predictions in all possible combinations.

These XAI algorithms are essential for helping stakeholders—from data scientists to end users—understand the reasoning behind AI decisions by clarifying the inner workings of complex models. LIME and SHAP become essential tools for opening the "black box," building confidence, and encouraging responsible AI development and application as the need for moral and accountable AI develops. In order to improve the interpretability of AI systems and enable well-informed decision-making in a world where AI is becoming more and more prevalent, this investigation explores the meaning, workings, and applications of LIME and SHAP.

2.2 Existing System :-

Here Main Aim was to identify the emotion from the EEG Signals. Finally, from all the signals we selected Anxiety because nowadays lots of people suffer from this and compared to other emotions this emotion has less dataset to work with.

Apart from this, also one of the datasets we requested for which is SEED - IV. This dataset contains the EEG as well as Eye data and output classes are 4 (happiness, Sadness, fear, Neutral)

To provide context and enhance our dataset collecting survey, we thoroughly examined existing datasets that exhibit similar patterns to our desired dataset. The purpose of this comparison analysis is to find datasets that exhibit commonalities in important elements such as structure, content, and attributes. By examining similar datasets, we acquire useful knowledge about prospective sources that are in line with the goals of our research. Table **??** shows, a carefully chosen assortment of datasets that demonstrate similar tendencies to our dataset of interest.

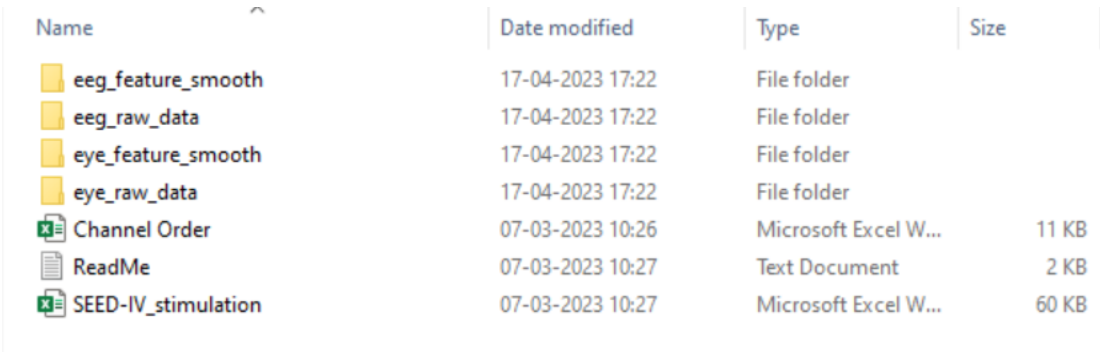
2.3 Dataset Acquisition Survey :-

We found a wide variety of datasets in our review that support the goals of our investigation. These datasets provide important tools for comprehending brain correlates and emotional processing, which forms the basis for our EEG data gathering activities. Through an analysis of these datasets' methods, characteristics, and content, we were able to learn a great deal about the most effective ways to present stimuli and gather data. Furthermore, considering our focus on gathering EEG data in response to anxiety-inducing stimuli, we acknowledge the significance of choosing suitable stimuli for evoking anxious responses. We seek to develop and use a useful experimental paradigm that successfully elicits anxiety in participants, building on the results of our literature research and providing our dataset with useful brain responses to this particular emotional state.

Some of the images that can describe the Dataset is as follows

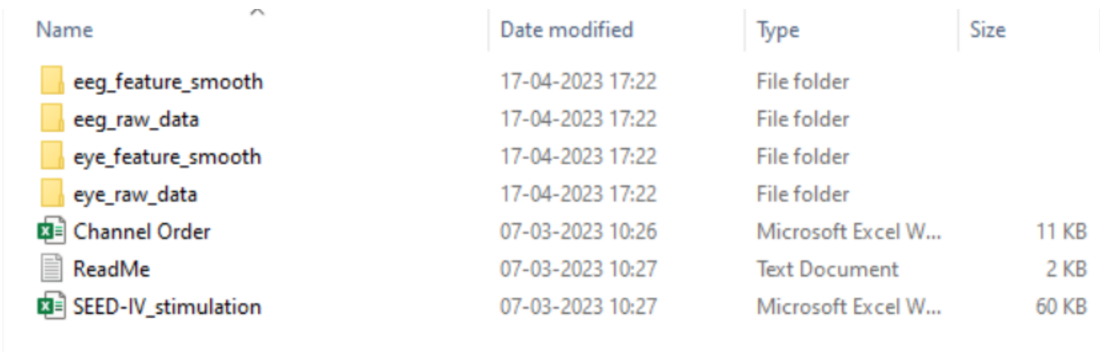
2.3.1 Implementation Survey

An overview of the state of the research on EEG-based emotion categorization algorithms can be obtained from the literature review, which synthesizes the conclusions and approaches of the publications indicated in the Table **??**.
Significant progress has been made in the field of EEG-based emotion categorization research, with studies using a variety of datasets and algorithms to extract emotional states



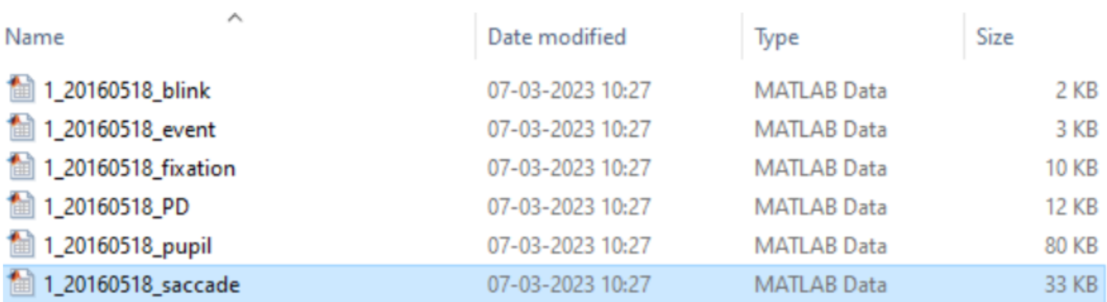
Name	Date modified	Type	Size
eeg_feature_smooth	17-04-2023 17:22	File folder	
eeg_raw_data	17-04-2023 17:22	File folder	
eye_feature_smooth	17-04-2023 17:22	File folder	
eye_raw_data	17-04-2023 17:22	File folder	
Channel Order	07-03-2023 10:26	Microsoft Excel W...	11 KB
ReadMe	07-03-2023 10:27	Text Document	2 KB
SEED-IV_stimulation	07-03-2023 10:27	Microsoft Excel W...	60 KB

Figure 2.3: Dataset Files
[10]



Name	Date modified	Type	Size
eeg_feature_smooth	17-04-2023 17:22	File folder	
eeg_raw_data	17-04-2023 17:22	File folder	
eye_feature_smooth	17-04-2023 17:22	File folder	
eye_raw_data	17-04-2023 17:22	File folder	
Channel Order	07-03-2023 10:26	Microsoft Excel W...	11 KB
ReadMe	07-03-2023 10:27	Text Document	2 KB
SEED-IV_stimulation	07-03-2023 10:27	Microsoft Excel W...	60 KB

Figure 2.4: Dataset Files



Name	Date modified	Type	Size
1_20160518_blink	07-03-2023 10:27	MATLAB Data	2 KB
1_20160518_event	07-03-2023 10:27	MATLAB Data	3 KB
1_20160518_fixation	07-03-2023 10:27	MATLAB Data	10 KB
1_20160518_PD	07-03-2023 10:27	MATLAB Data	12 KB
1_20160518_pupil	07-03-2023 10:27	MATLAB Data	80 KB
1_20160518_saccade	07-03-2023 10:27	MATLAB Data	33 KB

Figure 2.5: Eye Data

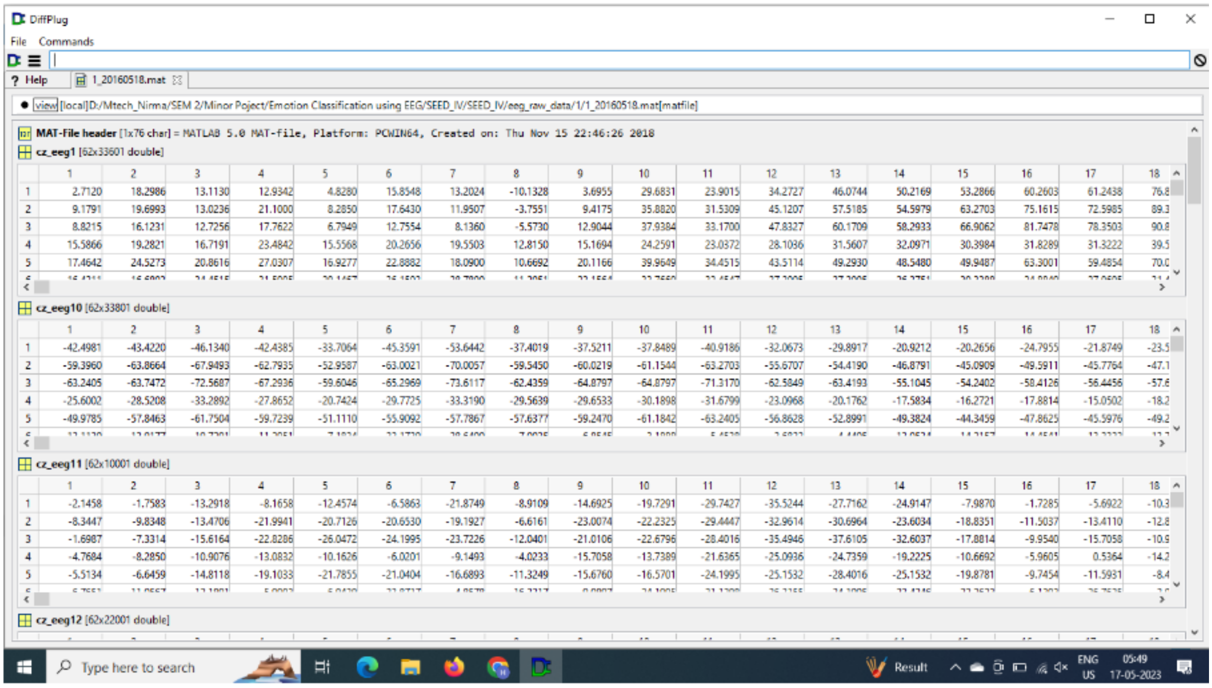


Figure 2.6: Example of eye data

from brain signals. These studies frequently make use of datasets such as MAHNOB-HCI, SEED, and DEAP, all of which provide different perspectives on the brain correlates of emotions. Emotions have been accurately classified using a variety of algorithms, from sophisticated deep learning architectures like CNN and DNN to more conventional machine learning techniques like SVM, KNN, and LDA.

By leveraging the insights gained from the review, our aim is to contribute to the broader understanding of emotional states, particularly anxiety, through the meticulous analysis of EEG datasets and the implementation of state-of-the-art machine learning algorithms. Our research has highlighted the critical aspects of data acquisition and the significance of selecting appropriate stimuli to invoke emotions. Additionally, the review has provided valuable understanding of the various popular ML algorithms extensively utilized in the domain of EEG dataset analysis. This knowledge forms the foundation for our current research efforts, allowing us to comprehensively address the less explored area of emotions and ML algorithm explainability of it with the use of XAI algorithm.

Sr.NO	Paper title	Published year	Method used	Classified emotions
1	DEAP: A Database for Emotion Analysis using Physiological Signals[8]	2011	Create Dataset Named DEAP 32 participants (22 male, 10 female) 40 videos participants also provided subjective ratings of their emotional states after watching each music video	arousal, valence and like/dislike
2	EmotionMeter: A Multimodal Framework for Recognizing Human Emotions (SEED)[10]	2019	EEG and eye movement data of 12 subjects. EEG data of another 3 subjects. Data was collected when they were watching film clips.15 Chinese movie clips of length of around 4 min	positive, negative, and neutral
3	SEED IV	2019	72 film clips. 3 sessions were performed on different days, and each session contained 24 trials. EEG signals and eye movements were collected with the 62-channel ESI Neuro Scan System and SMI eye-tracking glasses.	happiness, sadness, fear or neutral
4	SEED VIG	2017	constructed a system for virtual driving that places a sizable screen in front of a real car. In the automobile, participants can simulate driving in the actual world by participating in a driving simulation. Mostly Performed after lunch of afternoon	
5	SEED V		Provides EEG signals and eye movement features 20 subjects, including 10 males and 10 females. Figure 1	happy, sad, fear, disgust and neutral
6	Identifying similarities and differences in emotion recognition with EEG and eye movements among Chinese, German, and French People (SEED-GER and SEED-FRA)	2022	In SEED-FRA they included 8 France People in an experiment. In SEED-GER they included 8 German people in experiment	Positive,Negative and Neutral
7	Mahnob-HCI[9]		27 participants (19–40 years old, 11 male and 16 females) Each person watched 20 music videos 43.9–117 s long bandpass frequency filter (4–45 Hz) was applied	arousal, valence, and dominance.

Table 2.2: Literature Survey

Sr. No.	Paper title	Used dataset	Algorithm used	No. of classes	Accuracy
1	Emotion Classification Using EEG Brain Signals and the Broad Learning System[5]	DEAP and MAHNOB-HCI	DCNN, CNN, SVM, KNN, LDA, BLS	4	93.1% (DEAP) and 94.4% (MAHNOB-HCI)
2	Multi-Domain Feature Fusion for Emotion Classification Using DEAP Dataset[14]	DEAP	Support vector	9	65.92%
3	EEG-Based Emotion Classification Using a Deep NN and Sparse Autoencoders[15]	DEAP and SEED	CNN, ANN, BCI	2	89.49% (DEAP) and 96.77% (SEED)
4	An EEG Database and Its Initial Benchmark Emotion Classification Performance[16]	Database of 44 volunteers	ELM	4	94.72%
5	Automated Feature Extraction on AsMap for Emotion Classification Using EEG[17]	SJTU(SEED)	CNN	3	97.10%

Table 2.3: Literature Survey

Chapter 3

Project description

There are many benefits of emotion analysis from EEG data like Patient Monitoring, Psychological Therapy, Criminal detection, Disabled assistance, Security services, Robotics and many more.[18]

By taking this idea as the center we started working on it. There are many datasets available for detecting the emotion from the EEG data. We need a dataset which can be classified into 2 classes like anxiety or non-anxiety.[19]

flow diagram of project is :

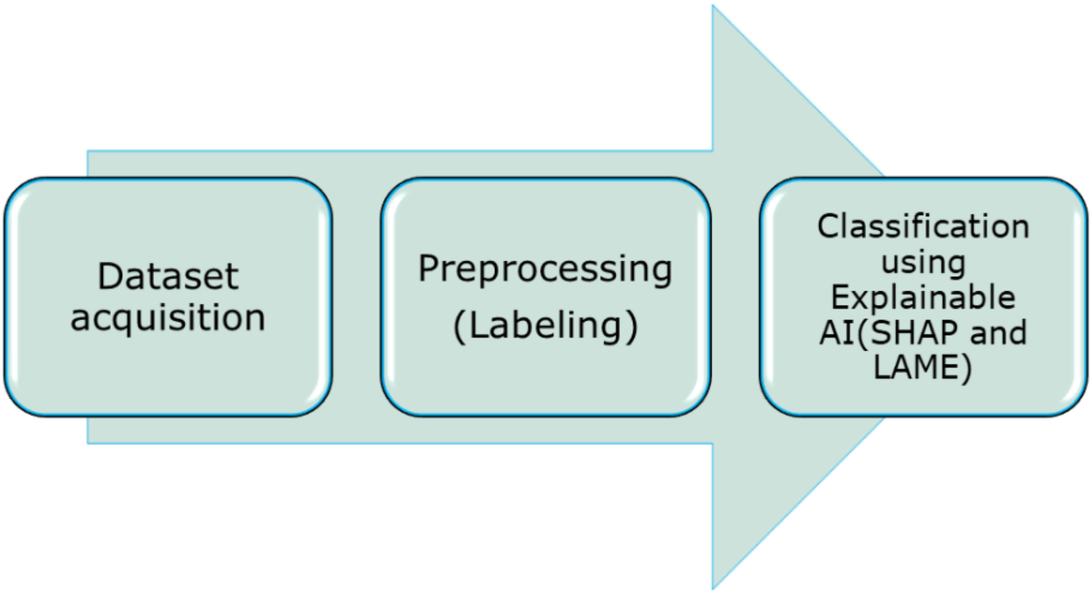


Figure 3.1: Flow chart of work
[20][17][21]

3.1 Dataset acquisition

NeuroSky Brainwave starter kit is used to collect the data.The NeuroSky Brainwave Starter Kit, featuring the NeuroSky MindWave Mobile EEG headset[22], is tailored to capture and assess brainwave patterns linked to attention and meditation levels. Employing a single-channel configuration, the headset focuses on gathering brainwave signals from the frontal lobe via the forehead region. Alongside this, the kit provides accessible developer tools and software for real-time visualization and analysis of the captured EEG signals, empowering researchers and developers to explore applications in cognitive psychology, neurofeedback, brain-computer interface development, and human-computer interaction. To establish a connection between the device and the computer, begin by activating the Bluetooth on both devices and pairing them. Upon successful pairing, ensure that the Neurosky headset is correctly worn, with one probe positioned on the hairy scalp and the other on the earlobe. Subsequently, open the software interface and initiate the data generation process by selecting the connect button. To complete this work steps are as below:

1. Choose the emotion that will be the focus for this project.

2. Collect the videos regarding selected emotion. These videos are selected from

youtube, references for this videos are mentioned in Table ??.

3. Understand and configure the EEG instrument for data collection. Selected EEG instrument is of Neurosky brainwave starter kit with 1 Channel. It has only one dry electrode to record EEG signals from the frontal lobe (Fp1) of the person's mind. with reference and ground electrodes situated in the ear clip. The active electrode, following the 10/20 system, monitors activity in the frontal lobe and is positioned slightly left from the mid line, while the left earlobe (A1 position) serves as the reference point. Data recorded at 256 Hz Sampling rate.

4. Collected data of 29 subjects(14 Males + 15 Females) between 22-27 age group. Mean age of all subjects is 24 years.

As mentioned above that selected classes are 2 which are anxiety or non-anxiety, once the classes were selected second step is to identify videos for the same that are able to make us feel like anxiety, so we selected hunting videos separated by one funny video of Tom and Jerry.

Sr. NO	**Video link**	**Video description**
1	https://www.youtube.com/watch?v=Bv4d5XHVgGg	first Anxious Video.
2	https://www.youtube.com/watch?v=t0Q2otsqC4I	Funny Video (up to 2:30 from starting)
3	https://www.youtube.com/watch?v=pkhE14Rou-E	Second Anxious Video

Table 3.1: Selected videos for subjects to show

By combining these videos, we made a 10-minute video to show to each subject and each video is separated by 5 seconds of white screen to give a break between changes of video.[23][22]

3.1.1 Used components and software in this project

An electroencephalogram device with three channels called the Narosky MindWave Starter Kit is used to measure and keep track of brainwave activity. It consists of a lightweight headset that the user can wear comfortably and three strategically placed sensors that pick up electrical signals from the user's forehead. These sensors record and send the

Figure 3.2: Used EEG Instrument in Project
[24]

electrical signals that the brain produces to the computer or other linked device for additional processing.[25]

Three channels of EEG data are available with the MindWave Starter Kit, providing a simple yet thorough insight of the brain's activity. Delta, theta, alpha, beta, and gamma waves are among the frequency bands covered by these channels. This information is essential for evaluating different cognitive states, such as relaxation, focus, and attention. The system is powered by Narosky's exclusive technology, which uses a dry sensor system instead of conductive gel or other preparations. With little setup time needed, this feature makes it convenient and user-friendly. A reference electrode is also included in the headset to guarantee accurate measurements and decrease unwanted noise interference.[26]

A computer or mobile device can wirelessly connect to the MindWave[25] Starter Kit through Bluetooth to monitor and analyze the user's brainwave activity in real-time. It includes software development tools and APIs that enable programmers to build original applications and incorporate the EEG data into their works. Due to its adaptability, the tool can be used for a variety of reasons, including neurofeedback, research into the brain-computer interface, training in meditation and relaxation, and educational purposes.

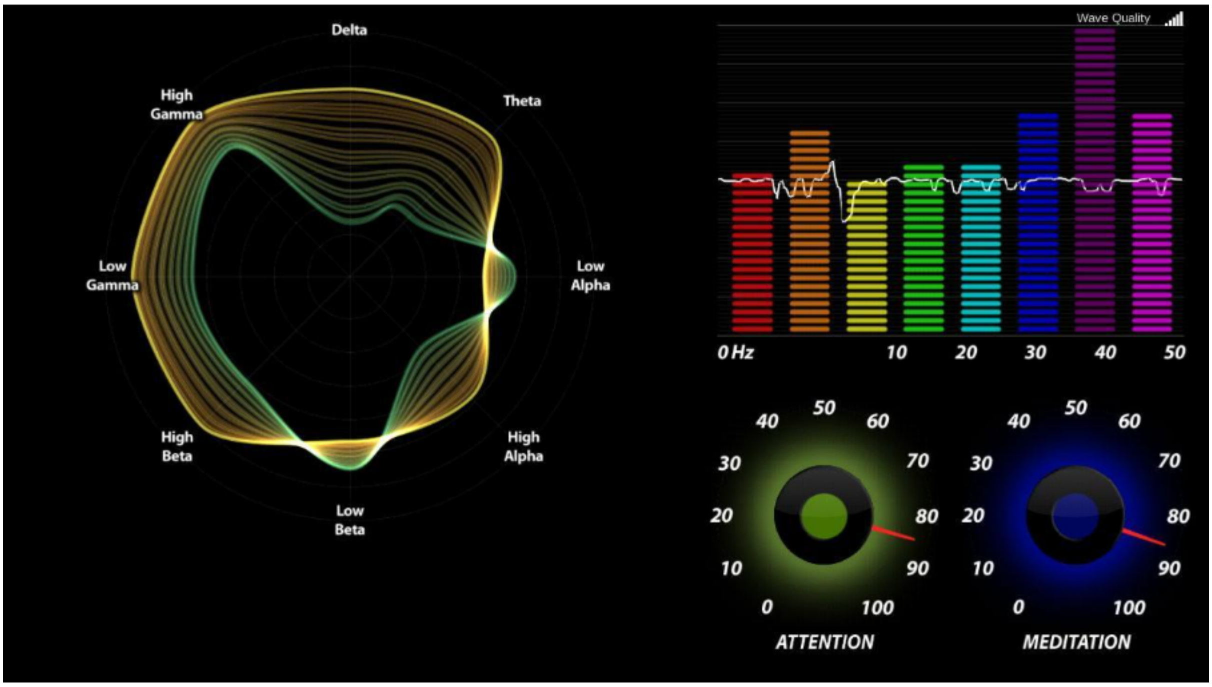


Figure 3.3: Used software in project

This shown is the flow diagram of video created by us which we used during experiment.

3.2 Data preprocessing :-

Data preprocessing involves several steps, including data cleaning, integration, transformation, and reduction.

For training and modeling, we focused on utilizing power spectrum data. Within this dataset, timestamps associated with anxiety-inducing videos were labeled as 0, those featuring a plain white screen were labeled as 1, and instances showcasing comedy videos of Tom and Jerry were marked as 2.

Label	Description
0	Non Anxious frames
1	Anxious screen

Table 3.2: Labels in dataset while preprocessing

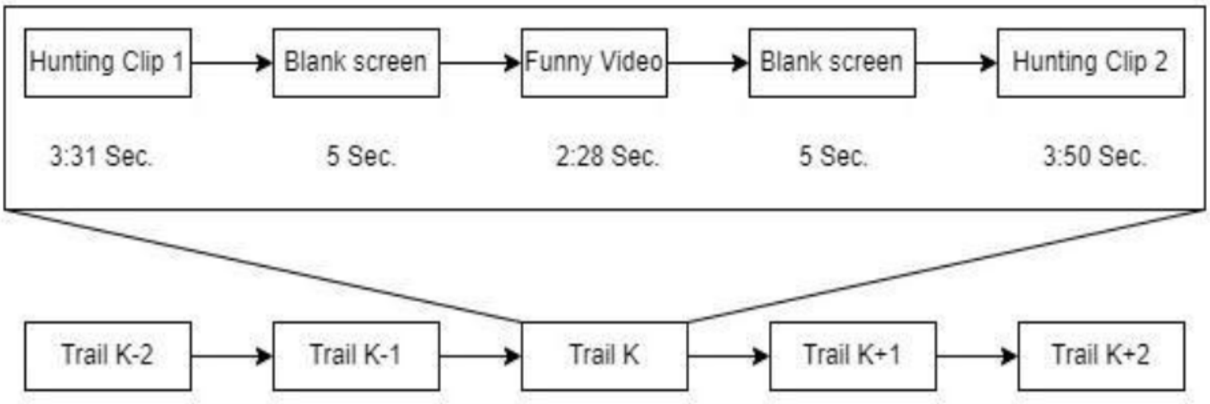


Figure 3.4: Flow chart of data acquisition in project
[10]

62.5	62.8	63	63.3	63.5	63.8	label
6.36E-14	6.35E-14	6.29E-14	6.23E-14	6.19E-14	6.17E-14	0
8.04E-13	7.95E-13	7.92E-13	7.88E-13	7.86E-13	7.85E-13	0
2.25E-14	2.03E-14	1.79E-14	1.59E-14	1.45E-14	1.37E-14	0
3.57E-14	3.58E-14	3.59E-14	3.57E-14	3.57E-14	3.57E-14	0
1.81E-13	1.78E-13	1.78E-13	1.77E-13	1.77E-13	1.77E-13	0
1.20E-13	1.17E-13	1.14E-13	1.12E-13	1.10E-13	1.09E-13	0
1.48E-14	1.01E-14	6.53E-15	3.65E-15	1.64E-15	4.38E-16	0
1.72E-14	1.67E-14	1.63E-14	1.59E-14	1.56E-14	1.55E-14	0
1.89E-14	1.86E-14	1.78E-14	1.71E-14	1.66E-14	1.63E-14	0
1.68E-14	1.38E-14	1.16E-14	9.63E-15	8.26E-15	7.44E-15	0

Figure 3.5: labeling of dataset

Due to the commencement and conclusion of videos, certain timestamps exhibited redundancy. To address this, any surplus timestamps at the end of a video were designated as label 1.

3.3 Model Training

We employed various machine learning models—SVM, Liner Regression, Decision Tree, and Random Forest—to train on the EEG dataset. After evaluating their performances, we found that Random Forest achieved the highest accuracy, reaching 80%. Consequently, we exclusively utilized the Random Forest model for subsequent applications of explainable AI tools. In the below image, it is shown that the random forest contains the best model accuracy among the other models.

Explainable AI (XAI) refers to the ability of artificial intelligence systems to pro-

Candidate	Decision Tree	LR	SVM	RF	Ensamble
bhaumik_22mcec06	0.7849	0.7990	0.8070	0.8663	0.8171
Dev_22MCED17	0.8323	0.9018	0.9080	0.9100	0.9141
dhruv_22mced01	0.6892	0.7568	0.7639	0.8012	0.7810
Divya_22mcec08	0.6952	0.7452	0.7554	0.7543	0.7635
Harsh_Maheshbhai_22mced04	0.7748	0.7984	0.7984	0.8506	0.8342
Harsh_nanavati_22mced12	0.7937	0.8703	0.8744	0.8795	0.8815
Het_pandya_22MCEC07	0.7719	0.8527	0.8648	0.8618	0.8708
Ishika_22MCEC10	0.7128	0.7454	0.7566	0.7800	0.7688
jainam_22mcec17	0.7129	0.7500	0.7500	0.7786	0.7489
jay_22MCEC18	0.6854	0.7575	0.7715	0.8036	0.7705
kaushal_22mced09	0.6946	0.7722	0.7865	0.7651	0.7824
khevna_22ftphde72	0.7020	0.7888	0.7949	0.8296	0.8061
Krinal_22mced06	0.6906	0.7582	0.7592	0.7838	0.7930
kuheli_22mcec04	0.6710	0.7488	0.7398	0.7617	0.7428
nidhi_21mcec07	0.6663	0.7528	0.7620	0.7467	0.7569
Nidhi_22ftphde70	0.6779	0.7574	0.7635	0.7503	0.7554
nikita_22mced11	0.7290	0.7331	0.7331	0.7958	0.7331
Priyal_22ftphde73	0.7138	0.7688	0.7291	0.7750	0.7505
rashi_22mcec12	0.6982	0.7663	0.7551	0.7805	0.7683
rutwik_22mcec01	0.6957	0.7677	0.7627	0.7880	0.7677
Sanket_22MCED07	0.6841	0.7754	0.7826	0.7559	0.7692
Selvi_22mcec16	0.6707	0.7573	0.7714	0.7714	0.7714
Shristi_22MCED05	0.7368	0.8089	0.8150	0.8323	0.8354
Shruti_22mces17	0.6988	0.7556	0.7596	0.7677	0.7576
tanisq_22mced16	0.7696	0.8919	0.9021	0.8654	0.8797
tirth_22mcec13	0.8092	0.8363	0.8454	0.8795	0.8655
vidhi_22mcec15	0.6368	0.7435	0.7374	0.7485	0.7425
vipasha_22mced14	0.7260	0.7553	0.7533	0.8291	0.7654
Yashesh_22mced13	0.7465	0.8099	0.8218	0.8505	0.8297
Average accuracy	**0.7169**	**0.7836**	**0.7870**	**0.8056**	**0.7939**

Table 3.3: Accuracy of Machine Learning Models for EEG Dataset

vide understandable explanations regarding their decisions or outputs. It aims to bridge the gap between complex AI models and human comprehension by making the decision-making process transparent and interpretable. XAI techniques enable users to comprehend why and how an AI system arrives at a specific conclusion or recommendation, fostering trust, accountability, and reliability in AI applications across various domains.

There are several tools and techniques used in Explainable AI (XAI) to enhance the interpretability and transparency of AI models. Here are some commonly used ones:

- **LIME (Local Interpretable Model-agnostic Explanations):** This tool generates locally faithful explanations for the predictions of any machine learning model.

- **SHAP (SHapley Additive exPlanations):** It uses cooperative game theory to explain the output of any machine learning model by attributing the prediction to different features.
- **Anchor:** It produces easy-to-understand rules that describe the local behaviour of the model for specific instances.
- **Saliency Maps:** These visualize the importance of input features by highlighting which parts of the input contributed most to the model's prediction.
- **Integrated Gradients:** This method assigns an importance score to each feature by integrating the model's predictions over a straight path from a baseline to the input.
- **Partial Dependence Plots (PDP):** While maintaining the same values for other features, it illustrates the connection between a feature and the anticipated result.

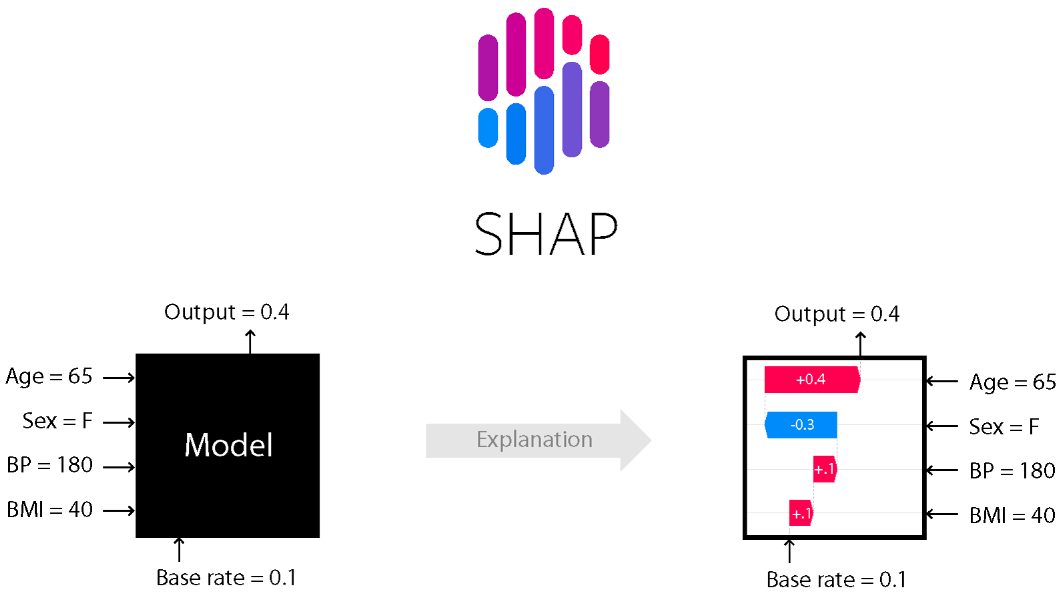


Figure 3.6: SHAP tools

3.3.1 SHAP

A game theoretic method called SHAP (Shapley Additive exPlanations) can be used to explain any machine learning model's output. The fundamental notion underlying Shap value-based explanations of machine learning models is to distribute credit for a model's output among its input features by applying the results of fair allocation from cooperative game theory.

The fact that Shapley values always add up to the difference between the result of the game with every player present and the result with none is one of its basic characteristics. This indicates that for machine learning models, the difference between the baseline (anticipated) and current model outputs for the explained prediction will always be the total of the SHAP values of all the input features.

3.3.2 LIME

Local Interpretable Model-agnostic Explanations (LIME). LIME focuses on elucidating the model's prediction for specific occurrences rather than offering a comprehensive knowledge of the model across the entire dataset.
LIME explainer can be set up using two main steps: (1) import the lime module, and (2) fit the explainer using the training data and the targets. During this phase, the mode is set to classification, which corresponds to the task being performed.
The code snippet below generates and displays a LIME explanation for the 8th instance in the test data using the random forest classifier and presenting the final feature contribution in a tabular format.

The result contains three main pieces of information from left to right: (1) the model's predictions, (2) features contributions, and (3) the actual value for each feature.
LIME is often favored for its focus on local explanations, while SHAP provides a more global view of feature importance.

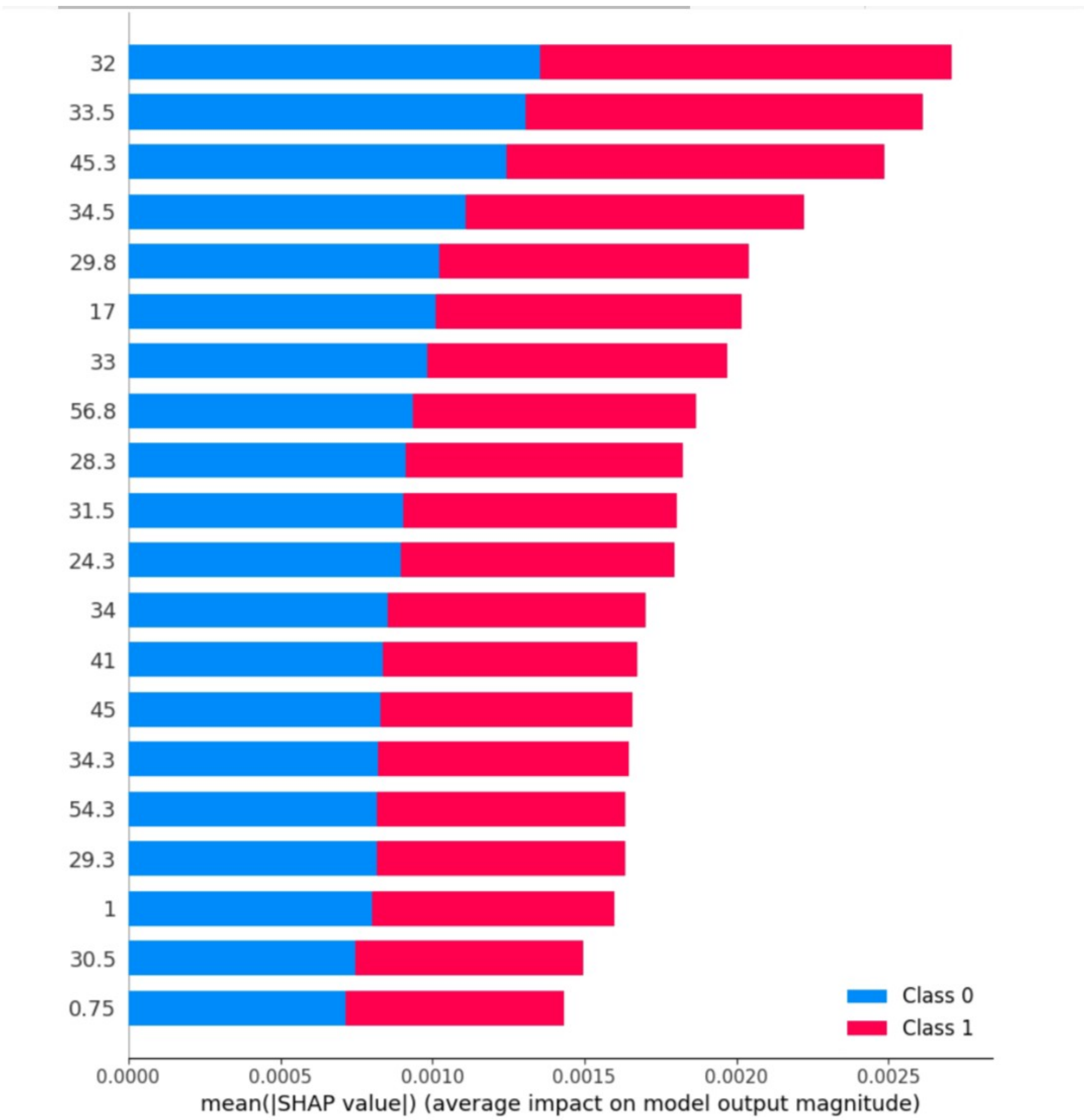


Figure 3.7: SHAP frequency

3.3.3 Explainable AI Use In EEG classification

Machine learning models have been used more and more in the past several years, especially in the fields of psychology and healthcare. A critical component in guaranteeing the transparency and interpret ability of these models is Explainable AI (XAI). This work investigates the usage of two popular XAI methods for EEG data analysis: SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations).The principal aim is to elucidate the roles played by distinct EEG frequency

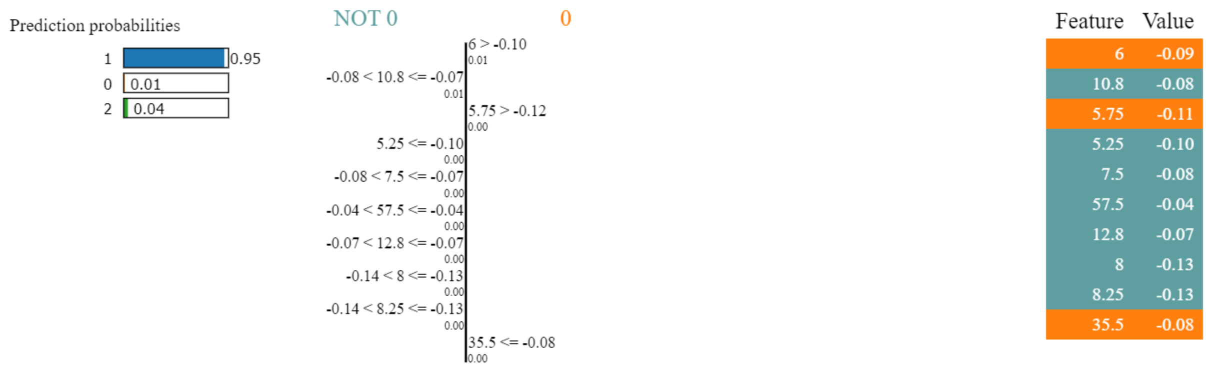


Figure 3.8: LIME tools

patterns affecting the recorded anxiety levels, so providing insight into the inner workings of the machine learning models utilized for the research. This investigation not only makes the brain correlates of anxiety more understood, but it also emphasizes how crucial interpretable AI techniques are to closing the gap between sophisticated model outputs and actual psychological experiences.

The use of LIME and SHAP for the analysis of EEG data is very beneficial in the context of our study. LIME's local interpretability can reveal how particular EEG frequency patterns affect predictions at the level of a single subject, offering insights into unique brain reactions to stimuli that cause fear. In the meantime, the global interpretability of SHAP provides a more comprehensive viewpoint on the cumulative significance of various EEG parameters throughout the full dataset, assisting in the discovery of recurring patterns associated with anxiety.[27]

Machine learning models have been used more and more in the past several years, especially in the fields of psychology and healthcare. A critical component in guaranteeing the transparency and interpret ability of these models is Explainable AI (XAI). This work investigates the usage of two popular XAI methods for EEG data analysis: SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations).The principal aim is to elucidate the roles played by distinct EEG frequency patterns affecting the recorded anxiety levels, so providing insight into the inner workings of the machine learning models utilized for the research. This investigation not only makes the brain correlates of anxiety more understood, but it also emphasizes how crucial interpretable AI techniques are to closing the gap between sophisticated model outputs and actual psychological experiences.

The use of LIME and SHAP for the analysis of EEG data is very beneficial in the context of our study. LIME's local interpretability can reveal how particular EEG frequency patterns affect predictions at the level of a single subject, offering insights into unique brain reactions to stimuli that cause fear. In the meantime, the global interpretability of SHAP provides a more comprehensive viewpoint on the cumulative significance of various EEG parameters throughout the full dataset, assisting in the discovery of recurring patterns associated with anxiety.

We utilized machine learning algorithms, including SVM, decision trees, and ensemble models, on our EEG dataset. Among these models, the random forest yielded the highest accuracy. Consequently, we leveraged the trained random forest model within explainable AI tools. These tools interpret the model's decisions, shedding light on the specific frequencies that led to the labeling of timestamps in our data.

Candidate	**F1**	**F2**	**F3**	**F4**	**F5**	**Accuracy**
bhaumik_22mcec06	51.5	57.3	55.5	57	34.5	**0.87**
Dev_22MCED17	33	37.8	46	34.8	46.3	**0.91**
Divya_22mcec08	20.3	0.25	0.5	26.3	3.75	**0.75**
Harsh_Maheshbhai_22mced04	8.25	8	28.8	16.3	20.3	**0.85**
Harsh_nanavati_22mced12	37.5	42.8	49	47	50.3	**0.88**
jainam_22mcec17	1.25	2.25	0.75	0.25	1.75	**0.78**
jay_22MCEC18	43.3	44	43.5	39	43.8	**0.8**
khevna_22ftphde72	0.25	0.5	2.25	2	1.5	**0.83**
Krinal_22mced06	12.5	0.5	0.25	31	12.3	**0.78**
nidhi_21mcec07	35.3	5.75	0.25	39.8	23	**0.75**
Nidhi_22ftphde70	21.3	4.5	45.5	5.25	52.5	**0.75**
nikita_22mced11	12.5	12.3	0.25	0.75	10	**0.8**
Priyal_22ftphde73	0.25	0.75	37.3	45.5	32.3	**0.77**
rashi_22mcec12	0.25	0.5	12.5	0.75	51.3	**0.78**
rutwik_22mcec01	46	32.3	43.5	32.5	42	**0.79**
Sanket_22MCED07	26.8	21	56.8	32.8	20.8	**0.76**
Selvi_22mcec16	8	33.5	48	47.8	34	**0.77**
Shruti_22mces17	3.5	4.25	7.25	33	16	**0.77**
tanisq_22mced16	41.3	43.3	35.5	37.8	41.5	**0.87**
tirth_22mcec13	42.8	47	42	48.8	40	**0.88**
vidhi_22mcec15	35	0.25	36.5	39.3	27.5	**0.75**
vipasha_22mced14	46	50.3	38.3	45.3	47	**0.83**
Yashesh_22mced13	0.25	42	0.5	36	50	**0.85**
dhruv_22mced01	38	46	50.25	40	48.75	**0.8**
Ishika_22MCEC10	3.75	3	3.25	0.25	25.75	**0.78**
kaushal_22mced09	37.25	12.5	6.25	7	34.5	**0.77**
kuheli_22mcec04	0.75	0.25	0.5	1.75	24.5	**0.76**
Shristi_22MCED05	7.75	31.75	9.75	29.75	15.75	**0.83**
Average accuracy						**0.803929**

Table 3.4: Top SHAP frequency

Chapter 4

Summary and Conclusion

4.1 Summary:

This project results in a dataset of 30 people from which 15 were male and the rest of it are male. Each person shows a video of 10 minutes which contains 2 videos of anxious feeling which are of 4 minutes each and in between 2 minutes of funny video contains a 5 sec. gap between each video with white blank screen. Many datasets of EEG are there on various portals but the videos containing anxiety are very less in number so this dataset will contribute to that for further researchers and analysis in the domain of anxiety. Also, data files contain reading of meditation as well as attention of the person continuously. The dataset was further trained using the Random Forest model. Subsequently, explainable AI tools like Lime and SHAP were employed to analyze and interpret the outcomes derived from the trained model.

4.2 Future Work:

n future research, a longitudinal study could be conducted to observe the changes in dominant EEG frequencies for anxi- ety over time. This would provide crucial insights into the sta- bility of these frequencies and their relationship with long-term emotional experiences, offering a better understanding of the progression of anxiety-related conditions and the effectiveness of interventions over time. Additionally, exploring gender-based differences in dominant EEG frequencies for anxiety could lead to more targeted and gender-specific approaches for anxiety management and treatment. Moreover, developing a real-time monitoring system using EEG to detect and classify anxiety could

frequency range	subsection	subsection count	Total count
Less thane 12	less than 12	33	33
12-30	12-20	12	29
	20-30	17	
30-100	30-40	34	83
	40-50	30	
	50-60	17	
	60-70	2	
	70-80	0	
	80-90	0	
	90-100	0	

Figure 4.1: Conclusion

have potential applications in mental health interventions and personalized therapy, enabling individuals to receive timely support based on real-time EEG data. Fur- thermore, conducting the experiment using a greater number of EEG channels could enhance the accuracy and depth of the findings, providing a more comprehensive understanding of dominant EEG frequencies for anxiety.

4.3 Conclusion:

From this results we came to a solution by analyzing dominating frequencies for each subject. Above figure shows count of frequency from all subjects data between the different ranges. After distributing the frequencies to ranges it is clear that highest count noted is in second and third range. 12-30(beta waves) and 30-100(Gamma waves) in these 2 frequency ranges majority of all counts come. Beta and lower gamma are responsible frequencies for the subjects when the saw anxious content or feel anxious.

Bibliography

[1] M. Teplan *et al.*, "Fundamentals of eeg measurement," *Measurement science review*, vol. 2, no. 2, pp. 1–11, 2002.

[2] N. H. Frijda, *The emotions*. Cambridge University Press, 1986.

[3] L. Abramson, R. Petranker, I. Marom, and H. Aviezer, "Social interaction context shapes emotion recognition through body language, not facial expressions.," *Emotion*, vol. 21, no. 3, p. 557, 2021.

[4] N. Mehendale, "Facial emotion recognition using convolutional neural networks (ferc)," *SN Applied Sciences*, vol. 2, no. 3, p. 446, 2020.

[5] S. Issa, Q. Peng, and X. You, "Emotion classification using eeg brain signals and the broad learning system," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 12, pp. 7382–7391, 2020.

[6] R. C. Solomon and L. D. Stone, "On "positive" and "negative" emotions.," *Journal for the theory of social behaviour*, vol. 32, no. 4, 2002.

[7] D. F. Klein, "Anxiety reconceptualized," *Anxiety: New research and changing concepts*, pp. 235–263, 1981.

[8] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis; using physiological signals," *IEEE transactions on affective computing*, vol. 3, no. 1, pp. 18–31, 2011.

[9] S. Petridis, B. Martinez, and M. Pantic, "The mahnob laughter database," *Image and Vision Computing*, vol. 31, no. 2, pp. 186–202, 2013.

[10] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki, "Emotionmeter: A multimodal framework for recognizing human emotions," *IEEE transactions on cybernetics*, vol. 49, no. 3, pp. 1110–1122, 2018.

[11] R. Agarwal, M. Andujar, and S. Canavan, "Classification of emotions using eeg activity associated with different areas of the brain," *Pattern Recognition Letters*, vol. 162, pp. 71–80, 2022.

[12] H. Alsuradi, W. Park, and M. Eid, "Explainable classification of eeg data for an active touch task using shapley values," in *International Conference on Human-Computer Interaction*, pp. 406–416, Springer, 2020.

[13] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks," *IEEE Transactions on autonomous mental development*, vol. 7, no. 3, pp. 162–175, 2015.

[14] M. Khateeb, S. M. Anwar, and M. Alnowami, "Multi-domain feature fusion for emotion classification using deap dataset," *Ieee Access*, vol. 9, pp. 12134–12142, 2021.

[15] J. Liu, G. Wu, Y. Luo, S. Qiu, S. Yang, W. Li, and Y. Bi, "Eeg-based emotion classification using a deep neural network and sparse autoencoder," *Frontiers in Systems Neuroscience*, vol. 14, p. 43, 2020.

[16] A. Seal, P. P. N. Reddy, P. Chaithanya, A. Meghana, K. Jahnavi, O. Krejcar, and R. Hudak, "An eeg database and its initial benchmark emotion classification performance," *Computational and mathematical methods in medicine*, vol. 2020, 2020.

[17] M. Z. I. Ahmed, N. Sinha, S. Phadikar, and E. Ghaderpour, "Automated feature extraction on asmap for emotion classification using eeg," *Sensors*, vol. 22, no. 6, p. 2346, 2022.

[18] J. Rodgers, K. Farquhar, D. Mason, S. Brice, S. Wigham, B. Ingham, M. Freeston, and J. R. Parr, "Development and initial evaluation of the anxiety scale for autism-adults," *Autism in Adulthood*, vol. 2, no. 1, pp. 24–33, 2020.

[19] M. N. Giannakos, K. Sharma, I. O. Pappas, V. Kostakos, and E. Velloso, "Multimodal data as a means to understand the learning experience," *International Journal of Information Management*, vol. 48, pp. 108–119, 2019.

[20] J. Wang and M. Wang, "Review of the emotional feature extraction and classification using eeg signals," *Cognitive robotics*, vol. 1, pp. 29–40, 2021.

[21] R.-N. Duan, J.-Y. Zhu, and B.-L. Lu, "Differential entropy feature for eeg-based emotion classification," in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 81–84, IEEE, 2013.

[22] D. J. Lancheros-Cuesta, J. L. R. Arias, Y. Y. Forero, and A. C. Duran, "Evaluation of e-learning activities with neurosky mindwave eeg," in *2018 13th Iberian conference on information systems and technologies (CISTI)*, pp. 1–6, IEEE, 2018.

[23] L. J. Julian, "Measures of anxiety," *Arthritis care & research*, vol. 63, no. 0 11, 2011.

[24] H. Al-Kaf, A. Khandoker, K. Khalaf, and H. F. Jelinek, "Neurosky mindwave mobile headset 2 as an intervention for reduction of stress and anxiety measured with pulse rate variability," in *2020 Computing in Cardiology*, pp. 1–4, IEEE, 2020.

[25] J. Katona, I. Farkas, T. Ujbanyi, P. Dukan, and A. Kovari, "Evaluation of the neurosky mindflex eeg headset brain waves data," in *2014 IEEE 12th international symposium on applied machine intelligence and informatics (SAMI)*, pp. 91–94, IEEE, 2014.

[26] X. Du, C. Ma, G. Zhang, J. Li, Y.-K. Lai, G. Zhao, X. Deng, Y.-J. Liu, and H. Wang, "An efficient lstm network for emotion recognition from multichannel eeg signals," *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1528–1540, 2020.

[27] M. S. Islam, I. Hussain, M. M. Rahman, S. J. Park, and M. A. Hossain, "Explainable artificial intelligence model for stroke prediction using eeg signal," *Sensors*, vol. 22, no. 24, p. 9859, 2022.

Harsh report

ORIGINALITY REPORT

5%	6%	0%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

7%

★ Submitted to Institute of Technology, Nirma University

Student Paper

Exclude quotes	On	Exclude matches	< 2%
Exclude bibliography	On		