A Project Report On

Emotion Classification on Physiological Data

Submitted By
Gaurav Kumar Singh (2K21/DSC/06)

MASTER OF TECHNOLOGY
IN
Data Science

Under the supervision of

Dr. Divyashikha Sethia

(Assistant Professor, Software Engineering, DTU)

DEPARTMENT OF SOFTWARE ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi – 110042

May 2023

DEPARTMENT OF SOFTWARE ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY (Formerly Delhi College of Engineering)
Bawana Road, Delhi - 110042

CANDIDATES DECLARATION

I, Gaurav Kumar Singh, Roll No. 2K21/DSC/06, student of M.Tech(Data Science), hereby declare that the project dissertation titled Emotion Classification on Physiological Data, which is submitted by me to the Department of Software Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not been submitted anywhere for the award of a degree, diploma, fellowship or other similar title or recognition to the best of my knowledge.

Gaurav Kumar Singh

Place: Delhi
Date: May 13, 2023

Gaurav Kumar Singh
Roll No. 2K21/DSC/06
Department of
Software Engineering,
Delhi Technological
University

DEPARTMENT OF SOFTWARE ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi - 110042

CERTIFICATE

I certify that the Project Dissertation titled **Emotion Classification on Physiological Data** which is submitted by **Gaurav Kumar Singh (Roll No. 2K21/DSC/06)**, Department of Software Engineering, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is a record of the Project work carried out by the student under my supervision. To the best of my Knowledge this work has not been submitted in part or full for any degree or diploma to this university or elsewhere.

Place: Delhi

Date:

Dr. Divyashikha Sethia
Assistant Professor,
Department of Software Engineering,
Delhi Technological University

ACKNOWLEDGEMENT

I am extremely grateful to Dr Divyashikha Sethia(Assistant Professor, Department of Software Engineering) and all the Department of Software Engineering employees at DTU. They all furnished us with enormous help and direction for the venture.

I would like to thank the University for giving us the research centres, framework, testing offices, and climate, which permitted us to work with next to no checks.

Likewise, I want to see the value in the help given to us by our lab partners, seniors, and companion bunch, who supported us with all the information they had regarding different subjects.

Gaurav Kumar Singh
2K21/DSC/06

CONTENTS

TABLE OF CONTENTS	PAGE NO
TITLE PAGE	1
DECLARATION	2
CERTIFICATE	3
ACKNOWLEDGEMENT	4
LIST OF FIGURES	5
LIST OF TABLES	6
ABSTRACT	7
CHAPTER 1 INTRODUCTION	8
CHAPTER 2 LITERATURE REVIEW	13
CHAPTER 3 RELATED WORK	20
CHAPTER 4 THE EMOTION RECOGNITION FRAMEWORK	25
CHAPTER 5 METHODOLOGY	31
CHAPTER 6 RESULT	38
CHAPTER 7 CONCLUSION AND FUTURE WORK	42
REFERENCES	44

LIST OF FIGURES

Figure Number	Figure Name	Page Number
1	Frequency Bands of EEG signal	8
2	Two-dimensional valence-arousal space	9
3	Centralized Machine Learning	10
4	Federated Learning	11
5	Delta Wave	14
6.	Theta Wave	14
7.	Alpha Wave	15
8.	Beta Wave	15
9.	Gamma wave	16
10	Wavelet decomposition of different bands	20
11.	Human emotions based on arousal and valence model	29
12.	Participants in experiment conditions during the short videos experiment recorded	31
13.	Proposed Workflow	32
14.	Overlapping window signal decomposition	35

LIST OF TABLES

Table Number	Table Name	Page Number
1	List of abbreviations used	21
2	List of features used in previous work done on the AMIGOS dataset	25
3	Previous models used and accuracies on the AMIGOS dataset	26
4	Coding of Valence, Arousal, and Dominance from 1-9 to 0-1	36
5	The accuracy obtained after applying ML/DL classifiers in this study	38
6	Comparison table of proposed work with existing work on the AMIGOS dataset	40

ABSTRACT

A person interacts with many different machines throughout their lifetime. Emotions also play a crucial and inevitable role in everyone's life, as they can trigger various thoughts, feelings, and behavioural responses. The Electroencephalogram (EEG) signal, which measures brain activity in the scalp region, is the most effective tool for tracing response changes. Emotion classification utilizing the physiological signal EEG has been conducted using the AMIGOS dataset. Before the emotion classification, the study extracted features from the EEG signals using the time domain feature Power Spectral Density (PSD) and time and frequency domain feature Wavelet entropy. Data cleaning and preprocessing were performed to prevent biased results caused by missing values among different users, which involved handling and addressing missing values. The study utilized Support Vector Machine, Artificial Neural Network (ANN), and Convolutional Neural Network with Overlapping and Non-Overlapping Sliding Window techniques, all based on machine learning and deep learning. Finally, the study classified emotions regarding Arousal, Valence, and Dominance.

Chapter 1

1.1 Introduction:

Estimating human emotions from electroencephalogram (EEG) signals is essential for creating a reliable Brain-Computer Interface (BCI) system. Rather than behavioural responses, physiological signals can represent difficult-to-hide alterations in the central nervous system strongly linked to people's emotional states. Many BCI studies have identified and recognized the user's affective state and used the results in various settings.

Brain activity produces different signals, including magnetic and electric ones [25]. This brain activity can be detected using various invasive and non-invasive methods. While non-invasive techniques do not require surgical intervention, invasive approaches involve implanting a specific device in the brain.

EEG is useful for getting brain waves from the scalp surface that correlate to various states. EEG is frequently employed in BCI research investigations because it is non-invasive to the research subject and poses no risk [25]. Fig. 1 shows the frequency bands of the EEG signal.

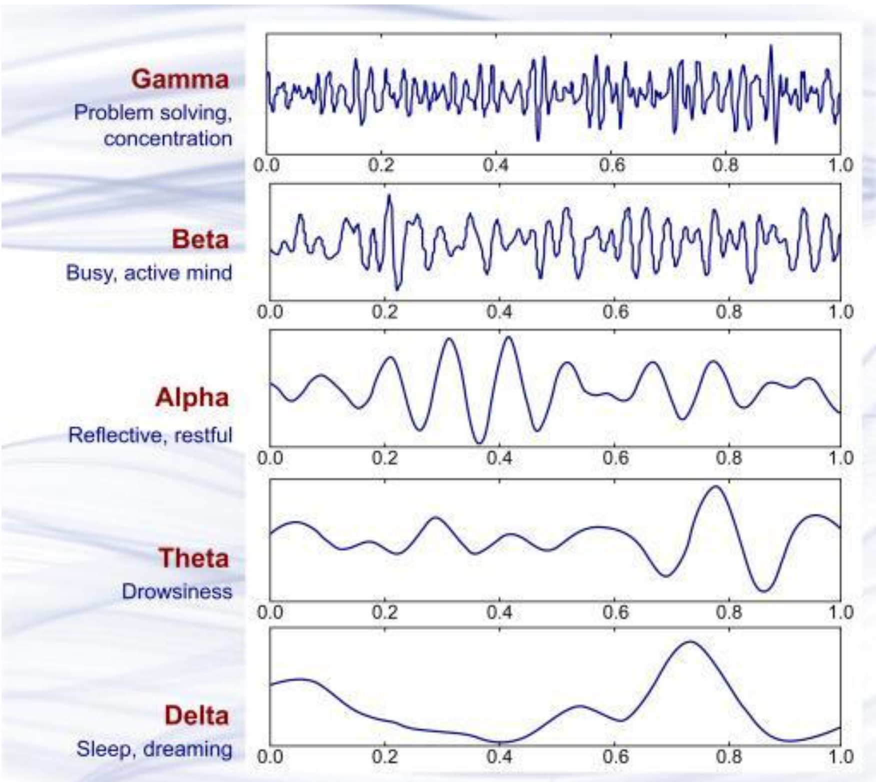

Fig. 1 Frequency Bands of EEG signal [19]

Discrete and Dimensional emotional models are employed to construct the emotional space. According to the discrete model, the emotional space consists of only a few basic discrete emotions, such as anger, happiness, fear, disgust, surprise, and sadness [20].in the dimensional model, emotions are defined as points in a dimensional space. The researchers have proposed two-three-dimensional models: Valence-Arousal and Valence-Arousal-Dominance [21].

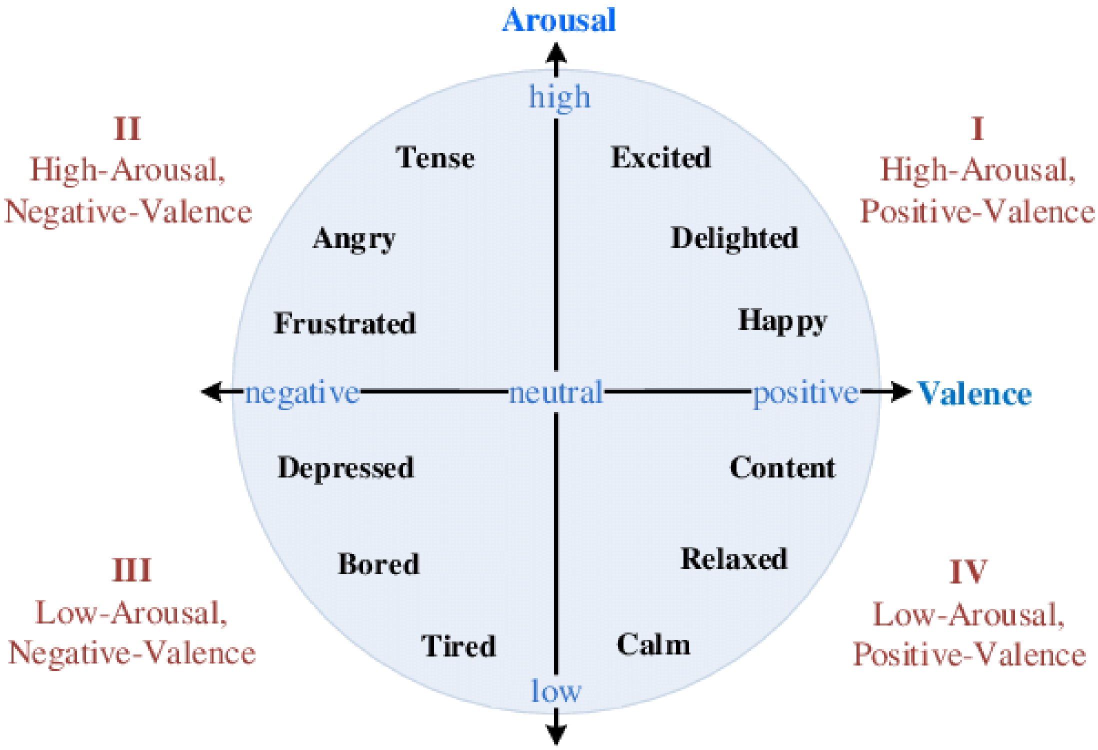

Fig. 2 Two-dimensional valence-arousal space [21]

The positive degree of emotion is called Valence [22], and the strength of emotion is called arousal. Dominance is the degree of perceived control over a person's emotional state.

1.2 Centralized Machine Learning

Endpoint devices storing data have significantly increased with the Internet of Things (IoT) advancement. Significant storage capacity and high-end resources are required to store and transfer this data to one location. Combining all the data from these devices and performing model training on a central server is called centralized machine learning, as shown in Fig. 3. A large volume of data is collected in centralized machine learning with powerful computing resources, resulting in a highly accurate model.

There is only a limited amount of data available, or poor-quality data exists; data is available on different devices across the globe. It would be challenging to break the barriers and use the data from all the devices without violating privacy concerns. Federated learning is one such method that provides solutions to given problems.

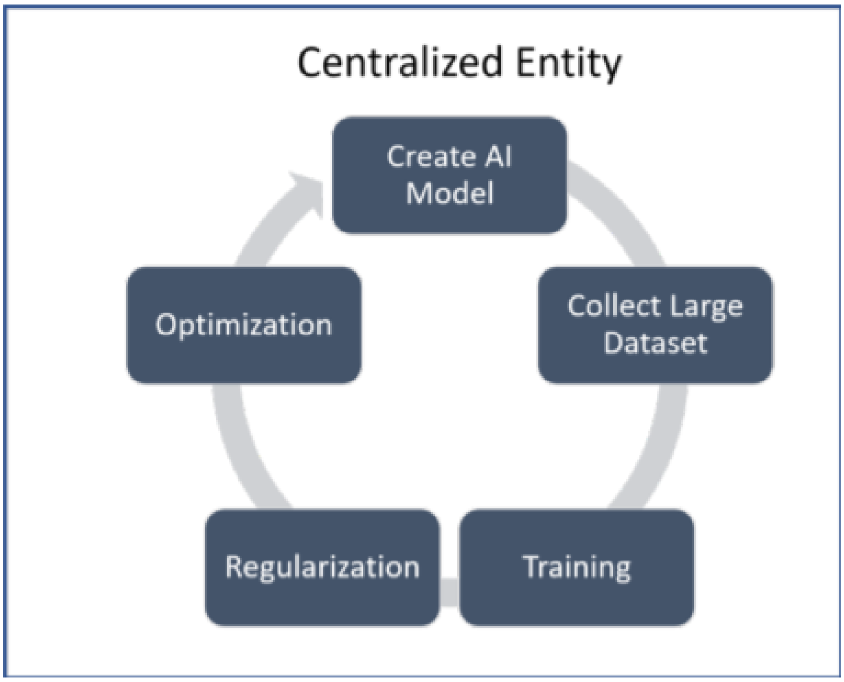

Fig. 3 Centralised Machine Learning [23]

1.3 Federated Learning:

Federated learning is a novel approach to machine learning that differs from traditional centralized training methods. Instead of relying on a central server to process data, Federated learning enables individual users located in different locations to train the model with their local data. This decentralized approach ensures that users can train the model without sharing their data with other participants, which helps to protect user privacy.

In Federated learning, each user's device acts as a node that processes the local data to update the global model. Unlike traditional centralized training methods, where the server processes all data, each edge device computes the updates on the raw data and shares only the updates with the central server rather than sharing the actual raw data. This approach helps to reduce the amount of data that needs to be transferred between devices and ensures that users retain control over their data.

To initiate the training process, the server shares variables from the global model with the local model on the client side. Then, the updated global model is returned to each client to enhance the accuracy of the classification model. FedAvg, an aggregation technique that employs Stochastic Gradient Descent (SGD) optimization, aggregates the updates from multiple clients. This approach ensures the resulting model is accurate and reliable while maintaining user privacy. Overall, Federated learning is a promising approach to machine learning that has the potential to revolutionize the field while also protecting user privacy.

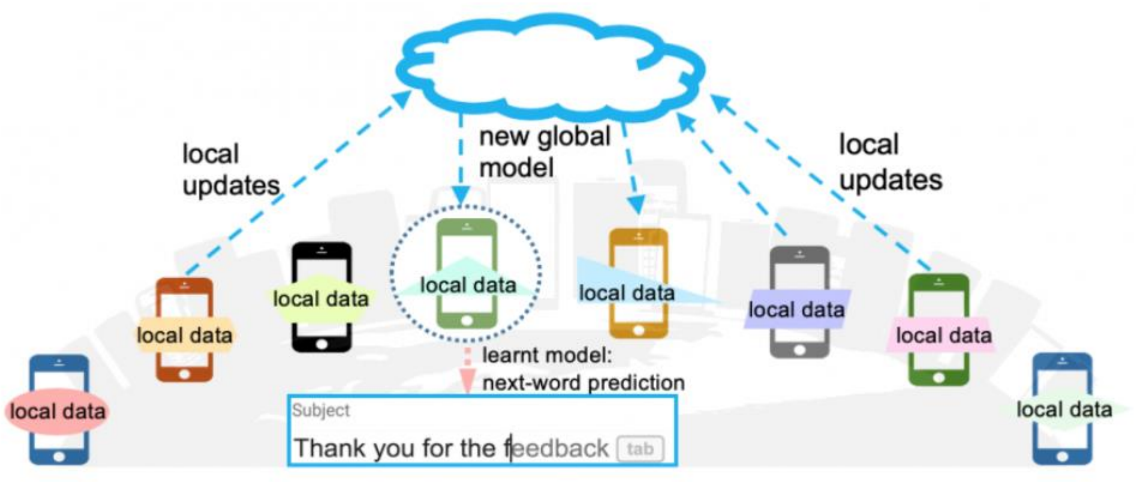

Fig. 4 Federated Learning [6]

1.4 Problem Statement:

Emotion is an individual's interpersonal state generated by a person's thoughts. Comprehending human emotions is crucial in various fields, including brain-computer interfaces, affective computing, and healthcare. With the recent development in technology, it has become easier to access and monitor the physiological signals of a person using wearable devices, smartphones, etc. Electroencephalogram (EEG) signals can accurately and inexpensively identify emotions due to their non-invasiveness and straightforward installation of the device in the brain to record brain activity [25]. Improving Human-Robot Interaction (HRI) involves enabling robots to recognize and react to the emotions of others, thus allowing them to act socially [26].

1.5 Objectives:

- To implement emotion recognition based on EEG signal using Overlapping Sliding Window (OSW) and Non-Overlapping Sliding Window (NOSW) based techniques.

- To implement a Support Vector Machine (SVM), Artificial Neural Network (ANN), and Convolutional Neural Network (CNN) on OSW and NOSW-based techniques.

- Compare the accuracies of both approaches and decide on the better approach for emotion classification on physiological data

Chapter 2: Literature Review

2.1 Electroencephalography (EEG):

The human brain is a complex organ that produces various signals, including electrical and magnetic signals. To study these signals, researchers employ different methods, including invasive and non-invasive approaches. Invasive methods require surgical intervention to implant specific devices in the brain, while non-invasive methods do not require such procedures. One of the most widely used non-invasive methods is Electroencephalography (EEG), which records the brain's electrical activity through electrodes placed on the scalp. This direct method allows researchers to monitor the current flow of neurons in the brain and generate EEG waves, which record voltage fluctuations over time. EEG is widely used in clinical and research settings due to its accessibility and safety. This method allows researchers to study brain activity in real-time and gain valuable insights into neurological disorders and brain function.

2.2 Types of Signals

In terms of signal types, EEG measures voltage as a function of time and is highly dependent on the degree of activity of the cerebral cortex [33][34]. EEG signals are classified based on various factors such as frequency, magnitude, and wave morphology. These characteristics can help classify the signals, along with their spatial distribution and reactivity. One popular method for classification is to group EEG waveforms into different frequency bands[35]. This approach can help researchers better understand the patterns and behaviour of EEG signals, which can ultimately lead to new insights and discoveries in the field. Five different frequency bands associated with a mental state can decompose EEG signals. These frequency bands include:

1. Delta (0.5-4Hz): Indicative of deep sleep or unconsciousness
2. Theta (4-8Hz): Indicative of drowsiness or meditative state
3. Alpha (8-12Hz): Indicative of relaxation or idling
4. Beta (12-30Hz): Indicative of alertness or focused concentration
5. Gamma (30-100Hz): Indicative of increased cognitive processing or heightened perception.

1. **Delta Waves (0.5-4 Hz):** Delta waves, which have a frequency of 0.5-4 Hz, make us feel rejuvenated after sleep. They are also a source of empathy and take away external awareness. Excessive delta waves in the brain can result in learning difficulties, brain damage, and cognitive impairment. Conversely, insufficient delta waves can indicate a lack of restorative sleep and an inability of the body to recharge and revitalize itself.

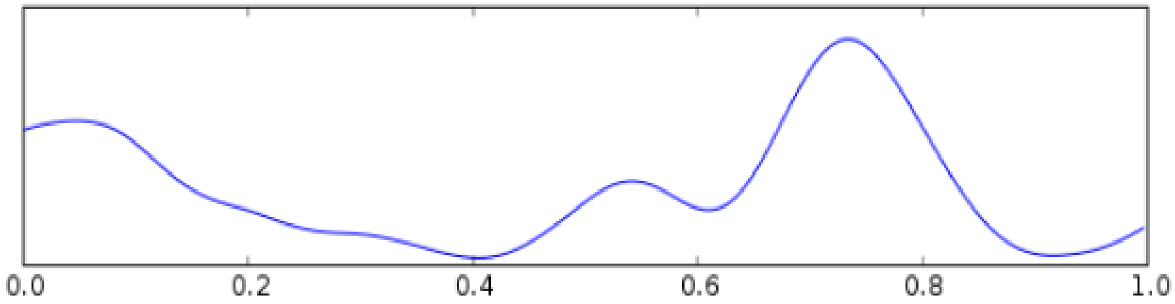

Fig. 5 Delta Wave [9]

2. **Theta Waves (4-8 Hz):** Theta waves, which have a 4-8 Hz frequency, are mainly emitted during sleep but are also present during deep meditation. Theta activity is often associated with learning, instinct, and memory. However, researchers have linked excessive theta activity to impulsivity, hyperactivity, depression, or inattentiveness. In contrast, a lack of theta waves indicates stress, poor emotional awareness, and anxiety. Theta waves are also associated with memory encoding during challenging tasks and can be observed in the early stages of sleep when one is feeling drowsy.

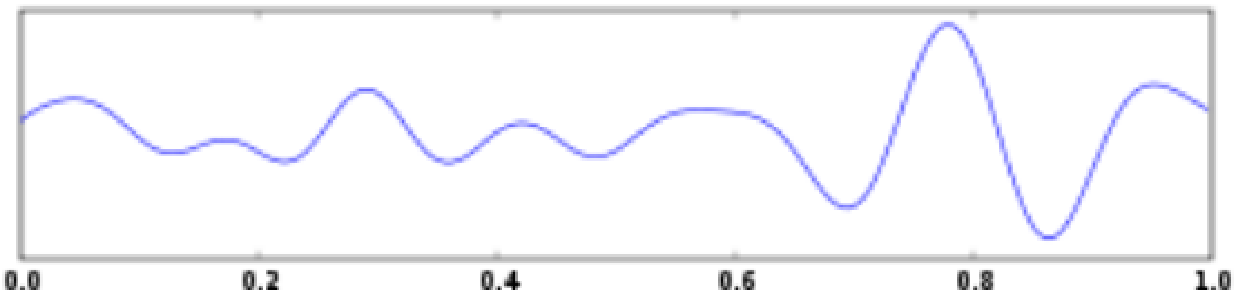

Fig. 6 Theta Wave [9]

3. **Alpha Waves (8-13 Hz):** Alpha waves, which have a frequency of 8-13 Hz, are responsible for mental coordination, learning, body/mind integration, and alertness. They can also help calm us down and relax the mind. On the other hand, too many alpha waves can lead to daydreaming and an inability to focus. Beta waves, which have a frequency of 13-30 Hz, are dominant when alert and engaged in cognitive activities. The right amount of beta waves can help us focus and efficiently perform tasks. However, excessive beta waves may lead to stress, anxiety, and an inability to relax or sleep. It is interesting to note that caffeine and other stimulants can increase beta activity in the brain.

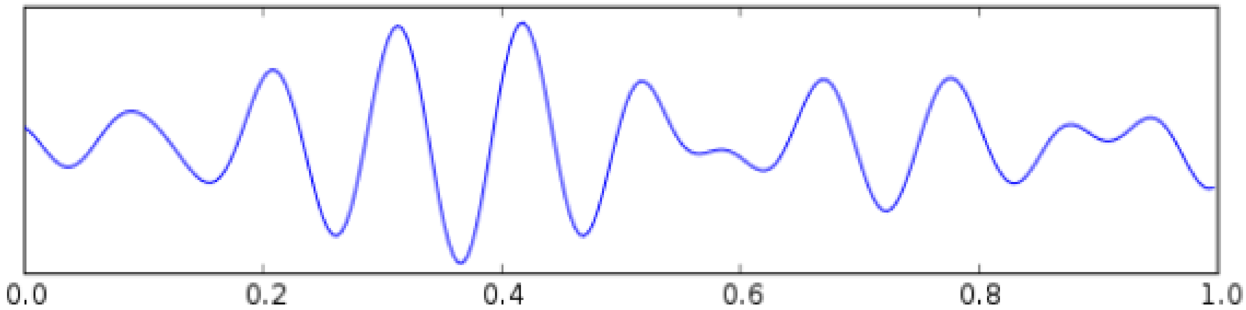

Fig. 7 - Alpha Wave [9]

4. **Beta Waves (13-30 Hz):** Understanding that beta brainwaves are crucial to our mental state is essential. When we are involved in cognitive activities related to the external world, these emotions dominate our minds. Beta waves are present when we are focused, alert, and actively solving problems or making decisions. They help us perform tasks with ease and concentration. However, too much beta can lead to stress, anxiety, and an inability to relax or sleep. Low beta waves are associated with daydreaming, depression, and poor cognition. Excessive adrenaline rush and arousal are caused by the high amount of beta waves, making it challenging to unwind. Consuming energy drinks, coffee, and other stimulants can potentially increase beta activity in the brain.

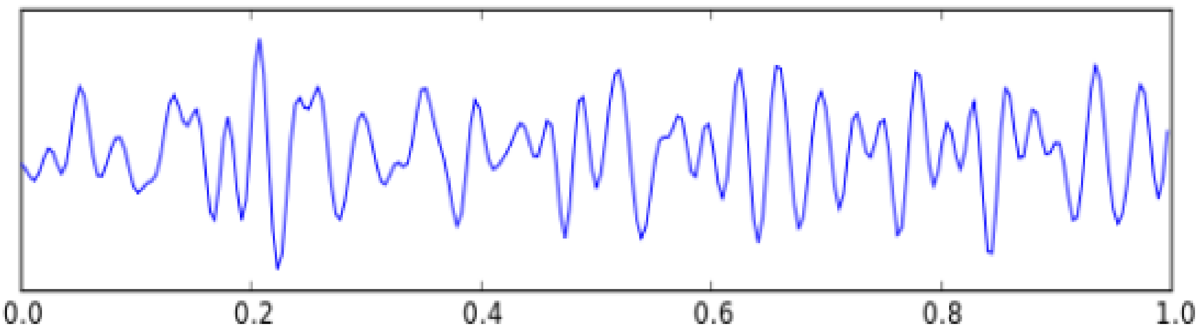

Fig. 8 – Beta Wave [9]

5. **Gamma Waves (30-60 Hz):** Gamma brainwaves are the fastest among all brainwave frequencies, and they indicate the simultaneous processing of information from multiple brain regions. During high-level cognitive processing, these waves become dominant and can be observed, such as combining different senses like sight and sound. Surprisingly, individuals with learning disabilities who have intellectual challenges tend to have lower gamma activity than those without such disabilities.

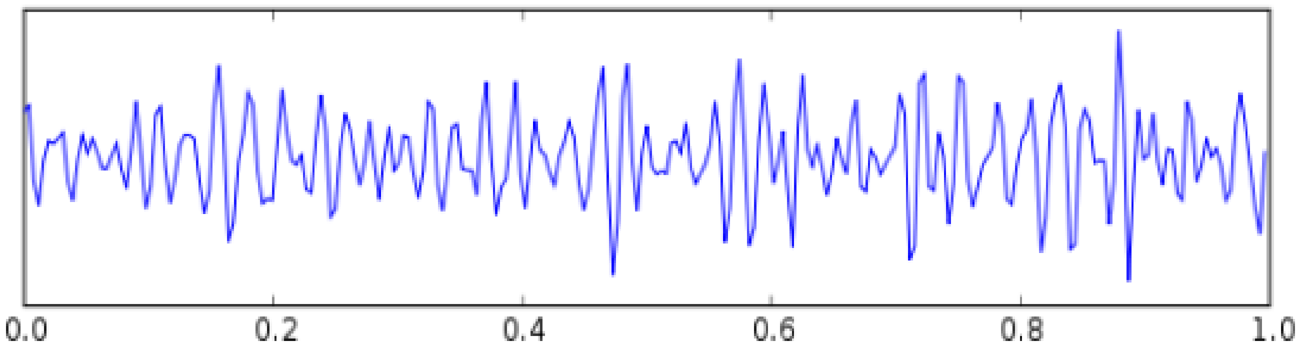

Fig. 9 – Gamma wave [9]

2.3 Artifacts:

Artifacts or noise are undesirable signals that contaminate the brain waves while measuring the EEG signals [46]. We can broadly categorize such artefacts into two categories.

a) Physiological Artifacts:

These various physiological artifacts are generated by the human body during different bodily activities such as eye movement and blinking, heart rate, head movement, jaw and tongue movement, and respiration. The electrooculogram (EOG) is a type of artifact caused by eye movement and blinking and is typically present below the frequency range of 4Hz. On the other hand, the electrocardiography (ECG) and electromyography (EMG) are generated by heart rate and muscle activity, respectively, and are present primarily above 30Hz. Out of all these physiological artifacts, EMG and EOG artifacts are typically regarded as being the most relevant for studies in human-computer interaction (HCI)[11][12][13][14].

b) Non-physiological artefacts:

Such artefacts originate from outside the human body. The primary sources of non-physiological artefacts are 50/ 60 Hz power line interference, variation in electrode impedance, dirt, and wire quality [45].

2.4 Features Used:

a) Power Spectral Density:

To accurately determine an EEG signal's Power Spectral Density (PSD), it is imperative to first convert the signal from the time domain to the frequency domain. Achieving this can be done using the Fast Fourier Transform (FFT) algorithm, a mathematical technique that enables the analysis of the signal in terms of its frequency content. Once the signal transforms into the frequency domain, it becomes possible to analyze its spectral characteristics in greater detail.

Power Spectral Density (PSD) is a critical measure used in signal processing to evaluate the frequency characteristics of a signal. It is a mathematical representation that allows for examining power distribution across various frequencies within a signal. One typically applies the Fourier Transform to the signal to obtain the Power Spectral Density (PSD), converting it from the time domain to the frequency domain and then calculating the squared magnitude of the resulting complex values. By squaring the magnitude, PSD captures the signal's power or energy content of each frequency component.

$$PSD(f) = |X(f)|^2$$

Where:

- PSD(f) represents the Power Spectral Density at frequency f
- X(f) denotes the Fourier Transform of the signal at frequency f

To calculate the PSD, Welch's method [24] is commonly used. This method involves dividing the signal into overlapping segments, computing a periodogram estimate for each segment, and then averaging the periodograms together to obtain an estimate of the PSD. This technique is beneficial for signals with non-stationary characteristics, as it allows for a more accurate and reliable estimate of the PSD.

PSD analysis provides crucial insights into a signal's frequency composition and underlying dynamics. It helps identify the presence and strength of different frequency components, such as delta, theta, alpha, beta, and gamma waves in EEG signals. PSD analysis can aid in developing effective treatment procedures and therapies by providing detailed information about a signal's frequency components. It can identify abnormalities in brain activity, characterize sleep patterns, investigate stimuli's effects on neural activity, and assess different physiological or pathological conditions' impact on signal properties.

The PSD signals undergo preprocessing to eliminate unwanted noise or components. Windowing functions, such as the Hamming, Hanning, and Blackman windows, reduce spectral leakage and distortions in the signal's frequency content. The Discrete Fourier Transform (DFT) of the windowed signal is then calculated using the Fast Fourier Transform (FFT) algorithm. The DFT values' squared magnitude for each frequency bin obtained from the FFT is calculated, and the PSD values are normalized by dividing them by either the total power of the signal or the total number of samples used in the calculation. The DFT or FFT algorithms are used in practical applications to efficiently compute the PSD of discrete-time signals.

It is important to note that PSD is a frequency domain feature, meaning that although it can identify every frequency present in the signal, we cannot pinpoint their specific locations. This limitation is because the frequency domain representation of a signal does not provide any information about the temporal characteristics of the signal. Therefore, it is essential to consider both time and frequency domain analysis techniques when analyzing EEG signals.

b) Wavelet Entropy:

Wavelet Entropy is a highly technical term commonly used in signal processing. This term refers to a feature calculated using the Discrete Wavelet Transform (DWT) method in the time-frequency domain. The DWT is a widely accepted signal processing technique which isolates distinct signal frequencies at various levels. In this study, researchers employed the Daubechies4 (db4) wavelet as the mother wavelet for the analysis.

The wavelet entropy originates from the wavelet transform, which decomposes a signal into various frequency components at multiple scales. It provides valuable information about the energy distribution or information across these scales, which can help analyze and understand signals.

We typically need to preprocess the signal by removing noise or unwanted components to calculate the wavelet entropy. Next, one applies the wavelet transform to decompose the signal into different scales or resolutions. Typically, researchers perform this using a specific wavelet basis, such as Daubechies, Haar, or Symlet. Once we have decomposed the signal, we calculate the normalized coefficients at each scale or resolution. Normalizing the coefficients ensures that their sum is equal to 1, which allows for a consistent entropy calculation across different scales. Finally, we calculate the wavelet entropy by taking the squared normalized coefficient ($D_i^{\wedge 2}$) at each scale (where D_i denotes the normalized coefficient at the ith scale or resolution), multiplying it by the logarithm of the squared coefficient, and then summing up these values for all scales. The resulting wavelet entropy value measures the signal's complexity or irregularity across different scales, with higher entropy values indicating higher complexity or irregularity and lower entropy values indicating more regular or predictable patterns.

Overall, wavelet entropy is a powerful tool in signal processing that can provide a wealth of information about the underlying structure of complex signals. The process involved in calculating wavelet entropy begins with the decomposition of the signal into multilevel wavelet coefficients. These coefficients are then used to calculate the wavelet energy by dividing the mother signal into two parts: cAn and cDn. The approximation coefficient, cAn, represents the low frequency at the nth level of the signal. In contrast, the coefficient at the nth level of the signal denotes the high frequencies, denoted by cDn. Analyzing these coefficients provides valuable insights into the signal's properties.

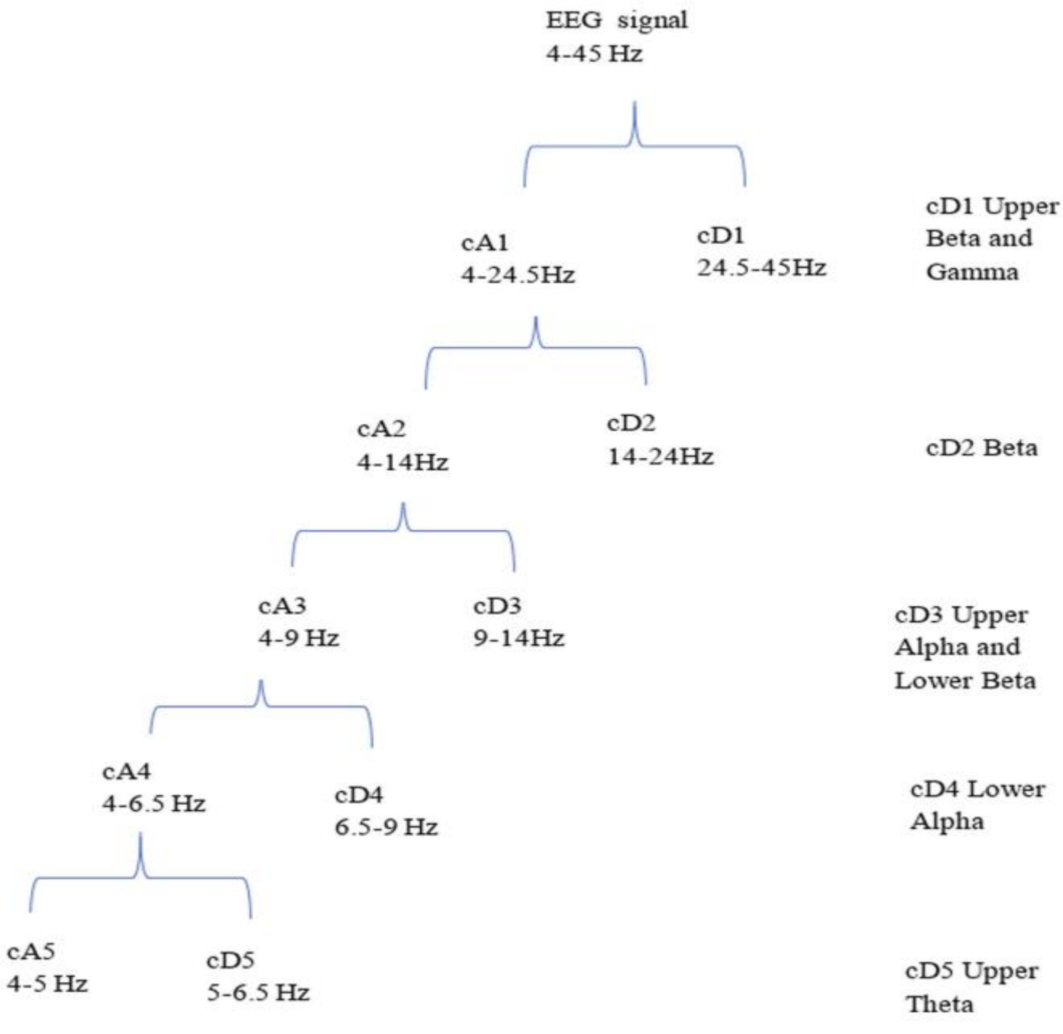

Fig.10 - Wavelet decomposition of different bands [31]

2.5 List of features used in previous work done on the AMIGOS dataset [10]

List of abbreviations used	
TFD	The Fractal Dimension
HA	Hjorth Activity
HM	Hjorth Mobility
HC	Hjorth Complexity
RMS	Root Mean Square
PSD	Power Spectral Density
DE	Differential Entropy
IBI	Inter beat intervals
AMP	Amplitude
RT	Rise Time
ST	Statistical features (mean, standard deviation)
DWT	Discrete Wavelet Transform
SWT	Stationary Wavelet Transform
TF	Multivariate time-frequency image

Table 1: List of abbreviations used [9][37][38][39]

Researchers have used various features for emotion recognition using EEG signals. Table 1 summarises the list of features used in previous work done on the AMIGOS dataset [10]. The description of all the features are as follows:

The Fractal Dimension: The fractal dimension measures the complexity or irregularity of a signal. In the context of EEG signals, the fractal dimension quantifies the self-similarity and intricacy of the recorded brain activity. It provides insights into the complexity of neural dynamics and can be used to study the organization and structure of the EEG signal. Topic A. et al. [9] have used this feature.

Hjorth Activity: Hjorth activity is a feature that characterizes the signal's overall amplitude or activity level. It measures the overall power or energy of the EEG signal and can provide information about the intensity of neural activity. To compute Hjorth Activity, one calculates the variance of the signal. Hjorth Activity features have been used by Shukla et al. [39] and Topic A. et al. [9]

Hjorth Mobility: Hjorth mobility [39][9] is a feature that describes the signal's rate of change or mobility. It quantifies the signal's variability or movement over time. In the context of EEG signals, Hjorth mobility can capture changes in brain activity dynamics and indicate how the signal transitions between different states or frequencies.

Hjorth Complexity: Hjorth complexity is a feature that characterizes the signal's irregularity or complexity. It combines measures of activity and mobility to provide insights into the signal's complexity dynamics. Hjorth complexity [39][9] can capture the non-linear and dynamic aspects of the EEG signal, reflecting the underlying neural processes.

Root Mean Square (RMS): Root mean square measures a signal's average power or amplitude. In the context of EEG signals, RMS calculates the square root of the average of the squared amplitudes of the signal. It provides information about the overall magnitude or intensity of the EEG signal. Topic A. et al. [9] have used this feature.

Power Spectral Density (PSD): Power spectral density is a frequency domain feature that characterizes power distribution across different frequencies in the EEG signal. It provides insights into the relative contribution of different frequency bands to the overall signal. PSD analysis helps identify the dominant frequency components and their respective power levels, which can be associated with different brain activities or states. Miranda-Correa et al. [10], . Topic A. et al. [9] and S. Siddharth et al[16 have used this feature

Differential Entropy: Differential entropy measures a signal's uncertainty or randomness. In the context of EEG signals, differential entropy quantifies the complexity and information content of the signal. It provides insights into the diversity and variability of the EEG signal, which can be related to different cognitive or emotional states. Topic A. et al. [9] have used this feature.

Interbeat Intervals (IBI): Interbeat intervals measure the time intervals between consecutive heartbeats. In the context of EEG signals, IBIs can provide insights into heart rate variability, which is associated with physiological and emotional states. Variations in IBIs can indicate changes in arousal levels or emotional responses. Santamaria-Granados et al. [37], L. Santamaria-Granados et al.[15] and S. Siddharth et al[16] have used this feature.

Amplitude: Amplitude refers to the magnitude or strength of the EEG signal. It measures the voltage or intensity of the electrical activity recorded from the brain. Amplitude can reflect the strength of neural activity and indicate the intensity or amplitude of underlying brain processes. Santamaria-Granados et al. [37], Shukla et al. [39] and L. Santamaria-Granados et al.[15] have used this feature.

Rise Time: Rise time measures the time it takes for a signal to increase from a certain threshold level to its peak value. In EEG signals, rise time can provide information about the speed or rate of brain activity change. It can be relevant for analyzing transient responses or rapid changes in neural activity associated with specific events or stimuli. Shukla et al. [39] and L. Santamaria-Granados et al.[15] have used this feature.

Statistical Features (Mean, Standard Deviation): Statistical features such as mean and standard deviation provide information about the central tendency and variability of the EEG signal. Mean represents the average value of the signal, while standard deviation quantifies the dispersion or spread of the data points. These statistical features can capture the overall level and variability of the EEG signal, which can be relevant for differentiating between different physiological or emotional states. Shukla et al. [39] have used this feature.

Discrete Wavelet Transform (DWT): The discrete wavelet transform is a mathematical technique that decomposes a signal into different frequency components at different scales. In the context of EEG signals, DWT can capture both localized and global frequency content. It provides a time-frequency representation that reveals how different frequency components contribute to the signal at different time points. Shukla et al. [39] have used this feature.

Stationary Wavelet Transform (SWT): The stationary wavelet transform is a variation of the wavelet transform that preserves the signal's temporal information. It decomposes the signal into different frequency components while retaining the original signal length. SWT is useful for analyzing non-stationary signals like EEG, where the signal properties may change over time. Shukla et al. [39] have used this feature.

Multivariate Time-Frequency Image: A multivariate time-frequency image represents the time-varying frequency content of multichannel signals such as EEG. It combines the time and frequency domains to visually represent how different frequency components evolve over time across multiple channels. This representation can facilitate the analysis of complex patterns and interactions between different brain regions. Padhmashree, V., & Bhattacharyya, A [38] have used this feature.

Chapter 3

Related Work

Researchers have used various approaches to identify emotion through physiological signals. Topic A. et al. [9] have proposed using EEG signal properties to create holographic and topographic feature maps, which they term HOLO-FM and TOPO-FM, respectively. In this study, researchers used CNN and Support Vector Machine for classification. Using HOLO-FM, the accuracy for Valence was 80.63%, and for arousal, it was 85.75%. Using TOPO-FM, the accuracy for Valence was 87.39%, while it was 90.54% for arousal on the AMIGOS dataset.

The AMIGOS dataset [10] was analyzed by Santamaria-Granados et al. [37] using DCNN, which includes electrocardiogram and galvanic skin response physiological signals. They achieved higher accuracy in classifying emotional states by utilizing advanced classical machine learning techniques to extract signal properties in time, frequency, and nonlinear domains. The potential of this study is significant as it can enhance our comprehension of how physiological signals offer the potential for identifying and categorizing emotional states.

A varied electrode placement profile for a different region of the brain also aids in improving accuracy. Corresponding to different brain areas, Padhmashree, V., & Bhattacharyya, A [38] have used four alternative electrode placement configurations. This study has utilized SVM, KNN, Naïve Bayes, Ensemble Random Forest, Ensemble Boosted Tree, and ResNet-18 techniques. The accuracy score is 96.68% for arousal, 96.34% for Valence, and 97.45% for Dominance on the AMIGOS dataset [10].

Shukla et al. [39] have used Electrodermal Activity (EDA) from the AMIGOS dataset for Emotion Recognition. The accuracy score for arousal is 85.75, and for Valence, 83.9%. Time-frequency and frequency domain features have found application in this context.

Menezes et al. [8] used statistical features, power spectral density, and higher order crossing (HOC) features with Support Vector Machine (SVM) and Random Forest, which gave an accuracy of 74% for Valence and 88% for arousal on the

DEAP dataset. Several researchers have utilized CNN to extract features automatically.

Features / Paper Title	TFD	HA	HM	HC	RMS	PSD	DE	IBI	AMP	RT	ST	DWT	SWT	TF
Topic A. et al. [9]	✓	✓	✓	✓	✓	✓	✓							
Miranda-Correa et al. [10]						✓								
Santamaria-Granados et al. [37]								✓	✓	✓				
Shukla et al. [39]		✓	✓	✓					✓	✓	✓	✓	✓	
Padhmashree, V., & Bhattacharyya, A [38]														✓
L. Santamaria-Granados et al.[15]								✓	✓					
S. Siddharth et al[16]						✓		✓						

Table 2: List of features used in previous work done on the AMIGOS dataset [10]

The identification of emotions from unprocessed EEG signals was carried out by Alhagry et al. [21] using a deep learning approach. They used a long short-term memory (LSTM) approach to extract features from the EEG signals and then passed them through a dense layer. The characteristics were subsequently classified into Valence, liking and low or high arousal. They tested their method on the DEAP dataset and achieved an average accuracy of 85.45%, 85.65%, and 87.99% for the Valence, Arousal, and liking categories, respectively.

Reference	Classification Method	Accuracy(%)	Year
Topic A. et a [9]	CNN+SVM	Using HOLO-FM: Valence: 80.63 Arousal: 85.75 Using TOPO-FM: Valence: 87.39 Arousal: 90.54	2021
Miranda-Correa et al. [10]	SVM	Valence: 56.4 Arousal: 57.7	2018
Santamaria-Granados et al. [37]	CNN	Arousal: 76 Valence: 75	2018
Shukla et al. [39]	SVM Classifier + Radial Basis Function (RBF)	Arousal: 85.75 Valence: 83.9	2019
Padhmashree, V., & Bhattacharyya, A [38]	SVM, KNN, Naive Bayes Ensemble Random Forest Ensemble Boosted Tree, ResNet-18	Arousal: 96.68 Valence: 96.34 Dominance: 97.45	2022
L. Santamaria-Granados et al.[15]	DCNN	Valence: 75 Arousal: 76	2018
S. Siddharth et al[16]	RGB heat-map	Arousal: 79.13 Valence: 83.02	2019
Garg, Shruti, et al. [31]	CNN+SVM	Arousal: 96.63 Valence: 95.87 Dominance: 96.30	2021

Table 3 Previous models used and accuracies on the AMIGOS dataset [10]

In recognizing emotions from physiological signals and facial expressions, various researchers have developed several models. The first model listed is from Topic A. et al. [9] and utilizes a combination of Convolutional Neural Network (CNN) and Support Vector Machine (SVM) algorithms. The model achieves an accuracy of 80.63% for Valence and 85.75% for arousal using HOLO-FM and 87.39% for Valence and 90.54% for arousal using TOPO-FM. Table 3 summarises these models and their respective accuracies on the AMIGOS dataset.

Miranda-Correa et al. [10] use an SVM algorithm and achieve an accuracy of 56.4% for Valence and 57.7% for arousal. The third model, from Santamaria-Granados et al. [37], utilizes a CNN algorithm and achieves an accuracy of 75% for Valence and 76% for arousal. Shukla et al. [39] use an SVM classifier with Radial Basis Function (RBF) and achieve an accuracy of 83.9% for Valence and 85.75% for arousal.

Padhmashree and Bhattacharyya [38] use several algorithms, including SVM, K-Nearest Neighbor (KNN), Naive Bayes, Random Forest Ensemble, Boosted Tree, and ResNet-18. The model achieves an accuracy of 96.34% for Valence, 96.68% for arousal, and 97.45% for Dominance. L. Santamaria-Granados et al. [15] use a Deep CNN algorithm and achieve an accuracy of 75% for Valence and 76% for arousal. S. Siddharth et al. [16] use an RGB heat map and achieve an accuracy of 83.02% for Valence and 79.13% for arousal.

The final model, from Garg, Shruti et al. [31], utilizes a combination of CNN and SVM algorithms and achieves an accuracy of 95.87% for Valence, 96.63% for arousal, and 96.30% for Dominance. It is worth noting that the models range from traditional machine learning algorithms to deep learning algorithms, with the most accurate model being a combination of several algorithms.

Chapter 4: The Emotion Recognition Framework

4.1 An Approach for Recognizing Emotions Based on EEG Signals Using the AMIGOS Dataset

Two distinct models of emotions are currently recognized: dimensional and discrete. The combination of Valence and arousal characterizes the dimensional model to gauge the lasting nature of an emotional state. On the other hand, the discrete model classifies emotions into a specific number of distinct emotions, which contrasts with the dimensional model since it typically describes the lasting nature of an emotional state by combining Valence and arousal.

Valence quantifies the degree of pleasantness or unpleasantness linked to an emotion., and it has a range that spans from unpleasant to pleasant. In contrast, arousal reflects the intensity of emotional experience, ranging from inactive, such as when one is bored, to active, such as when one is excited, along a continuous sequence. These emotion categories, Valence, arousal, and Dominance, are defined in more detail in reference [30].

Plutchik [9] proposes one of the most influential classifications of emotions, suggesting the existence of eight fundamental emotions.: sadness, fear, anger, surprise, anticipation, disgust, joy, and acceptance. These basic emotions can combine to form all other emotions; for instance, disappointment results from a mixture of surprise and sadness.

Furthermore, one can categorize emotions as positive, negative, or neutral. Positive emotions such as care and happiness are crucial for growth, development, and survival. In contrast, negative emotions, such as anger, sadness, fear, and disgust, typically operate automatically and briefly. The categorization of neutral emotions lacks a scientific foundation and is instead a theoretical or prescriptive model of negotiation, as per reference [10].

Finally, Figure 1 displays another emotion classification based on Valence, ranging from negative to positive, and arousal, ranging from high to low. For instance, the emotion of depression belongs to the category of negative Valence and low arousal. Understanding the different classifications of emotions can help individuals recognize and manage their emotional states, leading to better emotional well-being.

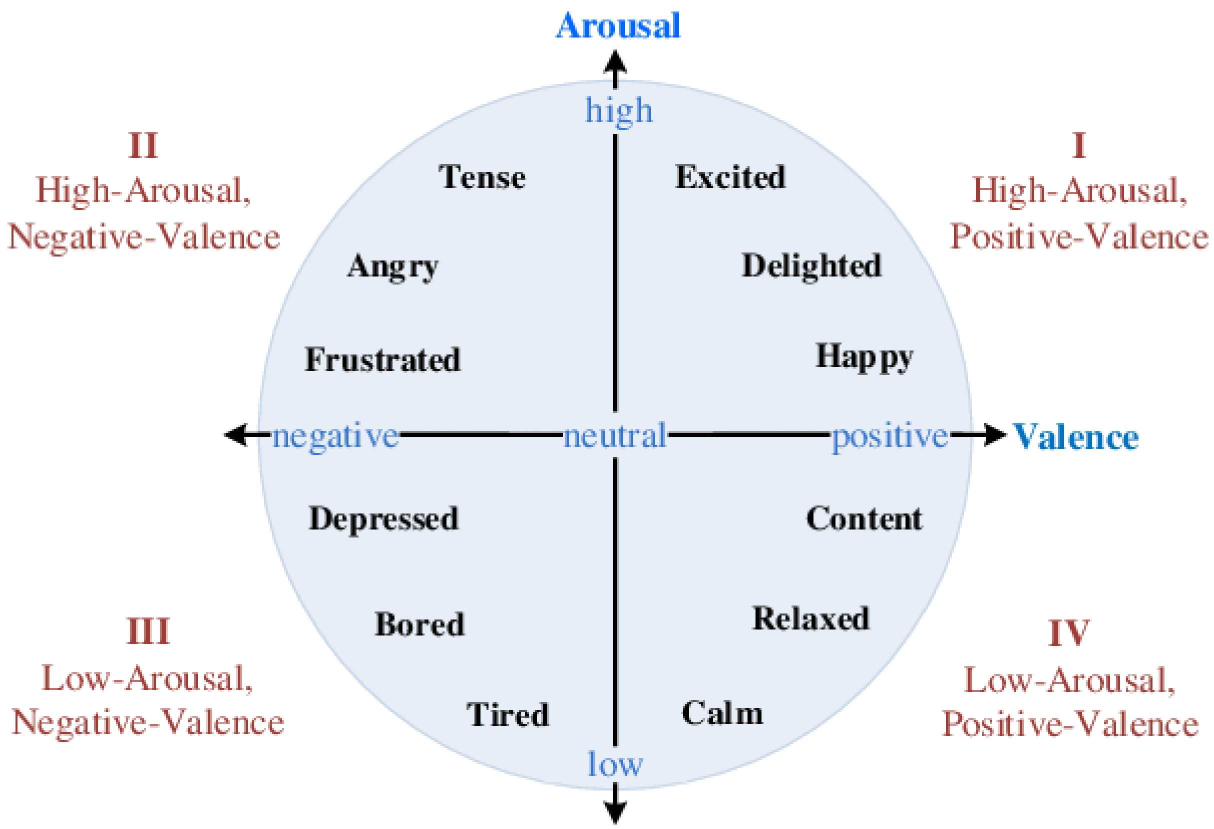

Fig 11 Human emotions based on arousal and valence model [21]

Recently, researchers have shown keen interest in utilizing the electroencephalogram (EEG) as the primary method for recognizing emotions from physiological signals. EEG is popular due to its high temporal resolution, safety, and ease of use. [43]. However, it is worth noting that EEG has low spatial resolution and is susceptible to artefacts that may arise from eye blinking, eye movements, muscle movements, heartbeats, and power line interference [12]. Despite these limitations, EEG electrodes can capture electrical stimulation on the skin surface caused by many active neurons in the brain, making it a particularly effective physiological stimulus. In addition to the EEG recordings, the dataset includes galvanic skin response (GSR) recordings, electromyography (EMG), blood pressure, eye activity, and temperature pulse.

The circumplex model of emotion, formulated by James Russell and Lisa Feldmann Barrett [32], proposes that arousal and Valence can be used as axes to represent emotions on a two-dimensional plane. According to this model, the vertical axis represents arousal, while the horizontal axis represents Valence. The circle's centre depicts a neutral valence and a moderate level of arousal. The valence dimension describes the degree of pleasantness or unpleasantness of a feeling. A positive valence indicates a happy emotion, whereas a negative valence suggests unhappiness. The arousal dimension illustrates how stimulating or calming an emotion is. A high-arousal emotion occupies the positive end of the vertical axis, while a low-arousal emotion occupies the opposing end.

EEG is a biological signal that measures the brain's electrical activity by placing electrodes on the scalp [13]. The popularity of the EEG is growing due to its accessibility, portability, affordability, and user-friendliness. The analysis of EEG signals is an interdisciplinary area that involves neuroscience, medical science, computer science, biomedical engineering, and health [14]. EEG-based emotion recognition finds applications in various fields, such as healthcare, e-learning, and entertainment. EEG has diverse uses, including psychology, online gaming, instant messaging, and assisting therapy.

4.2 Dataset

AMIGOS is the Affect, Personality, and mood research dataset in individual and group settings. It contains 14 channels of EEG signal, two channels of ECG signal, one channel of galvanic skin response (GSR), and frontal video (RGB) from two experiments [10]. This dataset aims to support research in affective computing, personality analysis, and mood recognition, specifically in understanding the dynamics of emotions and personality in group settings.

The dataset consists of data from 40 participants, including videos, audio, physiological signals, and self-reported data. The researchers collected the data during individual and group activities, including public speaking, group discussions, and socializing.

The authors also describe the annotation process of the dataset, which involved the labelling of the emotional states, personality traits, and moods of the participants. They explain using established psychological questionnaires and self-reported measures to gather this information. Additionally, they provide statistical analysis of the dataset, including the distribution of the annotations and the inter-annotator agreement.

The researchers collected the AMIGOS dataset [10] using a variety of sensors, including video cameras, microphones, electrocardiography (ECG) sensors, and electrodermal activity (EDA) sensors. The researchers instructed the participants to perform public speaking, group discussions, and socializing tasks. The researchers designed the tasks to induce different emotional states and capture affective behaviour dynamics in group settings.

During the annotation process, careful attention was paid to accurately label the participants' emotional states, personality traits, and moods. The Geneva Emotion Wheel was used to categorize the participants' emotional states into eight primary emotions and their corresponding subtypes. Additionally, researchers utilized the Big Five Personality Traits to accurately label the participants' personality traits: openness, conscientiousness, extraversion, agreeableness, and neuroticism. Furthermore, the Positive and Negative Affect Schedule (PANAS) was employed to evaluate the participants' moods, measuring their positive and negative affective states precisely. These methods obtained a detailed and comprehensive understanding of the participants' emotional, personality, and mood states.

The authors provide baseline results for three tasks: emotion recognition, personality trait prediction, and mood recognition. They compare three approaches for emotion recognition: audio-only, video-only, and multimodal. They use a support vector machine (SVM) classifier to classify the participants' emotional states. The results show that the multimodal approach outperforms the other two approaches, indicating that combining different modalities can lead to better emotion recognition. The authors note that the audio modality is more informative for recognizing negative emotions, while the video modality is more informative for recognizing positive emotions.

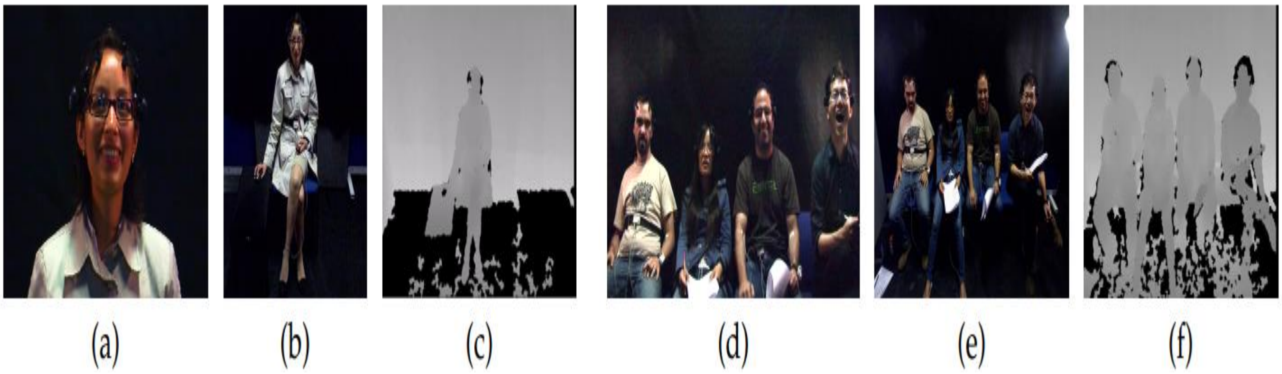

Fig. 12 [10] Participants in experiment conditions during the short videos experiment recorded in (a) Frontal HD video, (b) full body RGB video via Kinect, (c) full body depth video via Kinect, and a group of 4 participants during the long videos experiment recorded in (d) frontal HD video, (e) full body RGB video via Kinect and (f) full body depth video via Kinect.

For personality trait prediction, the authors use machine learning models to predict the Big Five Personality Traits of the participants. They compare three models: SVM, random forest (RF), and gradient boosting (GB). The results show that the RF and GB models outperform the SVM model, suggesting that ensemble models may be more suitable for predicting personality traits. The authors note that conscientiousness is the easiest to predict, while openness is the most challenging.

The authors use a deep learning model to classify the participants' positive and negative affective states for mood recognition. They use a convolutional neural network (CNN) to process the videos and a long short-term memory (LSTM) network to process the physiological signals. The results show that the deep learning model outperforms a baseline SVM model, indicating that deep learning models can be effective for mood recognition. The authors note that the physiological signals are more informative for recognizing negative affective states, while the videos are more informative for recognizing positive affective states.

Chapter 5

5.1 Methodology:

The Proposed Workflow is summarised as follows.

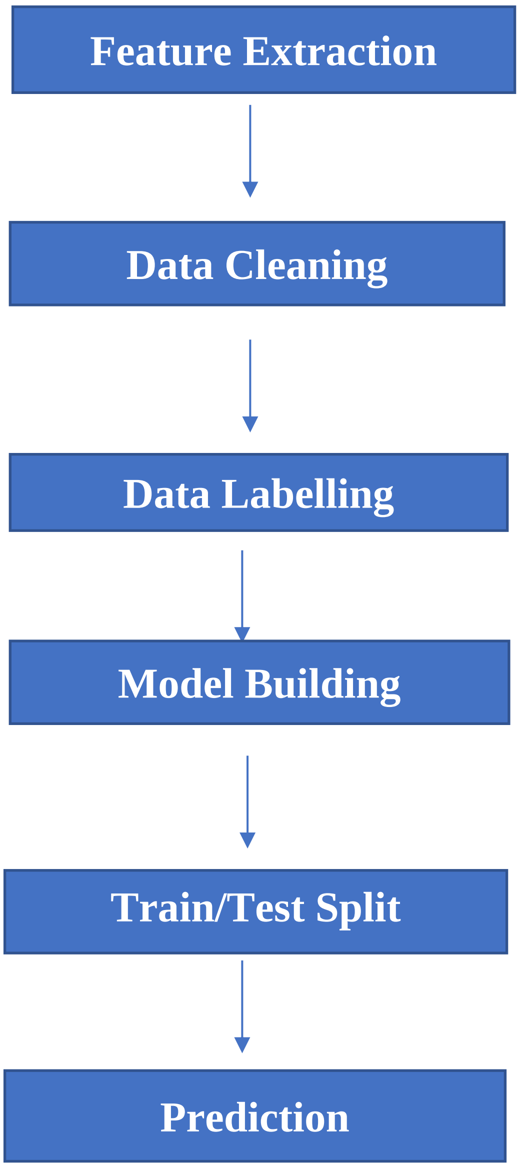

Fig. 13 Proposed Workflow

5.2 Feature Extraction:

In this study, researchers utilize two distinct features to analyze EEG signals: Power Spectral Density (PSD) in the frequency domain and Wavelet Entropy in the time-frequency domain. The PSD of all the frequency bands, including gamma, beta, alpha, theta, and delta, were extracted during the study.

To ensure the efficient and practical construction of the Deep Learning (DL) models, researchers employed an Overlapping Sliding Window (OSW) technique to amplify the emotion samples. The EEG signals generated from various experiments were partitioned into windows of size 512 with a shift of 32, as depicted in Figure 2. Any signals not covered by the 512 windows were either trimmed or disregarded for computation purposes.

The methodology adopted for this study facilitated the decomposition of the EEG signals into equal-length samples, which were easier to process. By employing OSW to partition the signals, extracting features from the decomposed samples was made more accessible and facilitated the processing of the signals. The careful attention given during the annotation process ensured that the data was effectively analyzed and accurate results were obtained.

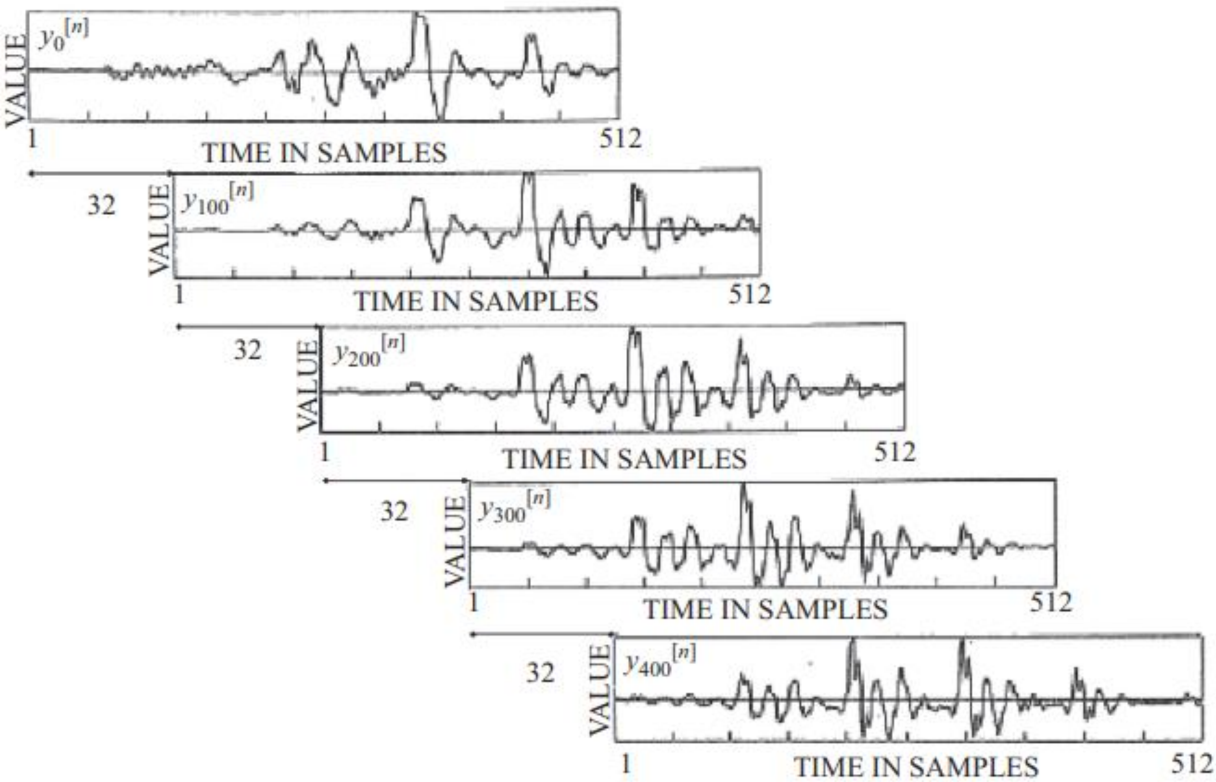

Fig. 14 Overlapping window signal decomposition [31]

Following the decomposition of signals into equally-sized samples using overlapping windows, discrete Fourier transform (DFT) [40] and discrete wavelet transform (DWT) [41] were employed to extract NBP and NWE features.

5.3 Data Cleaning and Preprocessing:

This work required adding values for different users from the dataset. Mishandling missing values can result in biased model results, further reducing accuracy. Since sufficient data points were available, any missing values were removed from the dataset before training the model.

5.4 Label Preprocessing:

In the AMIGOS dataset, we have scores for the labels from 1-9[10]. The threshold is at 4.5. If the score exceeds 4.5, it has been taken as high arousal/valence/dominance and labelled as 1. Similarly, if the score is less than or equal to 4.5, it has been labelled as 0. Table 3 shows the Coding of Valence, Arousal, and Dominance from 1-9 to 0-1

High Valence =1	Low Valence = 0	High Arousal =1	Low Arousal = 0	High Dominance =1	Low Dominance = 0
>4.5	≤4.5	>4.5	≤4.5	>4.5	≤4.5

Table 4 Coding of Valence, Arousal, and Dominance from 1-9 to 0-1 on the AMIGOS dataset [10]

5.5 Model Building:

The application of the supervised learning algorithm Support Vector Machine (SVM) has proven to be effective in both classification and regression problems [27]. This approach utilizes the Radial Basis Function (RBF) Kernel function to determine the classifier line. A regularisation parameter C regulates the trade-off between achieving a low training error and a low testing error. This parameter determines the classifier's generalization ability to new data [44]. In this study, researchers employed a regularization parameter of 100. The existing training data is mapped onto a higher-dimensional space using a nonlinear method. The algorithm then searches for an optimal separation hyperplane to differentiate between data points within this new dimension. Hyperplanes are used to differentiate between different data points.

On the other hand, the basic ANN model uses a training batch size of 32 and trains for 150 epochs with an initialized learning rate of 0.001. The model includes five layers: one input layer, three hidden layers, and one output layer. It is a multi-layer, fully connected neural network that utilizes an Exponential Linear Unit (ELU) as the activation function for each input and hidden layer. Finally, the output layer has one output with the sigmoid activation function.

Moreover, the CNN-based model training utilized a batch size of 32, a learning rate 0.001, and the "Adam" optimizer. The loss function used was Binary Cross Entropy. The model underwent training to differentiate between data points using these parameters accurately.

Description of each layer:

Layer 1:

Type: Input Layer
Layer Input size: (14,64)

Layer 2:

Type: Hidden Layer
Layer Input size: (64,128)

Layer 3:

Type: Hidden Layer
Layer Input size: (128,256)

Layer 4:

Type: Hidden Layer
Layer Input size: (256,512)

Layer 5:

Type: Output Layer
Layer Input size: (512,1)

5.6 Division of Dataset into train/test:

The dataset divided into 80% for training and 20% for testing.

Chapter 6 – Results

Model	**Features**	**Methodology**	**Accuracy**
SVM	PSD and DWT without feature fusion	Non-Overlapping Sliding Window	Arousal - 76.64 Valence – 67.41 Dominance- 64.8
ANN	Feature fusion of PSD and DWT	Non-Overlapping Sliding Window	Arousal – 83.22 Valence – 77.61 Dominance-80.12
SVM	Feature fusion of PSD and DWT	Non-Overlapping Sliding Window	Arousal – 80.82 Valence – 74.33 Dominance- 79.2
CNN	Feature fusion of PSD and DWT	Non-Overlapping Sliding Window	Arousal – 81.34 Valence – 75.48 Dominance- 78.67
ANN	Feature fusion of PSD and DWT	Overlapping Sliding Window	Arousal – 90.79 Valence – 91.39 Dominance- 92.18

Table 5 Accuracy obtained after applying ML/DL classifiers in this study on the AMIGOS dataset [10]

The evaluation aimed to assess three emotions: arousal, valence, and dominance. This assessment used two signal processing techniques, power spectral density (PSD) and discrete wavelet transform (DWT), as features. The researchers selected these techniques due to their capability to capture different signal aspects. PSD provides information about the power distribution across different frequencies, while DWT operates in the time-frequency domain, allowing for analysis of the signal's frequency and temporal characteristics. Two sliding window techniques, non-overlapping and overlapping, were employed to evaluate the models. The table illustrates the outcomes of several models tested for emotion recognition.

The first model tested was a support vector machine (SVM) that used PSD and DWT features without feature fusion. The model utilized the non-overlapping sliding window technique. The accuracy results of the SVM model were 76.64% for arousal, 67.41% for valence, and 64.8% for dominance. Feature fusion of PSD and DWT was then incorporated into the SVM model using the non-overlapping sliding window technique, which significantly improved accuracy. The SVM model with feature fusion achieved 80.82% for arousal, 74.33% for valence, and 79.2% for dominance.

The second model tested was an artificial neural network (ANN) that used PSD and DWT features with feature fusion. The model utilized the non-overlapping sliding window technique. The ANN model outperformed the SVM model with an accuracy of 83.22% for arousal, 77.61% for valence, and 80.12% for dominance. The third model tested was a convolutional neural network (CNN) that used PSD and DWT features with feature fusion. The model employed the non-overlapping sliding window technique. The CNN model achieved an accuracy of 81.34% for arousal, 75.48% for valence, and 78.67% for dominance.

Further testing was conducted on the ANN model with feature fusion of PSD and DWT using the overlapping sliding window technique. This model achieved the highest accuracy for all three emotions. The accuracy results for the overlapping sliding window technique were 90.79% for arousal, 91.39% for valence, and 92.18% for dominance.

In conclusion, the table overviews the models tested for emotion recognition. The results indicate that combining signal processing techniques with neural network models can potentiate emotion recognition in various applications, such as affective computing, human-robot interaction, and healthcare. Using the overlapping sliding window technique, the ANN model, with feature fusion of PSD and DWT, demonstrated the highest accuracy for all three emotions. These findings suggest that this model may benefit applications requiring reliable and accurate emotion recognition.

Reference	Classification Method	Accuracy(%)	Year
Topic A. et a [9]	CNN+SVM	Using HOLO-FM: Valence: 80.63 Arousal: 85.75 Using TOPO-FM: Valence: 87.39 Arousal: 90.54	2021
Miranda-Correa et al. [10]	SVM	Valence: 56.4 Arousal: 57.7	2018
Santamaria-Granados et al. [37]	CNN	Arousal: 76 Valence: 75	2018
Shukla et al. [39]	SVM Classifier and Radial Basis Function (RBF)	Arousal: 85.75 Valence: 83.9	2019
Padhmashree, V., & Bhattacharyya, A [38]	SVM, KNN, Naive Bayes Ensemble Random Forest Ensemble Boosted Tree, ResNet-18	Arousal: 96.68 Valence: 96.34 Dominance: 97.45	2022
L. Santamaria-Granados et al.[15]	DCNN	Valence: 75 Arousal: 76	2018
S. Siddharth et al[16]	RGB heat-map	Arousal: 79.13 Valence: 83.02	2019
Garg, Shruti, et al. [31]	CNN and SVM used	Arousal: 96.63 Valence: 95.87 Dominance: 96.30	2021
This Work	SVM, ANN and CNN	Arousal: 90.79 Valence: 91.39 Dominance: 92.18	

Table 5 Comparison table of proposed work with existing work on the AMIGOS dataset [10]

The table summarizes the classification accuracy of different methods used in affective computing. The studies focus on classifying arousal, Valence, and Dominance, three essential dimensions of emotions.

Topic A. et al. [9] used a CNN+SVM method with HOLO-FM and TOPO-FM features. The accuracy for arousal was 85.75% using HOLO-FM and 90.54%

using TOPO-FM. The accuracy for Valence was 80.63% using HOLO-FM and 87.39% using TOPO-FM.

Miranda-Correa et al. [10] used an SVM method for classification. The accuracy for arousal was 57.7%, and the accuracy for Valence was 56.4%. Santamaria-Granados et al. [37] used a CNN method for classification. The accuracy for arousal was 76%, and the accuracy for Valence was 75%. Shukla et al. [39] used an SVM Classifier with Radial Basis Function (RBF). The accuracy for arousal was 85.75%, and the accuracy for Valence was 83.9%. Padhmashree, V., & Bhattacharyya, A [38] used different classification methods, including SVM, KNN, Naive Bayes, Ensemble Random Forest, Ensemble Boosted Tree, and ResNet-18. The accuracy for arousal was 96.68%, the accuracy for Valence was 96.34%, and the accuracy for Dominance was 97.45%.

L. Santamaria-Granados et al. [15] used a DCNN method for classification. The accuracy for arousal was 76%, and the accuracy for Valence was 75%. S. Siddharth et al. [16] used an RGB heat-map method for classification. The accuracy for arousal was 79.13%, and the accuracy for Valence was 83.02%. Garg, Shruti, et al. [31] used a CNN+SVM method for classification. The accuracy for arousal was 96.63%, Valence was 95.87%, and the accuracy for Dominance was 96.30%.

Lastly, the present work employed SVM, ANN, and CNN and obtained accuracy rates of 90.79% for arousal, 91.39% for Valence, and 92.18% for Dominance with a PSD and DWT feature fusion using the overlapping sliding window technique.

In conclusion, the table shows that the accuracy of the classification methods for affective computing varies widely depending on the methodology, classification algorithm, and features used. The SVM method used by Miranda-Correa et al. [10] had the lowest accuracy for arousal and Valence. In contrast, the Padhmashree, V., & Bhattacharyya, A [38] study had the highest accuracy for all three dimensions of emotion. More recent studies achieve higher accuracy, likely due to advanced deep learning techniques and access to larger datasets. The table highlights the importance of choosing an appropriate method and features for affective computing to classify emotions accurately.

Chapter 7 – Conclusion and Future Work

This work explored the use of EEG signals for emotion recognition. The study highlights the potential of EEG signals as a non-invasive and objective approach to measuring human emotions. By implementing various signal processing techniques and machine learning algorithms, we achieved promising results in the classification of emotions such as arousal, Valence, and Dominance. One can apply the models trained in this work in various domains, including mental health, human-computer interaction, and entertainment. However, there are still challenges to overcome, such as handling missing data and improving the interpretability of the models. However, there is still room for improvement in accuracy and robustness, and future work can further enhance the applicability of this technology. Overall, the study represents a step towards developing a reliable and practical system for emotion recognition using EEG signals.

The average accuracy for Non-Overlapping Sliding- window-based techniques is lesser than the Overlapping Sliding Window based technique. Using the CNN-based model with the Overlapping Sliding Window technique can lead to further improvements in accuracy. This work incorporates Support Vector Machine (SVM), Artificial Neural Network (ANN), Convolutional Neural Network (CNN) based Machine, and Deep Learning methods with Overlapping and Non-Overlapping Sliding Window techniques.

Several avenues for future work could build upon the findings of this project. The multimodal approach for emotion recognition has the potential to expand by incorporating multiple physiological signal modalities. One potential application of the models trained in this study is designing automatic video recommendation systems to enhance individuals' moods based on the obtained results. Investigating the possibility of real-time emotion recognition using EEG signals can have practical applications in areas such as mental health monitoring and human-computer interaction.

Conducting more extensive user studies to collect EEG data from a larger and more diverse population can improve the generalizability of the emotion recognition model.

Although physiological data is inherently anonymous, there may be instances in which the signals are combined with additional modalities (such as audio or video) to identify the user's emotional state in emotion detection systems. Anonymity is more challenging to maintain since a higher level of data privacy protection is necessary, which will be challenging using centralized machine learning techniques.

Obtaining data from various sources and training the Machine Learning model is challenging due to multiple businesses and nations' strict data privacy policies. Traditional Machine Learning model training demands the centralization of all data. Federated Learning is a potential machine learning model that provides unique answers to various centralized learning difficulties because it protects data privacy and allows for data collection from several sources. Future work may include using a federated learning model using Overlapping and Non-Overlapping Sliding Window based methods.

References:

[1] Li, L., Fan, Y., Tse, M., & Lin, K. Y. (2020). A review of applications in federated learning. *Computers & Industrial Engineering*, 106854.

[2] Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. (2021). A survey on federated learning. *Knowledge-Based Systems*, *216*, 106775.

[3] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, *37*(3), 50-60.

[4] Yang, Q., Liu, Y., Chen, T., & Tong, Y. (2019). Federated machine learning: Concept and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *10*(2), 1-19.

[5] McMahan, H. B. (2021). Advances and open problems in federated learning. *Foundations and Trends® in Machine Learning*, *14*(1).

[6] Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., ... & Seth, K. (2017, October). Practical secure aggregation for privacy-preserving machine learning. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security* (pp. 1175-1191).

[7] McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017, April). Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics* (pp. 1273-1282). PMLR.

[8] Menezes, M. L. R., Samara, A., Galway, L., Sant'Anna, A., Verikas, A., Alonso-Fernandez, F., ... & Bond, R. (2017). Towards emotion recognition for virtual environments: an evaluation of EEG features on benchmark dataset. *Personal and Ubiquitous Computing*, *21*(6), 1003-1013.

[9] Topic, Ante, and Mladen Russo. "Emotion recognition based on EEG feature maps through deep learning network." Engineering Science and Technology, an International Journal 24.6 (2021): 1442-1454.

[10] Miranda-Correa, J. A., Abadi, M. K., Sebe, N., & Patras, I. (2018). Amigos: A dataset for affect, personality and mood research on individuals and groups. IEEE Transactions on Affective Computing, 12(2), 479-493.

[11] Daniel Olguın Olguın, Frantz Bouchereau, Sergio Martınez, Adaptive Notch Filter for EEG Signals Based on the LMS Algorithm, with Variable Step-Size Parameter. Conference on Information Sciences and Systems, The Johns Hopkins University, March 16–18, 2005.

[12] D.V. Moretti , F, Babiloni, F. Carducci, F. Cincotti, E. Remondini, P.M. Rossini ,S. Salinari, C. Babiloni, Computerized processing of EEG–EOG–EMG artefacts for multicentric studies in EEG oscillations and event-related potentials. International Journal of Psychophysiology 47 (2003) 199–216.

[13] Mehrdad Fatourechi, Ali Bashashati, Rabab K. Ward, Gary E. Birch, Invited review, EMG and EOG artifacts in brain-computer interface systems: A survey,International Federation of Clinical Neurophysiology, 118 (2007) 480-494.

[14] A Garcés Correa, E Laciar, H D Patiño, M E Valentinuzzi, Artifact removal from EEG signals using adaptive filters in cascade, 16th Argentine Bioengineering Congress and the 5th Conference of Clinical Engineering, Journal of Physics: Conference Series 90 (2007).

[15] L. Santamaria-Granados, M. Munoz-Organero, G. Ramirez-Gonzalez, E. Abdulhay, N. Arunkumar, Using the deep convolutional neural network for emotion detection on a physiological signals dataset (AMIGOS), IEEE Access 7 (2018) 57–67

[16]S. Siddharth, T.-P. Jung, T.J. Sejnowski, Utilizing deep learning towards multimodal bio-sensing and vision-based affective computing, IEEE Trans. Affect. Comput. (2019)

[17] Yekta Said Can and Cem Ersoy. 2021. Privacy-preserving Federated Deep Learning for Wearable IoT-based Biomedical Monitoring. ACM Trans. Internet Technol. 21, 1, Article 21 (February 2021)

[18] Arijit Nandi, Fatos Xhafa, A federated learning method for real-time emotion state classification from multimodal streaming, Methods, 2022, ISSN 1046-2023,https://doi.org/10.1016/j.ymeth.2022.03.005.

[19] Priyanka A. Abhang, Bharti W. Gawali, Suresh C. Mehrotra, Chapter 2 - Technological Basics of EEG Recording and Operation of Apparatus,Editor(s): Priyanka A. Abhang, Bharti W. Gawali, Suresh C. Mehrotra, Introduction to EEG- and Speech-Based Emotion Recognition, Academic Press,2016, Pages 19-50, ISBN 9780128044902, https://doi.org/10.1016/B978-0-12-804490-2.00002-6.(https://www.sciencedirect.com/science/article/pii/B9780128044902000026)

[20] Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., Xu, X., & Yang, X. (2018). A Review of Emotion Recognition Using Physiological Signals. *Sensors (Basel, Switzerland)*, *18*(7), 2074. https://doi.org/10.3390/s18072074

[21] Yu, Liang-Chih & Lee, Lung-Hao & Hao, Shuai & Wang, Jin & He, Yunchao & Hu, Jun & Lai, K. & Zhang, Xuejie. (2016). Building Chinese Affective Resources in Valence-Arousal Dimensions. 10.18653/v1/N16-1066.

[22] S. Gannouni, A. Aledaily, K. Belwafi and H. Aboalsamh, "Adaptive Emotion Detection Using the Valence-Arousal-Dominance Model and EEG Brain Rhythmic Activity Changes in Relevant Brain Lobes," in IEEE Access, vol. 8, pp. 67444-67455, 2020, doi: 10.1109/ACCESS.2020.2986504.

[23] Rodriguez, & Jesus. (2019). *The Challenges of Centralized AI. (2019)*. Retrieved April 22, 2023, from https://towardsdatascience.com/the-challenges-of-decentralized-ai-78bb44b7b69

[24] Welch, Peter D. The Use of the Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms. *IEEE Transactions on Audio and Electroacoustics* 15, no. 2 (1967): 70–73.

[25] Bansal, D., & Mahajan, R, EEG-Based Brain-Computer Interfaces: Cognitive Analysis and Control Applications. First edition (2019)

[26] Park S, Whang M. Empathy in Human-Robot Interaction: Designing for Social Robots. Int J Environ Res Public Health. 2022 Feb 8;19(3):1889.

[27] Dudzik, W., Nalepa, J., & Kawulok, M. (2021). Evolving data-adaptive support vector machines for binary classification. *Knowledge-Based Systems*, *227*, 107221.

[28] Vanwinckelen, G., & Blockeel, H. (2012). On estimating model accuracy with repeated cross-validation. In *BeneLearn 2012: Proceedings of the 21st Belgian-Dutch conference on machine learning* (pp. 39-44).

[29] Ali, M., Sarwar, A., Sharma, V., & Suri, J. (2019). Artificial neural network-based screening of cervical cancer using a hierarchical modular neural network architecture (HMNNA) and novel benchmark uterine cervix cancer database. *Neural Computing and Applications*, *31*(7), 2979-2993

[30] Koelstra, S.; Mühl, C.; Soleymani, M.; Lee, J.S.; Yazdani, A.; Ebrahimi, T.; Pun, T.; Nijholt, A.; Patras, I. DEAP: A Database for Emotion Analysis Using Physiological Signals. IEEE Trans. Affect. Comput. 2011, 3, 18–31.

[31] Garg, S., Patro, R. K., Behera, S., Tigga, N. P., & Pandey, R. (2021). An overlapping sliding window and combined features-based emotion recognition system for EEG signals. Appl. *Comput. Inform*.

[32] James A Russell and Lisa Feldman Barrett, Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant., Journal of Personality and social psychology 76 (1999), no. 5, 805.

[33] B. Schuller, S. Reiter, R. Mueller, M. Al-Hames, and G. Rigoll, "Speaker Independent Speech Emotion Recognition by Ensemble Classification," Proc. Sixth Int" l Conf. Multimedia And Expo pp. 864-867, 2005

[34] Ahmad Tauseef Sohaib & Shahnawaz Qureshi, Empirical study of various machine learning techniques for classifying emotional state from EEG Data, Blekinge Institute of Technology, Sweden, 2012.

[35] E. Ackerman and L.C. Gatewood, Mathematical Models in the Health Sciences: A Computer-Aided Approach, University of Minnesota Press, 1979

[36] W. Zheng, J. Zhu and B. Lu, "Identifying Stable Patterns over Time for Emotion Recognition from EEG," in IEEE Transactions on Affective Computing, vol. 10, no. 3, pp. 417-429, 1 July-Sept. 2019, doi: 10.1109/TAFFC.2017.2712143.

[37] Santamaria-Granados, L., Munoz-Organero, M., Ramirez-Gonzalez, G., Abdulhay, E., & Arunkumar, N. J. I. A. (2018). Using the deep convolutional neural network for emotion detection on a physiological signals dataset (AMIGOS). *IEEE Access*, *7*, 57-67.

[38] Padhmashree, V., & Bhattacharyya, A. (2022). Human emotion recognition based on time–frequency analysis of multivariate EEG signal. *Knowledge-Based Systems*, *238*, 107867.

[39] Shukla, J., Barreda-Angeles, M., Oliver, J., Nandi, G. C., & Puig, D. (2019). Feature extraction and selection for emotion recognition from electrodermal activity. *IEEE Transactions on Affective Computing*, *12*(4), 857-869.

[40] Bracewell RN, Bracewell RN. The Fourier transform and its applications. New York, NY: McGraw-Hill; 1986 Feb.

[41] Daubechies I. Ten lectures on wavelets, CBMS conf. Series Appl Math; 1992 Jan 1; 61.

[42] Ernst Niedermeyer and Fernando Lopes Da Silva, Electroencephalography–basic principles, clinical applications, and related fields, Urban & Schwarzenberg, 2020.

[43] Soraia Meneses Alarc˜ao and Manuel J. Fonseca, Emotions recognition using EEG signals: A survey, IEEE Transactions on Affective Computing PP (2017)

[44] Sumedha Aurangabadkar and MA Potey, Support vector machine based classification system for classification of sport articles, 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), IEEE, 2014, pp. 146–150.

[45] Turkey Alotaiby, Fathi E Abd El-Samie, Saleh A Alshebeili, and Ishtiaq Ahmad, A review of channel selection algorithms for EEG signal processing, EURASIP Journal on Advances in Signal Processing 2015 (2015), no. 1, 1–21.

[46] Xiao-Wei Wang, Dan Nie, and Bao-Liang Lu, Eeg-based emotion recognition using frequency domain features and support vector machines, International conference on neural information processing, Springer, 2011, pp. 734–743.

[47] Michal Teplan et al., Fundamentals of eeg measurement, Measurement science review 2 (2002), no. 2, 1–11

PAPER NAME

Emotion classification on Physiological Data 20th May 2023.docx

WORD COUNT

10063 Words

CHARACTER COUNT

58222 Characters

PAGE COUNT

47 Pages

FILE SIZE

1.1MB

SUBMISSION DATE

May 20, 2023 5:25 PM GMT+5:30

REPORT DATE

May 20, 2023 5:26 PM GMT+5:30

16% Overall Similarity

May 22, 2023

The combined total of all matches, including overlapping sources, for each database.

- 12% Internet database
- 8% Publications database
- Crossref database
- Crossref Posted Content database
- 11% Submitted Works database

Excluded from Similarity Report

- Bibliographic material
- Quoted material
- Cited material
- Small Matches (Less then 8 words)
- Manually excluded text blocks