A Comparative Analysis of Emotion Recognition Using EEG Signals With A Channel Selection Technique

by

Jonayed Ahmed Riduan
17101028
Most. Mahjabin
17101187
Nadia Tasnim Mim
17101487
Ridwane-ul-Islam
17101495
Md. Shahriar Rahman Rana
18201195

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science and Engineering

Department of Computer Science and Engineering
Brac University
January 2021

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.

4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Riduan

Jonayed Ahmed Riduan
17101028

Most. Mahjabin

Most. Mahjabin
17101187

Nadia Tasnim Mim

Nadia Tasnim Mim
17101487

Ridwane-ul-Islam

Ridwane-ul-Islam
17101495

Md. Shahriar Rahman Rana

Md. Shahriar Rahman Rana
18201195

Approval

The thesis/project titled "A Comparative Analysis of Emotion Recognition Using EEG Signals With A Channel Selection Technique" submitted by

1. Jonayed Ahmed Riduan (17101028)
2. Most. Mahjabin (17101187)
3. Nadia Tasnim Mim (17101487)
4. Ridwane-ul-Islam (17101495)
5. Md. Shahriar Rahman Rana (18201195)

of Spring, 2017 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of BSc in Computer Science and Engineering on January 11, 2021.

Examining Committee:

Supervisor:
(Member)

Zavid Parvez
Digitally signed by Zavid Parvez
DN: cn=Zavid Parvez, o=Brac University, ou=CSE, email=zavid.parvez@bracu.ac.bd, c=US
Date: 2021.01.15 17:22:38 +06'00'

Mohammad Zavid Parvez, PhD
Assistant Professor
Department of Computer Science and Engineering
Brac University

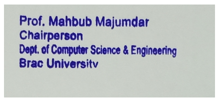
Co-Supervisor:
(Member)

Mohammed Abid Abrar
Lecturer
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Md. Golam Rabiul Alam, PhD
Associate Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)



Mahbubul Alam Majumdar, PhD
Professor and Dean, School of Data and Sciences
Department of Computer Science and Engineering
Brac University

Abstract

Emotion can be defined as the neurophysiological changes people experience due to significant internal or external occasions. This is a mental condition that can affect a person's behavior, mood, way of life, and relationship with others. As it directly affects one's life, emotion recognition is an important subject in the area of research field. In recent years, there has been a relentless effort to develop several models and datasets to detect human emotions and analyze them to understand the depth of complex human feelings and reduce error in the detection of emotions. To get better results in recognizing emotion, extensive research is needed to be done on the feature extraction methods and channel selection. While measuring the performance of different classification algorithms, it is very important to compare the results as well as preprocessing techniques. In this work, we extracted DWT wavelet features of the EEG channels from the DEAP dataset and used a statistical parameter Root Sum Square (RSS) to reduce the dimension of the features. Then we applied a channel selection algorithm on the preprocessed EEG data and selected ten channels with the highest average power. Finally, we classified positive and negative emotion related to valence and arousal using different classification algorithms (like KNN, RF, Bagging, Extra Tree, AdaBoost and MLP) for the selected EEG channels as well as for all EEG channels. The accuracy reports achieved for the selected channels were impressive; the highest test accuracy 67.58% for Valence was retrieved from the Bagging and Extra Trees classifier while MLP achieved the highest test accuracy result 63.67% for Arousal.

Keywords: Emotion recognition, DEAP, EEG, Channel Selection, DWT, RSS, PSD, KNN, RF, Bagging, Extra Tree, AdaBoost, MLP

Dedication

We want to dedicate our work to all the frontline fighters of Covid-19 throughout the world who dedicated their lives for the sake of humanity.

Acknowledgement

Firstly, all praise to the Almighty Allah for whom our thesis have been completed without any major interruption.

Secondly, to our supervisor Dr. Mohammad Zavid Parvez sir and co-supervisor Mohammed Abid Abrar sir for their kind support and advice in our work.

And finally to our parents without their throughout support it may not be possible. With their kind support and prayer we are now on the verge of our graduation.

Table of Contents

List of Figures

List of Tables

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

DT Decision Tree

DWT Discrete Wavelet Transformation

EEG Electroencephalogram

KNN K-Nearest Neighbour

MLP Multilayer Perceptron

PSD Power Spectral Density

RF Random Forest

RSS Root Sum Square

Chapter 1

Introduction

Emotion is a complex state of human beings. It consists of behavior, feelings, thoughts, reactions to internal or external stimuli, etc. Nowadays emotion recognition is a vital issue in this tech-based world. Researchers are trying to improve the Human-Computer Interaction (HCI) more appropriately using different technologies and methodologies like Artificial Intelligence, Machine Learning, Logistic Regression, Extreme Machine Learning, Deep Learning, Fast Fourier Transform, Neural Network, Principal Component Analysis (PCA), Power Spectral Deficiency (PSD), etc. It is required as people are attached so closely to different types of machines to solve daily problems. For instance, medical treatments, disease identification, e-learning, home and city automation, AI-based conversations, social media, video gaming, online shopping, recommendation systems, reducing accidents while driving vehicles, military and so on. On this note, appropriate emotion recognition is very important to make the machines adaptable to humans' emotional state and user-friendly in real time. And for this, focusing on various recent methods like channel selection and feature extraction to train the models is required.

1.1 Problem Statement

Accuracy in emotion mostly depends on which process the data has been collected and the selected features have been chosen to train the various models.
However, maximum human-computer interaction (HCI) systems can not interpret and also due to the lack of acknowledgement in emotion related information,emotional intelligence is absent. In most cases, human-computer interaction (HCI) systems fail to detect different stages of emotion and as a result, decision making is hampered.In the development of advanced HCI systems,computerized automatic recognition of a human's emotional state is remarkable.The idea of affective computing research came from here [19].
Researchers used different types of methodologies but did not get the best possible outcome. The outcomes were different for different parameters of emotional states. For this reason, there is still a vast scope to improve on this field.
Now-a-days, researchers are more interested in the progressively growing affective computing field since it is an important part of AI research. Recently, emotion related researches are covering following topics: Feature extraction algorithms to determine the characteristics of human emotions, identifying physiological signals based on correlation in order to find out the expected state of emotion, stimuli selec-

tion method is another approach, fusion of information using multimodal technique etc [19].

Emotion assessment methods mainly are divided into two terms which includes subjective & objective. [13], [14]. In order to achieve the subjective term, Self-Assessment Manikin (SAM) is highly used [13] and objective measures are mostly obtained from physiological cues which are basically executed from the physiological theories of emotion[15]

There is a vital relation between emotion recognition and affective computing. Emotion recognition is used in different interdisciplinary fields like artificial intelligence, computer science, cognitive neuroscience and in the study of psychology too. For the identification of human emotion many ways are used nowadays. Among these speech, facial expressions, behavior or psychological signals are noteworthy[1-4]. Researchers delineated the reliability of psychological signals as it is a subject independent approach[5]. The importance of EEG signals are much greater than the peripheral neural signals as it is generated from the central nervous system(CNS). Furthermore, along the change of emotion, EEG signals quickly capture or respond to the change. Furthermore, it is better in terms of reliable feature extraction for emotion recognition tasks[6, 7]. Electroencephalogram (EEG) signals are more insightful especially when any subject can not express his/her desired emotion[18]. Furthermore, due to the high speed factor, low cost and interaction capability with our central nervous system, most of the HCI/BCI system uses EEG signals[17].

A few databases have been constructed with a view to help the study of emotions in recent years. One of the most remarkable dataset here is DEAP which means Database for emotion analysis from physiological signals.[10] That database contains physiological signals and electroencephalograms, which allows researchers to study emotional states in the spaces of valence/arousal [9].

Since our body is capable of generating responses of various emotional states,hence we can say that if we use EEG signals as well as different physiological signals, the outcome would be appropriate enough to detect human emotions.[9] Therefore, it can be claimed that there is an existence of a complex relationship between EEG signals and emotional states among humans. On this note, we are planning to do some research work on the EEG signal dataset to get the best possible outcome.

Each emotional state was mapped to a region depending on the emotion model of arousal and valence which works on two-dimension. Among them, in the horizontal axis, valence of the emotion is represented and besides this some negative or positive emotions are also noticeable. Arousal of emotions are represented by the vertical axis which reflects the level of neurophysiological emotional activation [16].

Most of the research on this field valence and arousal space played an important role as a parameter of measuring emotional states of humans.

For the analysis purpose some additional algorithms were also used such as SVM,KNN,decision tree,PCA etc.[18]Hidden Markov Models (HMM) are one of the generative machine learning algorithms that represent states formed from dynamic signals which represent a certain class [8].

Deep convolutional neural networks (CNNs) provide superior performance in analysing multimedia tasks are regarded as the state-of-art models. To improve the recognition accuracy and for extracting more powerful features Deep CNN models are applicable [11].

However, the ML techniques have limitations as the techniques individually could not show a standard fixed output for all emotional state parameters of human beings. EEG channel selection and feature extraction method is another important factor in increasing the accuracy rate of emotion recognition. A study focusing on only one specific EEG channel F4 showed significant improvement on the accuracy level [40]. On the other hand, another study that uses all 32 channels of EEG signals mainly focuses on feature extraction and performed different experiments to measure the accuracy [39]. Different models perform better in different conditions. So it is necessary to focus on the performances of the models on different conditions. Furthermore, models perform better when trained with more features. The accuracy level improves to an extent with higher number of features [33]. Moreover, training models with both feature extraction and as well as channel selection increases the prediction accuracy of the model [41]. Therefore, more focus needs to be put on feature extraction and channel selection in order to predict some good results on different emotional states of humans.

1.2 Research Objectives

This research aims to recognize, predict and compare the emotional states (Valence and Arousal) by applying different machine learning and neural network algorithms like KNN, RF, Bagging, Extra Tree, AdaBoost and MLP using classified EEG signal data. We are planning to work on the DEAP dataset to compare different techniques of emotion classification. However, the detail list of the objectives is given below:

a) To know how EEG signals can be classified using modern techniques.
b) Extracting features using DWT and RSS.
c)Apply techniques for band selection and channel selection.
d) Executing different classification techniques on the data from selected channels and all channels.
e) Comparing the results of the accuracy of these classification techniques.
f) Visualising and analysing the retrieved results.
g) Comparing the accuracy results with other research works.
h)Suggesting some future works or directions to get the best possible outcome.

1.3 Thesis Outline

This report is arranged in the following order: a) In chapter 2, related works are discussed.
b) After that all the background studies are discussed in chapter 3.
c) Chapter 4 focuses on the dataset used in this study.
d) Chapter 5 explains the detailed methodology like feature extractions, channel selection etc.
e) Chapter 6 analyzes the retrieved results.
f) Finally, conclusion of the whole report is given in chapter 7 along with future works.

Chapter 2

Related Work

Emotion plays a very significant role for communicating among humans. In recent years, there has been extensive work on emotion recognition.

Torres-Valencia et al. used different classifiers in order to compare the EEG and physiological signals for recognising multimodal emotions. The paper identifies different emotional conditions in arousal and valence space using the Multimodal Emotion Recognition (MER) approach where results obtained from combination of different physiological signals are compared. To capture and analyze the physiological signals the paper proposed the use of Hidden Markov Models (HMM). The space of classification proposed here is the "Valence-Arousal space". Here they used the DEAP dataset to identify the physiological signals that bring related information related to emotion recognition. They experimented with different signals one by one and also combining different signals together by HMM model to find that differences between valence and arousal data happens due to the signals from EEG and physiological- GSR and rate of heart. The researchers used the Gaussian Mixture model to adjust HMM parameters. The arousal and valence range is 1 - 9 where 1 to 5 indicates low and 5 to 9 indicates high portion. The bio-physiological signals used in this work are: EEG, EOG, EMG, HR, GSR, Temperature and Respiratory signals. The obtained result showed that the classification accuracy (CA) is about 75% for the arousal of EEG and physiological signal. On the other hand, for the valence dimension the maximum CA (58.75%) is obtained using only EEG and also combining EEG, EOG, EMG signals together. The researchers suggest that for higher accuracy in identifying emotion, bringing differences in feature extraction and then combining generative models could be a method. Also, they suggest combining generative models with discriminative models (i.e. HMM-SVM) in the future to assess emotional states in arousal/valence space [9].

Different ML methods have been used in this research area. This recent paper showed some comparisons among multiple ML methods [namely Support Vector Machine (SVM), K-nearest neighbor (KNN), Linear Discriminant Analysis, Logistic Regression and Decision Trees] on the EEG and peripheral physiological signal dataset to detect particular emotion. For dimensionality reduction they used a technique where they tested with and without principal component analysis (PCA) . In addition, they used the DEAP dataset. DEAP dataset is a multimodal dataset which is used for affective state analysis of humans. In the experiment they tested accuracy, recall, precision and f1-score. So, finding the higher accuracy and F1 score to choose the best mode was the main research objective. Different classic ML techniques stated above achieved accuracy between 55 to 75%. whereas an F1 had an accuracy in between 70 to 86%. Among these models, KNN classification for "Liking" performed best with accuracy of 74.25%. as the best possible result obtained among all models and their parameters. Arranging in descending order KNN, SVM, Decision trees, Logistic Regression, and LDA achieved highest accuracy [18].

In this paper, the authors used EEG signals to focus the high risk for emotional disorders along with emotion type. They put forward a machine learning technique to find the high risky group for emotional disorders. Here they combined with the features of multiple cortex areas using a kernel SVM classifier. The EPrime 2.0 software was used to conduct the experiment. Besides, Neuro Scan recording system was used to collect EEG data of patients while watching videos. The support vector machine (SVM) was used to differentiate between low and high risk groups. They had chosen the kernel SVM instead of linear SVM due to limitations of data. Due to emotional disorder, irregularity in brain activity can be seen on bands especially on alpha band. An improved accuracy was found for the whole cortexes when integrated with multiple brain areas (95.20%). Their work still had some limitations like classifying only in two categories and not exploring different state's influence for classification. All of those are prime study in the field of emotional disorder and they aim to further development from this aspect [20].

The authors of this paper suggested a renewal system, the demonstration curve for the proposed method for emotion prediction using mL and only EEG signal experimenting on the DEAP and SEED data. For SEED data they reduced the data duration but used the full duration for DEAP data. They extracted features from the EEG signals to construct two coefficients which are entropy and correction. These coefficients were used for constructing the proposed curve which gives the information regarding the renewal process. They represented the proof of the rise of classification accuracy due to the coefficients. The entropy curve focuses on the uncertainty of emotional state at different time points and shows hints of interpreting human emotional activities during the experiments. They removed the falsely identified samples while doing the calculations.Encoders were used for obtaining the features. They obtained the highest classification accuracy of 62.63% and 74.85% for the EEG data of last 34 s and 75 s [12].

The researchers have presented the DEAP database for the analysis of human affective states. The dataset contains physiological signals of 32 participants. Among them, frontal face video of 22 participants were recorded. Here each participant rated their emotional response to 40 music videos on the basis of arousal, valence, and dominance. The rating also includes participants' liking and familiarity with the videos. They had used a novel semi-automatic stimuli selection method. This model validates the affective tags by analysing the ratings of participants. The classification method was challenging because of noisy signals. Moreover, they had insufficient number of self-assessments and individual physiological differences. However, the result showed remarkable correlations between the participant ratings and EEG frequencies. Increase in the performance was observed with the decision fusion of all the results. The researchers have also made the database publicly available for research purposes [10].

A structure named Multiple Instance Learning was introduced in this paper. This paper uses different SVM algorithms that represent time interims by catching the appearance of the related states and nonappearance of the related states while avoiding the continuous tagging of visceral reactions. They used two types of datasets - DEAP and consumer; for the experiments. The data was processed before the extraction of the physiological features using TEAP, a toolbox for Mat lab following the quantified model. Accuracy, Confusion Matrix, Macro-F1 and Receiver Operating Characteristic (ROC) are followed while performing the assessment of the MIL model. As a solution to valence and arousal activity the usefulness of the end-user centered environment and the accuracy of the suggested MIL methods was proved for the DEAP dataset. However, a huge difference among the obtained results was observed due to the difficulty of the tasks. EMDD-SVM gave the most correct result for arousal identification amidst providing a poor performance against valence identification. Despite the shortcomings, the MIL technique reliably foresees all through the unseen experiments for Valence and a convincing probable outcome for Arousal. Finally, in the near future their idea includes working more on the individual contrast on the DEAP dataset as well as extending their suggested approach in another scenario [21].

In the work, the authors explored the reliability of physiological signals as a means for emotion identification and discussed all the important steps like classification of feature-based multiclass, recording process of physiological data set etc. of an automatic recognition system. To gather a physiological data set they used an induction method to lead their experimentees to their actual state of emotion so that they can escape creating a forced particular set of emotion for the participant in the laboratory In order to correspond the suggested features and find the most accurate ones from different analogy domains with emotional states several physiological features were proposed. They also proved the suggestions from the classification outcome. The researchers identified 4 types of emotions that can be correlated to music using pLDA and developed an EMDC strategy. Then they performed an accuracy test for comparing their performance and achieved identification accuracy of 95 and 70 /
The authors of this paper suggested an emotion recognition technique using EEG signals. They measured EEG signals and classified them into five frequency ranges. Frequency ranges were classified based on the power spectrum density. They removed 0 to 4 Hz low frequencies to eliminate EEG artifacts. This method used the fast Fourier transform (FFT) to divide each EEG signal into five EEG sub-bands. The resulting calculations of the frequency ranges were based on the percentage of the selected range relative to the total range which were then compared to standard values from a Bayesian network, calculated from databases. Finally, text and a human face avatar was used to represent the outcome. While comparing results, the researchers found similar probability values in "anger" and "sadness" in contrast to other emotions. For future research they aimed to review human knowledge along with human emotions. Moreover, they hope to research BCI for real-time sensing which can be helpful for physically handicapped and elderly people [23].

The authors of this paper built a database related to visually aided emotion in order to test the connection of emotion between visual-aids and a person's interpretation. Auditory-optical features from the visual-aid and EEG channel features from the subjects were extracted for this database and various bagging and neural network classification models were considered to find the best suitable model. This paper used EEGLAB to extract the psd features of the frequency bands from the channels in order to utilize the classical analysis method available for frequency-domain. The outcome from this test proves the credibility of using this process while identifying visually aided emotion data that are of multi-mode variant as an increase of performance were observed for this variant compared to single-mode variant. This relatively small dataset was successful in obtaining better results when used in the MLP classification model. Despite the good outcome there were rooms for further revision like using the different coherent algorithms, Deep-CNN for powerful features as well as corporal signals for the development of the identification accuracy [11].

In this paper, the authors worked on the possibilities of applying visual and audio-visual stimulus to detect the human emotion through evaluating electroencephalograms (EEG). In this experiment they used EEG signals from five healthy subjects, evoking five different emotions (anger, sad, happy, disgust and surprise) on a single trial. To classify five emotions they derived qualitative characteristics procured from EEG signal where 'db4' wavelet function was used in alpha band. After preprocessing and normalization two statistical features namely energy and power were extracted. The average accuracy of classifying five emotions in visual stimulus for energy and power were 63.335% and 45.933% respectively. For auditory optical stimulus, average accuracy obtained for energy and power features are 64.673% and 67.33%. For emotion recognition, over visual stimulus, the auditory optical stimulus-based EEG recording gives the maximum classification accuracy of 67.33%. As improved classification accuracy can be achieved using 'db4' wavelet function, their method of extracting features from the alpha band was proven in identifying emotions from the EEG signals [17].

In this study the authors mainly reviewed corporal stimulus, feelings as well as briefly described some of the related research works of the corporal and electrical fields in feelings identification. This paper uses multi-mode dataset recorded using MIT Media Lab to identify if a person is glad, gloomy or flabbergasted. For the identification of these emotions, they performed several kinds of testing on this dataset. Despite eeg and emg proves to help classifying feelings, they are not really suitable for day to day life. Henceforth, several corporal stimuli were formed for this purpose. With the aim of bringing more accuracy in emotion classification in the not so soon days the authors would like to work on gaming and e related divisions [24].

In this paper, the authors proposed an EEG emotion wave based feature learning and classification technique using deep (CNN) based on some specific features and fusions of EEG signals in DEAP dataset in order to increase the accuracy. For classification they used different ML and NN models like bagging tree, SVM, LDA, BLDA and deep CNN. Since deep learning models can learn from raw samples large sets of data, while ignoring feature extraction and selection processes used in DEAP dataset.Their experimental outcome indicates that deep CNNS performance was better than any other identifiers. In order to achieve improvement in identifying emotions the authors hope for adopting more EEG features and as well as perform some synovitis analysis [33].

In this thesis researchers tried to prove that brain waves vary with emotion and to know more about Critical Frequency Bands And Channels, they introduced DBNs to develop models that work with EEG signals. Here they studied with positive, negative and neutral emotions.They received an accuracy rate of 86.65%. They also investigated different connected feelings that are related to the reduction of electrode sets and brain signals. For this purpose they followed weight arrangements which were retrieved from deep neural networks. For SVM, LR, and KNN they obtained accuracy of 86.08%, 83.99%, 82.70%, and 72.60% respectively. The final outcome showed that their selected set of electrodes shows better performance & also proved that DBN models obtain higher accuracy compared to KNN,LR,SVM processes [34].

In that particular paper authors stated that user's emotions can be recognised through EEG based Brain Computer Interfaces (BCI) devices.They used different algorithms such as emotion elicitation, signal acquisition,feature extraction & selection, different classification techniques,,performance evaluation etc.They also proved that negative emotional states are more challenging to detect compared to positive emotional states[35].

Another paper showed the statistical analysis of an EEG-based emotion recognition system using a self-organizing map for boundary detection. They worked with happiness, sadness, fear & placidity since these emotions were easily mappable from the affective responses. In order to use SAM scores for label classification, a self-organizing map was used. In order to identify the emotions from the EEG features K-fold cross validation and KNN were used. The proposed method improved the accuracy to 84.5% and it proved that the class difference helps rising the emotion identification accuracy using EEG signals [36].

A positive and negative emotion classification using five selected EEG channels using the DEAP dataset were proposed in this paper. They performed DWT and then used some analytical methods in order to excerpt the EEG features leaving the physiological data out. From there they determined five EEG channels from all the EEG channels vigorously that had outstanding performance while selecting the channels. K-fold cross validation, MLPNN and KNN was used in order to test the accuracy of this experiment. The outcome was achieved from the training set with the fixed environment. The classification performance was calculated for both of the algorithms. The average overall accuracies were 77.14% for MLPNN and 72.92% for KNN. In near future in order to consider the feelings related bands performance they plan on using all of the secondary bands individually on their proposed model [41].

This particular review paper put emphasis on the channel selection process & provided a detailed idea of selecting models. The whole paper covered different topics of channel selection approach such as selection techniques, seizure detection, motor imagery classification, feelings identification and separating cerebral tasks with selected channels, sleep state analysis, drug effects diagnosis etc. For identification of EEG channels they used various feature selection methods such as heuristic search, sequential search and random search. Subject evaluation was done following five strategies. In order to identify the accuracy of the feature vectors they used GMM, KNN and Parzen. From the classifications the most accurate result obtained was 80 percent. An identification method has been applied for evaluating the channel selection of candidates. Different channel selection techniques such as Filtering techniques,2 Wrapper techniques,Embedded techniques,Hybrid techniques, Human-based techniques were used to to reduce the feature pattern size.They also used SVM to classify the patterns into preictal and interictal states. They compared the result for all pairs of channels and pairs of 75 fixed channels in the ECoG database and the EEG database respectively. Basically, channel selection has been investigated with a variety of techniques. In future, the authors hope for improving these channel selection methods and also hope that new channel selection processes will be discovered [37].

Chapter 3

Background Study

The background about the EEG method is discussed in part 3.1. Part 3.2 describes the classification of valence and arousal, the PSD method used for EEG channel selection is discussed in 3.3, while 3.4 includes description of DWT method and 3.5 includes RSS method for feature extraction. Moreover, in part 3.6, 3.7, 3.8, 3.9, 3.10, 3.11 and 3.12 K-fold Cross Validation method, KNN, Random Forest, Bagging, Extra tree, AdaBoost and MLP have been described respectively.

3.1 EEG

The Electroencephalogram (EEG) is the method of recording electrical activity generated from the brain with the help of electrodes placed along the scalp surface. These electrodes are placed on the scalp following the 10/20 system. The main points of these electrodes are Frontal pole (Fp), central (C), Parietal (P), Occipital (O) and Temporal(T). The electrodes in the midline are marked with z. The points over the left hemisphere are marked with odd numbers whereas the points over the right hemisphere are marked with even numbers as shown in figure 3.1 [26].
EEG frequencies are categorized in four major bands named Theta, Alpha, Beta and Gamma[19]. Theta frequency wave is a manifestation of focal sub cortical lesions , alpha wave band can be best seen from the both posterior side of the head due

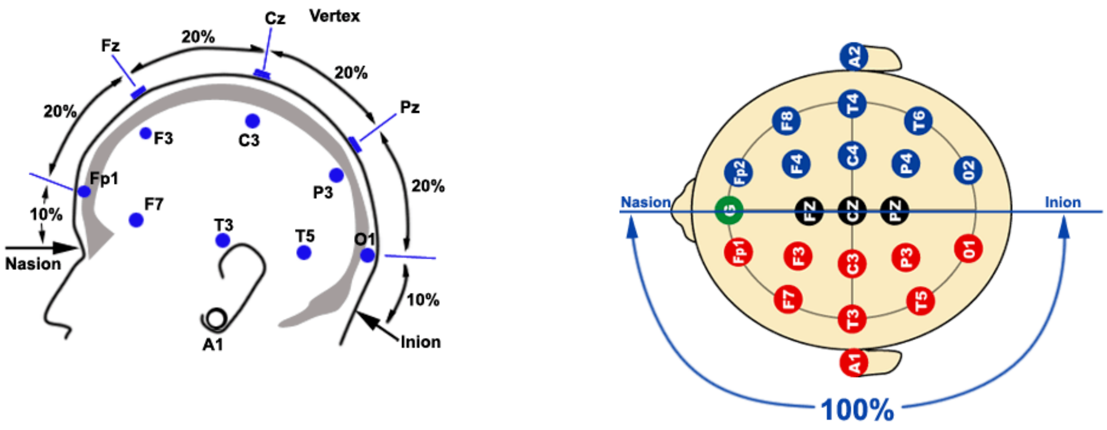


Figure 3.1: 10/20 system of electrode placement [26]

to having higher amplitude on the dominant side and beta band frequency is most prominent frontally and can be seen on both sides in a symmetrical distribution [26]. Gamma band is distributed throughout all the cerebral structures [27].

3.2 Valence and Arousal

Emotion can be classified into two dimensional space: valence and arousal. Generally valence describes the positivity and negativity of emotion while arousal describes the intensity and strength of the current state of the emotion [28]. The range of valence circulates from highly negative to highly positive whereas the range of arousal is from calming/ soothing to exciting/ agitating represented in fig 3.2 [29].

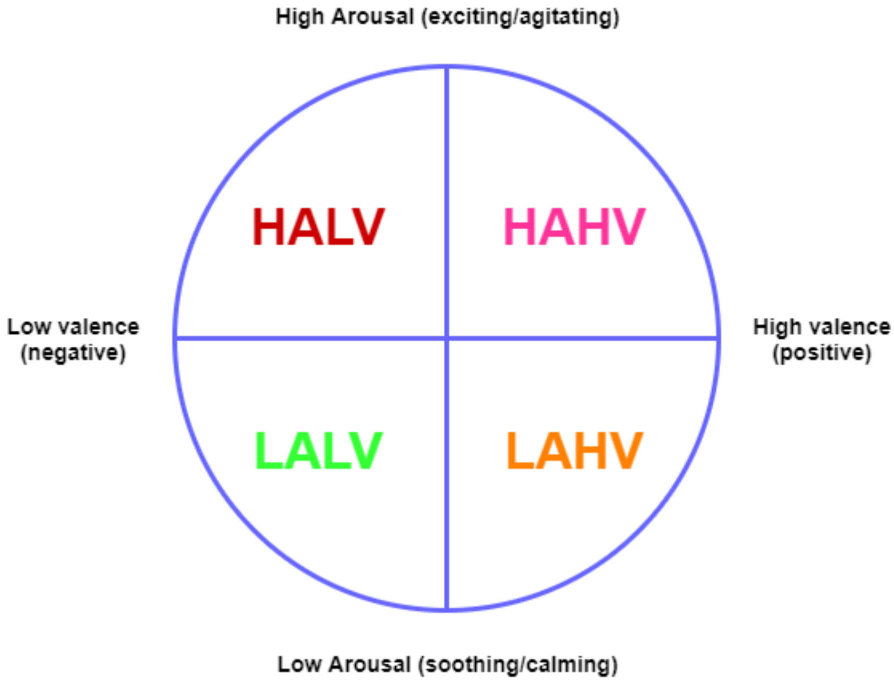


Figure 3.2: Arousal and Valence representation

3.3 PSD

Power Spectral Density (PSD) is mainly used to normalize frequency components. The PSD method is described below: Denote the ith signal x windowed, then zero-padded frame from the signal x by

$$x_i(k) \triangleq w(k)x(k+i\theta), \;\; k = 0, 1, ..., I-1, \;\; i = 0, 1, ..., Q-1$$

where, θ is set as the window hop size, and Q is the number of available frames. So that the periodogram of the ith block is defined as

$$P_{xi}, I(_{\omega_q}) = \frac{1}{I} \mid FFT_{K,q}(x_i) \mid^2 \triangleq \frac{1}{I} \mid \sum_{k=0}^{K-1} x_i(k) e^{-j2\pi kq/k|2}$$

Then the power spectral density is given by

$$\hat{S}_x^D(\omega_q) \triangleq \frac{1}{Q}\sum_{i=0}^{Q-1} P_{xi}, I(\omega_q)$$

3.4 DWT

Discrete wavelet transform (DWT) is useful for analysing, de-noising and compressing the signals. This decomposition method passes successive high and low pass filters of a time series with a down sampling rate of 2. The high pass filter is denoted as g(n) and the low pass filter is denoted as h(n). This decomposition returns "approximations" and "detailed" coefficients for high and low pass filters respectively. The approximation is disintegrated further and the procedure is repeated until the decomposition reaches the specified level. The whore procedure of DWT function is described below [30]: The scaling [ϕj, k(n)] and wavelet [ψj, k(n)] functions are denoted below:

$$\phi_{j,k}(n) = 2^{-j/2}h(2^{-j}n - k)$$
$$\psi_{j,k}(n) = 2^{-j/2}g(2^{-j}n - k)$$

Where n = 0,1,2,.....,M-1; j = 0,1,2,.....,J-1; k = 0,1,2,.....,$2^j - 1$; $j = 5$;
Here M is the length of the signal.
The function of Approximation coefficients (Ai) and detailed coefficients (Di) at level i are determined as following:

$$A_i = \frac{1}{\sqrt{M}}\sum_n x(n).\phi_{j,k}(n)$$

$$D_i = \frac{1}{\sqrt{M}}\sum_n x(n).\psi_{j,k}(n)$$

Wavelet energy for each decomposition level (i = 1, . . . , l) is determined as follows:

$$E_{Di} = \sum_{j=1}^{N} \mid D_{ij} \mid^2, \;\; i = 1, 2, 3, ..., l \qquad (1)$$

Here l = 4, reflects the level of decompositions

$$E_{Ai} = \sum_{j=1}^{N} \mid A_{ij} \mid^2, \;\; i = l \qquad (2)$$

Therefore, we can define total energy from Equations (5) and (6) as:

$$E_Total = (\sum_{i=1}^{l} E_{D_i} + E_{A_l})$$

3.5 RSS

Root Sum Square (RSS) is a statistical method for dealing with a series of values. In this method each value is squared, then the summation of the squared values is calculated. Finally, the root of the summation is taken.

$$X_{RSS} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Here x is the value and i denotes the term of the series.

3.6 K-fold Cross Validation

Cross validation, a resampling process to evaluate machine learning models based on a limited dataset [31]. Cross validation is a solution to the overfitting problem of the dataset. This process has a single parameter called K. Here the training set is split into a K number of sets. In K-fold cross validation the model is trained using a K-1 number of folds of the training dataset. To measure accuracy the trained model is validated using the test dataset which is created with the remaining part of the data. In figure 3.3 a representation of K-fold cross validation is shown for K=5[32].

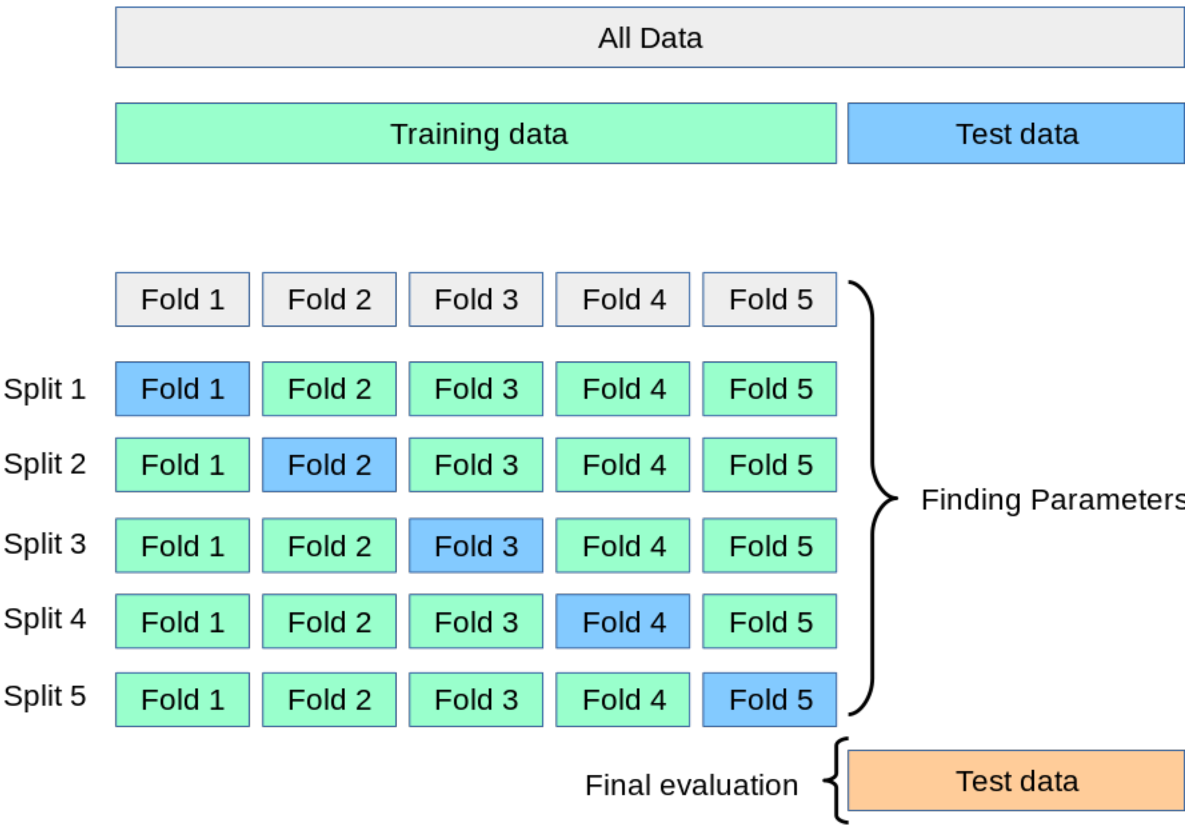


Figure 3.3: K-fold cross validation [32]

3.7 KNN

K-nearest-neighbour (KNN) extends the nearest-neighbour rule (NN) classifier. This classifier has a training set which consists of a training vector in a dimensional feature space and corresponding class label of the vector. The equation for training set is given below:

$$T = (x_i, y_i)_{i=1}^{N}$$

Given a query the classifier identifies the similar labelled target neighbours where the training set is arranged in an increasing manner in terms of euclidean distance. The distance function can be described as the following:

$$d(x^{'}, x_i^{NN}) = \sqrt{(x' - x_i^{NN})^T (x', x_i^{NN})}$$

After that the class label for the given query is predicted following the majority voting process of its neighbour [38]. The equation of class label is described below:

$$y^{'} = \underset{y}{argmax} \sum_{(x_i^{NN}, y_i^{NN}) \in T'} \delta(y = y_i^{NN})$$

3.8 Random Forest

Random forest is a supervised machine learning algorithm which can be used for both classification and regression. This algorithm uses original data to draw ntree bootstrap samples and grows an unpruned classification or regression tree for each sample where it randomly samples the mary of the predictors and chooses the best one. It provides an estimation of error rate from the training data. From the tree developed by the bootstrap data it predicts the data (OOB data) in each iteration. The model aggregates the OOB prediction and calculates the error rate [43]. A simplified visual representation of Random Forest is given in figure 3.4 below:

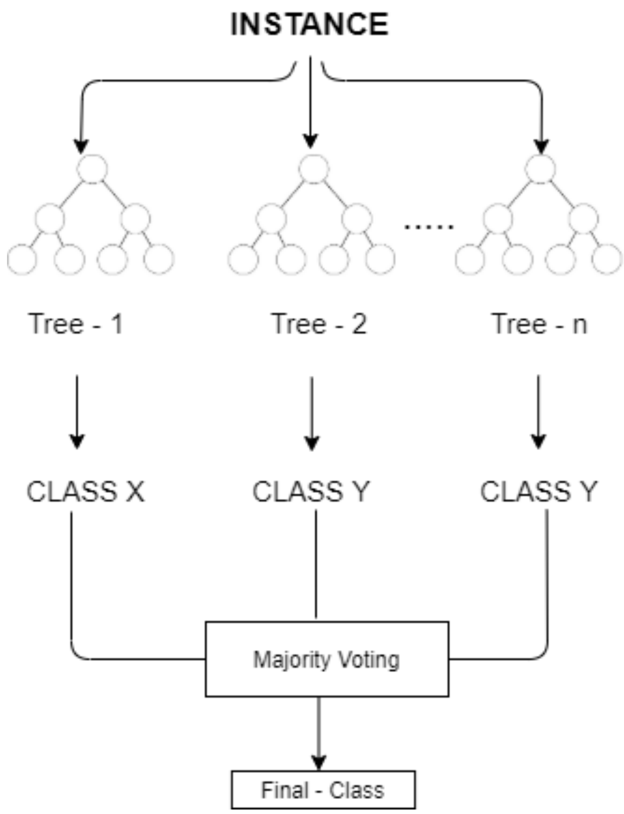


Figure 3.4: Random Forest

3.9 Bagging

Bagging is a classification method that generates multiple versions of predictors and uses these predictors to find an aggregated predictor. This method predicts a numerical outcome while averaging the versions and votes to predict a class. Bootstrap version of the learning sets forms the versions [44]. Averaging of the learning set versions' equation can be given as:

$$\phi_A(x) = E_{\mathcal{L}}\phi(x, \mathcal{L})$$

Prediction of class label by predictor can denoted as:

$$Q(j \mid x) = P(\phi(x, \mathcal{L}) = j)$$

The correct classification estimation can be represented by the following equation:

$$r = \int [\sum_j Q(j \mid x)P(j \mid x)]P_X(dx)$$

3.10 Extra Tree

Extra tree classifier fits a number of randomized decision trees in a classical top-down manner. It splits the node using randomly selected cut-points. Moreover, it uses the whole learning set to expand the tree. The final prediction from the trees are derived from majority voting and arithmetic average in regression problems[46].

3.11 AdaBoost

AdaBoost fits weak learners continuously on the improvised data version. This method initially fits the weighted training data to a classifier and obtains the probability estimates. These probability estimates are then re-normalized and fitted into the classifier again. Finally, the predictions derived from those learners are then combined to produce the ultimate prediction through majority voting denoted as[45]:

$$C(x) = \underset{y}{argmax} \sum_{m=1}^{M} h_k^{(m)}(x)$$

3.12 MLP

MLP is a multilayer perceptron neural network classifier. A simplest MLP classifier consists of three parts called input layer, hidden layer and output layer. In the hidden layer, the layers can be specified including the neuron element numbers in each layer. We can split the dataset into train and test data based on the given split ratio. Here train and test data is further scaled to make sure that the input data distribution is centered around zero with variance in the same order [42]. A simplified structure of MLP classifier with n hidden layers is given below in figure no 3.5.

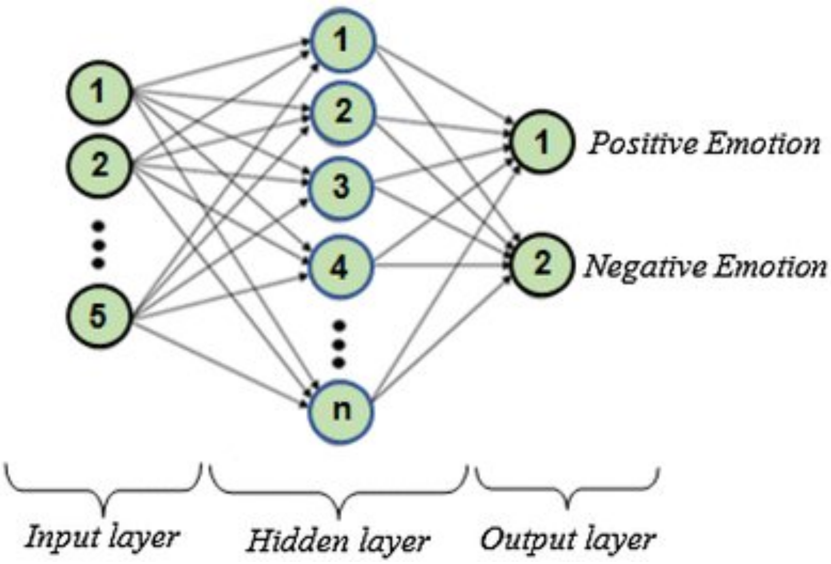


Figure 3.5: MLP Structure 5XnX2 [41]

Chapter 4

Dataset Description

This study is mainly focused on the electroencephalogram (EEG) signals that is used for identifying emotion. Hence, the dataset collected from the DEAP database was used for extracting the EEG features and perform experiments [4]. DEAP dataset is used for emotion analysis using EEG, physiological and video signals. It is a multimodal dataset which contains the EEG and peripheral physiological signals from 32 participants [25].

The DEAP dataset is divided into two parts:

1) In the dataset, they carried out an online self-assessment. In the assessment, they used 120 one-minute extracts of music videos which were voluntarily rated by 14-16 persons based on valence, arousal and dominance.

2) EEG and physiological signals recording. Where the number of total participants is 32. They were displayed 40 music videos in total which were about 1 minute long. After the end of the videos the participants had to rate each of the video in terms of the levels of valence, arousal, their likes/dislikes, presence of dominance and familiarity. There were video recordings of frontal faces of the first 22 participants among the 32 participants.

The dataset includes the following files:
a) Online ratings
b) Video list
c) Participant ratings
d) Participant questionnaire
e) Face video
f) Data original
g) Data preprocessed

File details are given below:

Online ratings: This file contains each of the rated videos used for the online self assessment. An online self-assessment tool was used to collect the video ratings. Participants were given the SAM mannequins of discrete 9-point scale to rate the valence, arousal and dominance . In addition, participants also rated the emotion they felt using an emotion wheel. This file contains online id number, valence, arousal, dominance, Wheel slice value & Wheel strength value.

Video list: The file contains a table that lists all the videos used for the online assessment and the experiment. It includes- online id number, experiment id number, tag from lastfm, artist name, song title, video link from youtube, highlighted_start, Num_ratings, VAQ_Online, AVG_x, STD_x, Q1_x, Q2_x, Q3_x, VAQ_Estimate : HA, HV; LA, HV; LA, LV; HA, LV. Here HA, HV, LA, LV indicates High Arousal, High Valence, Low Arousal, Low Valence respectively.

Participant ratings: This file lists the video ratings of the persons who participated in the experiment. The presentation software logged the start_time values. Valence, arousal, liking and dominance were rated by each of the participants after each trial based on a 9-point scale. They used a standard mouse to input the ratings. They visualized the video ratings for valence, arousal and dominance via the SAM mannequins. Thumbs up and thumbs down symbols were used for liking and for familiarity a 5 point scale was used at the end of the experiment. It contains the following information: participant_id no,trial no, experiment_id no, starting time, valence, arousal, dominance, liking and familiarity index. Participants 2, 15 and 23 do not have the ratings for familiarity.

Participant questionnaire: The file includes all the participants' except for participant 26th's responses to the question set that they had to fill in before they participated in the experiment. The questions were mostly multiple-choice questions that were self-explanatory.

Face_video.zip: The front face video recording of the first 22 participants among the 32 participants which were segregated into trials are contained in this Face_video.zip file. The synchronicity of each video is true for about 1/25 second without any human fault. The participants saw a red screen before and after the experiment and at that exact moment a marker was sent to the PC that were recording the EEG signals which helped in achieving synchronicity. Due to some technical issues several of the last videos are missing for the 3rd, 5th, 11th and 14th participants.

Data_original.zip: There are 32 original recrding files in .bdf format. These files were generated by the Actiview software where each of the 32 file were recorded at 512Hz frequency and at 48 recording channels. There are 32 EEG channels, 12 peripheral channels, 3 unused channels and 1 status channel. This type of files can be read using software toolkits such as EEGLAB that is used in Matlab and the BIOSIG toolkit. Two separate locations were used to record the data. Participants 1-22 were recorded in Twente and the rest of the participants, participant 23-32 were recorded in Geneva. The format has a few differences due to the revision of the hardware. The order of the EEG channels and the formation for GSR measure differs in each location. The EEG channels are named following the 10/20 system.

Data_preprocessed_matlab.zip and Data_preprocessed_python.zip: These files contain file named data_preprocessed_matlab.zip and also pickled python/numpy named data_preprocessed_python.zip formats. These files have been down-sampled

at 128 Hz frequency rate. They were preprocessed and segregated. This version of the data can be quickly tested with a classification or regression technique without having to go through the hassle of preprocessing the data beforehand which is very helpful. Each of the zip files has 32 .dat files for python or 32 .mat file for matlab for every participant. Total 40 channel layout has been used & the Preprocessing for channel (1-32) is:
a) Data down-sampling rate was 128Hz.
b) All the EOG artefacts of the channels were eliminated.
c) The applied bandpass frequency filter was 4.0-45.0Hz.
d) All the data for each channel was averaged so that they have common reference.
e) The EEG channels were reordered following the Geneva order.
f) They segmented the data into a 60 second trial where the 3 second pre-trial baseline was erased.
g) They reordered he trials from presentation order to the Experiment id order of the videos.

And also for channels (33-40):
a) Data down-sampling rate was 128Hz.
b) They segmented the data into a 60 second trial where the 3 second pre-trial baseline was erased.
c) They reordered he trials from presentation order to the Experiment id order of the videos.

Each participant file contains two arrays: data and labels. Data array contains video/trial x channel x data with the shape 40x40x8064 and the label array contains video/trial x label (valence, arousal, dominance, liking) with the shape 40x4.

4.1 Dataset Preprocessing

The valence and arousal target labels were classified as positive and negative class. The positive class was indicated by '1' whereas the negative class was indicated by '0'. For both valence and arousal data, where the ratings were in range of 1 to 9, a threshold value of '5' was considered to classify the data, where values less than 5 were classified as negative class. In contrast, values greater than or equal to 5 were classified as positive class. Total number of positive and negative emotions in valence is 724 and 556 respectively. On the other hand, Total number of positive and negative emotions in arousal is 754 and 526 respectively which is shown in Figure 4.1. Thus it can be depicted that the dataset is balanced for any classification task.

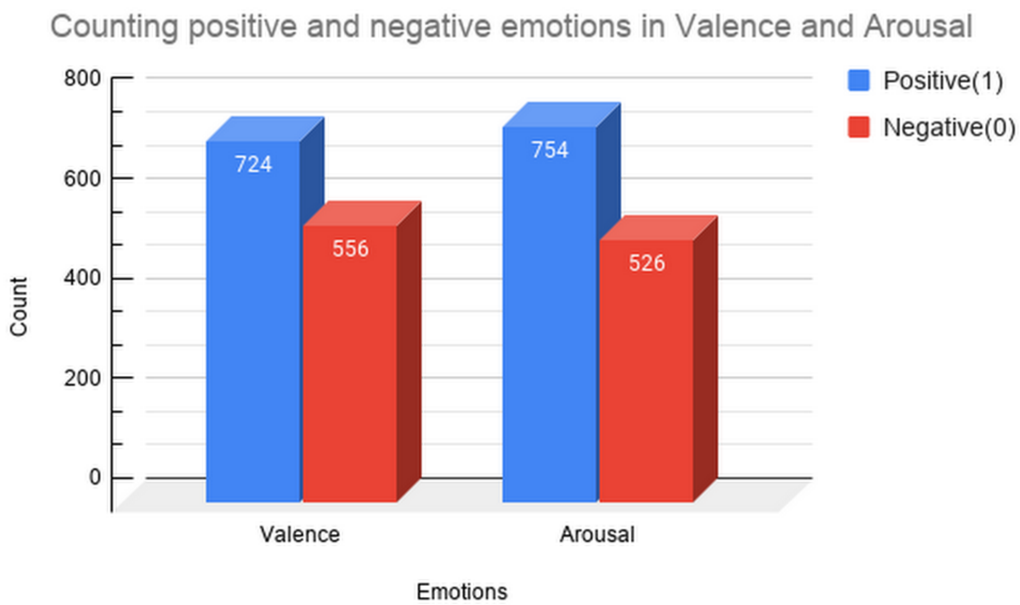


Figure 4.1: Counting positive and negative emotions in Valence and Arousal

Chapter 5

Methodology

In this study, mainly two emotional states which are valence and arousal were classified for eeg signals. First of all, feature vectors for all EEG channels were generated then used various classifiers. Similarly, a channel selection technique was applied to pick out ten selected EEG channels to classify the valence and arousal. For determining train set and test set split ratio, k-fold cross validation technique is used. Furthermore, best parameters for the classifiers were chosen by testing different combinations. Finally, test accuracy is compared with selected EEG channels and all EEG channels. For the machine learning classification task, Scikit-learn library[49] in python was used. In the later subsections, a detailed explanation of this methodology is given. The overall methodological illustration of this study is given in Figure 5.1.

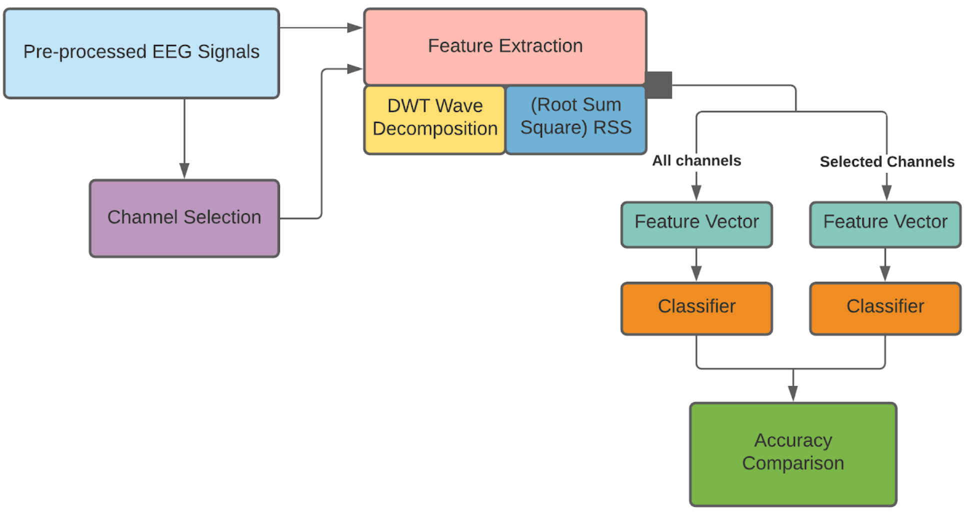


Figure 5.1: Proposed Methodology

5.1 Feature Extraction

For the feature extraction process, at first DWT decomposition were used where wavelet was db2 of Daubechies family and the level was four. As the EEG signals are non stationary data, wavelet transformation is more suitable to extract the effective features. Hence, the proposed methodology uses the Discrete Wavelet Coefficient as it depicts the degree of correlation between the signal and the wavelet function at various times[40].Therefore, four levels of detailed coefficients which are cD1, cD2, cD3 and cD4 were generated and one approximation coefficient was found as shown in Figure 5.2.

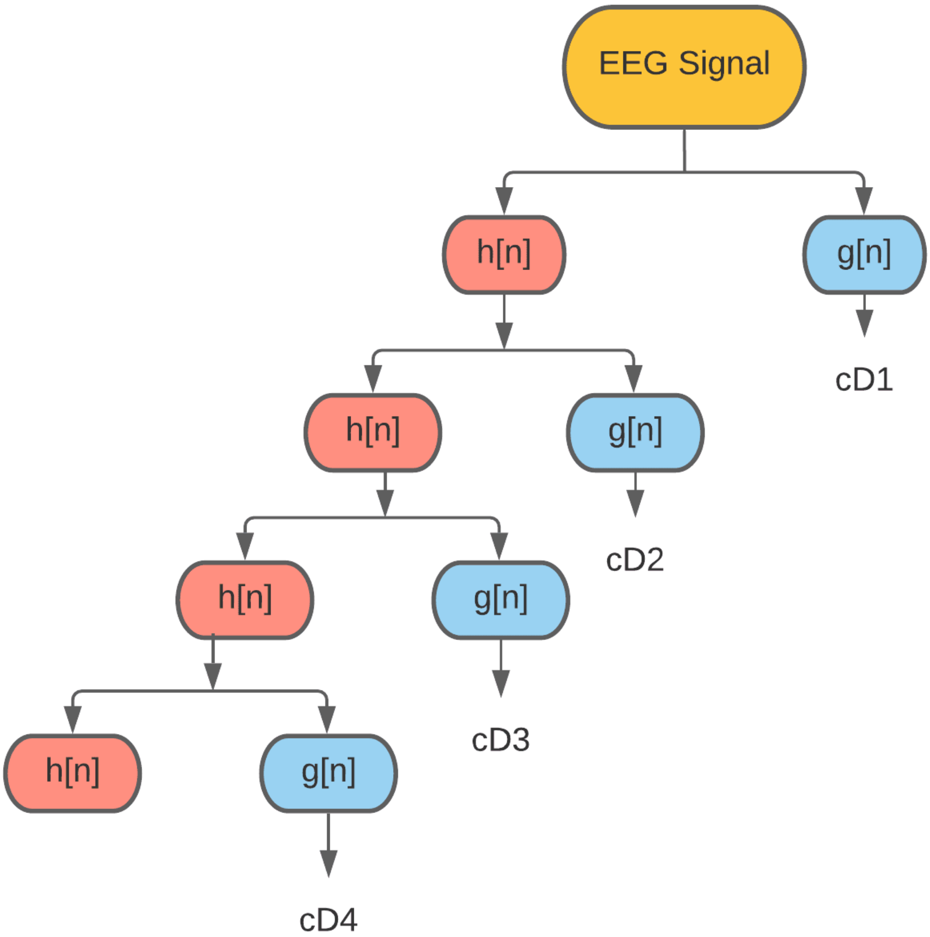


Figure 5.2: Multi-level Signal Decomposition by DWT with the low pass filter and the high pass filter h[n] and g[n]

Primarily, the dataset was preprocessed using sampling rate 128Hz. On this note, as shown in Table 5.1, four frequency bands were found for four detailed coefficients [41]. The next feature extraction tasks were performed on these four detailed coefficients. To reduce the dimension of these detailed coefficient feature vectors and training time of the models, a statistical method named Root Sum Square (RSS) was used as described in chapter Background Study. Hence, in this method four RSS features were extracted for each channel eeg data. According to our literature knowledge so far, this particular statistical method was not used on emotion recognition feature extraction tasks using EEG signals.
For generating more information from these detailed coefficients following steps are performed to add four more features for a particular eeg channel data.

a) RSS ratio of cD1 and cD2 was used.

b) RSS ratio of cD3 and cD4 was used.
c) RSS absolute difference between cD1 and cD2 was used.
d) RSS absolute difference between cD3 and cD4 was used.

Finally, the extracted feature dimension for a particular participant was 8x32x40 (features x channels x videos/trials), after reshaping the dimension which is 256x40 (features x videos/trials) for all EEG channels.
On the other hand, the feature vector for ten selected channels after reshaping the dimension is 80x40 (features x videos/trials).

5.2 Channel Selection

In the chapter Related Work, different methodologies regarding EEG channel selection are discussed.The cortical spectral power activity holds an important statistical significance in analyzing the EEG signals. However, this activity could vary consistently across the spot depending on the familiarity of the videos to the participants[50]. Therefore, the proposed methodology uses average band power and aims to select the most effective channels.In this work, a channel selection technique using average band power of a selected frequency band is proposed. First of all, the required band was selected before applying the channel selection algorithm.

5.2.1 Band Selection

Different plots of Frequency vs Power Spectral Density (PSD) of preprocessed eeg signals were observed. It is found that the effective band for eeg channel selection would be (4-45)Hz as shown in Figure 5.3 and Figure 5.4.

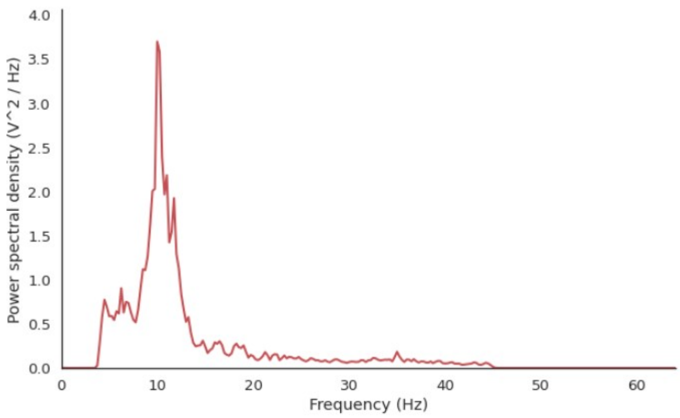


Figure 5.3: Welch's periodogram of P3 Channel for twentieth participant and eleventh trial/video

5.2.2 Selection Algorithm

After band selection, the average band power of a particular eeg channel data is calculated using SciPy welch method[48] in python. For individual participants a python dictionary was made for a particular trial/video where each channel name

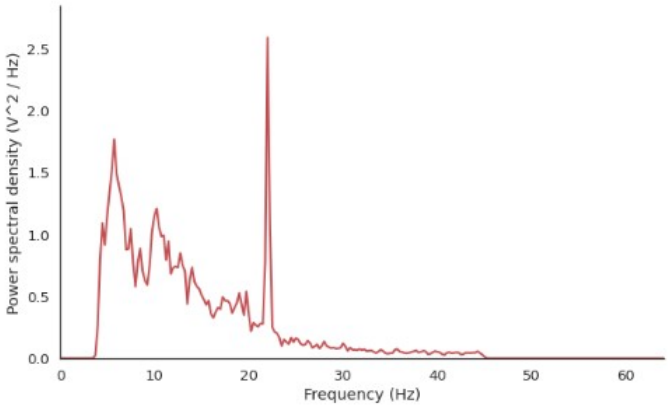


Figure 5.4: Welch's periodogram of FP1 Channel for first participant and first trial/video

Wavelet coefficients	Frequency ranges(Hz)
cD1	32-64
cD2	16-32
cD3	8-16
cD4	4-8

Table 5.1: Frequency ranges related to different wavelet coefficients

was as key and the average band power was as value.Now for all participants all the dictionaries were stored in a python list. Each dictionary was sorted in descending order. Then, the first ten channels were selected for each dictionary which means selecting channels on the basis of maximum average band power.

Finally, ten most appeared channels on this 1-10 ranking were selected for the emotion classification task. The selected channels were Fp2, T7, AF4, AF3, T8, FP1, F7, F8, P7, O1. For better observation, the flowchart of this channel selection technique is given in Figure 5.5 and Figure 5.6 for the selected channel locations based on 10-20 standard electrode location system applied on biosemi head cap setup.

5.3 Experimentation

The following experiments were performed to test the efficiency of the algorithms in terms of both all EEG channels and selected EEG channels. First of all k-fold cross validation technique was applied to find out the perfect train set and test set ratio for classification.

5.3.1 Cross validation Scores

Train set and test set splitting is an important task in machine learning for gaining accuracy. In this study, different combinations of k-folds were tested for the models K-Nearest Neighbour, Random Forest, Bagging, Extra Trees and AdaBoost. The criteria for testing these algorithms was test accuracy. The highest mean score was found 0.5002 for five fold cross validation which is depicted in Table 5.2. Hence,

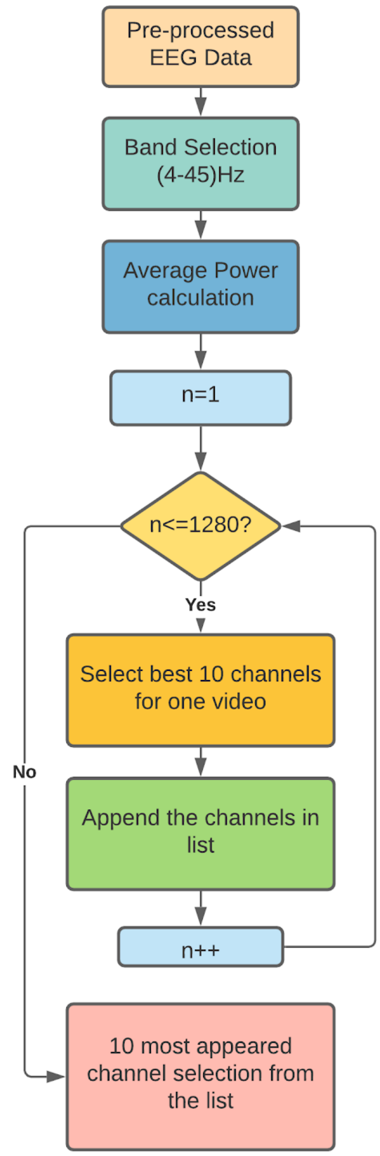


Figure 5.5: Channel selection algorithm

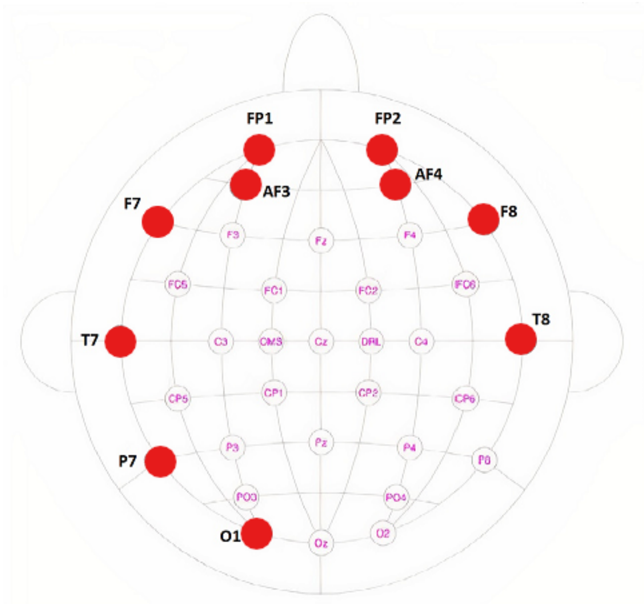


Figure 5.6: Selected channel locations based 10-20 standard electrode location system

four-fifths of the data was selected as the train set and one-fifth of the data was selected as the test set.

k-folds	Mean scores
5	0.5002
10	0.4572
15	0.4294

Table 5.2: k-fold cross validation scores

5.3.2 Performance Criteria

Test Accuracy was used in this study to evaluate the performance of the machine learning algorithms. It is a standard criteria which is frequently used in literature in terms of machine learning classification tasks.

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} \times 100\%$$

Where TP is True Positive which means the model correctly predicts the positive class and TN is the True Negative which indicates the model correctly predicts the negative class. Furthermore, False Positive (FP) refers that the model incorrectly predicts the positive class, whereas False Negative (FP) means the model incorrectly predicts the negative class [19]

5.3.3 Classification Algorithms

After that different classification algorithms of machine learning were used to compare their performance on the basis of test accuracy between all channel eeg data and selected channel eeg data. These are K-Nearest Neighbour, Random Forest, Bagging, Extra Trees, AdaBoost, and MLP classifier.

K-Nearest Neighbour(KNN)

In this work, the best values for k which is the number of neighbours is determined by plotting the test accuracy where k in range [1,21]. In the case of all eeg channels, the highest test accuracy was found for k=6 for valence classification and k=11 for arousal classification. Again, in the case of selected eeg channels, the highest test accuracy was found for k=6 for valence classification and k=3 for arousal classification as shown in Figure 5.7, 5.8, 5.9 and 5.10. After that, Minkowski, Manhattan, Euclidean, and Hamming distance were tried to get the best performance and Euclidean distance worked well with these features. Finally, the algorithm was chosen as 'auto' after giving trials with 'ball tree', 'kd tree' and 'brute force' algorithms.

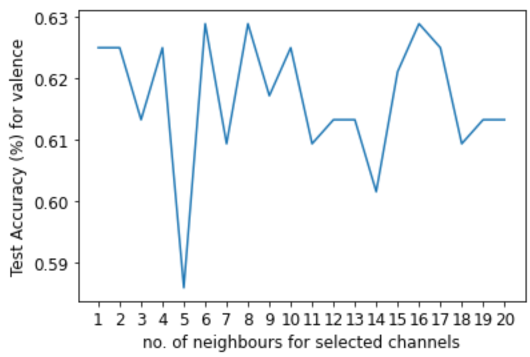


Figure 5.7: Neighbour value selection to classify valence for selected channels

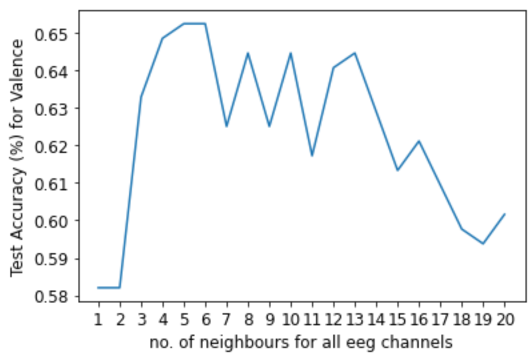


Figure 5.8: Neighbour value selection to classify valence for all channels

Random Forest(RF)

In this experiment, the best criterion for the RF classifier was found as 'gini'. The

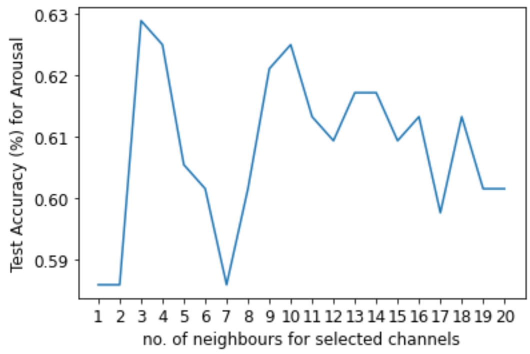


Figure 5.9: Neighbour value selection to classify arousal for selected channels

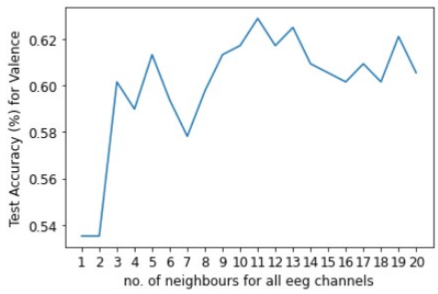


Figure 5.10: Neighbour value selection to classify arousal for all channels

other criterion was 'entropy', which could not obtain a better result compared to 'gini'. Also, the number of estimators (trees in the forest) was set as 150 as this value achieved the most accurate result. Finally, the number of random states was ten, which was found to be the optimum value for the experiment when compared to the other values.

Bagging

In this classifier, the Decision Tree Classifier performed better than other machine learning classifiers. Hence, DT was selected as the best estimator algorithm for bagging. Also, different values for random state were applied to get the efficient result. Random state value 10 was best among all other combinations. Finally, the number of estimators was 150 which performed very well in terms of other combinations.

Extra Trees

In Extra Trees (ET) classifiers the number of features to consider when looking for the best split (max features) was set as log2. Setting this value to log2 yielded the best result. Additionally, the number of random states was eight. Other values were also tested but the pre-mentioned value obtained the higher accuracy. Finally, the number of estimators was set as 150, which was in fact the same as the RF classifier. This combination of parameters performed better than other combinations that were tested.

AdaBoost

The AdaBoost classifier uses a base estimator to build the boosted ensemble. DecisionTreeClassifier was used as the base estimator for this classifier. Also, the number of random states was set as ten, which was the same as in case of the RF classifier. Similarly, the number of estimators was 150, which was similar to that value of the ET classifier. Finally, when compared to the results yielded by the other combinations, this combination of parameters showed better performance.

Multilayer Perceptron(MLP)

In this classifier, the activation function used was 'logistic'. Other options for activation functions were 'identity', 'tanh', and 'relu'. Also, the value for hidden layer sizes was set to 15. Additionally, the solver used for weight optimization was 'adam'. 'lbfgs' and 'sgd' were the other weight optimization solver options. Additionally, the number of random states was set as 15. Furthermore, the maximum number of iterations was set as 800 for the solver. The performance achieved by this combination of parameters outperformed the other combinations. As a result, this particular combination was selected.

Chapter 6

Results Analysis

In this study, test accuracy of the algorithms was the main criteria for evaluating the performance. As comparative analysis between all 32 EEG channels and selected 10 EEG channels is the main focus of this study, this work showed a notable difference while experimenting with selected channels as shown in Figure 6.1 and 6.2. In this experiment, different machine learning classifiers, ensemble methods and a neural network classifier were used to depict the efficiency of using selected channels.

6.1 Selected Channels

The highest test accuracy in this selected channel experiment was obtained by Bagging and ExtraTrees algorithm which was 67.58% for valence classification and MLP classifier performed better than the other algorithms with the highest test accuracy 63.67% for arousal classification.

6.2 All EEG Channel

Using all the EEG channels, the highest test accuracy in this selected channel experiment was obtained by KNN and ExtraTrees algorithm which was 65.23% for valence classification. On the other hand, KNN and Random Forest classifier performed better than the other algorithms with the highest test accuracy 62.89% for arousal classification.

From the above analysis it is obvious that the highest test accuracy for both valence and arousal classification is obtained by using the selected EEG channels. On this note, this proposed channel selection technique is much efficient for classification of valence and arousal in terms of test accuracy.

6.3 Comparative Study

Many literatures related to this emotion recognition were studied to portray the validation of this work. In terms of MLP classifier test accuracy, this proposed work outperformed some of the recent works with the same database as shown in Figure 6.3. By using selected channels both training time and test accuracy were significantly enhanced compared to using all EEG channels and single channel

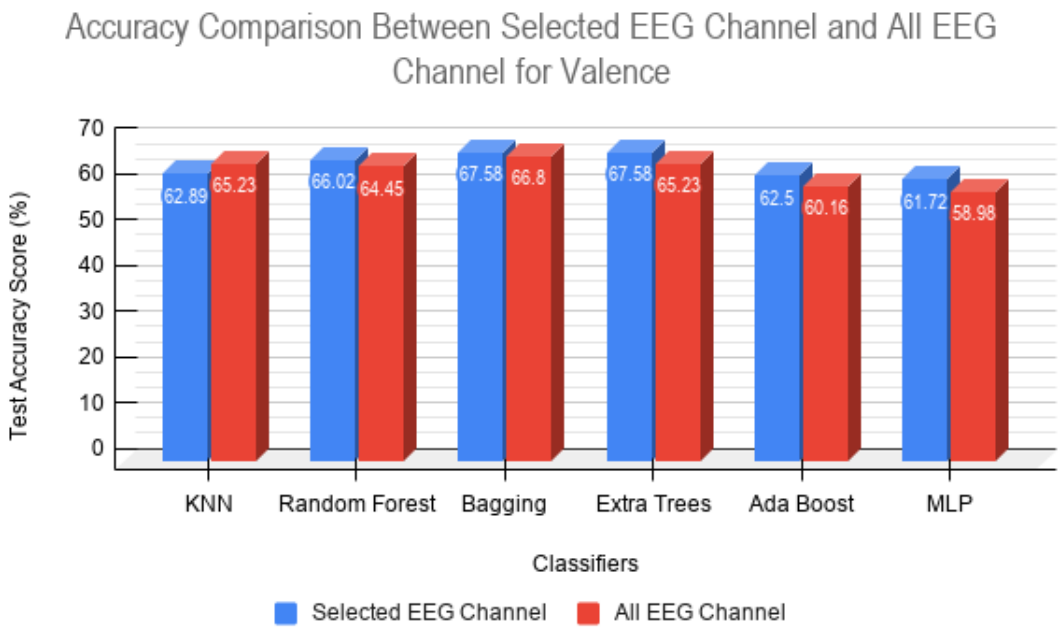


Figure 6.1: Test accuracy comparison for Valence

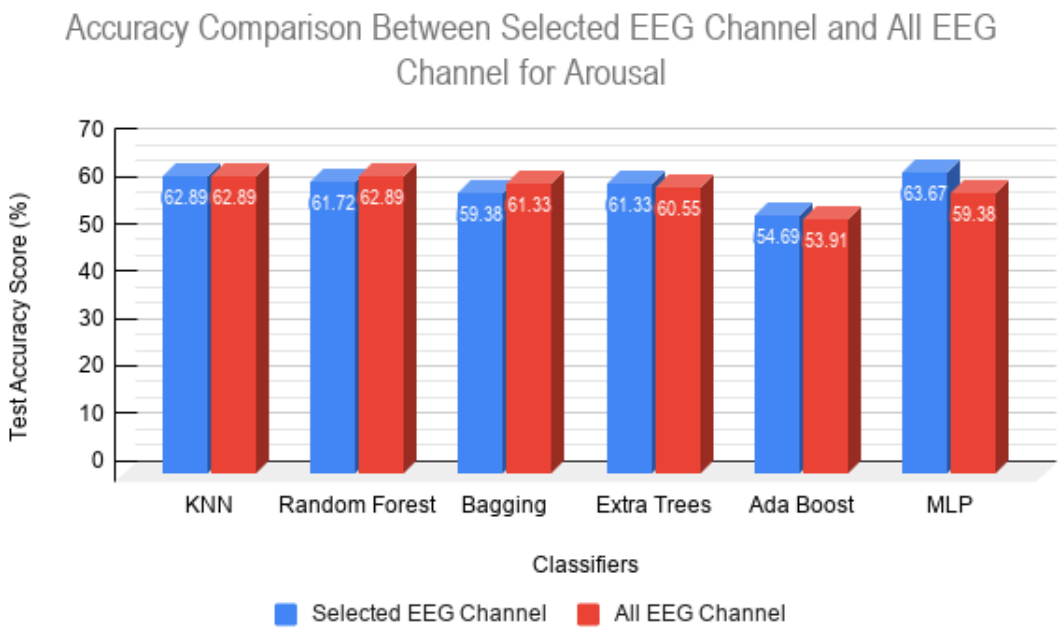


Figure 6.2: Test accuracy comparison for Arousal

only. Thus, the proposed channel selection technique in this study gives an optimal solution in terms of both training time and test accuracy.

Article	Database	MLP Classifier Accuracy (%)	Remarks
Zhang et al.[47]	DEAP	57.67 (classified four emotions with 12 channels) 58.75 (classified four emotions with all 32 channels)	As all 32 eeg channels were used, more time would be required while training[40]
Jirayucharoensak et al.[39]	DEAP	Classified two emotions (Valence accuracy is 53, Arousal accuracy is 52)	To classify emotions more work on mapping is required[40]
Pandey et al.[40]	DEAP	58.5 (Classified two emotions with single channel (F4) only)	Though training time is less but more accuracy can be obtained by adding more selected channels.
Proposed	DEAP	Classified two emotions using selected 10 eeg channels (Test accuracy for valence is 61.72 and test accuracy for arousal is 63.67)	As only selected channels were used, both training time and test accuracy improved significantly, compared to using all channels and single channel only.

Figure 6.3: Comparative analysis with recent studies

Chapter 7

Conclusion and Future Works

The aim of this study is to classify different emotions on the basis of audiovisual stimuli with the help of EEG signals and to compare the efficiency in terms of test accuracy of different machine learning algorithms using selected EEG channels and all EEG channels.

In this work, DEAP database was used which is publicly available to classify four types of emotions namely valence, arousal, dominance and likings. 32 participants preprocessed EEG data was taken to classify valence and arousal.

For the feature extraction task, DWT wave decomposition and Root Sum Square(RSS) were used and different machine learning classifiers, ensemble methods and MLP were used to classify two emotions.

A channel selection algorithm was proposed to determine the more effective channels related to emotion recognition. The best 10 selected channels with the highest average power of a selected band (4-45)Hz were found after applied to the channel selection algorithm and the selected channels were Fp2, T7, AF4, AF3, T8, FP1, F7, F8, P7, O1.

The performance of selected channels in terms of test accuracy was higher compared to all EEG channels as described in the results analysis section.

Furthermore, the validation of this study was found while comparing with some recent works on this database. Hence, this proposed channel selection technique was found more efficient in terms of emotion classification on valence and arousal. Finally, it can be depicted that, with selected channels the training time is saved and performance can be enhanced on this emotion recognition study.

In future studies, channel selection can be done using source localization technique of EEG signals. Also, other physiological signals like blood pressure, body temperature, galvanic skin response, respiratory rate and other signals can be used to analyse their role in determining emotion recognition.

7.1 Comments of the faculty members during Q/A session

01. What are the specific bands here? Asked about the frequency range alpha, beta, theta, gamma bands?
02. What is the channel selection methodology here? can you please explain?
03. Why KNN classifier performed better for all channels?
04. What is TP, TN, FP, FN here in the accuracy equation?

Bibliography

[1] Petrushin, V. (1999, November). Emotion in speech: Recognition and application to call centers. In Proceedings of artificial neural networks in engineering (Vol. 710, p. 22).

[2] Anderson, K., & McOwan, P. W. (2006). A real-time automated system for the recognition of human facial expressions. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 36(1), 96-105.

[3] Pantic, M., & Rothkrantz, L. J. M. (2000). Automatic analysis of facial expressions: The state of the art. IEEE Transactions on pattern analysis and machine intelligence, 22(12), 1424-1445.

[4] Yin, Z., Zhao, M., Wang, Y., Yang, J., & Zhang, J. (2017). Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. Computer methods and programs in biomedicine, 140, 93-110.

[5] Wang, X. W., Nie, D., & Lu, B. L. (2014). Emotional state classification from EEG data using machine learning approach. Neurocomputing, 129, 94-106.

[6] Petrantonakis, P. C., & Hadjileontiadis, L. J. (2011). A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG-based emotion recognition. IEEE Transactions on information technology in biomedicine, 15(5), 737-746.

[7] Li, X., Hu, B., Zhu, T., Yan, J., & Zheng, F. (2009, October). Towards affective learning with an EEG feedback approach. In Proceedings of the first ACM international workshop on Multimedia technologies for distance learning (pp. 33-38).

[8] Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE, 77(2), 257-286.

[9] Torres-Valencia, C. A., Garcia-Arias, H. F., Lopez, M. A. A., & Orozco-Gutiérrez, A. A. (2014, September). Comparative analysis of physiological signals and electroencephalogram (EEG) for multimodal emotion recognition using generative models. In 2014 XIX Symposium on Image, Signal Processing and Artificial Vision (pp. 1-5). IEEE.

[10] Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., ... & Patras, I. (2011). Deap: A database for emotion analysis; using physiological signals. IEEE transactions on affective computing, 3(1), 18-31.

[11] Xing, B., Zhang, H., Zhang, K., Zhang, L., Wu, X., Shi, X., ... & Zhang, S. (2019). Exploiting EEG signals and audiovisual feature fusion for video emotion recognition. IEEE Access, 7, 59844-59861.

[12] Qing, C., Qiao, R., Xu, X., & Cheng, Y. (2019). Interpretable emotion recognition using EEG signals. IEEE Access, 7, 94160-94170.

[13] Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry, 25(1), 49-59.

[14] Allen, J. J., Coan, J. A., & Nazarian, M. (2004). Issues and assumptions on the road from raw signals to metrics of frontal EEG asymmetry in emotion. Biological psychology, 67(1-2), 183-218.

[15] Al-Nafjan, A., Hosny, M., Al-Ohali, Y., & Al-Wabil, A. (2018, January). Recognition of affective states via electroencephalogram analysis and classification. In International Conference on Intelligent Human Systems Integration (pp. 242-248). Springer, Cham.

[16] Russell, J. A. (1980). A circumplex model of affect. Journal of personality and social psychology, 39(6), 1161.

[17] Murugappan, M., Juhari, M. R. B. M., Nagarajan, R., & Yaacob, S. (2009). An Investigation on visual and audiovisual stimulus based emotion recognition using EEG. International Journal of Medical Engineering and Informatics, 1(3), 342.

[18] Doma, V., & Pirouz, M. (2020). A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals. Journal of Big Data, 7(1), 1-21.

[19] Baratloo, A., Hosseini, M., Negida, A., & El Ashal, G. (2015). Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity. Emergency (Tehran, Iran), 3(2), 48–49.

[20] Deng, Y., Wu, F., Du, L., Zhou, R., & Cao, L. (2019, March). EEG-Based Identification of Latent Emotional Disorder Using the Machine Learning Approach. In 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC) (pp. 2642-2648). IEEE.

[21] Romeo, L., Cavallo, A., Pepa, L., Berthouze, N., & Pontil, M. (2019). Multiple Instance Learning for Emotion Recognition using Physiological Signals. IEEE Transactions on Affective Computing.

[22] Kim, J., & André, E. (2008). Emotion recognition based on physiological changes in music listening. IEEE transactions on pattern analysis and machine intelligence, 30(12), 2067-2083.

[23] Ko, K. E., Yang, H. C., & Sim, K. B. (2009). Emotion recognition using EEG signals with relative power values and Bayesian network. International Journal of Control, Automation and Systems, 7(5), 865.

[24] Wioleta, S. (2013, June). Using physiological signals for emotion recognition. In 2013 6th International Conference on Human System Interactions (HSI) (pp. 556-561). IEEE.

[25] Koelstra, S. (2020). DEAP: A Dataset for Emotion Analysis using Physiological and Audiovisual Signals. Retrieved 6 October 2020, from http://www.eecs.qmul.ac.uk/mmv/datasets/deap/readme.html

[26] Medicine.mcgill.ca. n.d. *EEG*. [online] Available at: <https://www.medicine.mcgill.ca/physio/vlab/biomed_signals/eeg_n.htm#:˜text=The%20electroencephalogram%20(EEG)%20is%20a,measured%20in%20microvolts%20(mV)>.

[27] Amo, C., de Santiago, L., Barea, R., López-Dorado, A., & Boquete, L. (2017). Analysis of Gamma-Band Activity from Human EEG Using Empirical Mode Decomposition. *Sensors (Basel, Switzerland)*,*17*(5), 989. https://doi.org/10.3390/s17050989

[28] Citron, F. M., Gray, M. A., Critchley, H. D., Weekes, B. S., & Ferstl, E. C. (2014). Emotional valence and arousal affect reading in an interactive way: neuroimaging evidence for an approach-withdrawal framework. *Neuropsychologia*, *56*(100), 79-89. https://doi.org/10.1016/j.neuropsychologia.2014.01.002

[29] Birkett, A., 2021. *Valence, Arousal, And How To Kindle An Emotional Fire — CXL*. [online] CXL. Available at: <https://cxl.com/blog/ valence-arousal-and-how-to-kindle-an-emotional-fire/#:˜text=Though%20the%20two%20terms%20are,of%20winning%20a%20big%20game>.

[30] Amin, H., Mumtaz, W., Subhani, A., Saad, M. and Malik, A., 2017. Classification of EEG Signals Based on Pattern Recognition Approach. *Frontiers in Computational Neuroscience*, 11.

[31] Brownlee, J., 2021. *A Gentle Introduction To K-Fold Cross-Validation*. [online] Machine Learning Mastery. Available at: https://machinelearningmastery.com/k-fold-cross-validation/#:˜text=Cross%2Dvalidation%20is%20a%20resampling,is%20to%20be%20split%20into.

[32] Scikit-learn.org. 2021. *3.1. Cross-Validation: Evaluating Estimator Performance - Scikit-Learn 0.24.0 Documentation*. [online] Available at: <https://scikit-learn.org/stable/modules/cross_validation.html>.

[33] J. X. Chen, P. W. Zhang, Z. J. Mao, Y. F. Huang, D. M. Jiang and Y. N. Zhang, "Accurate EEG-Based Emotion Recognition on Combined Features Using Deep Convolutional Neural Networks," in *IEEE Access*, vol. 7, pp. 44317-44328, 2019, doi: 10.1109/ACCESS.2019.2908285.

[34] W. Zheng and B. Lu, "Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks," in *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 3, pp. 162-175, Sept. 2015, doi: 10.1109/TAMD.2015.2431497.

[35] Torres, E. P., Torres, E. A., Hernández-Álvarez, M., & Yoo, S. G. (2020). EEG-Based BCI Emotion Recognition: A Survey. *Sensors* , *20*(18), 5083. MDPI AG. Retrieved from http://dx.doi.org/10.3390/s20185083

[36] Khosrowbadi R, Quek HC, Wahab A, Ang KK (2010) EEG based emotion recognition using self-organizing map for boundary detection. In: International conference on pattern recognition, pp 4242-4245

[37] Alotalby, T. , EI-Samie, F. , Alshebelll, S. , & Ahmed,I.(2015). A review of channel selection algorithms for EEG signal processing. EURASIP Journal on Advances in Signal Processing, 2015 (1). doi:10.1186/s13634-015-0251-9

[38] Gou, Jianping & Du, Lan & Zhang, Yuhong & Xiong, Taisong. (2011). A New Distance-weighted k -nearest Neighbor Classifier. J. Inf. Comput. Sci.. 9.

[39] Jirayucharoensak, Suwicha & Pan-ngum, Setha & Israsena, Pasin. (2014). EEG-Based Emotion Recognition Using Deep Learning Network with Principal Component Based Covariate Shift Adaptation. TheScientificWorldJournal. 2014. 627892. 10.1155/2014/627892.

[40] Pandey P., Seeja K.R. (2019) Emotional State Recognition with EEG Signals Using Subject Independent Approach. In: Mishra D., Yang XS., Unal A. (eds) Data Science and Big Data Analytics. Lecture Notes on Data Engineering and Communications Technologies, vol 16. Springer, Singapore. https://doi.org/10.1007/978-981-10-7641-1_10

[41] Ozerdem, Mehmet & Polat, Hasan. (2017). Emotion recognition based on EEG features in movie clips with channel selection. Brain informatics. 4. 10.1007/s40708-017-0069-3.

[42] Choudhury, K., 2021. *Deep Neural Multilayer Perceptron (MLP) With Scikit-Learn.* [online] Medium. Available at: <https://towardsdatascience.com/deep-neural-multilayer-perceptron-mlp-with-scikit-learn-2698e77155e>.

[43] Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. R news, 2(3), 18-22.

[44] Breiman, L. (1996). Bagging predictors. Machine Learning, 24(2), 123-140. doi:10.1007/bf00058655

[45] Zhu, Ji & Rosset, Saharon & Zou, Hui & Hastie, Trevor. (2006). Multi-class AdaBoost. Statistics and its interface. 2. 10.4310/SII.2009.v2.n3.a8.

[46] P. Geurts, D. Ernst., and L. Wehenkel, " Extremely randomized trees", Machine Learning, 63(1), 3-42, 2006.

[47] Zhang, J., Chen, M., Zhao, S., Hu, S., Shi, Z., & Cao, Y. (2016). ReliefF-Based EEG Sensor Selection Methods for Emotion Recognition. Sensors, 16(10), 1558. doi:10.3390/s16101558

[48] Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., . . . van Mulbregt, P. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods. doi:10.1038/s41592-019-0686-2

[49] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. & Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research, 12, 2825-2830.

[50] Vecchiato, G., Astolfi, L., Tabarrini, A., Salinari, S., Mattia, D., Cincotti, F., ... & Babiloni, F. (2010). EEG analysis of the brain activity during the observation of commercial, political, or public service announcements. Computational intelligence and neuroscience, 2010.