EMOTION RECOGNITION FROM EEG SIGNALS USING MACHINE LEARNING MODELS

A PROJECT REPORT

submitted by

AKSHAY K R
(Reg. No. TKM20EEII03)

to

the APJ Abdul Kalam Technological University
in partial fulfillment of the requirements for the award of the Degree

of

Master of Technology
in
Electrical and Electronics Engineering

with specialisation in

Industrial Instrumentation and Control

Department of Electrical and Electronics Engineering
TKM College of Engineering
KOLLAM - 691 005
KERALA

JULY 2022

DECLARATION

I undersigned hereby declare that the project report entitled "**Emotion recognition from EEG signals using machine learning models**", submitted for partial fulfillment of the requirements for the award of degree of Master of Technology in Electrical and Electronics Engineering with specialisation in Industrial Instrumentation and Control, APJ Abdul Kalam Technological University, Kerala is a bonafide work done by me under supervision of *Dr. Muhammed Shanir P P*, Supervisor, Assistant Professor, Department of Electrical and Electronics Engineering. This submission represents my ideas in my own words and where ideas or words of others have been included, I have adequately and accurately cited and referenced the original sources. I also declare that I have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Kollam
July 01, 2022

AKSHAY K R

DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

TKM COLLEGE OF ENGINEERING

KOLLAM - 691 005

CERTIFICATE

This is to certify that the report entitled " **Emotion recognition from EEG signals using machine learning models** " submitted by **AKSHAY K R** , **(Reg. No. TKM20EEII03)** of fourth semester to the APJ Abdul Kalam Technological University in partial fulfillment of the requirements for the award of the Degree of Master of Technology in Electrical and Electronics Engineering with specialisation in Industrial Instumentation and Control, is a bonafide record of the project work done by him under my guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

Dr. Muhammed Shanir P P
Project Supervisor
Assistant Professor
Department of EEE
TKM College of Engineering

Prof. Sumayya Jaleel
Project Coordinator
Assistant. Professor
Department of EEE
TKM College of Engineering

Prof. Shanavas T N
PG Coordinator
Associate Professor
Department of EEE
TKM College of Engineering

Dr. Sabeena Beevi
Head of Department
Associate Professor
Department of EEE
TKM College of Engineering

Acknowledgment

A lot of effort and hard work has been put into this project in course of its presentation. However, it would not have been possible without the kind support and help of many individuals and other sources. I would like to extend my sincere thanks to all of them. I take this opportunity to express my deep sense of gratitude and sincere thanks to all who helped me to complete this project report successfully.

I thank God Almighty for his infinite grace and help without which this project would not have reached its completion.

I express my special thanks to *Dr. T A Shahul Hameed*, Honourable Principal of TKM College of Engineering for his timely support extended throughout the work.

I thank *Dr. Sabeena Beevi*, Head of the Department, Department of Electrical and Electronics Engineering and *Prof. Shanavas T N*, Assocoiate Professor, PG Co-ordinator, Department of Electrical and Electronics Engineering for their support and cooperation.

I am deeply indebted to *Prof. Sumayya Jaleel*, Assistant Professor, Project Coordinator, Department of Electrical and Electronics Engineering for her care, help and support.

I am greatly thankful to my supervisor *Dr. Muhammed Shanir P P* , Assistant Professor, Department of Electrical and Electronics Engineering for his motivation, supervision, assistance and helpful suggestions for my entire journey of this work.

I am deeply indebted to *Dr. Imthias Ahamad T P*, Professor, Department of Electrical and Electronics Engineering, *Prof. Amal A*, Assistant Professor, Department of Electrical and Electronics Engineering and *Prof. Sumod Sundar*, Assistant Professor, Centre for Artificial Intelligence, for their excellent support, positive criticism and valuable comments.

I would like to extend my thanks to *Mr. Lijin K L*, *Mr. Alan Dutt*, *Mr. Sanjay Steephen*, *Mr. Abdul Basith*, *Mr. Farzin Ahamed*, *Dr. Manesh Rasheed* and *Dr. Shailesh*, they deserves a mention for the wonderful friend they have been, for being my stress relief throughout this lengthy journey.

The list of acknowledgement would be woefully incomplete without my family. My 'Achan' and 'Amma', my sister, *Akshara* and my uncle. Your boundless love and encouragement assured me a peaceful, conducive atmosphere to work in, without which I wouldn't be at the finish line now, holding my work with pride. You've been my greatest source of strength and inspiration all these years, holding my back all the time, which can never be repaid with a thank you. I only offer my love.

AKSHAY K R

Abstract

The objective of this work is to identify emotion from EEG signals, that represent the brain activity of individuals. With the rapid advancement of machine learning algorithms and numerous real-world applications of brain-computer interface for regular people, emotion categorization from EEG data has recently gained a lot of attention. Researchers previously have little knowledge of the specific interactions between distinct EEG characteristics and various emotional states. The computer may peer into the user's head to assess the mental state with the use of EEG-based emotion identification. This work is executed with DEAP dataset with 32 channels for EEG recording, it achieves a better classification accuracy with different machine learning models. In the subject wise experiment an average best accuracies of 91.26%, 92.83% and 94.99%, and in the subject dependent experiment, the best accuracies of 78.5%, 82.77% and 92.73% is obtained for the random forest, XGBoost, and KNN classifiers respectively.

Contents

List of Tables

List of Figures

Abbreviations

AI	Artificial Intelligence
ANN	Artificial Neural Network
API	Application Programming Interface
CNN	Convolution Neural Network
DASM	Differential Asymmetry
DNN	Deep Neural Network
ECG	Electrocardiogram
EEG	Electroencephalogram
EOG	Electrooculogram
ERO	Event Related Oscillations
ERP	Event Related Potential
ESL	Extreme Sparse Learning
FFT	Fast Fourier Transform
FN	False Negative
FP	False Positive
KNN	K - Nearest Neighbor
LBP	Local Binary Pattern
LSTM	Long Short Term Memory
NLD	Normalised Length Density
NSI	Non Stationary Index

PSD Power Spectral Density

RASM Rational Asymmetry

RF Random Forest

SVM Support Vector Machine

TN True Negative

TP True Positive

Notations

E_i	Evaluation result of i^{th} iteration
$l\left(\hat{y}_i, y_i\right)$	Training loss function
$\Omega\left(f_k\right)$	Regularization term
$p\left(x\right)$	Probability of x point
x_i	x coordinate of i^{th} point
y_i	y coordinate of i^{th} point
$\hat{y}_i{}^t$	Prediction value at iteration t

Chapter 1

INTRODUCTION

Emotions are the mental state brought on by neuro-physical changes, variously associated with thoughts, feelings, behavioural responses, and a degree of pleasure or displeasure. Simply, the emotion can be determined as the complex combination of thinking, feeling and behaviour [1].

Emotions are important symbols of human intelligence; as an important symbol of intelligence of artificial intelligence is a machine that can understand human emotions. The influence of emotions in our life is very huge as it directly relates with our thinking ability, process of handling situation, actions and communication etc. [2]. For the quality service and better delivery of those services by a machine, estimation and analysis of human emotions are necessary.

The current approaches to emotion recognition rely heavily on data, and it can often be difficult to get the annotated data required to train machine learning algorithms [3]. The following datasets are available for the task of classifying various emotion types from multimodal sources such as texts, audio, videos, or physiological signals: HUMAINE [4], SEMAINE [5], Belfast database [6], IEMOCAP [7], DEAP [8], DREAMER [9], eNTERFACE[10] etc.

These days, a growing number of intelligent systems use emotion detection models to enhance their communication with people. This is significant because it enables systems to give high-quality services and change their answers and behavioural patterns in reaction to human emotions, making interactions feel more natural [11]. Emotion recognition is utilised nowadays for many things, some of which people don't even imagine. The following are some instances where emotion recognition is advantageous:

- There are businesses that employ artificial intelligence (AI) -based Human Resource (HR) assistants with emotion recognition Application Programming Interface (API) capabilities. By analysing voice modulations, facial expressions, and keywords, the system aids in identifying whether the applicant is sincere and genuinely interested in the position. A report is then generated for the human recruiters' final evaluation.

- Systems have been introduced recently and are installed in customer service centres. The customer's emotions can be compared before and after entering the centre using artificial intelligence-enhanced cameras to gauge their level of satisfaction with the service they received. Additionally, if the score is poor, the system can give the staff members advice on how to raise the level of service.

- Businesses are now employing emotion recognition to forecast their financial results based on audience emotional reactions. Apple also introduced Animoji, a new feature in their iPhones that allows an emoji to replicate a person's facial emotions.

- Video games are evaluated in order to gather user feedback and ascertain whether the companies' objectives have been met. The feelings a user is having in the moment can be recognised using emotion recognition during these testing phases, and their input can be used to improve the final product.

- Face emotion recognition is definitely being used by the healthcare sector nowadays. They use it to determine whether a patient requires medication or to help doctors decide which patients to see first. etc.

The emotion classification techniques are broadly divided into two, using physiological signals and non-physiological signals [12]. Physiological signals includes electrocardiogram (ECG) [13], electroencephalogram (EEG) [12], [13], pulse variability [14] etc., where as the non-physiological signals includes the facial expression [15], speech signals [16] etc. The non-physiological methods are all based on the emotions expressed externally and may not be able to identify the innermost feeling, so this techniques are not reliable as the other.

The first electroencephalogram (EEG) was recorded in 1924 by neurologist Hans Berger using an electrode taken from a human brain. The EEG revealed wave patterns representing the electrical signal activities of human brains. The EEG signal can now be utilised to more accurately interpret states of the cognitive process and behaviour, such as selective attention, working memory, and mental computations, thanks to improvements in processing power and EEG technology [17]. An electroencephalogram (EEG) is essentially a device that uses tiny metal discs (electrodes) placed to the scalp to detect the electrical activity in a human brain.

EEG is a key term commonly used to describe the long-term recording of the brain's spontaneous electrical activity using a number of scalp electrodes [18]. Diagnostic applications typically concentrate on the spectrum content of EEG or event-related potentials. The first looks into potential fluctuations that are tied to a specific event, like "stimulus onset" or "button press." The latter examines the different cerebral oscillations (often referred to as "brain waves") that can be seen in the frequency domain of EEG recordings.

It is mentioned that the EEG signals are measured or recorded using the help of electrodes, but we cannot arbitrarily place the electrodes, there is some specific set of rules for that [8]. The 10-20 system, often known as the International 10-20 system of electrode placement, is a widely accepted technique for describing and applying the placement of scalp electrodes [19]. The 10-20 system is based on the connection between an electrode's position and the cerebral cortex area beneath it and is mentioned in the Figure 1.2. Each site contains a letter (to indicate the lobe) and a number or another letter (indicating the position of the hemisphere) to identify it. Frontal, Temporal, Central, Parietal, and Occipital are represented by the letters F, T, C, P, and O. Odd numbers (1,3,5,7) refer to the left hemisphere, while even numbers (2,4,6,8) refer to the right. The central electrode is indicated by the z. Also take note that the closer a location is to the centerline, the smaller the number.

As you can see in the Figure 1.3, the brain is divided into different regions Frontal, Temporal, Central, Parietal, and Occipital, the rule for electrode placement for recording the EEG signals make sure that the information from all the part of regions are obtained.

The visual examination of long-term EEG data is tedious and time-consuming, and it can also result in the identification of erroneous critical points. Therefore, a variety of machine learning algorithms have been used to the categorization problem. Based on the rules of the employed algorithms, the machine will perform this task effectively and without restriction.

Figure 1.1: Electrode placement for EEG recording in real life

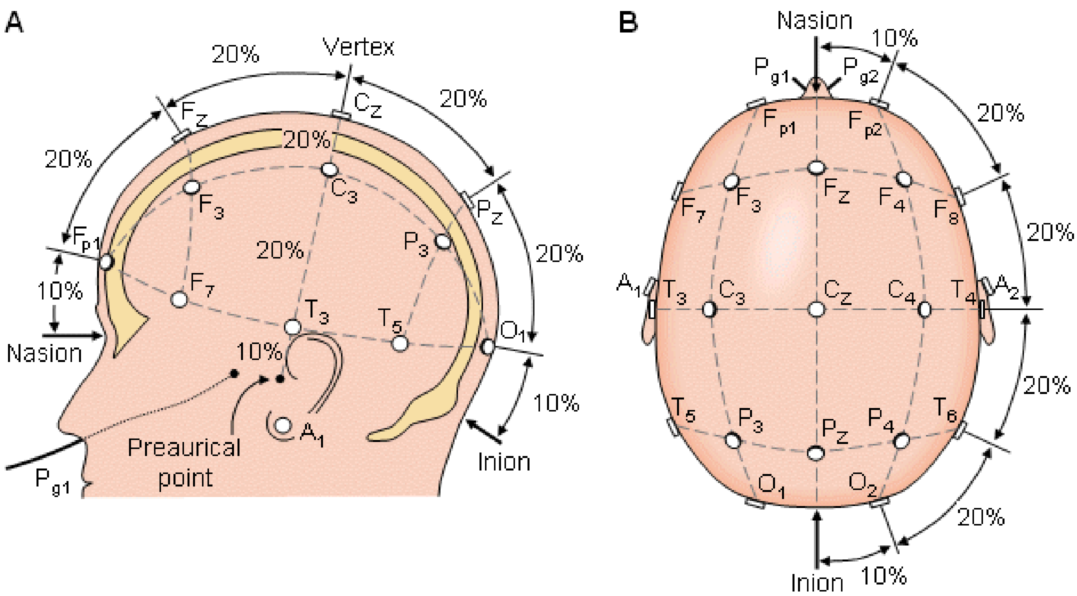

Figure 1.2: The International 10-20 system of electrode placement for EEG recording

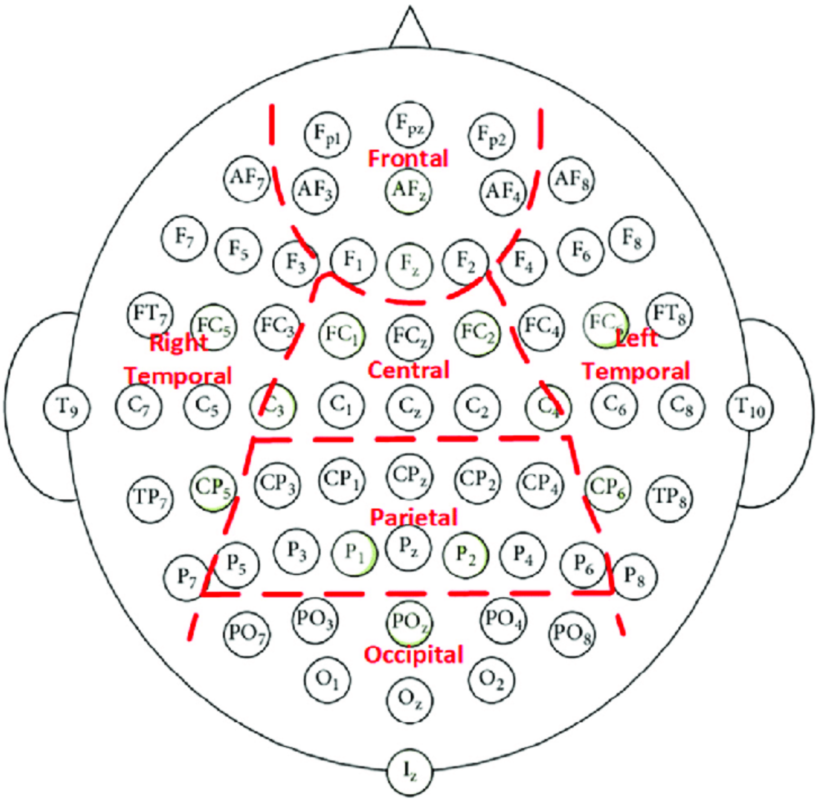

Figure 1.3: Electrodes placed in the different regions of the brain

1.1 Objective

These days, a growing number of intelligent systems use emotion detection models to enhance their communication with people. This is significant because it enables systems to give high-quality services and change their answers and behavioural patterns in reaction to human emotions, making interactions feel more natural.

Emotion recognition is not flawless and has flaws and difficulties. Datasets are labelled by individuals, and different people can read and interpret emotions in different ways. This presents one of the obstacles. Additionally, certain evident visual cues, such as wrinkled brows, can indicate emotions other than anger, and other clues may contain more subtly expressed signs of anger. Also emotion recognition using EEG signals are very challenging as we are dealing with the signals that have the amplitude in the range of micro-volt. The objective of this work is to develop a feature extraction and classification technique that can able to capture the emotions by analysing every instances in the EEG signals. And the model should be able to classify the emotion non-specifically from the channels.

1.2 Organisation of the report

The literature survey done for this work is explained in the chapter 2. In the chapter 3, the dataset used in this work, DEAP dataset, is explained along with the preprocessing and the classification methods. Also the experimental methods used for the classification of the emotion is explained. In chapter 4, the obtained results are mentioned and discussed, and finally in the chapter 5, the conclusion and the future works of this thesis is described.

Chapter 2

LITERATURE REVIEW

Previous chapter deals with the overall information about the emotions and its recognition methods, current chapter focuses on in-depth understanding of methods by analysing the different works that have been conducted by many authors all around the world, on the area of emotion recognition. The works related to thesis such as different classification methods, different classes etc. is been studied in this chapter.

The two main categories of emotion categorization techniques are physiological technique, which uses physiological signals like EEG, ECG, pulse variability etc. and non-physiological technique, which uses non-physiological signals such as facial expression, speech signals etc.The non-physiological approaches are all dependent on how emotions are presented on the outside and may not be able to pinpoint the deepest emotion, making them less trustworthy than the physiological approach.

2.1 Works conducted with non-physiological signals

Table 2.1 contains the investigated information of some works done with non-physiological signals for the emotion classification. Facial features for the emotion classification problem is taken by Shojaeilangari et al. [20] using the dataset CK+, ECK+. Extreme sparse learning method is used in their work and the results are compared with the previous works done in LBP methods and obtained comparitively better accuracy. [22]-[25] also uses the facial features for their emotion classification with different datasets and using various classifiers. Speech signals [21],[26] are used in different works using neural network and machine learning classifier mod-

Table 2.1: Comparison of different works with the facial and speech signals

Authors	Signal	Method	Dataset	Remarks
Shojaeilangari et al.[20]	Facial features	Extreme Sparse Learning (ESL)	CK+,ECK+	Results are compared with the Local Binary Pattern (LBP) based works.
Yuanlu Kuang and Lijuan Li[21]	Speech	Hidden Markov Models (HMM) and Artificial Neural Network (ANN)	Berlin database	
Cruz et al.[22]	Facial features	SVM	CK+, MMI	One-Shot Emotion Score
Li et al.[23]	Facial features	AdaBoost	CK+	Eye-point alignment, Gabor filter, active shape models
Poursaberi et al.[24]	Facial Expression	KNN	MMI	Gauss-Laguerre wavelets and 18 fiducial facial point features
Miao et al.[25]	Facial expression	Supervised kernel mean matching	CK+,MMI	Eye-point alignment, local binary patterns
Moschona[26]	Speech	Supervised learning algorithms	EmoDB, CASIA, IEMOCAP	Multi-modal emotion recognition

els.

2.2 Works conducted with EEG signals

Table 2.2: Comparison of different works done on electroencephalogram signals

Author	Method	Classes	Dataset
Li et al.[12]	Time and Frequency	4	DEAP Dataset
Horlings[27]	Activity, Mobility, and Complexity	2	-
Schaaff et al.[28]	Frequency features	3	-
Frantzidis et al.[29]	Time, frequency features	2	DEAP Dataset
Murugappan et al.[30]	DWT	2	DEAP Dataset
Brown et al.[31]	Frequency features	2	-
Hosseini et al.[32]	HOS	2	-
Li et al.[1]	Time and Frequency	2	DEAP Dataset
Gupta et al.[33]	FAWT	4,2	SEED
Nakisa et al.[34]	Time, frequency, and Time-Frequency features	4	SEED
Ackermann et al.[35]	-	3	DEAP Dataset

Different classes such as 2, 3, or 4 different classes were taken into consideration in each work. The classes are evaluated using a variety of rules, such as a threshold value that, when exceeded,

results in one feeling and, when below it, in another emotion. The majority of the works took into account emotions based on valence and arousal. Low valence and high valence or low arousal and high arousal are both possible in a two-class problem. Combination of both low and high of valence and arousal are used for the four-class problem.
Table 2.2 shows the works done with EEG signals using different datasets and different feature extraction methods. Also different classifiers are used for performing the experiment. SVM classifiers are used in the works [27]-[28]. Li et al. [12] used CLRNN for their 2 class problem while Hosseini et al. [32] used KNN and SVM for their 2-class problem. Gupta et al. [33] and Nakisa et al. [34], in their 4-class problem, used PNN and random forest respectively.

Kroupi et al. [36] suggested a method for emotion classification based on correlation. They evaluated the coefficient for the Spearman correlation to distinguish between the different emotion classes among the feature vectors that include the non-stationary index (NSI), the power spectral density (PSD), and the normalised length density (NLD).The labelled data of EEG signals are used to classify emotions using a variety of factors, most of which are nonlinear [27], [37]–[40]. The fractal and multifractal dimension [37]-[38], the Hjorth parameter [27], differential entropy [39]-[40] are used to classify emotions.

Frantzidis et al. [29] provide a two-level emotion classification system. Event-related oscillations (ERO) and event-related potential (ERP) are employed as features, whereas support vector machines (SVM) and MDA are used as classifiers. The resulting accuracy rates are 81.3 and 79.5 percent, respectively. The short-time Fourier transform (STFT) is utilised as the extraction method by Zheng et al. [39] and the features for emotion categorization include the PSD, density estimate (DE), differential asymmetry (DASM), and rational asymmetry (RASM). The overall level of accuracy is 86.65 percent. Murugappan et al. [30] achieved an overall accuracy of 71.3 percent with a k-NN classifier employing the discrete wavelet transform (DWT) as the feature extraction technique and signal/wave power as the feature.

Yin et al. [14] consider the time-frequency domain characteristics collectively. Fast Fourier Transform (FFT) is utilised to extract the frequency features and the time-frequency features. The accuracy for the 2-class problem with the highest accuracy was 85.21 percent, while the accuracy for the 4-class problem with the best accuracy was 82.01 percent. The spectral-domain methods' time-domain limitations, which include their inability to provide information about frequency fluctuations over time and the distribution of energy throughout the wave's various frequencies, led to the development of time-frequency methods. Convolution neural network (CNN)-based techniques are also suggested for automated human emotion recognition [12], [41].

Most works share one or more traits, and all algorithms depend on certain features. The feature extraction process is carried out to quantify the data. For perfectly capturing the dynamics and qualities of the signal, there is no set of guidelines or standards. So there's a risk of overfitting, and irrelevant characteristics could lead to incorrect outputs that negatively impact the outcome. Every channel contributes data to the emotion analysis process, so even just one channel should be sufficient to get the desired outcome.

2.3 Summary

In this chapter the literature review of the several emotional recognition technique with the non-physiological technique and with the proposed method (using EEG signals) also investigated. And observed that all works are based on feature extraction and these features are used to quantify the information. The issues may rise due to this is also observed. Another significant finding is that the accuracy of the model appears to be declining as the number of categorization classes grows in the literature. The next chapter deals with the methods and techniques used in our work.

Chapter 3

METHODOLOGY

The investigation of the different methodology, classifiers and works including their performance is done in the previous chapter. This chapter deals with the method with which the work is done, also discusses the dataset, emotion classification, classifiers etc. which are used in the work.

3.1 Proposed Method

The block diagram of the proposed method is mentioned in the Figure 3.1. First of all, the dataset is preprocessed, the dataset may contain some irrelevant channels that do not add any meaningful information, in that case we must eliminate those channels, the dataset using in this work uses 40 channels, on those 40 channels only the 32 channels are the relevant ones. After that, the trial wise data is used for emotion mapping and we need to get the data for each channels in a trial. So a channel separation is done and these channels are the mapped to an emotions formulated (4-classes/emotions). Then these data is fed into the classifiers (random forest, xgboost and KNN classifiers are used in this work) for training and testing and finally the classification problem is executed.

The work is mainly divided into two experiments. One is subject specific and the other is subject dependent. In the first experiment, we use the data of that particular subject for the classification of emotion and we will be having a sample of 1280 (ie. 40 trials x 32 channels) so that the training data used is the 70% of 1280 samples and the remaining is used for testing purpose. The latter mentioned techniques involves the use of various random trials of each subject for its learning purpose and use those expertise to predict the test result. For the training purpose 70% of the total data (70% x 32 subject x 40 trials x 32 channels) is used and for the testing remaining 30% is used (30% x 32 subjects x 40 trials x 32 channels) specifically for the test which doesnot involves cross validation. Apart from this, cross-validation technique is performed to estimate and improve the model accuracy. The cross validation technique used here is the Kfold cross validation with 10 folds.

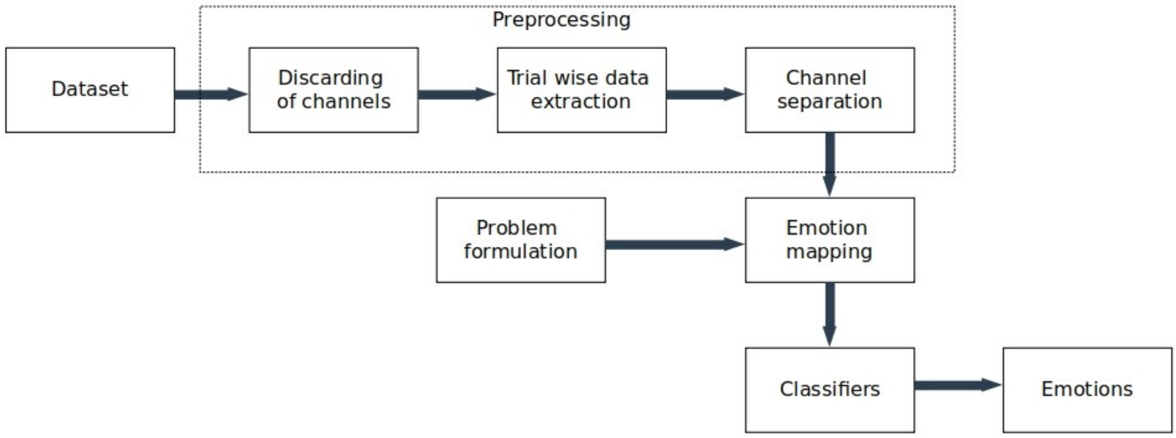

Figure 3.1: Block diagram of the proposed method

3.2 Dataset

The DEAP dataset consists of 32 subjects, 16 of them are women and the remaining 16 are men. These 32 participants watched a one-minute video, and the EEG readings during that time were captured with the aid of 32 scalp electrodes. In a similar manner, each participant is shown 40 videos of one-minute duration. The data is first collected at 512 Hz sampling frequency and then down-sampled to 128 Hz.

The dimension of the data and the sampling rate are mentioned in the Table 3.1 and the graphical version of the DEAP dataset is mentioned in the Figure 3.2. The dataset consists of the recordings of the physiological signals (EEG and EOG) for a length of 63 sec. It is expected that during the initial viewing of the videos, emotion labelling has not begun. The 3-s duration is therefore regarded as the baseline recording. There are a total number of 8064 samples for this time duration (63 sec x 128 Hz). For the emotion labelling, the participants are asked

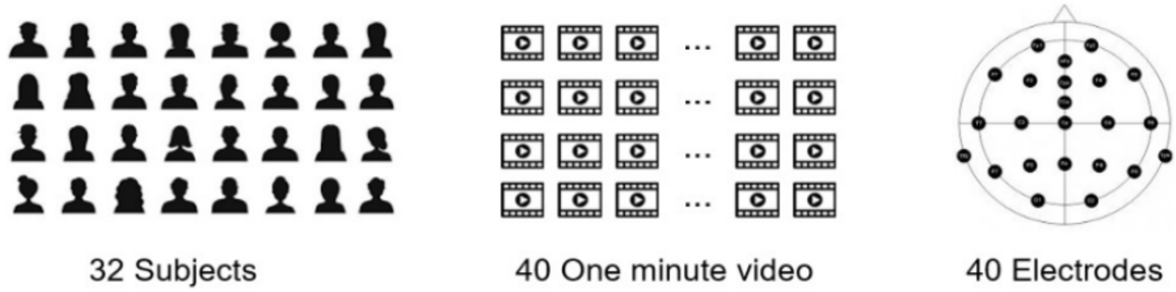

Figure 3.2: Representation of DEAP dataset

Table 3.1: Data dimension and sampling rate of DEAP dataset

Data	40x40x8064 Video/trials x channel x data
Labels	40x4 video/trials x lables (valence, arousal, dominance, liking)
Sampling rate	128 Hz

to write down the levels of valence (1–9), like/dislike (1–9), dominance (1–9), and familiarity (1–5), as well as arousal (1–9). The resulting signals are downsampled at a frequency of 128 Hz, and the artefacts are eliminated using a bandpass filter (4-45 Hz).

Each trial is the outcome of all 40 channels that are available in the experiment setup for recording physiological signals. The 8 channels/electrodes are utilised to collect EOG data, while the remaining 32 are inserted into the scalp to collect EEG signals. Therefore, a single individual will undergo a total of 40 trials, and likewise there will be a data of 32 subjects.

3.3 Pre-processing

As mentioned earlier the DEAP dataset contains both EEG and Electrooculogram (EOG) signals, where the 32 electrodes for the experiments records the EEG signals while the remaining records EOG. The nature of the EEG necessitates proper preprocessing and the removal of noise and artefacts. Electrooculogram (EOG) data (eye-generated artifacts) had been eliminated from the DEAP dataset. EEG measurements from 40 trials over 32 channels and 32 subjects make up the utilised data. With the use of the Valence-Arousal model, one can determine how each trial corresponds to a certain emotion. Each trial will last one minute, therefore there will be 32 one-minute EEG measurements total for each session. The data is initially divided into distinct trials (40 trails x 32 subject), which is a total of 1280 trials, with each trial containing readings from 32 channels (32 channels x 40 trials). In a trial, the same emotion is mapped to each channel reading.

In total, there are 51200 samples of channel data in the dataset (32 subjects x 40 trials x 40 channels), but only 40960 of those samples include the relevant information needed for this study (32 subjects x 40 trials x 32 channels). For a single subject, it is 1280 channel data samples.

3.4 Problem formulation and emotion mapping

Its been mentioned that, after each video trials, the participants are asked to write down the levels of valence (1–9), like/dislike (1–9), dominance (1–9), and familiarity (1–5), as well as arousal (1–9) for their emotional state recognition. From this values, considering the values of valence and arousal only, a certain set of rule is formulated for the classification of emotions.

The value is in the range of 1 to 9, where 9 represents the highest value and the 1 represents the lowest value. For eg., consider the value of arousal is 9, then the person is in a state of high excitement or very alert. Like this, 4 emotion classes are formulated with the values of valence and arousal. By considering 5 as a threshold value, above which the valence and arousal is considered to be high and below which they are considered to be low. By the combination of the 4 group, four quadrant can be constructed. Each quadrant represents a combination which in-turn represents an emotion. The combinations and emotions considered are:

- low arousal and low valence (sad)
- high arousal and low valence (angry)
- low arousal and high valence (calm)
- high arousal and high valence (happy)

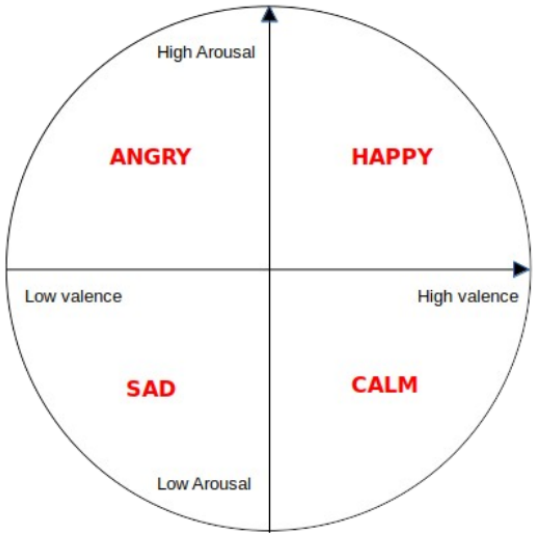

Figure 3.3: Valence-Arousal model

The four quadrants of the valence-arousal space [21], shown in Figure.3.3 This work addresses the 4-class classification problem according to the produced emotions in each of these quadrants. The labelled EEG signals associated with the four classes of DEAP-dataset are shown in Figure. 3.4.

So, the problem definition is: Based on the EEG readings obtained from the dataset, categorise these four emotion types.

3.5 Classifiers

A classifier in machine learning is an algorithm that automatically groups data into one or more "classes" or categories. Classes are also called as targets. Applications for classification algorithms include speech recognition, cancer tumour cell identification, email spam detection, drug categorization, and biometric identification. The classifiers used in this work are for the formulated classes based on the EEG data from the samples. Random forest, XG Boost and KNN are the three classifiers useed in this work for the emotion categorization.

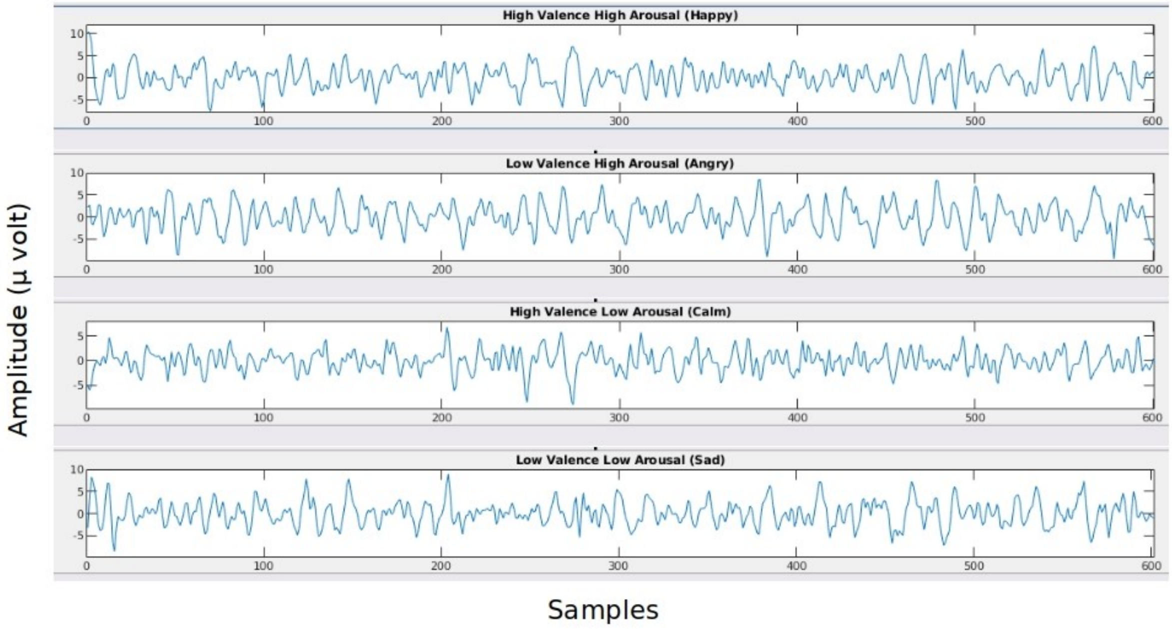

Figure 3.4: EEG signals associated with the four classes

3.5.1 Random Forest

Random forests is also called as random decision forests. It is an ensemble learning method, ie. it uses multiple learning algorithms to achieve better performance, for regression, classification and for all the tasks that uses decision tree at training phase. The problem of decision trees' overfitting issue is corrected by the random forest, basically, the Random Forest Algorithm's ability to handle data sets with both continuous variables, as in regression, and categorical variables, as in classification, is one of its most crucial qualities. It produces superior outcomes for categorization tasks.

The ensemble method employed by random forest is bagging, sometimes referred to as Bootstrap Aggregation. A random sample is chosen from the data set using bagging. As a result, each model is created using the samples (Bootstrap Samples) that the Original Data gave, with a replacement process known as row sampling. Bootstrap refers to this stage of row sampling with replacement. Each model is currently trained independently, producing results. After merging the outputs of all the models, the final decision is made based on a majority vote. Aggregation is the process of aggregating all the results and producing a result based on a majority vote.

The figure 3.5 depicts the working of random forest and the steps involved are,

Step 1: From a data set with k records, n random records are selected at random and used in the Random Forest algorithm.

Step 2: For each sample, a unique decision tree is built.

Step 3: An output will be produced by each decision tree.

Step 4: For classification and regression, the final result is evaluated using a majority vote or an average.

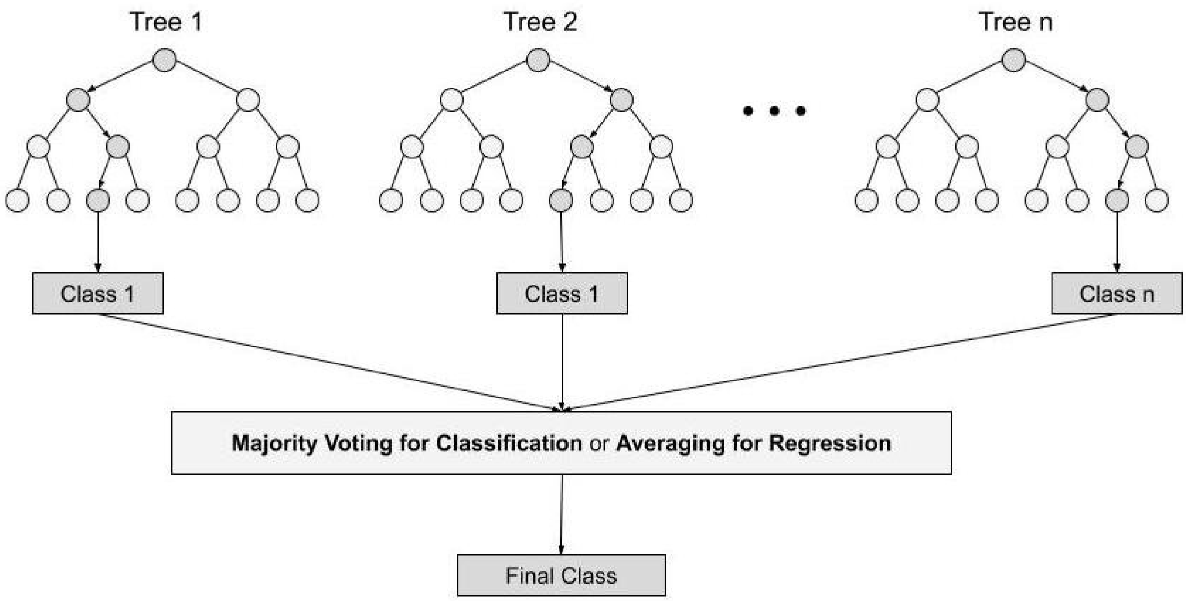

Figure 3.5: Random Forest classifier

The mathematics of Random Forest:

Gini index:

The Gini index is a tool used in the CART (Classification and Regression Tree) technique, which uses decision trees to classify data and determine its impurity or purity. It generates a binary split, which is then used by the CART algorithm. Equation 3.1 shows how to calculate the gini index. An attribute is low Gini index is preferred as the root node

$$GiniIndex = 1 - \sum_{j} \mathrm{P}_J \tag{3.1}$$

Information Gain:

Entropy in the data set and the attribute entropy are used to calculate information gain, which tells us how much information a feature gives us about a class. Entropy measures how much randomness or impurity is present in the provided data, the formula for entropy and information gain is mentioned in the equation 3.2 and 3.3 respectively. It is utilised to select the decision tree's root node for data segmentation.

$$Entropy = -\sum p\left(x\right)\log p\left(x\right) \tag{3.2}$$

$$InformationGain = Entropy(S) - [(WeightedAvg)xEntropy(eachfeature)] \tag{3.3}$$

The features which is having high information gain is chosen as the root node. The advantages of the random forest classifiers are,

- It helps to increase accuracy and decrease overfitting in decision trees.
- It is adaptable to problems involving classification and regression.
- Both categorical and continuous values can be used with it.
- It automates filling in data's missing values.
- Data normalisation is not necessary because a rule-based methodology is used.

3.5.2 XGBoost

A variation of the gradient tree boosting method put out by Friedman [50] is known as extreme gradient boosting (XGBoost). XGBoost is a distributed gradient boosting library that has been developed to be very effective, adaptable, and portable. Under the Gradient Boosting framework, it implements machine learning techniques. With the use of XGBoost, many data science issues can be quickly and accurately solved using parallel tree boosting, sometimes referred to as GBDT or GBM.

XGBoost is a technique for group learning. It might not always be enough to rely solely on a single machine learning model's output. A methodical approach to combining the prediction capacity of various learners is provided by ensemble learning. A single model that provides the combined output from multiple models is the outcome. The base learners, sometimes referred to as the ensemble's models, may come from the same learning algorithm or from other learning algorithms.

Let $(\mathrm{x}_1, \mathrm{y}_1), (\mathrm{x}_2, \mathrm{y}_2), ..., (\mathrm{x}_n, \mathrm{y}_n)$ be a set of inputs and corresponding outputs. The tree ensemble algorithm uses *K* addative functions, each representing a CART, to predict the output. The predicted output, shown in equation 3.4, is given by the sum of each individual function prediction,

$$\hat{y}_i = \sum_{k}^{K} f_k\left(x_i\right), f_k \in F \tag{3.4}$$

where $f \in F$ is the space of CARTs. Thus, the objective is to approximate the functions by minizming the following regularized objective function, equation 3.5, given a set of parameters θ :

$$\mathrm{obj}(\theta) = \sum_{i}^{n} l\left(\hat{y}_i, y_i\right) + \sum_{k}^{K} \Omega\left(f_k\right) \tag{3.5}$$

where the first term $l\left(\hat{y}_i, y_i\right)$ represents the training loss function that measures the difference between the predicted output and the actual output. The training loss function can be measured using different types of error, such as Mean Squared Error (MSE), given by the equation 3.6:

$$MSE = \sum_{i}^{n} \left(y_i - \hat{y}_i\right)^2 \tag{3.6}$$

and Logistic Loss, given by the equation 3.7:

$$\text{Logistic Loss } = \sum_{i}^{n} \left[y_i \ln\left(1 + e^{-\hat{y}_i}\right) + \left(1 - y_i \ln\left(1 + e^{\hat{y}_i}\right)\right)\right] \tag{3.7}$$

The second term $\Omega\left(f_k\right)$ is the regularization term, which penalizes the complexity of the model to avoid overfitting. In XGBoost the regularization term is given by the equation 3.8:

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda\|w\|^2 \tag{3.8}$$

where T is the number of leaves and the second term is the L2 norm of leaf scores. During training, the model is trained additively, by optimizing for one tree at a time. Let $\hat{y}_i{}^t$ be the prediction value at iteration t, the additive procedure is :

$$\begin{aligned}
&\hat{y}_i{}^0 = 0 \\
&\hat{y}_i{}^1 = f_1\left(x_i\right) = \hat{y}_i{}^0 + f_1\left(x_i\right) \\
&\hat{y}_i{}^1 = f_1\left(x_i\right) + f_2\left(x_i\right) = \hat{y}_i{}^1 + f_2\left(x_i\right)
\end{aligned}$$

...

$$\hat{y}_i^t = \sum_{k=1}^{t} f_k\left(x_i\right) = \hat{y}_i^{t-1} + f_t\left(x_i\right) \tag{3.9}$$

The equation 3.9 gives the formula to calculate the prediction value at iteration t. The tree added at each step is the tree that optimizes the objective function. The objective function can be rewritten as

$$\begin{aligned} obj^{(t)} &= \sum_i l\left(\hat{y}_i, y_i\right) + \sum_k^K \Omega\left(f_k\right) \\ &= \sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{t-1} + f_t\left(x_i\right)\right) + \Omega\left(f_t\right) \end{aligned} \tag{3.10}$$

The objective function, as shown in equation 3.10, can be further simplified using a second-order approximation, into a function that can be used as a scorer. The score function is then used to determine how good the tree structure is.

To further prevent overfitting. XGBoost implements the shrinkage introduced by Friedman. The shrinkage variable scales the feature weights by a factor of η, also called learning rate. Furthermore, XGBoost also supports row subsampling and column subsampling, two techniques used to control bias and variance in Random Forest.

3.5.3 K-Nearest Neighbour

One of the simplest machine learning algorithms, based on the supervised learning method, is K-Nearest Neighbour. The K-NN algorithm makes the assumption that the new case and the existing cases are comparable, and it places the new instance in the category that is most like the existing categories. A new data point is classified using the K-NN algorithm based on similarity after all the existing data has been stored. This means that utilising the K-NN method, fresh data can be quickly and accurately sorted into a suitable category. Although the K-NN approach is most frequently employed for classification problems, it can also be utilised for regression.

The K Nearest Neighbor algorithm, which is used for both regression and classification (most frequently), belongs to the domain of supervised learning. It is a flexible approach that may also be used to resample datasets and impute missing values. The K Nearest Neighbor method, as its name suggests, uses K Nearest Neighbors to forecast the class or continuous value for a new data point.

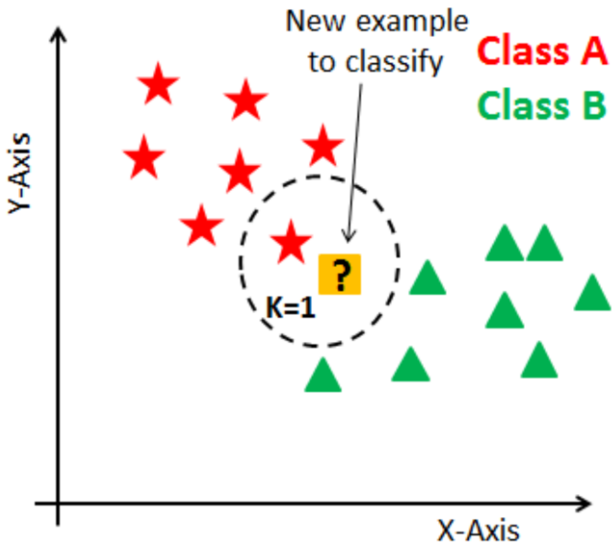

Figure 3.6: KNN classifier

K value determines how many neighboring points have to be considered, depending on the number of points and the distance between them, the method operates by calculating the distance between these points' mathematical values. Figure 3.6 shows the working of KNN classifier. It determines the chance that each data point will be similar to the test data by first calculating the distance between each data point and the test data. The highest probability points are used for classification. The distance function can be Hamming, Minkowski, or Euclidean. Euclidean is used as a default method.

If $(\mathrm{x}_1, \mathrm{y}_1), (\mathrm{x}_2, \mathrm{y}_2)$ are the coordinates of two points, then the formula for Euclidean distance, D is

$$D = \sqrt[2]{(\mathrm{x}_1 - \mathrm{x}_2)^2 + (\mathrm{y}_1 - \mathrm{y}_2)^2} \tag{3.11}$$

The input x is then assigned to the class with the highest probability once the distance D has been calculated by the equation 3.11. The formula for probability is shown in equation 3.12:

$$P(y = j \mid X = x) = \frac{1}{K} \sum_{i \in \mathcal{A}} I\left(y^{(i)} = j\right) \tag{3.12}$$

3.6 Kfold Cross-Validation

A method for assessing a machine learning model's performance is cross-validation. In real-world applications of machine learning, cross-validation is frequently utilised. Comparing and choosing the best model for the particular predictive modelling issue is helpful. The process contains a single parameter, K, that designates how many groups should be created from a given data sample. As a result, the process is frequently referred to as K-fold cross-validation. When a particular number for K is selected, it may be substituted for K in the model's reference, such as when k=10 is used to refer to cross-validation by a 10-fold factor.

It is a well-liked technique since it is easy to comprehend and typically yields a less biassed or overly optimistic assessment of the model skill than other techniques, including a straightforward train/test split.

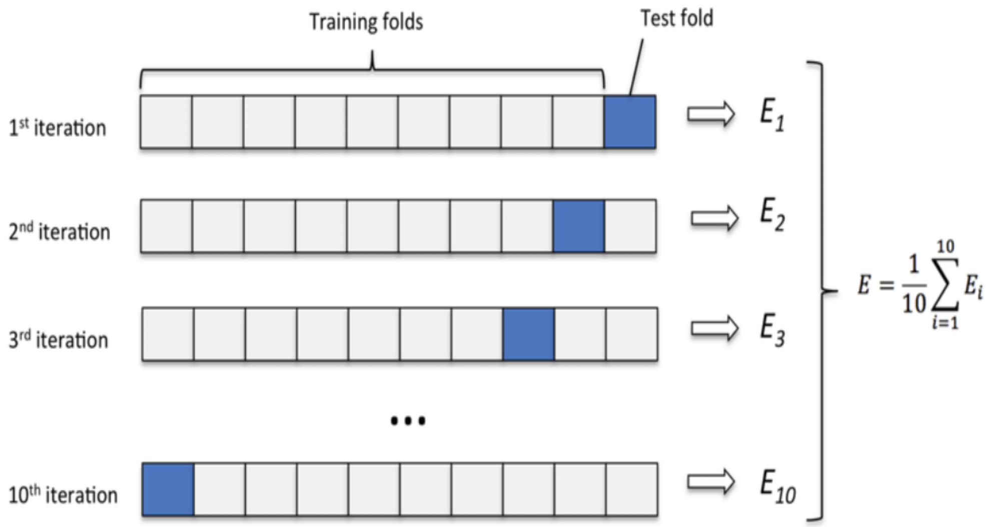

Figure 3.7: KFold cross validation technique

Figure 3.7 shows the working of KFold cross validation technique in the provided dataset. The procedure or the steps of KFold technique is,

- Randomly shuffle the dataset.
- Create k groups from the dataset.

- For every distinct group:
  - The group should be used as a holdout or test data set.
  - As a training data set, use the remaining groupings.
  - Fit a model to the training data, then assess it against the test data.
  - Keep the evaluation result, but throw away the model.
- Using a sample of model evaluation ratings, summarise the model's expertise.

3.7 Performance evaluation

An N x N matrix called a confusion matrix is used to assess the effectiveness of a classification model, where N is the total number of target classes. In the matrix, the actual goal values are contrasted with those that the machine learning model anticipated. This provides us with a comprehensive understanding of how well our categorization model is functioning and the kind of mistakes it is making.

Information retrieval requires both precision and sensitivity, so both are very important. Also it considers mainly the positive classes when compared with the negative class. Only the true positive (TP), false positive (FP), false negative (FN) are used in precision and sensitivity, it doesnot considers true negative (TN).

- True positive - Predicted class and actual class both are positive
- True negative – Predicted class and actual class both are negative
- False positive – Predicted positive but actual class is negative
- False negative - Predicted class is negative but actual class is positive.

Figure 3.8 represents a 4x4 confusion matrix showing the TP, TN, FP, FN exclusively for class 0. From this various parameters for performance evaluation can be calculated. Precision determines the percentage of the truly positive out of all the positive predicted, whereas the sensitivity gives the predicted positive percentage out of all the total positive. F1-score is the harmonic

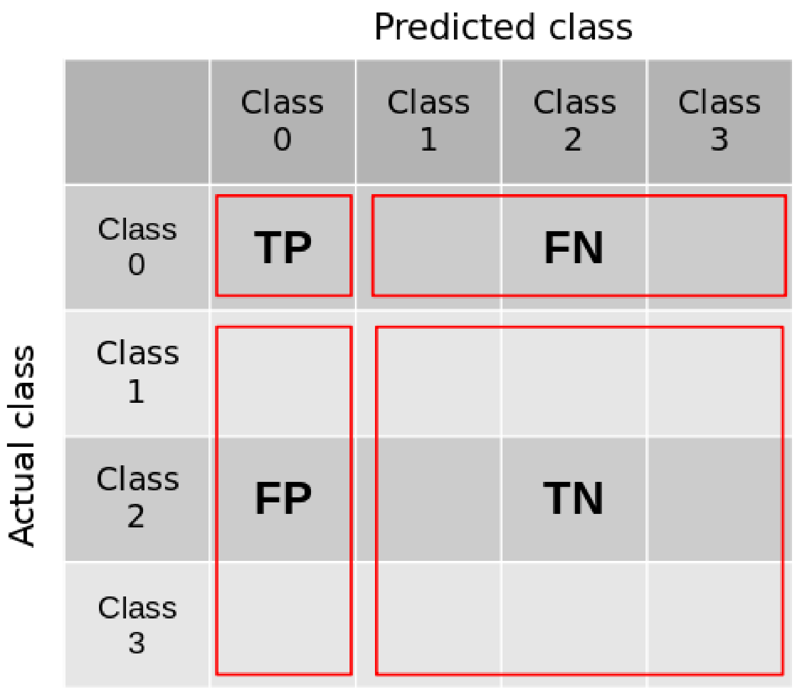

Figure 3.8: A 4X4 Confusion matrix showing the TN, TP, FN, FP

mean of precision and sensitivity. The likelihood of false negatives and false positives is revealed by specificity and sensitivity. The fundamental reliability of a test is shown by statistical evaluations of accuracy and precision. These phrases do not interchangeably refer to the sources of variability. A test method may be accurate (really measuring what it is designed to measure) without being precise (reliably reproducible in what it measures), or vice versa.

Measurements such as accuracy, sensitivity, specificity, f1-score, and precision have been used to validate the suggested emotion classification algorithm's performance. The mathematical representation of these parameters are shown in the equation 3.13 to 3.17

$$Precision = \frac{TP}{TP + FP} \tag{3.13}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{3.14}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{3.15}$$

$$Specificity = \frac{TN}{FP + TN} \tag{3.16}$$

$$F1 - score = \frac{2TP}{2TP + FP + FN} \quad (3.17)$$

3.8 Subject specific

The DEAP dataset comprises of 32 volunteers' physiological recordings, with each subject's records taken into account separately for the experiment. Such that the experiment using this method will consist of 32 overall trials, that is 32 different results will be obtained. As mentioned earlier, the DEAP dataset have the recordings of electroencephalogram (EEG) and electrooculogram (EOG), this work is based on the EEG signals therefore extracting of the channels which are used to record the EEG is the first task. Each trial from the subjects is split after the necessary channels that record the EEG have been extracted. The trials described here are the physiological readings taken while watching the one-minute video. There will be 40 trials because each subject will see a total of 40 videos.

This project involves a four-class problem, thus the four classes of emotions must be determined based on the valence and arousal levels that the participant indicated after each video trial. Even though each trials are having different channels to record the physiological signals, the emotions are assigned to each trials based on the entire channel data.

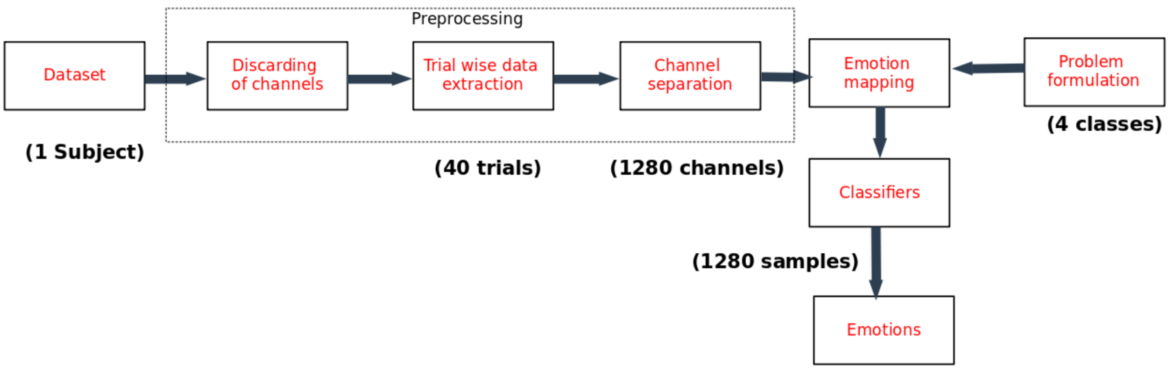

Figure 3.9: Block diagram of subject specific method

For the subject specific method, the data of each subjects are considered separately. So for the entire experiment to perform, we have 32 different datasets each consists of 40 trials which haves the one minute samples from 32 channels. Figure 3.9 shows the block diagram of the subject specific method. First block is the dataset, here, it is the data of 1 subject which consist the physiological signals' recordings for the 40 one minute video trials. The preprocessing block is the next block, and the first goal is the elimination of undesirable channels.

The emotion classes are formulated for each trials, therefore the trial wise extraction is carried out in the next block. Right now, there are 40 samples are there. In the next block each trial data is splitted to respective number of channel data. The emotion formulated is then embeded to each channel data. There are 32 channels for each trial, therefore 1280 samples will be there after this stage.

After that, the data is used for training or learning purpose by different classifiers. The data is splitted into 70-30 for training and testing purposes respectively. And while performing the cross validation technique, the data is entirely fed to the classifier for the model evaluation. 32 models will be obtained for each classifier method, since we are considering each subjects' data separately. 32 different models with 32 different results or accuracy. Each results are analysed to find which subjects have lowest accuracy, which one have highest accuracy and how many of the subjects are having accuracy greater than the average accuracy etc.

3.9 Subject dependent

This approach takes into account the physiological signals recorded from all 32 subjects at once. Compared to the prior approach, this approach uses just one model for each classifier, as opposed to 32 separate models. The block diagram for this approach is shown in Figure 3.10; while the blocks used in this method resemble those used in the previous method, there are differences in the input and the dimension of the data.

A single dataset containing the data from 32 subjects will make up the first block. Following that, this will be passed to the preprocessing block, where the channels that record EOG are discarded. Then a trial-by-trial extraction is performed, totaling 1280 trials (i.e., 32 subject x 40 trials). The following block divides these trials once again according to the number of channels; following this block, the model will have 40960 samples of data (i.e., 32 subject x 40 trial x 32

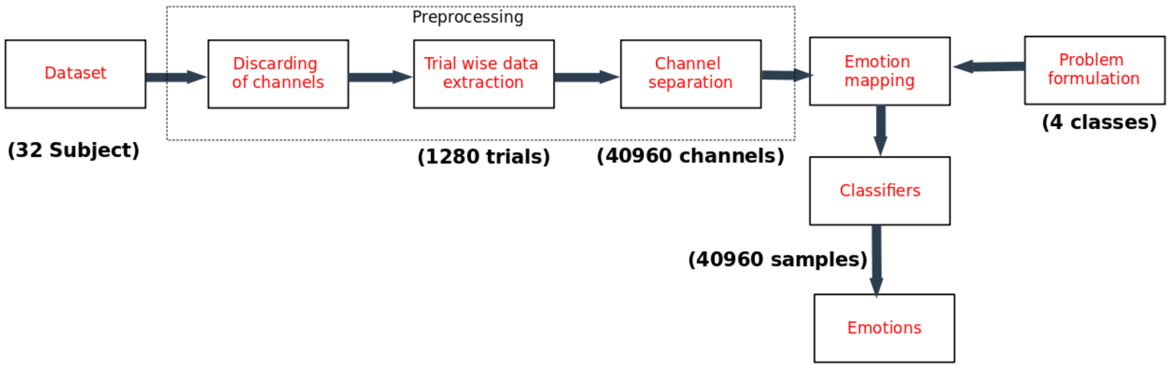

Figure 3.10: Block diagram of subject dependent method

channels).

The data is then used by various classifiers for learning or training purposes. For purposes of training and testing, the data is divided 70-30 %. Additionally, the input is totally given to the classifier for the model's evaluation while the cross validation technique is being used. Analysis of the output from each classifier is done, and comparisons between various methods are made. On the basis of the results produced by the confusion matrix for each class, it is also necessary to undertake a performance study of the various classifier models.

Since the subject specific technique trains and tests using the same person's data, it is anticipated that the results will be better than those of the subject dependent method.

3.10 Summary

This chapter discussed about the methods and materials such as dataset, different classifiers and problem formulation etc, are studied in this chapter, along with the experimentation procedures. Two methods are followed in this work, one is subject specific and the other is subject dependent. The performance evaluation indices was another focus area of this chapter. The next chapter deals with the verification of the results obtained using this method.

Chapter 4

RESULTS AND DISCUSSION

In the previous chapter, the experimentation methodologies are mentioned along with the performance evaluation method. The accuracies of the different ML models and their performance for each classes are verified in this chapter. As mentioned above, mainly two experiments are conducted to understand the efficiency of the classifiers with the proposed model.

1. Subject-specific
2. Subject dependent

In the former one, the channel data of that specific subjects are only considered for learning and testing purpose whereas in the later one, randomly chosen channels of all subjects are used for the training/learning stage and testing stage.

4.1 Subject-specific results

As the result of the investigation done in this work, which is, how well the different classifiers work on the DEAP dataset, with the proposed method, to classify the emotion, performance of various classifiers are evaluated. The accuracies obtained from the classifiers are noted with and without cross-validation technique.

4.1.1 Results obtained from Random forest classifier

Table 4.1: Random classifier accuracies of 32 subjects

Subject	Without CV	With CV	Subject	Without CV	With CV
1	88.8%	91.4%	17	88.4%	89.7%
2	91.7%	92.1%	18	91.4%	91.6%
3	88.5%	90.5%	19	89.6%	91%
4	96.6%	97.4%	20	88.5%	91.6%
5	99.7%	99.5%	21	88.8%	91.2%
6	82.8%	87.7%	22	95.3%	96.7%
7	72.1%	74.6%	23	74.0%	83.4%
8	88.5%	90.1%	24	96.4%	97.1%
9	94.0%	96.2%	25	94.8%	93.4%
10	92.7%	94.3%	26	95.3%	94.9%
11	97.1%	95.7%	27	88.5%	89.6%
12	87.8%	89.3%	28	92.4%	92.0%
13	89.3%	90.1%	29	87.8%	89.1%
14	86.7%	90.6%	30	92.7%	94.8%
15	89.6%	92.1%	31	86.5%	88.7%
16	79.4%	82.2%	32	90.4%	91.7%

From the Table 4.1, it can be observed that the use of cross validation improves the accuracy of our model. The average accuracy of 89.58% is increased to 91.26% with the use of Kfold cross-validation technique with 10 folds. The lowest and best accuracy of 72.1% and 99.7% respectively shifts to 74.6% and 99.5% respectively. We can see a an improvement in accuracy of the lowest but a slight down for the best accuracy, but considering the overall performance we can see an improvement.

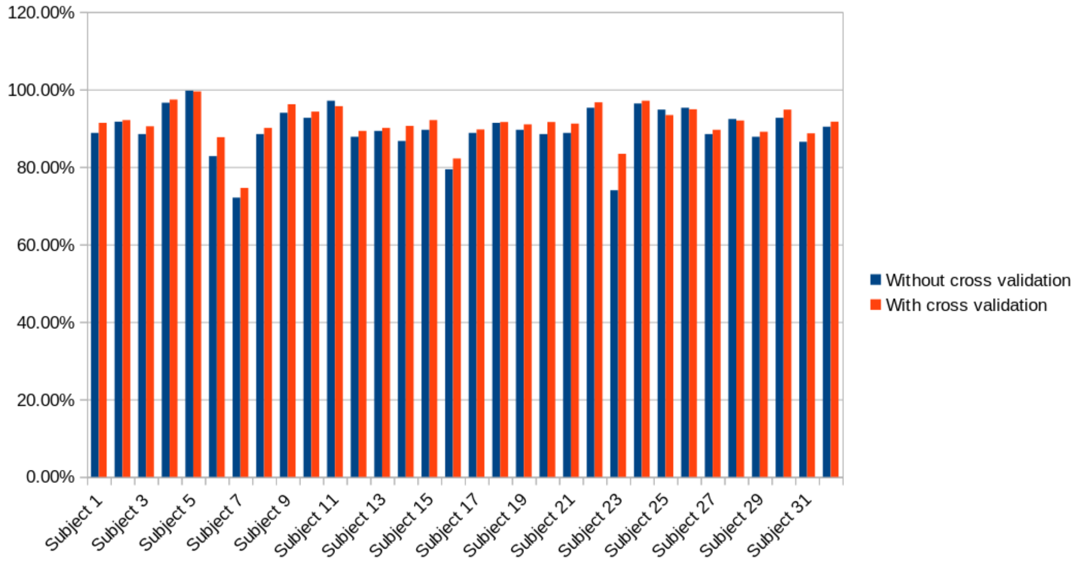

Figure 4.1: Comparison of random forest classifier accuracies' of 32 subjects

4.1.2 Results obtained from XGBoost classifier

For the XG Boost classifiers, the accuracies obtained for each subjects is mentioned in the Table 4.2. Compared with the random forest classifier, XG Boost gives good accuracy. The lowest and the best accuracy for this classifier are 81.25% and 98.44% respectively. It have an average accuracy of 92.83%. Results of the subjects are graphically shown in Figure 4.2.

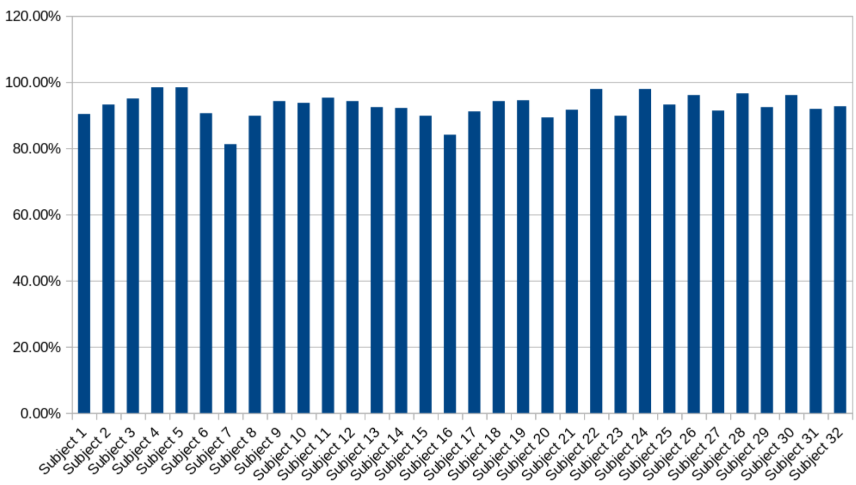

Figure 4.2: Comparison of xgboost classifier accuracy of 32 subjects

Table 4.2: XGBoost classifier accuracy of 32 subjects

Subject	Accuracy in %	Subject	Accuracy in %
1	90.36	17	91.15
2	93.23	18	94.27
3	95.05	19	94.53
4	98.44	20	89.32
5	98.44	21	91.67
6	90.62	22	97.92
7	81.25	23	89.84
8	89.84	24	97.92
9	94.27	25	93.23
10	93.75	26	96.09
11	95.31	27	91.41
12	94.27	28	96.61
13	92.45	29	92.45
14	92.19	30	96.09
15	89.84	31	91.93
16	84.11	32	92.71

4.1.3 Results obtained from KNN classifier

While investigating the accuracy of KNN classifier, the results obtained were impressive and they are listed in the Table 4.3.

Table 4.3: KNN classifier accuracy of 32 subjects

Subject	Without CV	With CV	Subject	Without CV	With CV
1	99.22%	99.3%	17	92.19%	91.5%
2	94.92%	95.5%	18	95.31%	96.6%
3	96.48%	97.3%	19	97.27%	94.9%
4	99.61%	99.0%	20	94.53%	94.9%
5	96.48%	95.8%	21	96.09%	95.3%
6	96.48%	96.5%	22	92.19%	94.2%
7	85.94%	92.2%	23	96.48%	96.8%
8	93.75%	92.7%	24	96.88%	97.4%
9	93.36%	94.5%	25	96.09%	93.3%
10	88.28%	89.6%	26	94.14%	95.2%
11	95.70%	95.4%	27	98.43%	98.4%
12	93.75%	95.1%	28	92.19%	93.8%
13	92.97%	95.1%	29	94.14%	96.3%
14	94.53%	93.7%	30	94.14%	94.1%
15	92.19%	90.9%	31	95.70%	97.6%
16	86.72%	91.5%	32	93.36%	95.3%

The average accuracy of the KNN classifier, with number of neighbors = 3, is increased from 91.64% to 94.99% with the use of Kfold cross-validation technique with 10 folds. The lowest and best accuracy of 85.94% and 99.61% respectively shifts to 91.5% and 99.3% respectively. We can see a an improvement in accuracy of the lowest but a slight down for the best accuracy, but considering the overall performance we can see a better improvement in the result. Comparing with any other models, we can see KNN performs well.

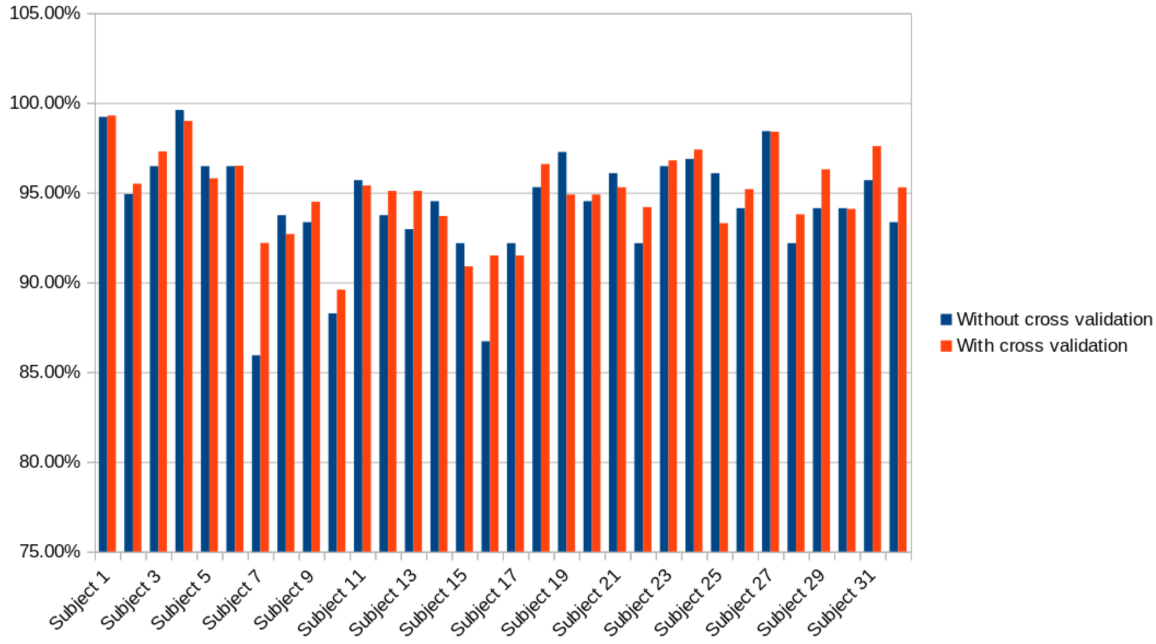

Figure 4.3: Comparison of KNN classifier accuracy of 32 subjects

Comparison of the accuracies obtained from the different classifiers are listed in the Table 4.4. It contains the details of the lowest accuracy, highest accuracy obtained and the mean of all 32 results obtained while conducting the experiment with different classifiers.

The Figure 4.4 - 4.6 shows the confusion matrix of different classifier methods and based on this, the performance of the model is evaluated and tabulated in Table 4.5.

Table 4.4: KNN classifier accuracies' of 32 subjects

Classifiers	Lowest accuracy	Highest accuracy	Average accuracy
Random forest	72.1%	99.7%	89.58%
Random forest (10-fold)	74.6%	99.5%	91.26%
XGBoost	81.25%	98.44%	92.83%
KNN	85.94%	99.61%	91.64%
KNN (10-fold)	91.5%	99.30%	94.99%

4.2 Subject dependent results

Based on the subject dependent experiments, which uses the random samples of each subject for training as well as testing purpose, the obtained results are tabulated in the table 5.6.

Table 4.5: Comparison of subject dependent results for different classifiers.

Classifier	Accuracy in %
Random forest	75.3%
Random forest (10-fold)	78.5%
XGBoost	82.77%
KNN	90.31%
KNN (10-fold)	92.73%

As of the obtained result, we can understand that the KNN classifier has the better accuracy. Also the computational speed of KNN is much faster than any other classifiers opted. The classification result of the different classifiers are also investigated to understand the accuracies of different classes. Figure 5.4 depicts the bar graph representation of table 5.6.

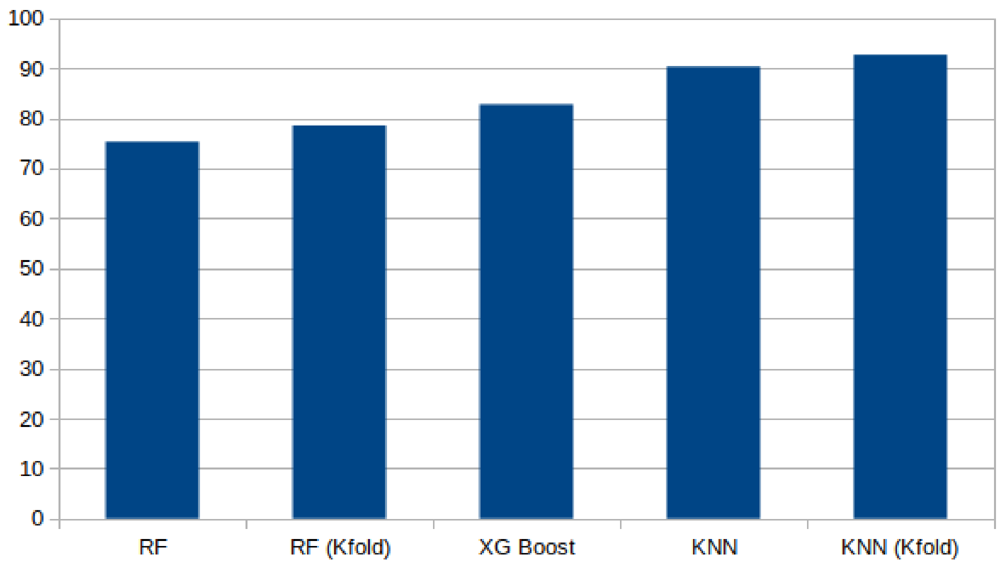

Figure 4.4: Comparison of classifier result for subject dependent experiment

4.3 Performance analysis

4.3.1 Subject specific

After taking the readings from each classifier, evaluation of the model performance is done by using the result of subject 5 (a randomly chosen subject), that is by analysing the confusion matrix obtained from different classifier and evaluated by finding precision, sensitivity, f1-score, specificity and accuracy of the classes. The classes 0,1,2 and 3 are the names given for each emotion in confusion matrix which represents the classes sad, angry, calm and happy respectively.

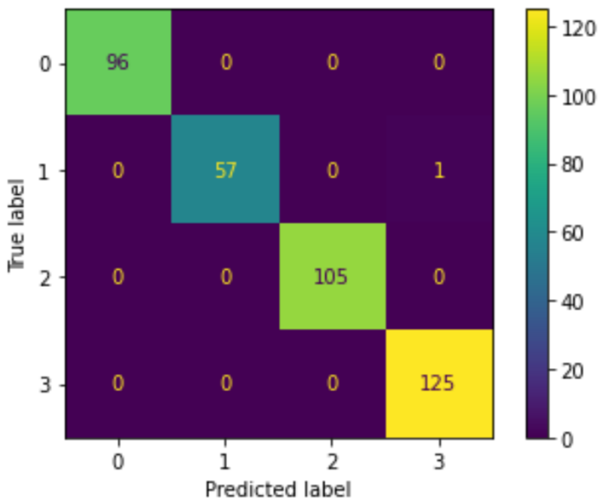

Figure 4.5: Confusion matrix of random forest

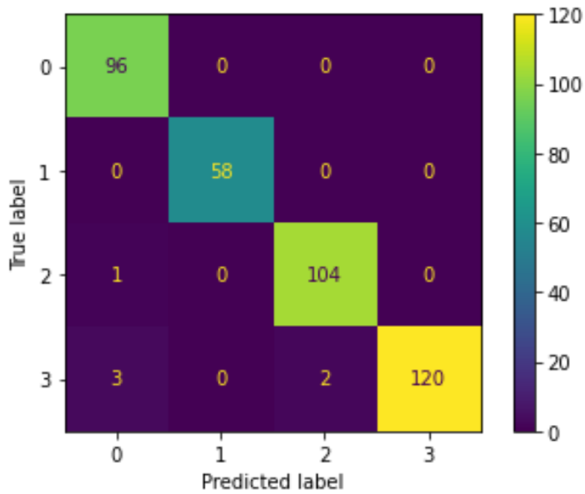

Figure 4.6: Confusion matrix of XGBoost

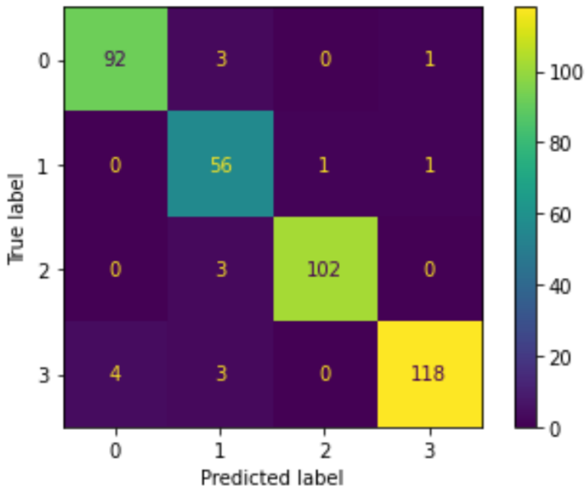

Figure 4.7: Confusion matrix of KNN

Table 4.6: Performance analysis of Subject 5 based on the obtained confusion matrix.

		Sad	Anger	Calm	Happy
Random forest	Precision	1.00	1.00	1.00	0.99
	Sensitivity	1.00	0.98	1.00	1.00
	f1-score	1.00	0.99	1.00	1.00
	Specificity	1.00	1.00	1.00	0.99
	Accuracy	1.00	0.99	1.00	0.99
XG Boost	Precision	0.96	1.00	0.98	1.00
	Sensitivity	1.00	1.00	0.99	0.96
	f1-score	0.0.98	1.00	0.99	0.98
	Specificity	0.99	1.00	0.99	1.00
	Accuracy	0.99	0.99	1.00	0.92
KNN	Precision	0.96	0.86	0.99	0.98
	Sensitivity	0.96	0.97	0.97	0.94
	f1-score	0.96	0.91	0.98	0.96
	Specificity	0.99	0.97	0.99	0.99
	Accuracy	0.98	0.97	0.97	0.98

4.3.2 Subject dependent

After the classification with the data of 32 subjects, the obtained confusion matrix is used to evaluate the performance of the model on different classifier algorithm. The testing and training data used consist of all 32 subjects, therefore the result will be a generalised one compared with the previous method. Also, it is expected that the results will not reach upto the level of subject specific method because it deals with the data of a unique individual.

From the obtained confusion matrix (Figure 4.8 - 4.10), the performance of the models are evaluated by finding precision, sensitivity, f1-score, specificity and accuracy of the classes. And is tabulated in the Table 4.7

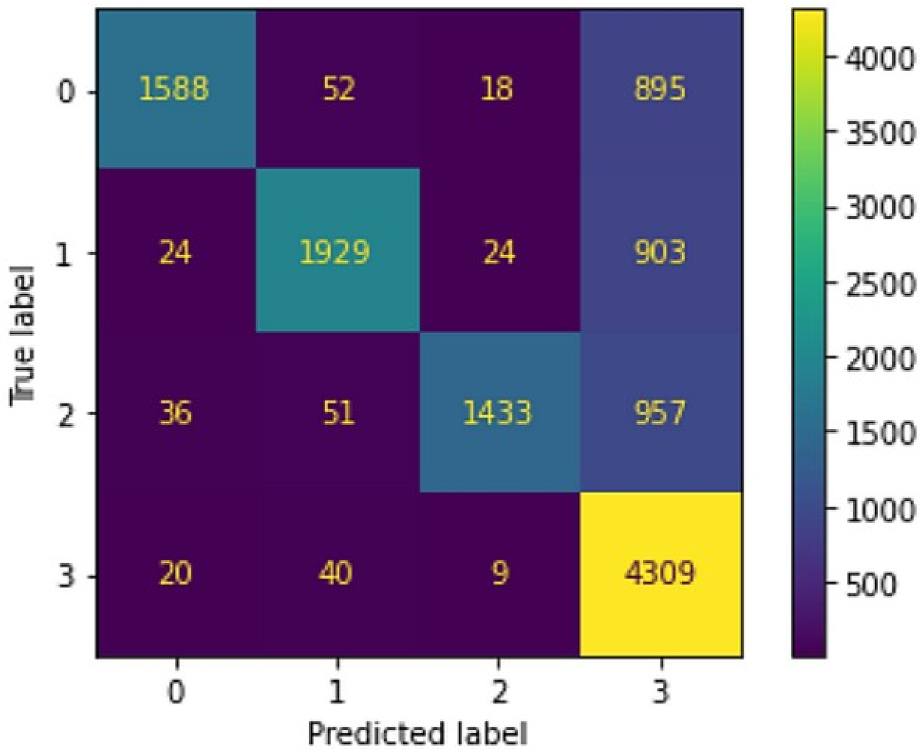

Figure 4.8: Confusion matrix of random forest

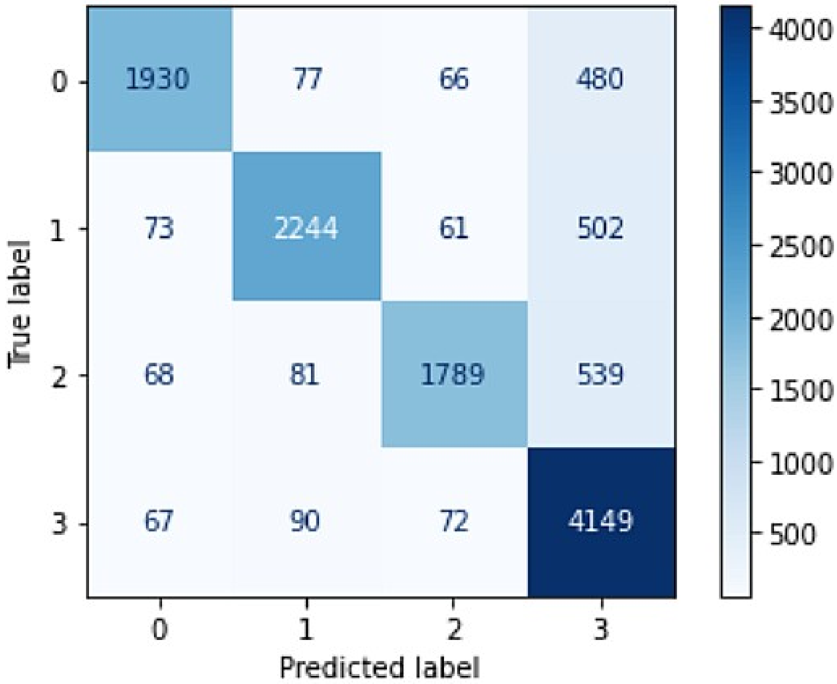

Figure 4.9: Confusion matrix of XG Boost

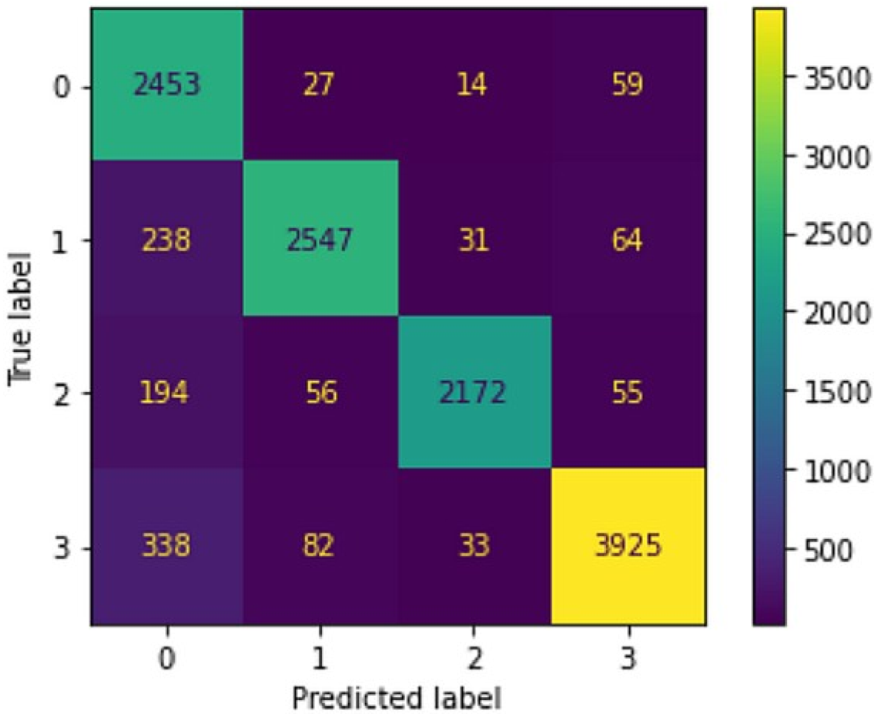

Figure 4.10: Confusion matrix of K-Nearest Neighbor

Precision determines the percentage of the truly positive out of all the positive predicted, whereas the sensitivity gives the predicted positive percentage out of all the total positive. F1-score is the harmonic mean of precision and sensitivity. The likelihood of false negatives and false positives is revealed by specificity and sensitivity. The fundamental reliability of a test is shown by statistical evaluations of accuracy and precision.

Considering the precision values of different classes, class ‘sad’ has best and class ‘happy’ has the lowest precision score in the random forest classifier. Same is the case for XG Boost classifier, KNN has a different classification result for precision, best value is for class ‘calm’ and lowest value is for class ‘sad’. Like this, all results for the parameters chosen for the performance analysis are summarised in the table 5.7.

Table 4.7: Performance analysis of the models based on the obtained confusion matrix.

		Sad	Anger	Calm	Happy
Random forest	Precision	0.95	0.93	0.97	0.61
	Sensitivity	0.62	0.67	0.58	0.98
	f1-score	0.75	0.78	0.72	0.75
	Specificity	0.99	0.98	0.99	0.65
	Accuracy	0.91	0.91	0.91	0.77
XG Boost	Precision	0.9	0.9	0.9	0.73
	Sensitivity	0.76	0.78	0.72	0.95
	f1-score	0.82	0.84	0.80	0.83
	Specificity	0.98	0.97	0.98	0.81
	Accuracy	0.93	0.93	0.93	0.86
KNN	Precision	0.76	0.94	0.97	0.96
	Sensitivity	0.96	0.88	0.88	0.9
	f1-score	0.85	0.91	0.92	0.93
	Specificity	0.92	0.98	0.99	0.97
	Accuracy	0.93	0.96	0.97	0.95

4.4 Comparison with previous works

The accuracy of the proposed model is compared with the previous works that have been done on this area by other research enthusiasts

.

Table 4.8: Comparison of the proposed models with the previous works

Author	Classifier	Classes	Accuracy
Li et al.[1]	CLRNN	2	74.12%
Horlings[13]	SVM	2	81%
Frantzidis et al.[17]	SVM	2	81.3%
Murugappan et al.[18]	SVM	2	71.3%
Brown et al.[29]	KNN	2	82%
Hosseini et al.[30]	SVM, KNN	2	82%
Schaaff et al.[28]	SVM	3	66.7%
Ackermann et al.[24]	Random forest	3	55%
Gupta et al.[22]	PNN	4	71.43%
Nakisa et al.[23]	Random forest	4	67.47%
Proposed work 1	**Random forest**	**4**	**78.50**%
Proposed work 2	**XG Boost**	**4**	**82.77**%
Proposed work 3	**KNN**	**4**	**92.73**%

As we can see that by varying the number of classes may affect the accuracy of the model. Accuracy decreases as the number of classes is increased. The proposed technique is tested in 4 class model and obtained a very impressive result.

4.5 Summary

This chapter verified the result obtained from the different experimentation conducted. Two methods are performed which is subject specific and subject dependent, KNN, XGBoost and random forest classifiers are considered in both of these methods. The Kfold cross validation technique is used for both the random forest and KNN but for XGBoost classifier, that test is not conducted. Because, while performing Kfold cross validation, the computational power required is more, and my system donot have that much power to perform the test. Performance of the model is also analysed using various parameters and the comparison of the proposed model with previous works is tabulated to highlight the upper-hand of the proposed work. Each model selected in this work is comparable with any previous work as it can provide a better performance in categorization task.

Chapter 5

CONCLUSION AND FUTURE SCOPE

The proposed system avoided the manual feature extraction technique and to avoid the extracted feature to quantify the information, as there are no rules for extracting feature and classification of emotion. 4-class problems are considered here. Compared with other works, the proposed model gives good performance and accuracy in the considered problem. For the proposed model, the performance of the machine learning algorithm with the DEAP dataset were analysed and selected the best performing classifiers. KNN, XGBoost and random forest classifiers are found to be the better performing models. The results obtained were impressive and very much high compared with other works. The average accuracy of 91.26% for random forest, 92.83% for XG Boost and 94.99% for KNN is obtained for subject independent experiment. For the sub- ject dependent test, the best accuracy of 78.5% is obatined for random forest, 82.77% for XG Boost and 92.73% for KNN is obtained. The result of this work is that we can efficiently classify emotions using information from any part of the brain and with greater classification accuracy than in previous studies.

The future scope of the project is to improve the performance further by finding the better performing feature-classifier combination.

REFERENCES

[1] Panksepp, Jaak (2005). Affective neuroscience: the foundations of human and animal emotions ([Reprint] ed.). Oxford [u.a.]: Oxford Univ. Press. p. 9. ISBN 978-0195096736.

[2] Martınez-Miranda, J. and Aldea, A., 2005. Emotions in human and artificial intelligence. Computers in Human Behavior, 21(2), pp.323-341.

[3] Z. Madhoushi, A. R. Hamdan and S. Zainudin, "Sentiment analysis techniques in recent works," *2015 Science and Information Conference (SAI),* 2015, pp. 288-291, doi: 10.1109/SAI.2015.7237157.

[4] Paolo Petta; Catherine Pelachaud; Roddy Cowie, eds. (2011). Emotion-oriented systems the humaine handbook. Berlin: Springer. ISBN 978-3-642-15184-2.

[5] McKeown, G.; Valstar, M.; Cowie, R.; Pantic, M.; Schroder, M. (January 2012). "The SEMAINE Database: Annotated Multimodal Records of Emotionally Colored Conversations between a Person and a Limited Agent". *IEEE Transactions on Affective Computing.* 3 (1): 5–17. doi:10.1109/T-AFFC.2011.20

[6] Douglas-Cowie, Ellen; Campbell, Nick; Cowie, Roddy; Roach, Peter (1 April 2003). "Emotional speech: towards a new generation of databases". *Speech Communication.* 40 (1–2): 33–60.

[7] Busso, Carlos; Bulut, Murtaza; Lee, Chi-Chun; Kazemzadeh, Abe; Mower, Emily; Kim, Samuel; Chang, Jeannette N.; Lee, Sungbok; Narayanan, Shrikanth S. (5 November 2008). "IEMOCAP: interactive emotional dyadic motion capture database". *Language Resources and Evaluation.* 42 : 335–359. doi:10.1007/s10579-008-9076-6

[8] S. Koelstra et al., “DEAP: A database for emotion analysis; Using physiological signals,” *IEEE Trans. Affect. Comput.,* vol. 3, no. 1, pp. 18–31, 2012, doi: 10.1109/TAFFC.2011.15.

[9] Katsigiannis, Stamos; Ramzan, Naeem (January 2018). "DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals From Wireless Low-cost Off-the-Shelf Devices". *IEEE Journal of Biomedical and Health Informatics.* 22 (1): 98–107. doi:10.1109/JBHI.2017.2688239

[10] Martin, O.; Kotsia, I.; Macq, B.; Pitas, I. The eNTERFACE’05 Audio-Visual Emotion Database. Icdew ’06. *IEEE Computer Society.* pp. 8–. doi:10.1109/ICDEW.2006.145

[11] Dzedzickis, A., Kaklauskas, A., Bucinskas, V. "(2020). Human Emotion Recognition: Review of Sensors and Methods." *Sensors (Basel, Switzerland),* 20(3), 592. https://doi.org/10.3390/s20030592

[12] Y. Li, J. Huang, H. Zhou, and N. Zhong, “Human emotion recognition with electroencephalographic multidimensional features by hybrid deep neural networks,” *Appl. Sci.*, vol. 7, no. 10, 2017, doi: 10.3390/app7101060.

[13] Y. L. Hsu, J. S. Wang, W. C. Chiang, and C. H. Hung, “Automatic ECG-Based Emotion Recognition in Music Listening,” *IEEE Trans. Affect. Comput.*, vol. 11, no. 1, pp. 85–99, 2020, doi: 10.1109/TAFFC.2017.2781732.

[14] Z. Yin, Y. Wang, L. Liu, W. Zhang, and J. Zhang, “Cross- subject EEG feature selection for emotion recognition using transfer recursive feature elimination,” *Front. Neurorobot.,* vol. 11, no. APR, pp. 1–16, 2017, doi: 10.3389/fnbot.2017.00019.

[15] M. Perusquía-Hernández, M. Hirokawa, and K. Suzuki, “A Wearable Device for Fast and Subtle Spontaneous Smile Recognition,” *IEEE Trans. Affect. Comput.,* vol. 8, no. 4, pp. 522–533, 2017, doi: 10.1109/TAFFC.2017.2755040.

[16] S. Shojaeilangari, W. Y. Yau, K. Nandakumar, J. Li, and E. K. Teoh, “Robust Representation and Recognition of Facial Emotions Using Extreme Sparse Learning,” *IEEE Trans. Image Process.*, vol. 24, no. 7, pp. 2140–2152, 2015, doi: 10.1109/TIP.2015.2416634.

[17] R. Lotfidereshgi and P. Gournay, “Biologically inspired speech emotion recognition,” *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, pp. 5135–5139, 2017, doi: 10.1109/ICASSP.2017.7953135.

[18] Niedermeyer E.; da Silva F.L. (2004). Electroencephalography: Basic Principles, Clinical Applications, and Related Fields. Lippincott Williams Wilkins ISBN 978-0-7817-5126-1

[19] “Report of the committee on methods of clinical examination in electroencephalography: 1957,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 10, no. 2, pp. 370–375, 1958, doi: https://doi.org/10.1016/0013-4694(58)90053-1.

[20] S. Shojaeilangari, W. -Y. Yau, K. Nandakumar, J. Li and E. K. Teoh, ”Ro- bust Representation and Recognition of Facial Emotions Using Extreme Sparse Learning,” *in IEEE Transactions on Image Processing*, vol. 24, no. 7, pp. 2140- 2152, July 2015, doi: 10.1109/TIP.2015.2416634.

[21] Yuanlu Kuang and Lijuan Li, ”Speech emotion recognition of decision fusion based on DS evidence theory,” *2013 IEEE 4th International Conference on Software Engineering and Service Science,* 2013, pp. 795-798, doi: 10.1109/IC- SESS.2013.6615425.

[22] A. C. Cruz, B. Bhanu and N. S. Thakoor, ”One shot emotion scores for facial emotion recognition,” *2014 IEEE International Conference on Image Process- ing (ICIP)*, 2014, pp. 1376-1380, doi: 10.1109/ICIP.2014.7025275.

[23] Y. Li, S. Wang, Y. Zhao and Q. Ji, ”Simultaneous Facial Feature Tracking and Facial Expression Recognition,” *in IEEE Transactions on Image Processing,* vol. 22, no. 7, pp. 2559-2573, July 2013, doi: 10.1109/TIP.2013.2253477.

[24] Poursaberi, A., Noubari, H.A., Gavrilova, M. et al. Gauss–Laguerre wavelet textural feature fusion with geometrical information for facial expression iden- tification. *J Image Video Proc* 2012, 17 (2012). https://doi.org/10.1186/1687- 5281-2012-17

[25] Y. Miao, R. Araujo and M. S. Kamel, ”Cross-Domain Facial Expression Recog- nition Using Supervised Kernel Mean Matching,” *2012 11th International Conference on Machine Learning and Applications,* 2012, pp. 326-332, doi: 10.1109/ICMLA.2012.178.

[26] D. S. Moschona, ”An Affective Service based on Multi-Modal Emotion Recog- nition, using EEG enabled Emotion Tracking and Speech Emotion Recogni- tion,” *2020 IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia)*, 2020, pp. 1-3.

[27] C. Paper, “Emotion recognition using brain activity Emotion recognition using brain activity,” no. April, pp. 1– 6, 2016.

[28] K. Schaaff and T. Schultz, “Towards emotion recognition from electroencephalographic signals,” *Proc. - 2009 3rd Int. Conf. Affect. Comput. Intell. Interact. Work. ACII* 2009, 2009, doi: 10.1109/ACII.2009.5349316.

[29] C. A. Frantzidis, C. Bratsas, C. L. Papadelis, E. Konstantinidis, C. Pappas, and P. D. Bamidis, “Toward emotion aware computing: An integrated approach using multichannel neurophysiological recordings and affective visual stimuli,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 3, pp. 589–597, 2010, doi: 10.1109/TITB.2010.2041553.

[30] M. Murugappan, N. Ramachandran, and Y. Sazali, “Classification of human emotion from EEG using discrete wavelet transform,” *J. Biomed. Sci. Eng.,* vol. 03, no. 04, pp. 390–396, 2010, doi: 10.4236/jbise.2010.34054.

[31] L. Brown, B. Grundlehner, and J. Penders, “Towards wireless emotional valence detection from EEG,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 2188–2191, 2011, doi: 10.1109/IEMBS.2011.6090412.

[32] S. A. Hosseini, M. A. Khalilzadeh, M. B. Naghibi-Sistani, and V. Niazmand, “Higher order spectra analysis of EEG signals in emotional stress states,”*Proc. - 2nd Int. Conf. Inf. Technol. Comput. Sci. ITCS 2010*, pp. 60–63, 2010, doi: 10.1109/ITCS.2010.21.

[33] V. Gupta, M. D. Chopda, and R. B. Pachori, “Cross- Subject Emotion Recognition Using Flexible Analytic Wavelet Transform from EEG Signals,” *IEEE Sens. J.,* vol. 19, no. 6, pp. 2266–2274, 2019, doi: 10.1109/JSEN.2018.2883497

[34] B. Nakisa, M. N. Rastgoo, D. Tjondronegoro, and V. Chandran, “Evolutionary computation algorithms for feature selection of EEG-based emotion recognition using mobile sensors,” *Expert Syst. Appl.,* vol. 93, pp. 143–155, 2018, doi: 10.1016/j.eswa.2017.09.062.

[35] P. Ackermann, C. Kohlschein, J. Á. Bitsch, K. Wehrle, and S. Jeschke, "EEG-based automatic emotion recognition: Feature extraction, selection and classification methods," *2016 IEEE 18th Int. Conf. e-Health Networking, Appl. Serv. Heal.* 2016, 2016, doi: 10.1109/HealthCom.2016.7749447.

[36] E. Kroupi, A. Yazdani, and T. Ebrahimi, "EEG correlates of different emotional states elicited during watching music videos," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics),* vol. 6975 LNCS, no. PART 2, pp. 457–466, 2011, doi: 10.1007/978-3-642-24571-8-58.

[37] S. Paul, A. Mazumder, P. Ghosh, D. N. Tibarewala, and G. Vimalarani, "EEG based emotion recognition system using MFDFA as feature extractor," *Proc. 2015 Int. Conf. Robot. Autom. Control Embed. Syst. RACE 2015*, no. February, pp. 18–22, 2015, doi: 10.1109/RACE.2015.7097247.

[38] N. Thammasan, K. Moriyama, K. I. Fukui, and M. Numao, "Continuous Music-Emotion Recognition Based on," *IEICE Trans. Inf. Syst.*, vol. E99D, no. 4, pp. 1234–1241, 2016.

[39] W. L. Zheng and B. L. Lu, "Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks," *IEEE Trans. Auton. Ment. Dev.*, vol. 7, no. 3, pp. 162–175, 2015, doi: 10.1109/TAMD.2015.2431497.

[40] W. L. Zheng, J. Y. Zhu, and B. L. Lu, "Identifying stable patterns over time for emotion recognition from eeg," *IEEE Trans. Affect. Comput.,* vol. 10, no. 3, pp. 417–429, 2019, doi: 10.1109/TAFFC.2017.2712143.

[41] Y. Yang, Q. Wu, M. Qiu, Y. Wang, and X. Chen, "Emotion Recognition from Multi-Channel EEG through Parallel Convolutional Recurrent Neural Network," *Proc. Int. Jt. Conf. Neural Networks,* vol. 2018-July, pp. 352– 359, 2018, doi: 10.1109/IJCNN.2018.8489331.

[42] A. D. Wibawa et al., "Physiological Pattern of Emotion in Elderly Based on Pulse Rate Variability Features: A preliminary study of e-Health monitoring system," *Proc. 2019 6th Int. Conf. Instrumentation, Control. Autom. ICA* 2019, no. August, pp. 121–126, 2019, doi: 10.1109/ICA.2019.8916684.

[43] G. K. P. Veeramallu, Y. Anupalli, S. K. Jilumudi, and A. Bhattacharyya, "EEG based automatic emotion recognition using EMD and Random forest classifier," *2019 10th Int.*

Conf. Comput. Commun. Netw. Technol. ICCCNT 2019, pp. 1–6, 2019, doi: 10.1109/ICCCNT45670.2019.8944903.

[44] S. Parui, A. Kumar, R. Bajiya, D. Samanta, and N. Chakravorty, "Emotion recognition from EEG signal using XGBoost algorithm," *2019 IEEE 16th India Counc. Int. Conf. INDICON 2019 - Symp. Proc.*, pp. 1–4, 2019, doi: 10.1109/INDICON47234.2019.9028978.

[45] N. Pusarla, A. Singh, and S. Tripathi, "Ensemble algorithms for EEG based emotion recognition," *26th Natl. Conf. Commun. NCC 2020*, no. 1, pp. 48–51, 2020, doi: 10.1109/NCC48643.2020.9056002.

[46] J. A. Russell, "A circumplex model of affect," *J. Pers. Soc. Psychol.*, vol. 39, no. 6, pp. 1161–1178, 1980, doi: 10.1037/h0077714.

[47] T. Song, W. Zheng, C. Lu, Y. Zong, X. Zhang, and Z. Cui, "MPED: A multi-modal physiological emotion database for discrete emotion recognition," *IEEE Access,* vol. 7, no. c, pp. 12177–12191, 2019, doi: 10.1109/ACCESS.2019.2891579.

[48] V. M. Joshi and R. B. Ghongade, "IDEA: Intellect database for emotion analysis using EEG signal," *J. King Saud Univ. - Comput. Inf.* Sci., no. xxxx, 2020, doi: 10.1016/j.jksuci.2020.10.007.

[49] P. Pandey and K. R. Seeja, "Subject independent emotion recognition from EEG using VMD and deep learning," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 5, pp. 1730–1738, 2022, doi: 10.1016/j.jksuci.2019.11.003.

[50] Jerome H. Friedman. "Greedy Function Approximation: A Gradient Boosting Machine". *In: The Annals of Statistics* 29.5 (Oct. 2001), pp. 1189–1232 ISSN : 00905364.

List of Publications

[1] Akshay K R, Dr. Muhammed Shanir P P, "Emotion recognition from EEG signals using Machine Learning models," *International Conference on Multimedia, Signal Processing and Communication Technologies (IMPACT)*, 2022.- Processing