

Emotion Analysis using Machine Learning Model and Deep Learning Model on DEAP Dataset

By

Anita Hasan

17301221

Fahim Abrar

21341028

Eshaan Tanzim Sabur

16101255

Iftehaj Muntasir

17301223

Sumaia Sadia Nafisha

17201030

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

Department of Computer Science and Engineering
Brac University
October 2021

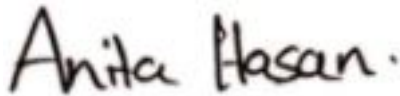
© 2021. Brac University
All rights reserved.

Declaration

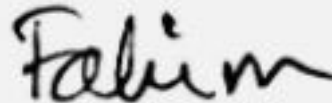
It is hereby declared that

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:



Anita Hasan
17301221



Fahim Abrar
21341028



Eshaan Tanzim Sabur
16101255



Iftehaj Muntasir
17301223



Sumaia Sadia Nafisha
17201030

Approval

The thesis/project titled “Emotion Analysis using Machine Learning Model and Deep Learning Model on DEAP Dataset” submitted by

1. Anita Hasan (17301221)
2. Fahim Abrar (21341028)
3. Eshaan Tanzim Sabu (16101255)
4. Iftehaj Muntasir (17301223)
5. Sumaia Sadia Nafisha (17201030)

Of Summer, 2021 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on October 2, 2021.

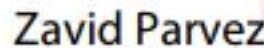
Examining Committee:

Supervisor:
(Member)



Moin Mostakim
Lecturer
Department of CSE
BRAC University

Co-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=BD
Date: 2021.10.02 23:59:27 +11'00'

Dr. Mohammad Zavid Parvez
Assistant Professor
Department of CSE
BRAC University

Head of Department:
(Chair)

Sadia Hamid Kazi
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

Emotion has a significant influence on how you think and interact with others. It serves as a link between how you feel and the actions you take, or you could say it influences your life decisions on occasion. Since the patterns of emotions and their reflections vary from person to person, their inquiry must be based on approaches that are effective over a wide range of population regions. To extract features and enhance accuracy, emotion recognition using brain waves or EEG signals requires the implementation of efficient signal processing techniques. Various approaches to human-machine interaction technologies have been ongoing for a long time, and in recent years, researchers have had great success in automatically understanding emotion using brain signals. In our research, several emotional states were classified and tested on EEG signals collected from a well-known publicly available dataset, the DEAP Dataset, using SVM (Support Vector Machine), KNN (K-Nearest Neighbor), and an advanced Neural Network model RNN (Recurrent Neural Network) trained with LSTM (Long Short Term Memory). The main purpose of this study is to use improved ways to improve emotion recognition performance using brain signals. Emotions, on the other hand, can change with time. As a result, the changes in emotion through time are also examined in our research.

Keywords: Deap Dataset; Machine Learning; EEG; Prediction; Emotionomics

Dedication

This Thesis is in honor of our parents, who have always inspired us to learn. They gave us the strength to grow and develop ourselves. We may not be able to complete our studies without their invaluable support. They have always encouraged us with all endeavors, and have constantly loved us unconditionally. We would also like to dedicate this thesis to our friends who supported us to do better, as well as the participants who assisted us.

Acknowledgement

First and foremost, our praises and appreciation to the Almighty for his mercies during our thesis work, which enabled us to complete it successfully. Secondly, we would want to express our gratitude to our cherished family members, to whom we will be eternally grateful for their love, care and support. We would want to express our heartfelt gratitude to our supervisor, Moin Mostakim, for all of his assistance and persistent guidance. We also appreciate our Co- Supervisor Dr. Mohammad Zavid Parvez's assistance. They have trusted us to do this research alongside them. Without their direction and support, none of this would be possible. Finally, we would want to express our gratitude to all of our faculty members and staff of CSE Department. They have helped us to grow and develop ourselves by providing us a great learning and teaching atmosphere.

Table of Contents

Declaration	i
Approval	ii
Ethics Statement	iv
Abstract	iv
Dedication	v
Acknowledgment	vi
Table of Contents	vii
List of Figures	ix
List of Tables	x
Nomenclature	xi
1 Introduction	1
1.1 An overview to Emotion Recognition and its approaches	1
1.2 Problem Statement	2
1.3 Research Aims and Objectives	2
1.4 Thesis Orientation	2
2 Literature Review	4
2.1 Variety of Analyses with EEG Signals	4
2.2 DEAP Dataset	6
3 Background Study	8
3.1 FFT	8
3.2 RNN and LSTM	8
3.3 DWT	10
4 Proposed Methodology	11
4.1 Materials	11
4.2 Data Visualization	13
4.3 Welch's method	17
4.3.1 Topographical Mapping:	18
4.4 Cross-Validation and Splitting Dataset	18

4.5	Confusion matrix	18
4.6	Machine Learning	19
4.7	Support Vector Machine	19
4.8	K-Nearest Neighbour	19
5	Implementation and Results	21
5.1	Welch's feature Extraction	21
5.2	Topographical Mapping Results	24
5.2.1	Theta Band	25
5.2.2	Alpha Band	25
5.2.3	Beta Band	26
5.2.4	Gamma Band	27
5.3	Analysis of Machine Learning Model Using FFT	28
5.3.1	Arousal-Accuracy	28
5.3.2	Arousal-F1 Score	28
5.3.3	Valence-Accuracy	29
5.3.4	Valence-F1 Score	29
5.3.5	Valence results based on EEG regions and band waves	30
5.4	Implementation with RNN and FFT	35
5.5	Implementation with DWT	41
5.5.1	Arousal-Accuracy	41
5.5.2	Arousal-F1 Score	41
5.5.3	Valence-Accuracy	41
5.5.4	Valence-F1 Score	41
6		45
6.1	Conclusion	45
6.2	Future Work Plan	45
	Bibliography	50

List of Figures

3.1	LSTM cell structure	10
4.1	Sensor position on head to collect data	12
4.2	Data rows Plotting of first participant	14
4.3	Box Plot of Valence and Arousal	15
4.4	Box Plot on Channels	16
4.5	Figure: Distribution table of a confusion matrix	19
5.1	Theta band on Welch's Periodogram	22
5.2	Alpha band on Welch's Periodogram	22
5.3	Beta bandgram on Welch's Periodogram	23
5.4	Gamma bandgram on Welch's Periodogram	23
5.5	Power Spectral Density across the channels	24
5.6	The change in voltage with respect to time with EEG signal	24
5.7	FIR of theta band	25
5.8	Voltage topographical map (theta band)	25
5.9	FIR of alpha band	25
5.10	Voltage topographical map (alpha band))	26
5.11	FIR of beta band	26
5.12	Voltage topographical map (beta band)	26
5.13	Voltage topographical map (gamma band)	27
5.14	Voltage topographical map (gamma band)	27
5.15	Confusion matrix of valence with respect to "theta" band and "central" EEG regions using KNN algorithm	31
5.16	Confusion matrix of valence with respect to "beta" band and "left" EEG regions using KNN algorithm	32
5.17	Confusion matrix of valence with respect to "gamma" band and "right" EEG regions using KNN algorithm	34
5.18	Information of LSTM	35
5.19	Parameter Information of LSTM	36
5.20	Epoch Vs Loss	38
5.21	Epoch vs Accuracy	38
5.22	Epoch Vs Loss (From 51-100 Epochs)	39
5.23	Epoch Vs Accuracy (From 51-100 Epochs)	39
5.24	Epoch Vs Loss (From 101-150 Epochs)	40
5.25	Epoch Vs Accuracy (From 101-150 Epochs)	40
5.26	Confusion matrix of valence with K-NN Classifier	43
5.27	Confusion matrix of arousal with K-NN Classifier	44

List of Tables

4.1	Pandas (Python framework) description of the dataset	13
4.2	Number of trials per each group	14
4.3	Number of trials on each group based on Russell’s circumplex	14
4.4	DEAP Dataset Labels with Emotional States	15
4.5	Table 3.2.6: Details of Trial Transformation	16
5.1	Band waves and emotions	21
5.2	Accuracy of band power values on Arousal using SVM and KNN Classifiers	28
5.3	F1-score of band power values on Arousal using SVM and KNN Classifiers	28
5.4	Accuracy of band power values on Valence using SVM and KNN Classifiers	29
5.5	F1-score of band power values on Arousal using SVM and KNN Classifiers	29
5.6	Valence accuracy results based on EEG regions and EEG bands	30
5.7	Valence F1-score results based on EEG regions and EEG bands	30
5.8	TP, TN, FP, FN Distribution of valence with respect to “theta” band and “central” EEG regions using KNN	31
5.9	Distribution of different metrics on valence with respect to “theta” band and “central” EEG regions using KNN	32
5.10	TP, TN, FP, FN Distribution of valence with respect to “beta” band and “left” EEG regions using KNN	33
5.11	Distribution of different metrics on valence with respect to “beta” band and “left” EEG regions using KNN	33
5.12	Distribution of valence with respect to “gamma” band and “right” EEG regions using KNN	33
5.13	Distribution of different metrics on valence with respect to “gamma” band and “right” EEG regions using KNN	34
5.14	Accuracy result of Arousal using SVM and K-NN	41
5.15	F1-score result of Arousal using SVM and K-NN	41
5.16	Accuracy result of Positive Valence using SVM and K-NN	41
5.17	F1 Score result of Positive Valence using SVM and K-NN	42
5.18	Distribution of different metrics on valence using KNN with DWT	42
5.19	TP, TN, FP, FN Distribution of valence using K-NN	43
5.20	Table: Distribution of different metrics on arousal using KNN with DWT	44
5.21	TP, TN, FP, FN Distribution of valence using K-NN	44

Nomenclature

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

BCI Brain-Computer Interface

CNS Central nervous system

DEAP Dataset for Emotion Analysis using Physiological Signals

EEG Electroencephalogram

FFT Fast Fourier transform

LSTM Long Short Term Memory

ML Machine Learning

RNN Recurrent neural network

SD Standard Deviation

Chapter 1

Introduction

1.1 An overview to Emotion Recognition and its approaches

Emotion is defined as a person's conscious or unconscious behavior that indicates our response to a situation. It is an essential part of human interaction and behavior. Emotion is interconnected with a person's personality, mood, thoughts, motivation, and a variety of other aspects. Emotions are mental states that cause physical and psychological changes in people. Fear, happiness, wrath, pride, anger, panic, despair, grief, joy, tenseness, surprise, confidence, enthusiasm are the common emotions are all experienced by humans [47]. The experience can be both positive or negative. In the light of this, physiological indications such as heart rate, blood pressure, respiration signals, and Electroencephalogram (EEG) signals might be useful in properly recognizing emotions. Emotion recognition has always been a major necessity for humanity, not just for usage in fields like computer science, artificial intelligence, and life science, but also for assisting those who require emotional support. For a long time, experts couldn't figure out a reliable way to identify true human emotion. One method was to use words, facial expression, behavior, and image to recognize one's emotions [3], [5], [10], [41]. Researchers found that subject answers are unreliable for gauging emotion; people are unable to reliably express the strength and impact of their feelings. Furthermore, it is simple to manipulate self-declared emotions, resulting in incorrect findings. Another preferable approach was to utilize software to predict human emotion based on facial expressions, but due to the limited amount of facial expressions and the presence of people with expressing difficulties, the results were inaccurate. On the other hand, there were no such human-machine interface devices to identify those human emotions. As a result, researchers had to shift their focus to approaches that do not rely on subject reactions. The development of Brain-Computer Interface (BCI) and Electroencephalogram (EEG) signals demonstrated more accurate methods for detecting human emotions. It introduced an involuntary approach to get more accurate and reliable results. Previous approaches had limitations in terms of face expression. True emotions, on the other hand, are frequently hidden and suppressed in images, resulting in inaccurate emotion prediction. Involuntary signals, on the other hand, are uncontrollable and detect people's true feelings. They have the ability to express genuine emotions. Affective computing using brain waves or EEG signal tries to create effective methods and systems for recognizing and analyzing human

emotion. The advancement of a reliable human emotion recognition system using EEG signals could help people regulate their emotions and open up new possibilities in fields like education, entertainment, and security and might aid people suffering from Alexithymia or any other psychiatric disease.

1.2 Problem Statement

An electroencephalogram (EEG) is a frequent [36] and dependable technique for measuring changes in brain activity that are detected on the scalp and recorded by a device with a grid of electrodes. EEG data represents the central nervous system's (CNS) brain oscillations and is directly connected to a range of higher-level cognitive functions, including emotion. EEG-based emotion identification has shown more promise than facial expression- and speech-based approaches because natural brain oscillations cannot be purposely concealed. Due to large-scale readings, however, capturing EEG signal data might become complicated. Depending on the manner of feature extraction and analysis, large-scale readings can be increased even more. Utilizing EEG datasets with variety of ML or deep learning models can often prove to be really time consuming. For example, we employed FFT, RNN with LSTM to get a good RNN model by fitting Emotiv Epoch+ in our research, it required a long time to construct a more accurate RNN model. However, the significance of employing EEG signals to recognize or analyze emotional performance cannot be overstated. Involuntary signals appear to be more precisely mirror to detect emotional states, according to experimental researches.

1.3 Research Aims and Objectives

In our research, we want to achieve the following goals:

- Explore the DEAP dataset and its components to have a broader idea about the dataset and how emotion states work
- Utilize both machine learning models and deep learning models to extract features from the dataset and train and test the dataset
- Evaluate the changes in human emotion over time
- Compare all important Machine learning classifiers and metrics to find the best outcome for the prediction.

1.4 Thesis Orientation

The following chapters of the paper are covered in the order shown below:

- The first chapter introduces us to human emotion recognition systems, EEG signals, and their relationship. The chapter also discusses the difficulties that anyone may encounter when using EEG signals to detect emotion. Our goals and objectives are also stated in this chapter.

- In Chapter 2, similar work in this field is discussed in depth, as well as existing methodologies used by researchers. We discussed the results of different types of related research and the used algorithms, techniques and models to reach that conclusion.
- We addressed the some learning techniques we utilized throughout our research in Chapter 3, as well as how they were used in our study. We discussed about FFT and DFT, the relation between these two techniques. We also discussed how FFT can help in the research of EEG signal area. We also gave a brief idea about RNN and LSTM.
- The proposed model and methodologies we will use in this paper are described in detail in Chapter 4. The chapter also covers the train-test splitting ratio and dataset validation, as well as a detailed description of the machine learning and deep learning models we will use in our research.
- The test results and related discussions are presented in Chapter 5. The chapter includes a comparison of various measurement metric approaches. The implementation of machine learning algorithms and neural network model (RNN) is showed here.
- Finally, Chapter 6 brings our research to an end by discussing the thesis, as well as its limitations and future plans.

Chapter 2

Literature Review

There are two parts to this chapter. The first section will discuss various types of studies that focus on various analyses involving EEG data, as well as their conclusions and significance. The second section goes through the previous studies and how they performed on the DEAP Dataset.

2.1 Variety of Analyses with EEG Signals

It appears that, as artificial intelligence advances, physiological signals will be able to be exploited to do emotional computing. Physiological signals like ECG, EEG, EMG, GSR, TEMP, RSP, PPG can be detected by wearable devices. There is increasing evidence that these signals contain information about one's emotions. It has been shown that EEG signals may be utilized to extract various types of features from a range of different investigations. The EEG research community is expanding its reach into a number of different fields. In her research, Vanitha V. et al. [19] aims to connect stress and EEG, and how stress can have both beneficial and bad effects on a person's decision-making process. She also discusses how stress affects one's interpersonal, intrapersonal, and academic performance. She also argues that stress can cause insomnia, lowered immunity, migraines, and other physical problems. Liu et al. [18] proposed a fractal-based algorithm to identify and visualize emotions in real time. They found that gamma band could be used to classify emotion. For emotion recognition, the authors analyzed different kinds of EEG features to find the trajectory of changes in emotion. They then proposed a simple method to track the changes in emotion with time. In this paper, the authors built a bimodal deep auto encoder and a single deep auto encoder to produce shared representations of audios and images. They also explored the possibility of recognizing emotion in physiological signals. Two different fusion strategies were used to combine eye movement and EEG data. The authors tested the framework for cross modal learning tasks. The authors introduce a novel approach that combines deep learning and physiological signals. BCI can identify and distinguish brain waves of people conducting different tasks, according to Fakhruzzaman et al. [14]. Their implementation was based on an automative EPOC to see if the classifier SVM could distinguish between two major training activities: moving the left hand and moving the right foot, as well as four other combinations of the same two actions with the addition of noise, such as nodding or moving the right foot while moving the left hand. EEG data was used to examine patients' sleep patterns, enabling for more

accurate detection and identification of environmental factors when subjects had better or poorer sleep cycles by Pouya Bashivan [13] and Stanislas Chambon [28]. In this study, Convolutional Neural Networks were used to extract time-invariant features, and bidirectional-Long Short-Term Memory was used to automatically predict the sleep transition stages, which resulted in the development of various sleep improvement mobile applications that assist people in developing different exercises to get the best sleep. The researchers made use of data from EEG epochs. They employed a two-step training procedure as well as two publicly available sleep datasets to conduct their research. First, the model was trained to learn filters for extracting time-invariant features from raw channel EEG epochs, which was the first stage. Second, utilizing sequence residual learning, the researchers were able to incorporate the sleep stage transition rules into the recovered features from a sequence of EEG epochs. In his article, Thomas J. et al. [33] points out that CNNs require images as raw data and to transform EEG signals to 2D projection map images. Many other articles have used different classification algorithms to explain differing degrees of success, ranging from 50 to 85 percent. Belakhdar et al. [17] obtained a maximum accuracy of 86.5 percent when detecting drowsiness using the MIT-IH Polysomnographic dataset in their study. Jin et al. [2] while analyzing emotions reported promising results, claiming that combining FFT, PCA, and SVM yielded results that were about 90 percent accurate. As a result, rather than the complexity of the classification algorithm used, the feature extraction stage determines the accuracy of any model. As a result, categorization systems can offer consistent accuracy and recall. In the article [12], the authors also explored several techniques and feature types from time domain, frequency field, statistical features and time frequency features with emotion recognition system. Physiological data were utilized in another research [8] to assess stress. Skin behavior was observed for 5 days for 18 individuals using wrist sensors and their mobile telephone activity, such as call, SMS and location. In the research of Yishu [46], the analysis suggested that individual emotions be recognized by integrating six EEG signal statistical characteristics from the temporal domain as EEG features. A procedure to remove noisy and superfluous canals by use of PCA and ReliefF algorithms was carried out. SVM then was built and used with DEAP to recognize emotions. The results from the research achieved an overall precision rating of 81.87 percent. Google recently completed a project titled FaceNet [16], in which they trained a deep convolutional network on facial photos in the Wild (LFW) dataset and obtained 99.63 percent accuracy. Xing et al. [39] developed a stacked autoencoder (SAE) to breakdown EEG data and classify them using an LSTM model. - The observed valence accuracy rate was 81.1 percent, while the observed arousal accuracy rate was 74.38 percent.

2.2 DEAP Dataset

The DEAP Dataset was utilized by the following writers to analyze emotion states. The impact of identifying emotion states accuracy of brain signal with varied bands of frequencies and number of channels is explained by Hongpei et al. [31]. The K-nearest neighbor Classifier is employed. In gamma frequency ranges, accuracy of around 95 percent is attained, and accuracy rose as the number of channels grew from 10, 18, 18, and 32. Zamanian et al. [34] retrieved Gabor and IMF along with time-domain features and attained an accuracy of 93 percent for 3 and 7 channels using multiclass SVM as a classifier. Chao et al. [29] investigated a deep learning architecture, reaching an arousal rate of 75.92 percent. and 76.83 percent for valence states. Using SVM, Liu et al. [32] categorized the data using time and frequency domain characteristics and attained 70.3 percent and 72.6 percent accuracy, respectively. Mohammadi et al. [24] classified arousal and valence using Entropy and energy of each frequency band and reached an accuracy of 84.05 percent for arousal and 86.75 percent for valence. Xian et al. [22] utilized MCF with statistical, frequency, and nonlinear dynamic characteristics to predict valence and arousal with 83.78 percent and 80.72 percent accuracy, respectively. Among those who have contributed to this work are Maria et al. [23] explored power characteristics based on Russell’s Circumplex Model and applied to SVM with results of 88.4 percent for Valence and 74 percent for Arousal. Singh et al. [26] utilized an SVM classifier to divide emotions into four quadrants based on ERP and latency data. The accuracy rate for single trails ranged from 62.5 percent to 83.3 percent, while for multi-subject trails, they achieved a categorization rate of 55 percent for 24 subjects. Ang et al. [21] developed a wavelet transform and time-frequency characteristics with ANN classification method. For joyful feeling, the classification rate was 81.8 percent for mean and 72.7 percent for standard deviation y . The performance of frequency domain characteristics for sad emotions was 72.7 percent. Krishna and colleagues [27] utilized the Tunable-Q wavelet transform to get sub-bands of EEG data obtained from 24 electrodes by watching 30 video clips. They achieved a classification accuracy of 84.79 percent. According to the survey, researchers have created several ML algorithms that use various characteristics and have an accuracy of 74 to 90 percent. A machine learning algorithm [40] is created in this suggested approach to categorize emotion states into four groups utilizing the channel fusion method using data accessible from the DEAP and SEED-IV data bases independently. The author [45] did a comparison research on domain adaptation strategies on two emotional EEG datasets, as well as a preliminary investigation on cross-dataset emotion recognition, in this work. To summarize, they presented an emotion identification system for EEG data based on the AdaBoost ensemble learning algorithm. You-Yun Lee and Shulan Hsieh [11] conducted a research in which they utilized video snippets to categorize emotions as positive, neutral, or negative [22]. Following that, the researchers examined brain connection as a result of watching video with various levels of emotional categorization. Three indices are used to assess brain connectivity: correlation, coherence, and phase synchronization index. The degree of the link between two brain locations is referred to as correlation. Coherence describes how closely two brain locations operate together at a given frequency. Finally, the phase synchronization index describes the similarity of two signals’ phases. It was discovered that negative emotion showed stronger connections with the occipital site than

neutral or positive emotion in terms of theta and alpha bands, which are distinct frequency ranges. Positive emotion exhibited stronger temporal connections than neutral emotion, particularly in the right hemisphere of the brain. Negative emotions had stronger coherence in the theta, alpha, and beta bands than positive emotions, with the biggest difference in the right parietal and occipital areas, as shown in Figure 3. At each frequency range, positive emotion was more synchronized than negative emotion, especially in the frontal area. Negative states exhibited greater correlation and coherence than positive states overall, particularly in the occipital and temporal areas. This study found a strong relationship between various EEG signals for films with variable perceived emotional output. Yimin Hou and Shuaiqi Chen [37] investigated the connection between EEG signals and emotional states generated by music in another study. It was discovered that when an individual listened to emotion-evoking music, there was more energy of various frequencies in the frontal area. Whereas listening to cheerful music, theta and alpha energy levels were notably high in the occipital area, while beta energy levels were greater around the forehead. Listening to sorrowful music increased the activity of alpha signals. Alhagry et al. [20] developed a deep learning technique for identifying emotions from raw EEG data that used long-short term memory (LSTM) neural networks to learn features from EEG signals and then classified these characteristics as low/high arousal, valence, and liking. The DEAP data set was used to evaluate the -e technique. -The method's average accuracy was 85.45 percent for arousal and 85.65 percent for valence.

Therefore, we present an emotion recognition approach for EEG data based on the collective learning method in this work. To begin, we will attempt to summarize the EEG data and then use FFT and DWT to extract features from the preprocessed EEG signals. We will attempt to analyze the FFT and DWT findings using various machine learning model. We will also try to apply RNN and LSTM to particular EEG channels and various band waves, with the goal of training or creating a model with higher accuracy for learning, testing, and training.

Chapter 3

Background Study

3.1 FFT

The discrete Fourier transform (DFT) is a technique that may be used to convert certain types of function sequences to other sorts of representations. FFT is a mathematical procedure that computes the discrete Fourier transform (DFT) of a sequence. Alternatively, think of the discrete Fourier transform as a transformation that converts the structure of a waveform's cycle into sine components. A Fast Fourier transform may be applied to a variety of signal processing applications, including audio and video processing. It may be useful to have this skill if you are reading things such as sound waves or utilizing image-processing tools. A Fast Fourier transform may be used to solve a variety of different types of equations or to graphically depict a range of various types of frequency activity. Fourier analysis is a signal processing technique used to convert digital signals (x) of length (N) from the time domain to the frequency domain (X) and vice versa. This approach may be used to investigate both continuous and discrete time signals when it comes to temporal signals. The Fourier Transform (FFT) mathematical approach is used to calculate the Discrete Fourier Transform (DFT) or its inverse of a sequence. In terms of results, it is identical to explicitly evaluating the DFT specification, but it does so at a much quicker rate. When estimating the Power Spectral Density of an EEG signal, FFT is a technique that is widely utilized (PSD). PSD is an abbreviation for Power spectral distribution (spectral energy distribution) at a specific frequency. It can be computed directly on the signal using FFT, or indirectly by altering the estimated autocorrelation sequence.

3.2 RNN and LSTM

As the first algorithm to recall its input, RNN (Recurrent Neural Networks) is appropriate for machine learning challenges involving sequential data due to its internal memory. Some algorithms, such as this one, have been important in advancing deep learning so rapidly. Because they are the only neural network type with an internal memory, RNNs are an effective and useful class of neural networks. Deep learning techniques like recurrent neural networks are relatively new. They were created in the 1980s, but their full potential has only just been discovered. RNNs have risen to prominence as computing power has improved, data volumes have exploded, and long short-term memory (LSTM) technology became available in the 1990s. RNNs

may be incredibly precise in forecasting what will happen next because of their internal memory, which allows them to retain key input details. The reason they're so popular is because they're good at handling sequential data kinds like time series and voice. Recurrent neural networks have the advantage over other algorithms in that they can gain a deeper understanding of a sequence and its context. A short-term memory is common in RNNs. When linked with an LSTM, they have a long-term memory as well (more on that later). The use of an example is another good teaching strategy for the notion of memory in a recurrent neural network: Put the word "neuron" into a feed-forward neural network, and see what happens. Individual letters and symbols are removed one at a time. In other words, by the time it gets to the letter "r," the neural network has already forgotten about the letters "n," "e," and "u," making it almost hard to anticipate the next character. A recurrent neural network, on the other hand, can recall such characters because of its internal storage. Produced data is sent back into the network via a process known as cloning. Recurrent neural networks, in a nutshell, enrich the present by incorporating memories from the recent past. Consequently, an RNN is fed the present state as well as recent past. Due to the data sequence providing important information about what will happen next, an RNN may do jobs that other algorithms are unable to complete [48].

Long short-term memory networks (LSTMs) are a sort of recurrent neural network extension that expands memory effectively. As a result, it's well-suited to learning from big experiences separated by long periods of time. RNN extensions that increase memory capacity are known as long short-term memory (LSTM) networks. The layers of an RNN are built using LSTMs. RNNs can either assimilate new information, forget it, or give it enough importance to alter the result thanks to LSTMs, which assign "weights" to data. The layers of an RNN, which is sometimes referred to as an LSTM network, are built using the units of an LSTM. With the help of LSTMs, RNNs can remember inputs for a long time. Because LSTMs store data in a memory comparable to that of a computer, this is the case. The LSTM can read, write, and delete information from its memory. This memory can be thought of as a gated cell, with gated signifying that the cell decides whether to store or erase data (i.e., whether to open the gates) based on the value it assigns to the data. To allocate importance, weights are utilized, which the algorithm also learns. This basically means that it learns over time which data is critical and which is not. [48] Long-Short-Term Memory Networks (LSTMs) are recurrent neural network subtypes (RNN). Because typical RNNs are trained via back-propagation through time (BPTT), which introduces the vanishing/exploding gradient problem, learning long sequences can be problematic. The RNN cell is replaced by a gated cell, such as an LSTMs cell, to address this. Figure 3.1 depicts the architecture of an LSTM cell [51]. These gates determine which data must be stored in memory and which data does not. The addition of memory to an LSTM cell allows it to keep track of previous activities. For LSTMs, cell health is crucial. The LSTM may modify the state of the cell by removing or adding information using three gates. The first gate is a forget gate, which uses a sigmoid layer to decide which information to erase from the cell state [20].

$$f_t = \sigma \times (W_f \times [h_{t-1}, x_t] + b_f) \quad (3.1)$$

The second gate is an input gate with a sigmoid layer that determines which values should be updated and a tanh layer that creates a vector of newly updated values.

$$i_t = \sigma \times (W_i \times [h_{t-1}, x_t] + b_i) \quad (3.2)$$

$$\bar{C}_t = \tanh(W_c \times [h_{t-1}, x_t] + b_c) \quad (3.3)$$

Finally, the output of the current state will be calculated using the updated cell state and a sigmoid layer that determines which parts of the cell state will be output,

$$C_t = f_t * C_{t-1} + i_t * \bar{C}_t \quad (3.4)$$

$$o_t = \sigma(W_o \times [h_{t-1}, x_t] + b_o) \quad (3.5)$$

where σ is the sigmoid activation function that squashes numbers into the range (0,1), \tanh is the hyperbolic tangent activation function that squashes numbers into the range (-1,1). [51]

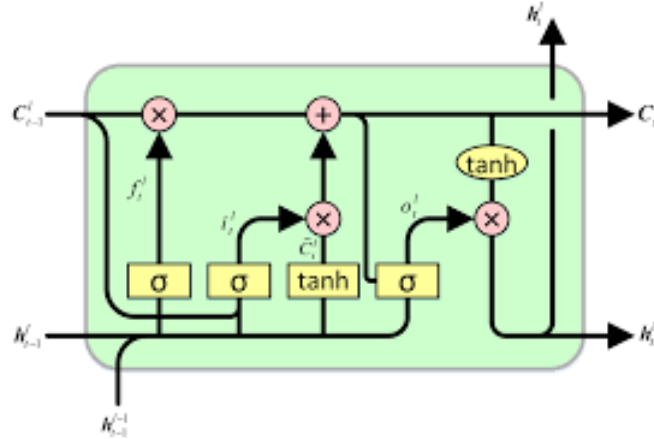


Figure 3.1: LSTM cell structure

3.3 DWT

The Discrete Wavelet Transform (DWT) is a versatile signal processing technique that may be used in a wide range of applications [7]. DWT can denoise and extract characteristics from a wide range of signals, including physiological (EEG, EMG, EOG, ECG, BVP, and others), voice, vibration, acoustic, and biological data. A discrete wavelet transform (DWT) splits a signal into several sets, each of which is a time series of coefficients reflecting the temporal development of the signal in the relevant frequency band [4].

Chapter 4

Proposed Methodology

4.1 Materials

For our research, we have chosen the DEAP [49] dataset. The DEAP dataset for emotion classification is freely available on the internet. A number of physiological signals found in the DEAP dataset can be utilized to determine emotions. It includes information on four main types of states: valence, arousal, dominance, and liking. Due to the use of various sample rates and different types of tests in data gathering, the DEAP Dataset is an amalgamation of many different data types. EEG data was gathered from 32 participants, comprising 16 males and 16 women, in 32 channels. The EEG signals were collected by playing 40 different music videos, each lasting 60 seconds, and recording the results. Following the viewing of each video, participants were asked to rate it on a scale of one to nine points. According to the total number of video ratings received, which was 1280, the number of videos (40) multiplied by the number of volunteers (40) yielded the result (i.e. 32). Following that, the signals from 512 Hz were downsampled to 128 Hz and denoised utilizing bandpass and lowpass frequency filters, as well as a lowpass frequency filter. 512 Hz EEG signals were acquired from the following 32 sensor positions (according to the worldwide 10-20 positioning system): Fp1, AF3, F3, F7, FC5, FC1, C3, T7, CP5, CP1, P3, P7, PO3, O1, Oz, Pz, Fp2, AF4, Fz, F4, F8, FC6, FC2, Cz, T8, CP2, P4, P8, PO4, and O2 (Figure: 4.1)

Sensor positions (eeg)

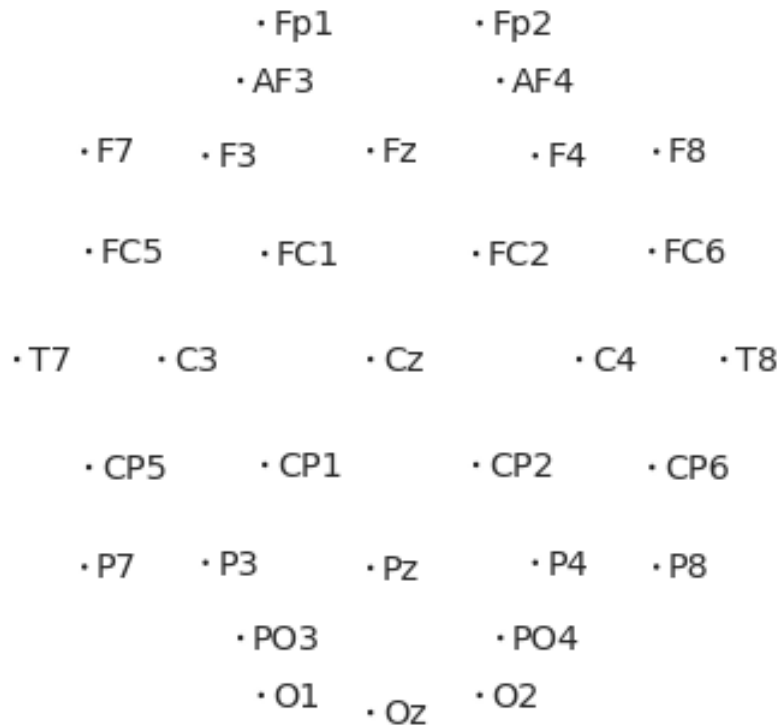


Figure 4.1: Sensor position on head to collect data

The international 10–20 system specifies the placement of the electrodes that will be placed on the human skull in order to detect electroencephalogram (EEG) data. The numerical values "10" and "20" indicate that the distance between adjacent electrodes is 10 percent or 20 percent of the distance between the front and rear of the skull, or between the right and left sides of the skull, respectively, and that the distance between adjacent electrodes is 10 percent or 20 percent of the distance between the front and rear of the skull, or between the right and left sides of the skull. It was also possible to take a frontal face video of each of the 22 participants. Several signals, including EEG, electromyograms, breathing region, plethysmographs, temperature, and so on, were gathered as 40 channel data during each subject's 40 trials, with each channel representing a different signal. EEG data is stored in 32 of the 40 available channels for research, which is a significant amount.

4.2 Data Visualization

Emotions can be classified into various categories, the most prominent of which are as follows: When it comes to negative emotions (such as anger and anxiety), disgust, shame, fear, and sadness are instances. When it comes to good emotions (such as affection and amusement), happiness, joy, pleasure, pride, and relief are examples. Arousal-valence space is an alternative to using continuous values to define emotions. Arousal relates to the intensity of an emotion, i.e., the power of the related emotional state, whereas Valence refers to the amount to which an emotion is positive or negative. For this research, we focused on valence and arousal. There were 1240 trials number of experiments in this study to obtain valence and arousal ratings. We gathered information from ratings on a scale of one to nine. The dataset defined the values from one to nine. To begin with, we plotted 40 data rows of 100 emotion

	Valence	Arousal
count	1240	1240
mean	5.252435	5.144210
sd	2.136497	2.031844
min	1.00	1.00
25% or First Quartile or Q ₁	3.80	3.68
50% or Median or Q ₂	5.04	5.165
75% or Third Quartile or Q ₃	7.05	6.94
max	9.00	9.00

Table 4.1: Pandas (Python framework) description of the dataset

distribution points in numerical number of the first participant to visualize the data. The goal was to observe how the distribution of valence and arousal was distributed.

We extracted valence and arousal ratings from the dataset. The combination of Valence and Arousal can be converted to emotional states: High Arousal Positive Valence (Excited, Happy), Low Arousal Positive Valence (Calm, Relaxed), High Arousal Negative Valence (Angry, Nervous) and Low Arousal Negative Valence (Sad, Bored). We have analyzed the changes in emotional state along with the number of trials for each group by following Russell’s circumplex model. The question might rise about the way to classify the dataset. Russell’s circumplex model can help classify the DEAP dataset. Russell’s methodology for visualizing the scale with the real numbers 0–10, the DEAP dataset employs self-assessment manikins (SAMs) [1]. 1–5 and 5–9 were chosen as the scales based on self-evaluation ratings [9], [15], [35]. The label was changed to “positive” if the rating was greater than or equal to 5, and to “negative” if it was less than 5. We utilized a different way to determine “positive” and “negative” values. The difference in valence and arousal was rated on a scale of 1 to 9. It is not a good idea, in our opinion, to categorize the dataset using a mean value because a particular “number” may express differently for different users. As a result, we used median values to discriminate between “positive” and “negative”



Figure 4.2: Data rows Plotting of first participant

integers. We looked to see if each trial had a positive or negative valence, as well as a positive or negative arousal level. In general, values larger than the median are seen as "positive," while those less than the median are regarded as "negative."

Positive Valence	661
Negative Valence	579
High Arousal	620
Low Arousal	620

Table 4.2: Number of trials per each group

High Arousal Positive Valence	353
Low Arousal Positive Valence	308
High Arousal Negative Valence	267
Low Arousal Negative Valence	312

Table 4.3: Number of trials on each group based on Russell's circumplex

As a result, four labels have been created: high arousal low valence (HALV), low arousal high valence (LAHV), high arousal high valence (HAHV), and low arousal low valence (LALV).

A box plot is a graphical representation of numerical data groupings through their quartiles used in descriptive statistics. Lines extending from the boxes (whiskers) can be seen in box plots, suggesting variability outside the upper and lower quartiles. A boxplot is a standardized data visualization approach that uses a five-number

No.	Proposed Labels	Emotional States
1.	High Arousal Low Valence	Calm, Relaxed
2.	High Arousal High Valence	Happy, Excited
3.	Low Arousal High Valence	Angry, Nervous
4.	Low Arousal Low Valence	Sad, Bored

Table 4.4: DEAP Dataset Labels with Emotional States

summary: minimum, maximum, sample median, first, and third quartiles.

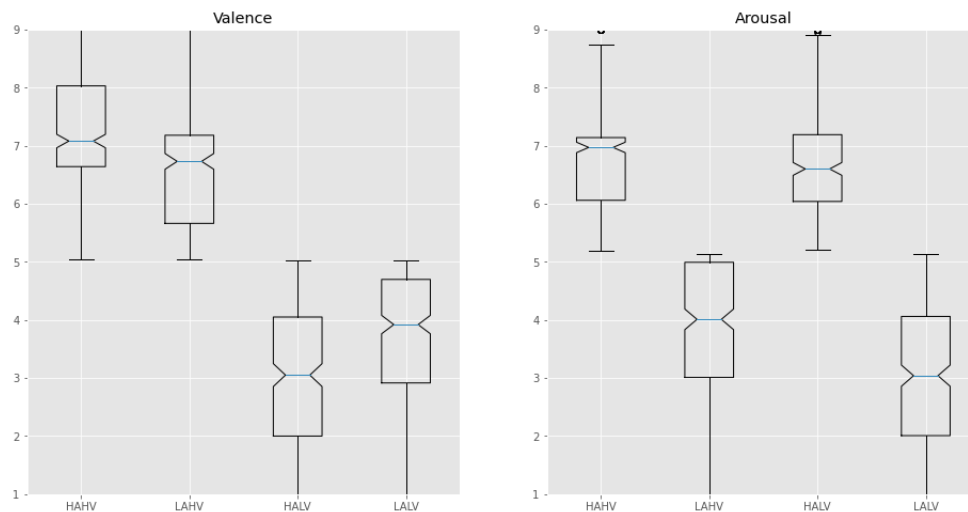


Figure 4.3: Box Plot of Valence and Arousal

The mean value is shown by the blue line in the boxplot. It specifies the location with the greatest number of values. In the boxplot, we can observe that there have been four groups of means in terms of valence: HAHV (mean=7.22), LAHV (mean=6.57), HALV (mean=3.08), and LALV (mean=3.58). On the other side, we can also obtain values from the standpoint of arousal. HAHV has a mean of 6.86, LAHV has a mean of 3.8, HALV has a mean of 6.8, and LALV has a mean of 3.11.

One-hot encoding is a technique for transforming categorical data into a format that machine learning algorithms can use to enhance prediction accuracy. To encode category information, it employs a one-hot numeric array. Each category variable level is compared to a preset reference level [50]. A single variable [25]te with n observations and d different values is transformed into d binary variables with n observations each using a single hot encoding. The experiment uses 1 as a positive and 0 as a negative to study the dataset in our research.[6] For our research, we have transformed the trial description on a scale from 0 to 1. In the tables, the transformed values are shown.

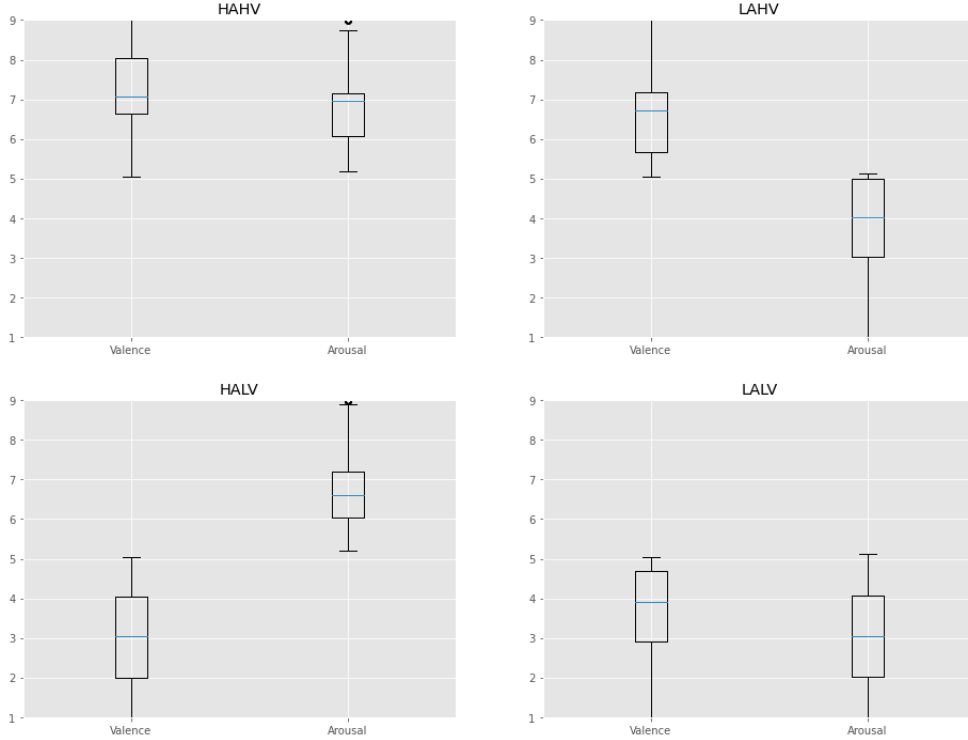


Figure 4.4: Box Plot on Channels

	Positive Valence	High Arousal
count	1240	1240
mean	0.533065	0.50
std	0.499107	0.500202
min	0.00	0.00
25% or First Quarryile or Q_1	0.00	0.00
50% or Median or Q_2	1.00	0.50
75% or Third Quarntile or Q_3	1.00	1.00
max	1.00	1.00

Table 4.5: Table 3.2.6: Details of Trial Transformation

In the DEAP dataset, there are 40 channels which include 32 EEG channels and breathing region, GSR, temperature, plethysmographs are the rest. From all of the channels, 40 x 8064 data are collected. Extracting features from EEG data can be done in a variety of methods. Periodogram and power spectral density calculations and combining band waves of various frequencies are required for feature extraction. Periodogram and PSD calculations can also be done in a variety of methods. PSD can be calculated using the periodogram [36]. It is determined by the modulus squared of the signal's Fourier transform and is made of a frequency decomposition:

$$S(f) = \frac{\Delta t}{N} \left| \sum_{n=0}^{N-1} x_n e^{\frac{-2\pi i k n}{N}} \right|^2 \quad (4.1)$$

Where $S(f)$ = PSD of x_n

Δt = space between samples
 x_n = input sequence
 N = elements in input sequence

4.3 Welch's method

In our research, we used Welch's approach for extracting features. The Welch technique is nowadays commonly used to predict the power spectrum. The Welch method is [38] a modified segmentation scheme for calculating the average periodogram. Generally the Welch method of the PSD can be described by the equations below, the power spectra density, $P(f)$ equation is defined first. Then, for each interval, the Welch Power Spectrum, $P_{\text{welch}}(f)$, is given as the mean average of the periodogram.

$$P(f) = \frac{1}{MU} \left| \sum_{n=0}^{M-1} x_i(n)w(n)e^{-j2\pi f} \right|^2 \quad (4.2)$$

$$P_{\text{welch}}(f) = \frac{1}{L} \sum_{i=0}^{L-1} P(f) \quad (4.3)$$

The power spectral density (PSD) shows how a signal's power is distributed in the frequency domain. Among the PSD estimators, Welch's method and the multitaper approach have demonstrated the best results [30]. The input [43] signal $x[n]$, $n = 0, 1, 2, \dots, N-1$ is divided into a number of overlapping segments. Let M be the length of each segment, using $n=0, 1, 2, \dots, M-1, M$.

$$x_i = x\left[i \times \frac{M}{2} + n\right] \quad (4.4)$$

where $n=0, \dots, M-1, i=0, 1, 2, \dots, N-1$

Each segment is given a smooth window $w(n)$. In most cases, we employ the Hamming window at a time. The Hamming window formula for each segment is as follows:

$$w(n) = 0.54 - 0.46 \cos\left[\frac{2n\pi}{M}\right] \quad (4.5)$$

Here,

$$U = (1/M) \sum_{n=0}^{M-1} w^2(n) \quad (4.6)$$

denotes the mean power of the window $w(n)$. So,

$$MU = \sum_{n=0}^{M-1} w^2(n) \quad (4.7)$$

denotes the energy of the window function $w(n)$ with length M .

It is to be noted that, in the 2nd equation L denotes the number of data segment.

4.3.1 Topographical Mapping:

The use of topographical mapping to show EEG data is quite useful. Voltage activity will be examined in our study. The black dots (fig p) correspond to the approximate physical placements of each electrode on the scalp, allowing humans to respond because it lets us to see changes in data at a single or multiple time points [44]. This approach is a particularly powerful visualization method.

4.4 Cross-Validation and Splitting Dataset

Cross-validation is a technique in applied machine learning that is used to assess the competence of a machine learning model on unknown data. In other words, a small sample size is used to evaluate how the model would perform in general when used to forecast data. In this study, we use a dataset ratio of 0.70 for training and 0.30 for testing, with a random state value of 42. We utilized a standard scaler to balance the model before fitting any machine learning model since variables recorded at different scales do not always contribute equally to model fitting and trained function, which can lead to bias. As a consequence, we obtained a balanced model that may be used to apply machine learning techniques to this prospective situation. The idea of our research is to maximize the output of the result but when we are limiting the train-test by setting ratio, every data point we use for training the dataset are lost for the testing, and vice versa. To avoid this dilemma, we can split the dataset into “k” numbers of equal size. In k-fold cross validation, k separate learning experiments are needed. For our research, the splitting value, “k” is equal to 5. For validation, ”Accuracy” is the most popular metric. However, a model’s performance cannot be judged based only by the accuracy. So, we have used other metrics, such as - precision, recall, and f-score. The metrics were calculated using the mean of metrics for all the folds.

4.5 Confusion matrix

When describing the performance of a classification model, a confusion matrix is often used to characterize the model’s performance. In terms of delivering probabilistic results, it’s a fantastic model to use. TP, FP, TN, and FN are the four possible combinations of anticipated and actual values that can be found in a table: T, F, and N. It is useful for a variety of calculations, including the calculation of Recall, Precision, Specificity, Accuracy, and F1-score. With the help of a confusion matrix, we can evaluate four measurement metrics

$$\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{Total Predictions}} \quad (4.8)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.9)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4.10)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4.11)$$

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Figure 4.5: Figure: Distribution table of a confusion matrix

$$F = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (4.12)$$

4.6 Machine Learning

Machine learning is a technique for teaching a dataset to recognize specific patterns. We can improve the accuracy of our predictions by training a model over a set of data and providing it with a classifier algorithm to use to train over and learn from those data. We used SVM and k-NN classifier to train, test and predict from the datasets. We utilized the mentioned classifiers by using different metrics in our research.

4.7 Support Vector Machine

A support vector machine (SVM) [42] is a type of supervised machine learning method that is commonly used in classification and regression models. SVM is powerful machine learning in managing nonlinear data for solving regression and classification problems when using an appropriate kernel function. The classifier of the SVM is called SVC(Support Vector Classifier). SVM chooses the best separator to classify the vectors or data points. SVM classifier uses various types of kernels. For FFT analysis, we used linear kernel and for DWT analysis we used sigmoid kernel. The equation for linear kernel is written below:

$$F(x, x_j) = \text{sum}(x \cdot x_j) \quad (4.13)$$

The equation for sigmoid kernel is written below:

$$F(x, x_j) = \tanh(\alpha x a y + c) \quad (4.14)$$

4.8 K-Nearest Neighbour

According to the authors [42], Nearest Neighbor (NN) is a simple and widely used classifier for pattern recognition tasks. K-Nearest Neighbor (KNN) came from NN concept, is a technique for solving regression and classification issues. The number of k in a k-NN classifier has a significant impact on the classification outcome, however determining the optimal value of k is difficult. The weighted k-NN algorithm's

approach is built by introducing a neighbor weight that varies exponentially based on the neighbor's square Euclidean distance. Since the value of k in our study is 5, it will look for the 5 nearest neighbors to that data point, which may provide the best classification result. The equation of the distance is:

$$Distance(x, y) = \sqrt{\sum_i (x_i - y_i)^2} \quad (4.15)$$

Chapter 5

Implementation and Results

5.1 Welch's feature Extraction

In our research, we tried to come up with a relation among EEG channel, time and voltage using Welch's Periodogram (Figure 5.2-5.5) with the help of theta, alpha, beta and gamma waves. The band waves identify the following emotions. From

Band Waves	Frequency (Hz)	Features or emotions
Theta	4 – 8	Drowsiness, Emotional Connection, Intuition, Creativity
Alpha	8 – 16	Reflection & Relaxation
Beta	16 – 32	Concentration, Problem Solving, Memory
Gamma	32 – 64	Cognition, Perception, Learning, Multi-tasking

Table 5.1: Band waves and emotions

Welch's method, we get the following results. These figures show us the relationship with power spectral density over band wave's frequencies. From the periodogram figures, we get to see the peak points of the signal. Suppose, for theta band, the peak point is somewhere in between 5-7Hz, for alpha band, the peak point is somewhere in between 10-12 Hz. For beta band, the peak point is somewhere in between 16-19 Hz. And, for gamma band, it can not be detected with eye easily. The data and information are collected from the EEG signal for observing the relationship among the signals, power spectral density and band waves.

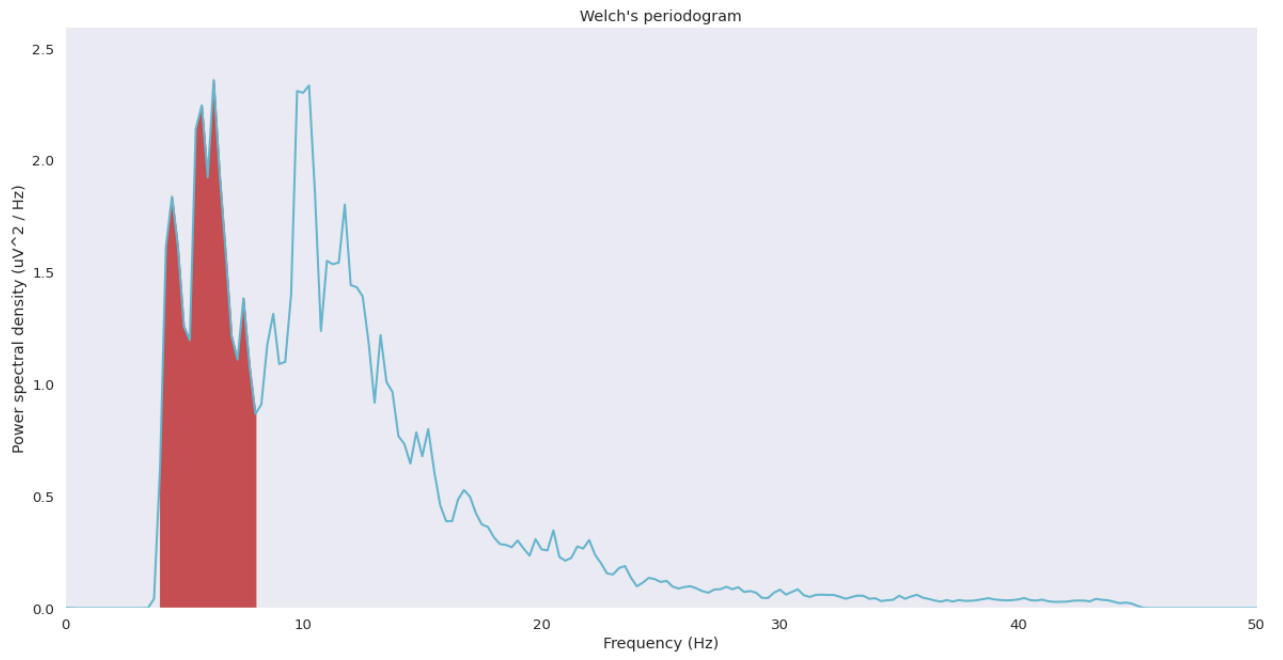


Figure 5.1: Theta band on Welch's Periodogram

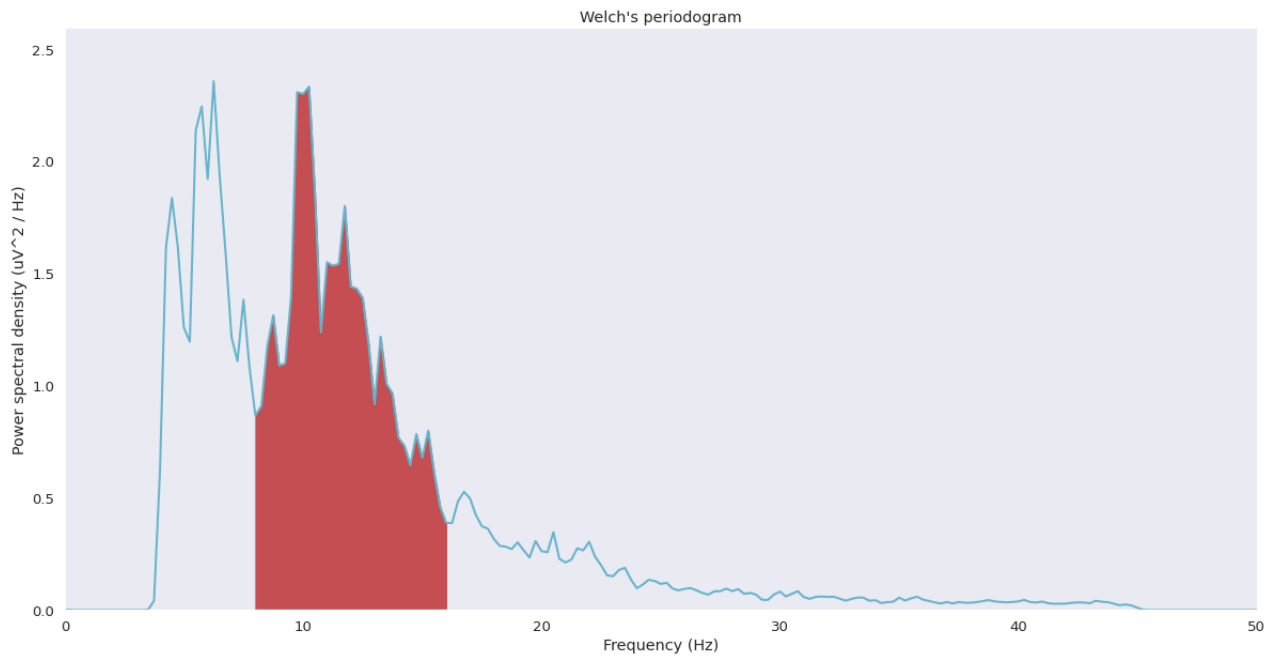


Figure 5.2: Alpha band on Welch's Periodogram

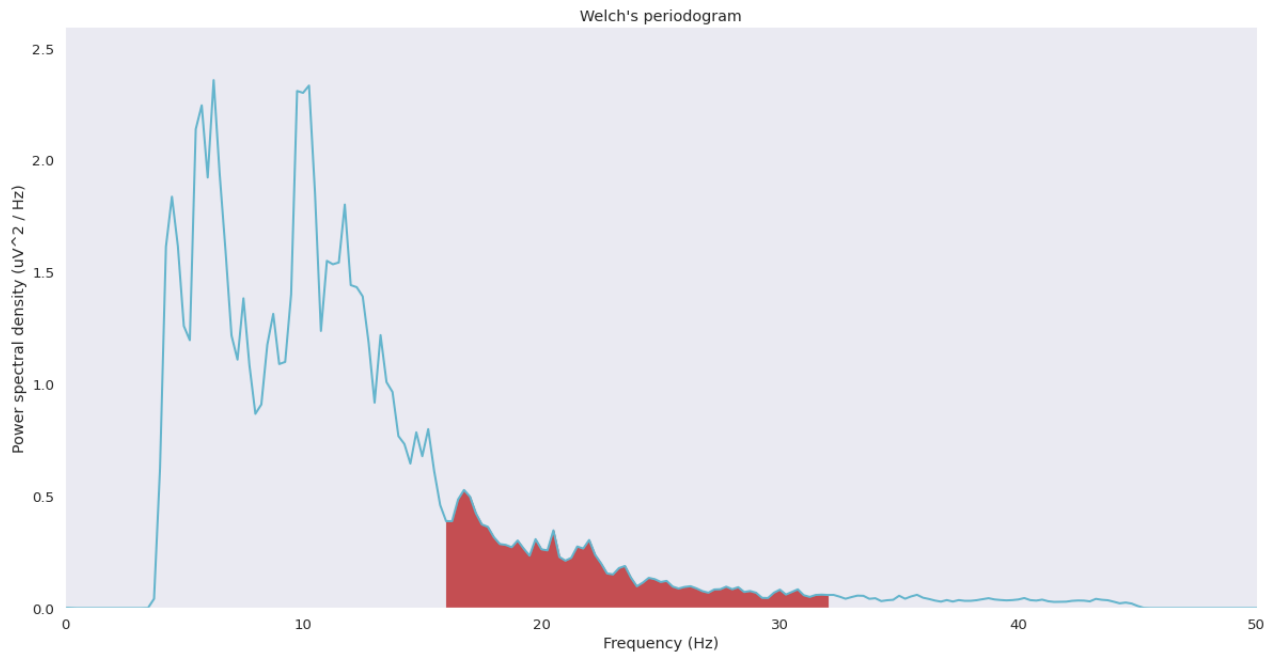


Figure 5.3: Beta bandgram on Welch's Periodogram

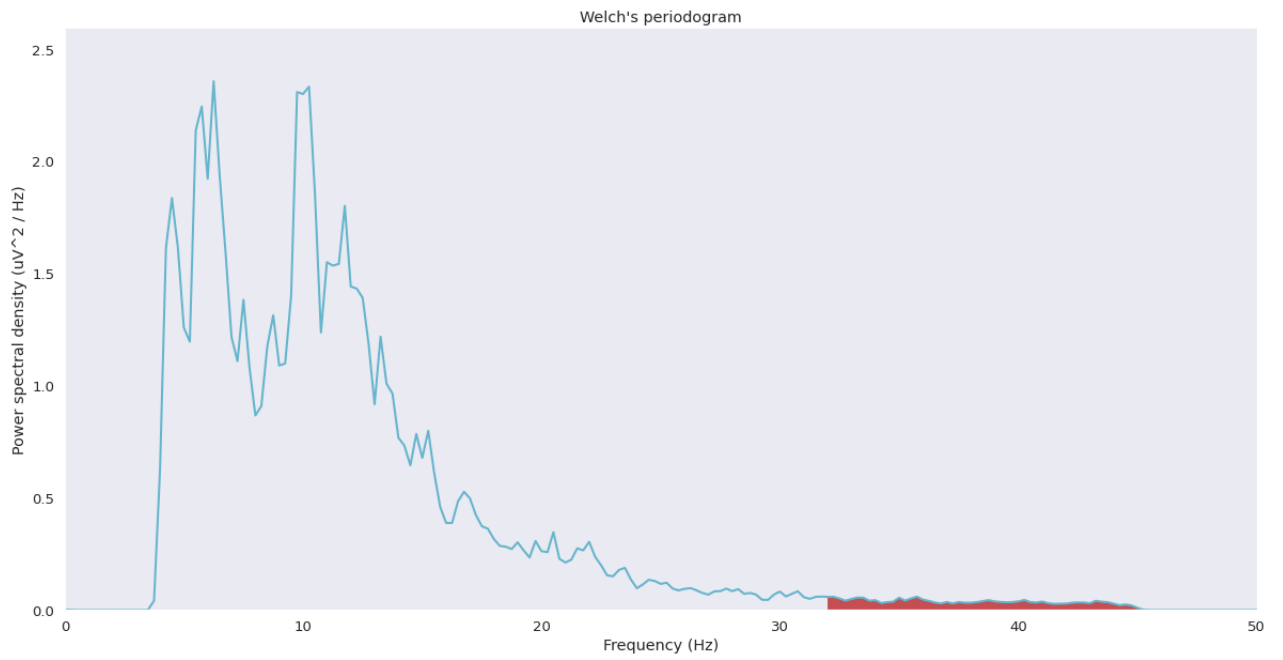


Figure 5.4: Gamma bandgram on Welch's Periodogram

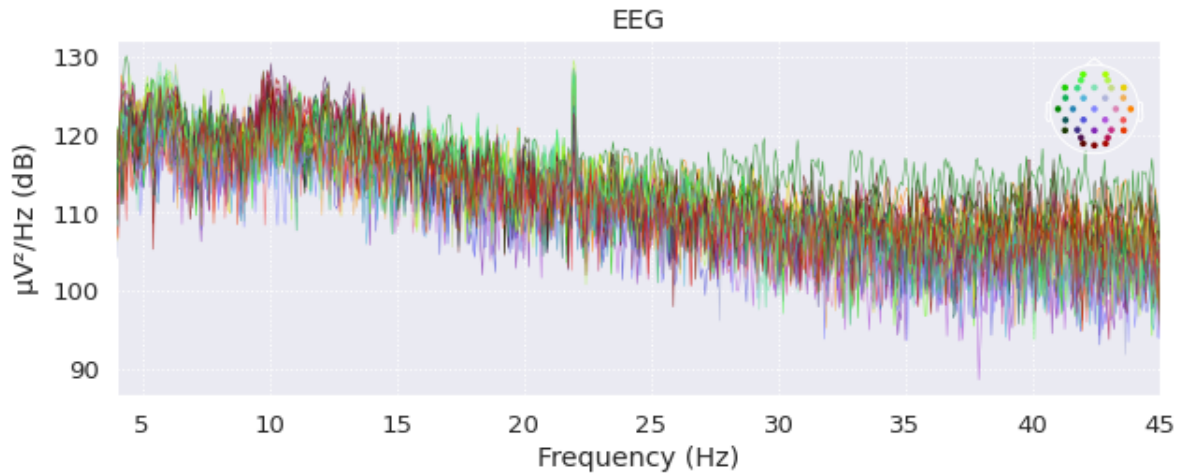


Figure 5.5: Power Spectral Density across the channels

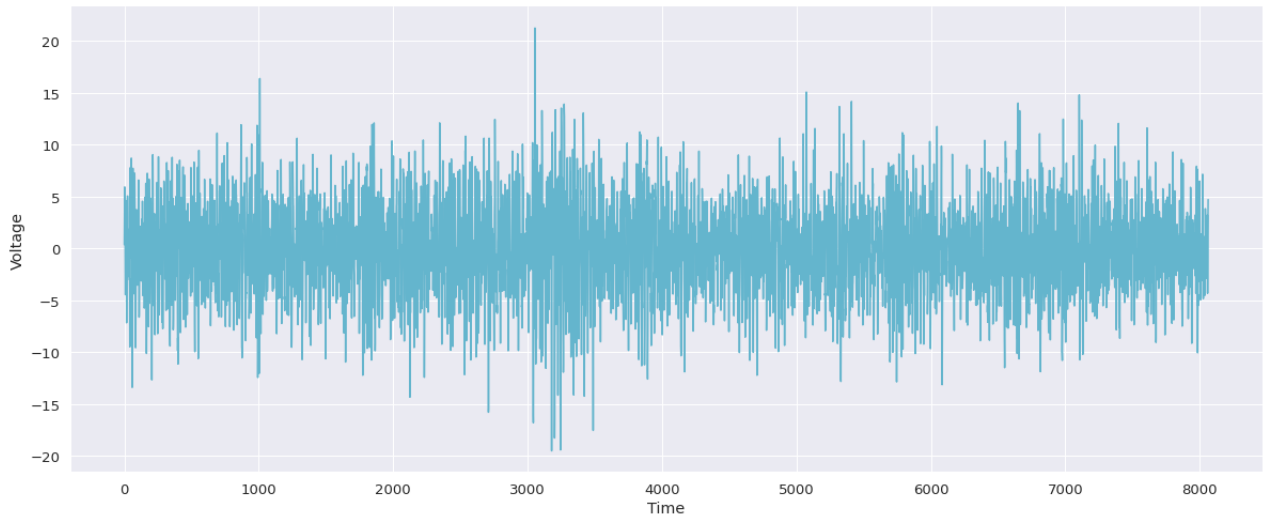


Figure 5.6: The change in voltage with respect to time with EEG signal

5.2 Topographical Mapping Results

Human emotions aren't always consistent. At the same moment, we can experience a variety of emotions. With the use of some specific FIR Filter parameters, we applied topographical mapping on EEG signals to represent the change in emotion for several time points utilizing different frequency ranges that denote bandgram such as theta, alpha, beta, and gamma.

We used the same time points, 0.153 second, 0.173 second, 0.213 second, 0.233 second, 0.253 second, and 0.273 second, to assess the changes based on different frequency ranges.

5.2.1 Theta Band

```
Setting up band-pass filter from 4 - 8 Hz

FIR filter parameters
-----
Designing a one-pass, zero-phase, non-causal bandpass filter:
- Windowed time-domain design (firwin) method
- Hamming window with 0.0194 passband ripple and 53 dB stopband attenuation
- Lower passband edge: 4.00
- Lower transition bandwidth: 2.00 Hz (-6 dB cutoff frequency: 3.00 Hz)
- Upper passband edge: 8.00 Hz
- Upper transition bandwidth: 2.00 Hz (-6 dB cutoff frequency: 9.00 Hz)
- Filter length: 213 samples (1.664 sec)
```

Figure 5.7: FIR of theta band

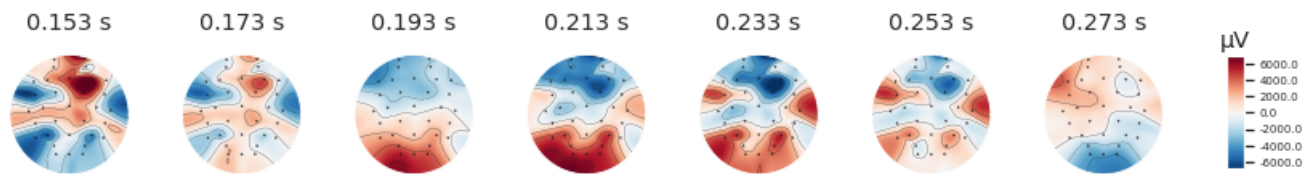


Figure 5.8: Voltage topographical map (theta band)

5.2.2 Alpha Band

```
Setting up band-pass filter from 8 - 16 Hz

FIR filter parameters
-----
Designing a one-pass, zero-phase, non-causal bandpass filter:
- Windowed time-domain design (firwin) method
- Hamming window with 0.0194 passband ripple and 53 dB stopband attenuation
- Lower passband edge: 8.00
- Lower transition bandwidth: 2.00 Hz (-6 dB cutoff frequency: 7.00 Hz)
- Upper passband edge: 16.00 Hz
- Upper transition bandwidth: 4.00 Hz (-6 dB cutoff frequency: 18.00 Hz)
- Filter length: 213 samples (1.664 sec)
```

Figure 5.9: FIR of alpha band

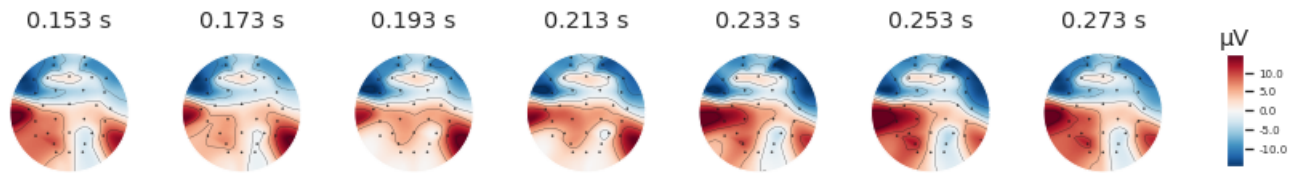


Figure 5.10: Voltage topographical map (alpha band))

5.2.3 Beta Band

```
Setting up band-pass filter from 16 - 32 Hz

FIR filter parameters
-----
Designing a one-pass, zero-phase, non-causal bandpass filter:
- Windowed time-domain design (firwin) method
- Hamming window with 0.0194 passband ripple and 53 dB stopband attenuation
- Lower passband edge: 16.00
- Lower transition bandwidth: 4.00 Hz (-6 dB cutoff frequency: 14.00 Hz)
- Upper passband edge: 32.00 Hz
- Upper transition bandwidth: 8.00 Hz (-6 dB cutoff frequency: 36.00 Hz)
- Filter length: 107 samples (0.836 sec)
```

Figure 5.11: FIR of beta band

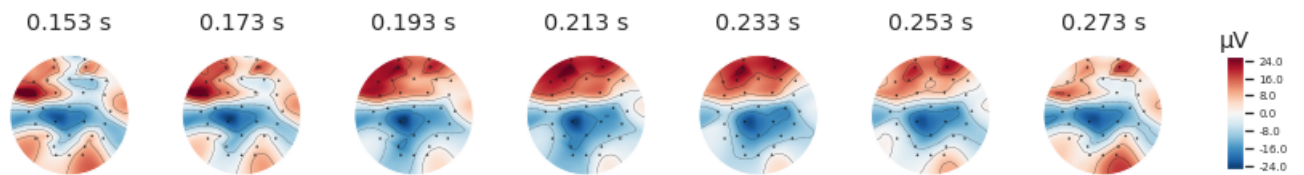


Figure 5.12: Voltage topographical map (beta band)

5.2.4 Gamma Band

```
Setting up band-pass filter from 32 - 64 Hz

FIR filter parameters
-----
Designing a one-pass, zero-phase, non-causal bandpass filter:
- Windowed time-domain design (firwin) method
- Hamming window with 0.0194 passband ripple and 53 dB stopband attenuation
- Lower passband edge: 32.00
- Lower transition bandwidth: 8.00 Hz (-6 dB cutoff frequency: 28.00 Hz)
- Upper passband edge: 63.99 Hz
- Upper transition bandwidth: 0.01 Hz (-6 dB cutoff frequency: 64.00 Hz)
- Filter length: 42241 samples (330.008 sec)
```

Figure 5.13: Voltage topographical map (gamma band)

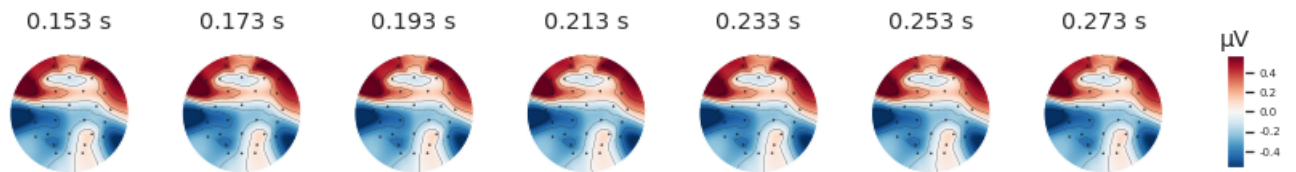


Figure 5.14: Voltage topographical map (gamma band)

5.3 Analysis of Machine Learning Model Using FFT

After working on Welch's feature extraction method and topographical mapping, we calculated the mean of accuracy, standard deviation and the time taken to get the results from data visualization using FFT. For our research, we have processed new datasets with 6 EEG regions and four band power values such as theta, alpha, beta and gamma. We have divided the EEG regions based on left (Fp1, AF3, F7, FC5, T7), right (Fp2, AF4, F8, FC6, T8), frontal (F3, FC1, Fz, F4, FC2), parietal (P3, P7, Pz, P4, P8), occipital (O1, Oz, O2, PO3, PO4) and central (CP5, CP1, Cz, C4, C3, CP6, CP8) sensor positions. The research calculates mean, std, min, first quartile, median, third quartile and max values of 1240 trials of the mentioned six regions based on four band power values. For this research, we used SVM and K-NN classifiers. SVM classifier used "linear" kernel in this research.

5.3.1 Arousal-Accuracy

	SVM	KNN
Mean of accuracy	58.52%	62.32%
STD	0.033537092407755495	0.01809672118602032
Time in minutes	2.8121	0.1343

Table 5.2: Accuracy of band power values on Arousal using SVM and KNN Classifiers

5.3.2 Arousal-F1 Score

	SVM	KNN
Mean of F1	58.35%	63.28%
STD	0.05551135016067932	0.019482042202837713
Time in minutes	2.76777	0.1281

Table 5.3: F1-score of band power values on Arousal using SVM and KNN Classifiers

5.3.3 Valence-Accuracy

	SVM	KNN
Mean of accuracy	56.79%	56.92%
STD	0.03921589438580304	0.04104184725930891
Time in minutes	3.3126	0.1471

Table 5.4: Accuracy of band power values on Valence using SVM and KNN Classifiers

5.3.4 Valence-F1 Score

	SVM	KNN
Mean of F1	65.61%	60.48%
STD	0.0437543076694958	0.04882980750657068
Time in minutes	3.1484	0.1253

Table 5.5: F1-score of band power values on Arousal using SVM and KNN Classifiers

Moreover, the research tries to compare the accuracy rate and F1-score of valence label based on top EEG regions, bands, EEG region per each band and bands with highest scores per each EEG region using K-NN classifier.

5.3.5 Valence results based on EEG regions and band waves

Valence Accuracy Results in “%”							
	Left	Frontal	Right	Central	Parietal	Occipital	
Theta	55.49	57.80	56.07	64.16	58.96	58.96	
Alpha	56.65	58.38	55.49	63.01	62.43	56.65	
Beta	56.65	61.27	58.96	56.65	61.27	54.34	
Gamma	58.96	56.65	57.23	53.18	58.96	55.49	

Table 5.6: Valence accuracy results based on EEG regions and EEG bands

Valence F1-Score Results in “%”							
	Left	Frontal	Right	Central	Parietal	Occipital	
Theta	60.91	64.04	65.45	70.48	65.02	66.98	
Alpha	63.77	63.27		61.31	68.00	68.90	
						63.05	
Beta	62.69	68.25	65.02	61.93	67.63	58.64	
Gamma	64.68	63.41	64.42	59.70	64.32	61.69	

Table 5.7: Valence F1-score results based on EEG regions and EEG bands

To observe confusion matrix, we have worked on top combinations for valence. In the first experiment, we plot confusion matrix of valence with respect to “theta” band and “central” EEG regions using KNN algorithm.

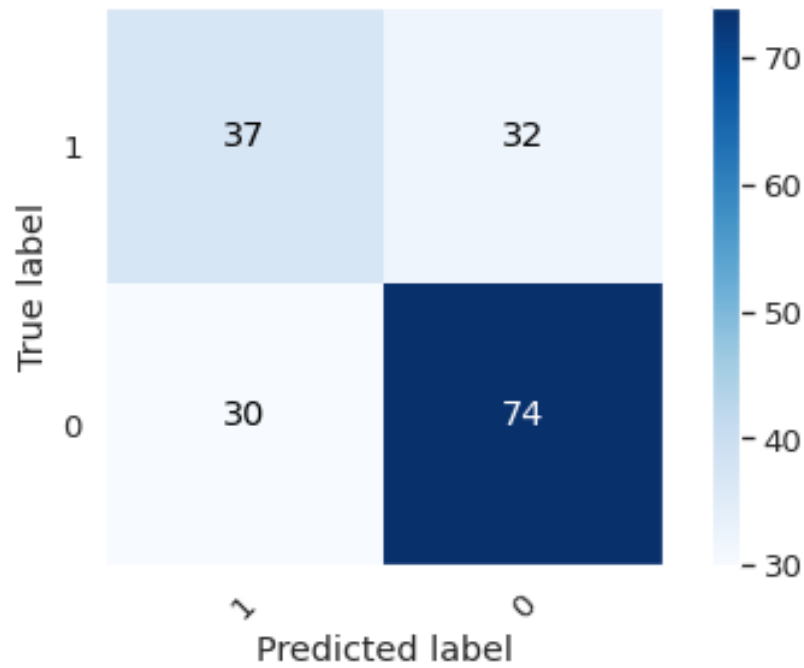


Figure 5.15: Confusion matrix of valence with respect to “theta” band and “central” EEG regions using KNN algorithm

N=173	Positive	Negative
True	37	74
False	32	30

Table 5.8: TP, TN, FP, FN Distribution of valence with respect to “theta” band and “central” EEG regions using KNN

	Precision	Recall	F1-Score	Support
0	0.55	0.54	0.54	69
1	0.70	0.71	0.70	104
Accuracy			0.64	173
Macro Avg	0.63	0.62	0.62	173
Weighted Average	0.64	0.64	0.64	173

Table 5.9: Distribution of different metrics on valence with respect to “theta” band and “central” EEG regions using KNN

We have also observed 173 distribution of valence with respect to “beta” band and “left” EEG regions using KNN algorithm.

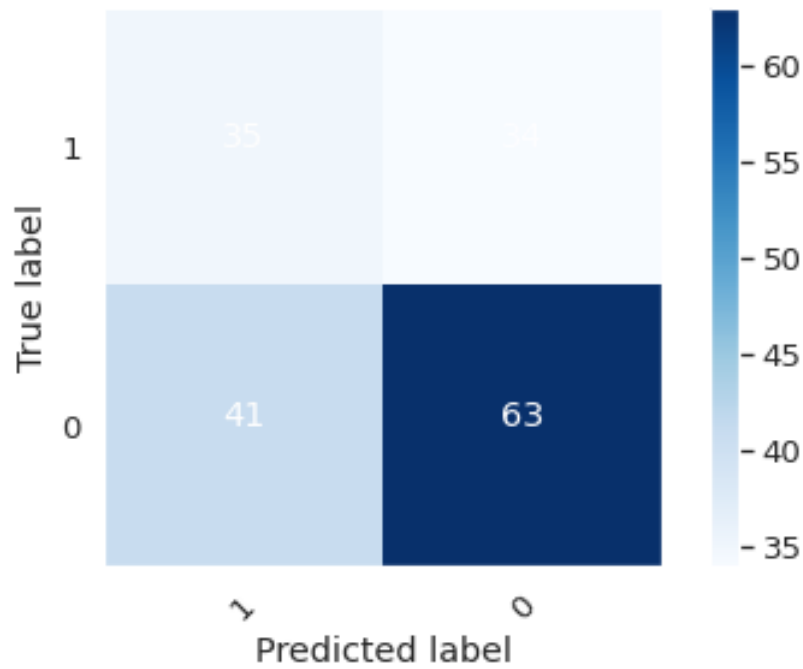


Figure 5.16: Confusion matrix of valence with respect to “beta” band and “left” EEG regions using KNN algorithm

N=173	Positive	Negative
True	35	63
False	34	41

Table 5.10: TP, TN, FP, FN Distribution of valence with respect to “beta” band and “left” EEG regions using KNN

	Precision	Recall	F1-Score	Support
0	0.46	0.51	0.48	69
1	0.65	0.61	0.63	104
Accuracy			0.57	173
Macro Avg	0.56	0.56	0.55	173
Weighted Average	0.57	0.57	0.57	173

Table 5.11: Distribution of different metrics on valence with respect to “beta” band and “left” EEG regions using KNN

We have also observed 173 distribution of valence with respect to “gamma” band and “right” EEG regions using KNN algorithm.

N=173	Positive	Negative
True	32	67
False	37	37

Table 5.12: Distribution of valence with respect to “gamma” band and “right” EEG regions using KNN

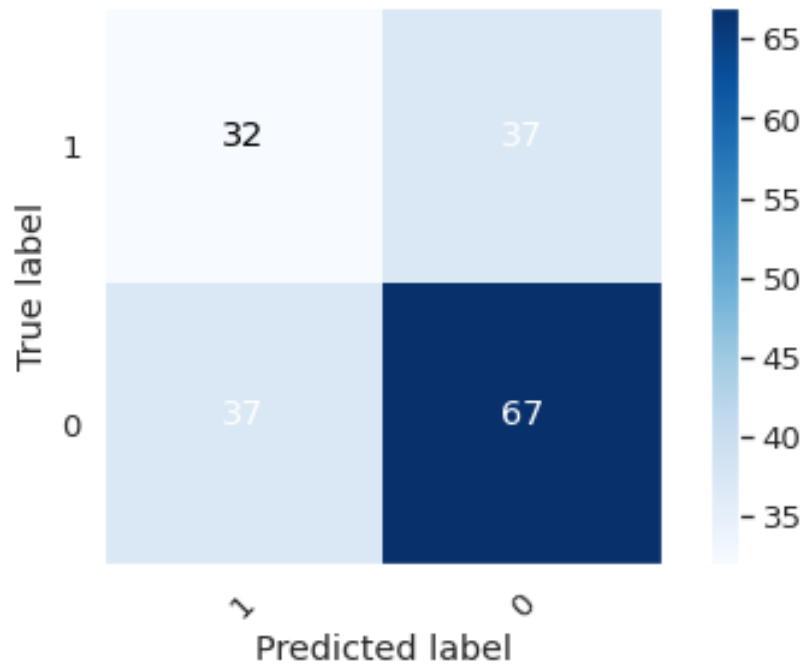


Figure 5.17: Confusion matrix of valence with respect to “gamma” band and “right” EEG regions using KNN algorithm

	Precision	Recall	F1-Score	Support
0	0.46	0.46	0.46	69
1	0.64	0.64	0.64	104
Accuracy			0.57	173
Macro Avg	0.55	0.55	0.55	173
Weighted Average	0.57	0.57	0.57	173

Table 5.13: Distribution of different metrics on valence with respect to “gamma” band and “right” EEG regions using KNN

5.4 Implementation with RNN and FFT

For this research, during the FFT processing, we employed meta data for the purpose of doing a meta vector analysis. Raw data was split over a time span of 2 seconds, with each slice having a 0.125-second interval between it. A two-second FFT of channel j was carried out in different frequencies in a sequence. Emotiv Epoch+ was fitted with a total of 14 channels, which were carefully selected. The number of channels is [1,2,3,4,6,11,13,17,19,20,21,25,29,31]. The number of bands is 5. band = [4,8,12,16,25,45]. A band power of 2 seconds on average is used. The window size was 256 with a step size of 16, with each update occurring once every 0.125 seconds. The sampling rate was set to 128 hertz. The FFT was then performed on all of the subjects using these settings in order to obtain the required output. Neural net-

Model: "sequential"		
Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, None, 512)	1193984
batch_normalization (Batch Normalization)	(None, None, 512)	2048
dropout (Dropout)	(None, None, 512)	0
lstm_1 (LSTM)	(None, None, 256)	787456
batch_normalization_1 (Batch Normalization)	(None, None, 256)	1024
dropout_1 (Dropout)	(None, None, 256)	0
lstm_2 (LSTM)	(None, None, 128)	197120
batch_normalization_2 (Batch Normalization)	(None, None, 128)	512
dropout_2 (Dropout)	(None, None, 128)	0
lstm_3 (LSTM)	(None, None, 64)	49408
batch_normalization_3 (Batch Normalization)	(None, None, 64)	256
dropout_3 (Dropout)	(None, None, 64)	0
lstm_4 (LSTM)	(None, 32)	12416
batch_normalization_4 (Batch Normalization)	(None, 32)	128

Figure 5.18: Information of LSTM

works and other forms of artificial intelligence require a starting collection of data, referred to as a training dataset, that serves as a foundation for subsequent application and use. This dataset serves as the foundation for the program's developing information library. Before the model can interpret and learn from the training data, it must be appropriately labeled. The lowest value of the data is 200 and the greatest value is above 2000, which means that trying to plot it will result in a lot of irrelevant plots, which will make conducting the analysis tough. The objective of machine learning is to create a plot and then optimize it further in order to obtain

lstm_4 (LSTM)	(None, 32)	12416
batch_normalization_4 (Batch Normalization)	(None, 32)	128
dropout_4 (Dropout)	(None, 32)	0
dense (Dense)	(None, 10)	330
activation (Activation)	(None, 10)	0
=====		
Total params: 2,244,682		
Trainable params: 2,242,698		
Non-trainable params: 1,984		

Figure 5.19: Parameter Information of LSTM

a pattern. And if there are significant differences between the plotted points, it will be unable to optimize the data. As a result, in order to fix this issue, the values have been reduced to their bare minimum, commonly known as scaling. The values of the data will not be lost as a result of scaling; instead, the data will be optimized to the point where there is little difference between the plotted points. StandardScaler is the name given to this technique. In order to achieve this, StandardScaler must transform your data into a distribution with a mean of zero and a standard deviation of one. When dealing with multivariate data, this is done feature-by-feature to ensure that the data is accurate (in other words independently for each column of the data). Because of the way the data is distributed, each value in the dataset will be deducted from the mean and then divided by the standard deviation of the dataset. Categorical Data refers to information that has a finite number of possible values. All machine learning models are mathematical models that require numbers to operate on. This is one of the key reasons for pre-processing categorical data prior to feeding it to machine learning models. In our scenario, we were unable to apply regression because we are attempting to classify our data. We will transform our data to categorical in order to undertake classification. After that, we divided the data set into two parts: a training data set and a testing data set. Training will be carried out on 75% of the data, and testing will be carried out on 25% of the data. A total of 456768 data were used in the training process. A total of 152256 data were used in the testing. RNN has been kept sequential. The first layer LSTM of sequential model takes input of 512. The second layer takes input of 256. The third and fourth layer takes an input of 128 and 64. And, the final layer LSTM of sequential model takes input of 10. Since we are conducting classification where we will need 0 or 1 that is why sigmoid has been used. The activation functions used are relu and for the last part sigmoid. The rectified linear activation function, abbreviated ReLU, is a piecewise linear function that, if the input is positive, outputs the value directly; otherwise, it outputs zero. Batch normalization was used. Batch normalization is a method for training extremely deep neural networks in which the inputs to a layer are standardized for each mini-batch. This results in a stabilization of the learning process and a significant drop in the total of training epochs required for training deep networks. Through randomly dropping out nodes while training, a single model can be utilized to simulate having a huge variety of distinct network designs.[2] This is referred to as dropout, and it is an extremely computationally

efficient and amazingly successful regularization technique for reducing overfitting and improving generalization error in all types of deep neural networks. In our situation, dropout rates began at 30%, increased to 50%, then 30%, 30%, 30%, and eventually 20%. Previously, we worked with three-dimensional datasets; however, when we converted to a dense layer, we obtained a one-dimensional representation in order to make a prediction. RMSprop was used as the optimizer with a learning rate of 0.001, a rho value of 0.9, and an epsilon value of 1e-08. RMSprop calculates the gradient by dividing it by the root of the moving (discounted) average of the square of the gradients. This application of RMSprop makes use of conventional momentum rather than Nesterov momentum. Additionally, the centered version calculates the variance by calculating a moving average of the gradients. As we can see, accuracy increases very gradually in this case, and learning rate plays a major part. If we increased the learning rate, accuracy would also increase rapidly, and when optimization is reached, the process would reverse, with accuracy decreasing at a faster rate. That is why the rate of learning has been reduced. When one zero is removed, the accuracy decreases significantly. As our loss function, we utilized the Mean Squared Error. The Mean Squared Error (MSE) loss function is the most basic and extensively used loss function, and it is typically taught in introductory Machine Learning programs. To calculate the MSE, take the difference between your model's predictions and the ground truth, square it, and then average it across the whole dataset. The MSE can never be negative since we are constantly squaring the errors. To compute loss, we utilized mean squared error. Because of the squaring portion of the function, the MSE is excellent for guaranteeing that the trained model does not contain any outlier predictions with significant mistakes. Because of this, the MSE places greater emphasis on outlier predictions with large errors. We tried our best to reduce the percentage of value loss and increase the accuracy rate. We saved the model and kept track by every 50 epochs. In the first picture, we can see that for the first 50 epochs the training loss 0.1588 and validation loss reduced to 0.06851 and 0.06005. And the training accuracy rate increased from 9.61 percent to 45.784 percent and validation accuracy increased to 53.420 percent. For the second 50 epochs, the training loss reduced to 0.06283 and the validation loss reduced to .05223 where the training accuracy increased to 51.661 percent and validation accuracy increased to 60.339 percent. For the third 50 epochs, the training loss reduced to 0.05992 and the validation loss reduced to .04787 where the training accuracy increased to 54.492 percent and validation accuracy increased to 64.413 percent. After 200 epochs the ratio started to change at a very slow rate. We ran 1000 epochs and got the training accuracy rate of 69.21% and the validation accuracy rate was 78.28%.

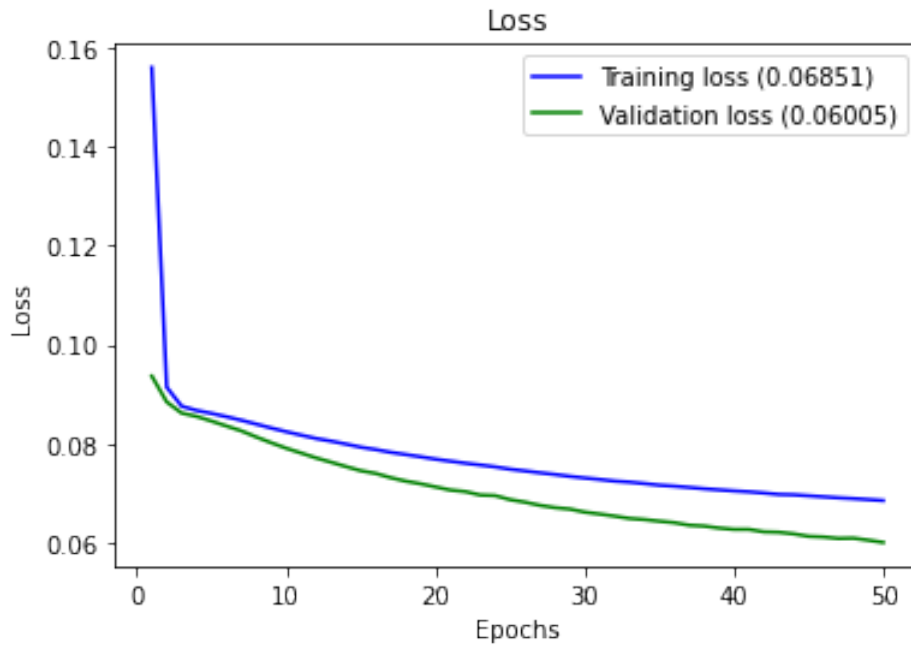


Figure 5.20: Epoch Vs Loss

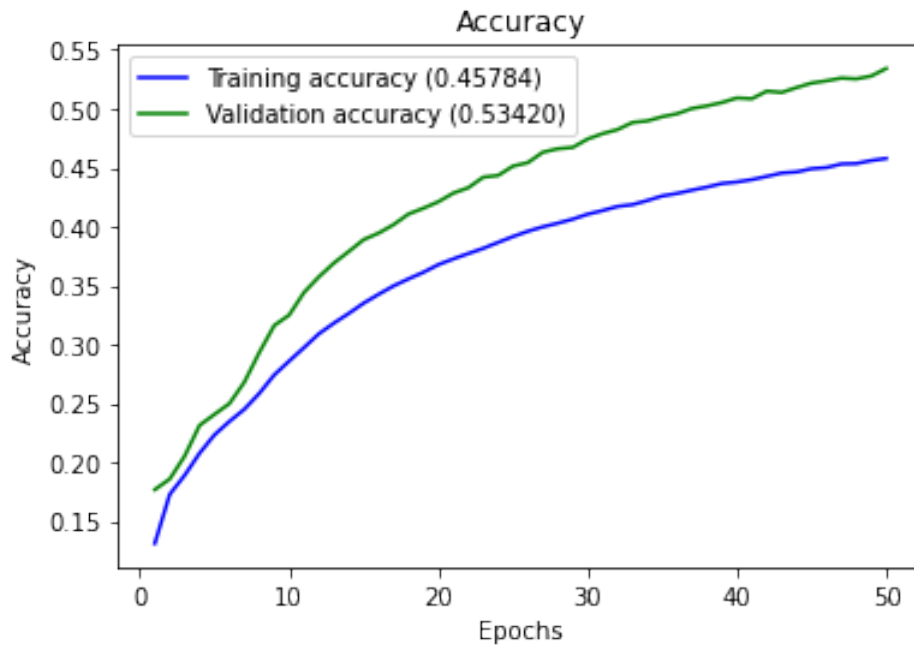


Figure 5.21: Epoch vs Accuracy

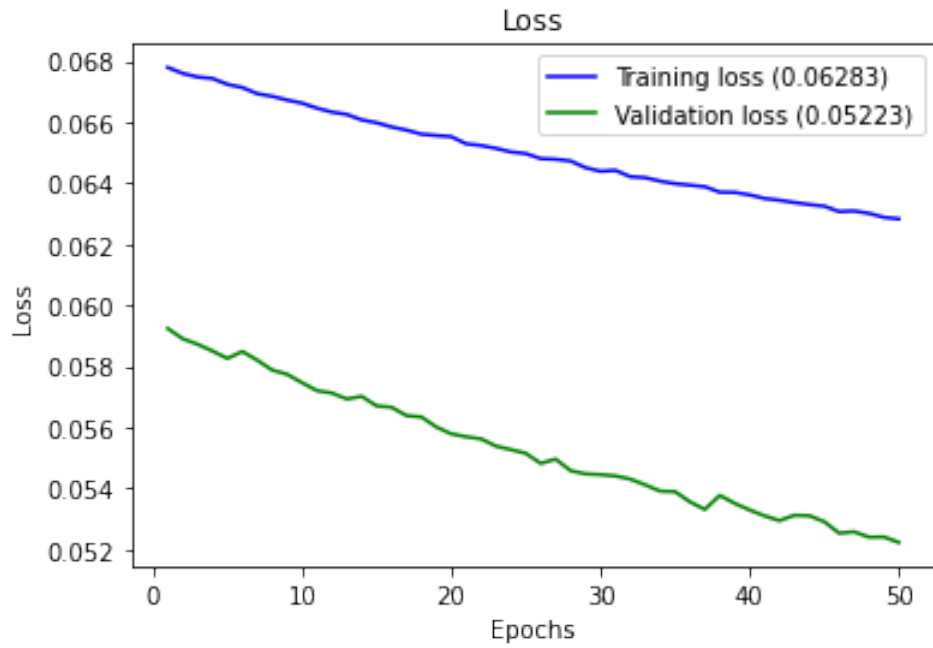


Figure 5.22: Epoch Vs Loss (From 51-100 Epochs)

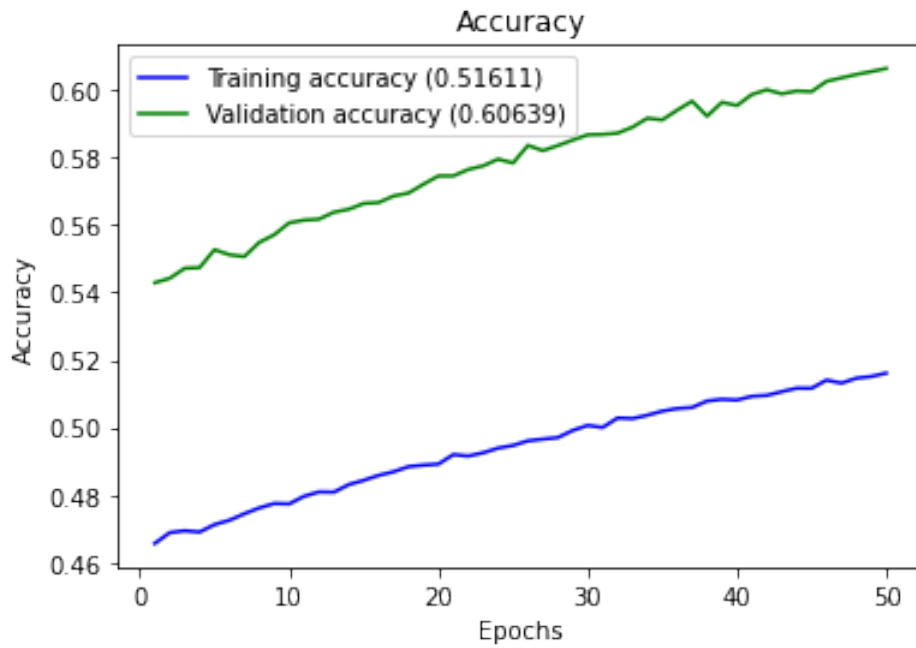


Figure 5.23: Epoch Vs Accuracy (From 51-100 Epochs)

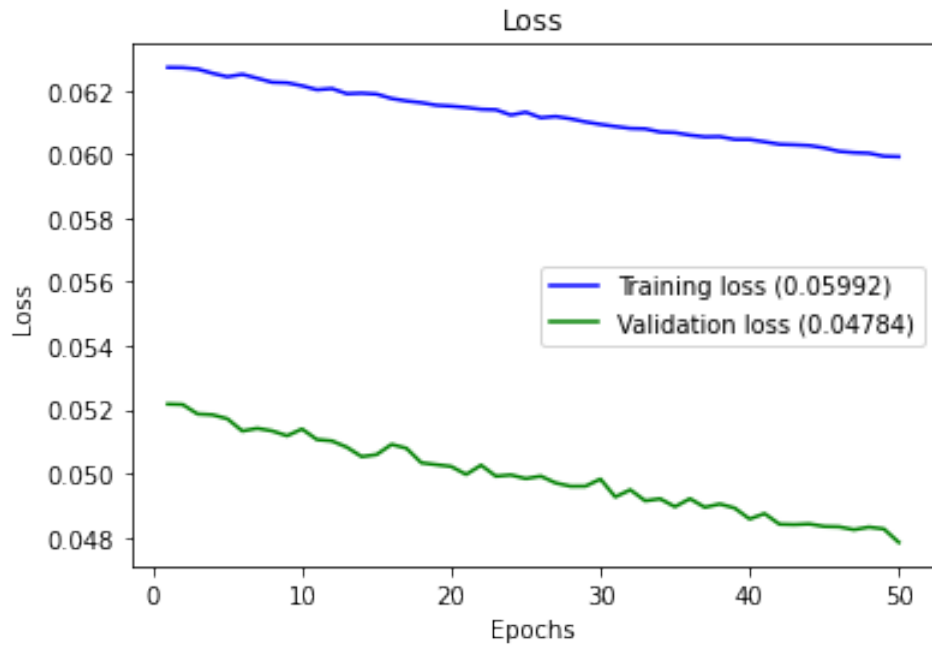


Figure 5.24: Epoch Vs Loss (From 101-150 Epochs)

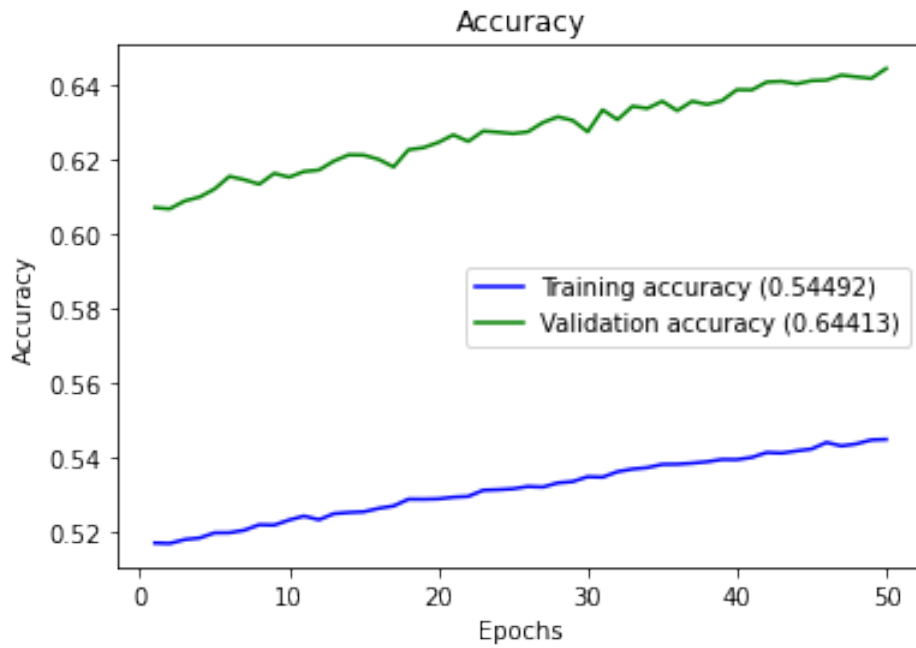


Figure 5.25: Epoch Vs Accuracy (From 101-150 Epochs)

5.5 Implementation with DWT

We tried to explore some result from DWT transformation with STD, min-max standard scaler.

Using SVM with sigmoid kernel and K-NN, we get the following results:

5.5.1 Arousal-Accuracy

	SVM	KNN
Mean of accuracy	52.21%	52.73%
STD	0.15954184532165527	0.003194570541381836

Table 5.14: Accuracy result of Arousal using SVM and K-NN

5.5.2 Arousal-F1 Score

	SVM	KNN
Mean of accuracy	52.21%	52.73%
STD	0.16068744659423828	0.0840451717376709

Table 5.15: F1-score result of Arousal using SVM and K-NN

5.5.3 Valence-Accuracy

	SVM	KNN
Mean of accuracy	51.17%	54.55%
STD	0.1488192081451416	0.0040628910064697266

Table 5.16: Accuracy result of Positive Valence using SVM and K-NN

5.5.4 Valence-F1 Score

	SVM	KNN
Mean of accuracy	51.17%	54.55%
STD	0.14807367324829102	0.002973318099975586

Table 5.17: F1 Score result of Positive Valence using SVM and K-NN

Using K-NN classifier, we compared the accuracy rate, precision, recall and F1-score of valence and arousal labels.

	Precision	Recall	F1-Score	Support
0	0.49	0.40	0.44	89
1	0.50	0.58	0.53	90
Accuracy			0.49	179
Macro Avg	0.49	0.49	0.49	179
Weighted Average	0.49	0.49	0.49	179

Table 5.18: Distribution of different metrics on valence using KNN with DWT

In the experiment, we showed 179 distribution of confusion matrix of valence for different metrics on KNN classifier.

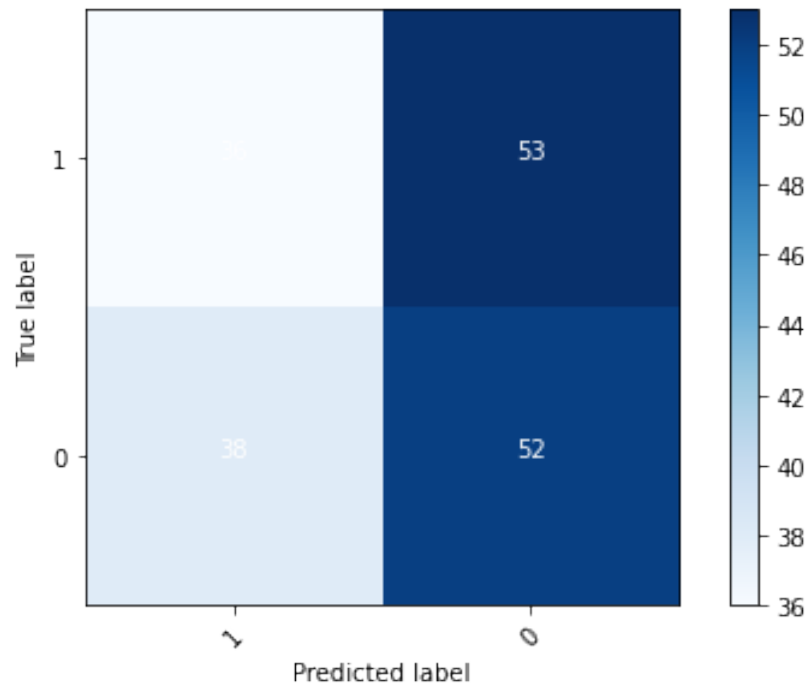


Figure 5.26: Confusion matrix of valence with K-NN Classifier

N=179	Positive	Negative
True	36	52
False	53	38

Table 5.19: TP, TN, FP, FN Distribution of valence using K-NN

	Precision	Recall	F1-Score	Support
0	0.57	0.61	0.59	89
1	0.58	0.54	0.56	90
Accuracy			0.58	179
Macro Avg	0.58	0.58	0.58	179
Weighted Average	0.58	0.58	0.58	179

Table 5.20: Table: Distribution of different metrics on arousal using KNN with DWT

In the experiment, we showed the confusion matrix.

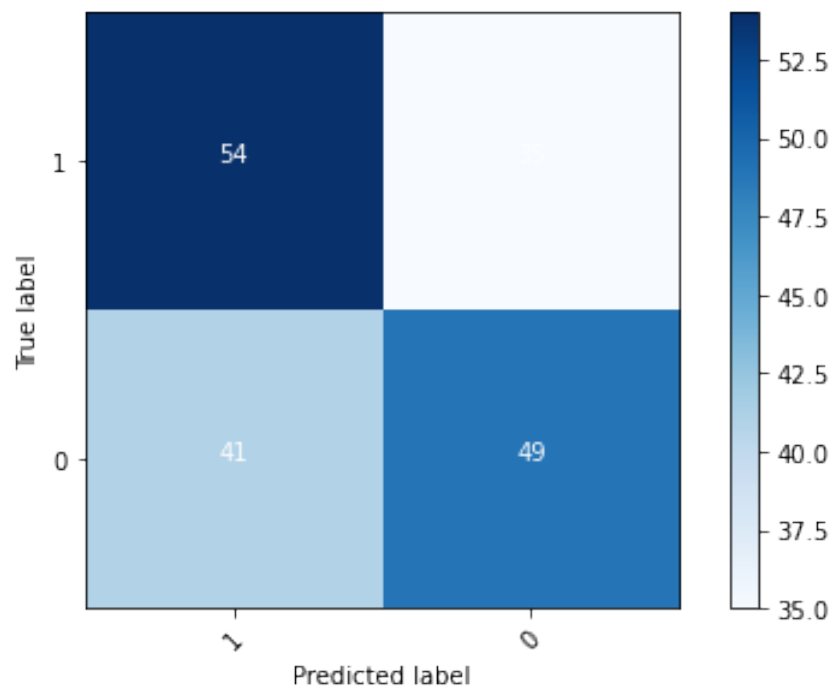


Figure 5.27: Confusion matrix of arousal with K-NN Classifier

N=179	Positive	Negative
True	54	49
False	35	42

Table 5.21: TP, TN, FP, FN Distribution of valence using K-NN

Chapter 6

6.1 Conclusion

To summarize, in this research, we describe the EEG-based emotion recognition challenge, as well as existing and proposed solutions to this problem. Emotion detection by the use of EEG waves is a relatively new and exciting area of study and analysis. With this study, we hope to acquire more meaningful information for emotion recognition from a variety of features and combine it in a useful way for future research. To identify and evaluate on numerous emotional states using EEG signals acquired from the DEAP Dataset, SVM (Support Vector Machine), KNN (K-Nearest Neighbor), and RNN (Recurrent Neural Network) trained with LSTM (Long Short-Term Memory) are used in conjunction with LSTM (Long Short-Term Memory). According to the findings, the suggested method is a very promising option for emotion recognition, owing to its remarkable ability to learn features from raw data in a short period of time. When compared to typical feature extraction approaches, it produces higher average accuracy over a larger number of people.

6.2 Future Work Plan

Our long-term goal is to enhance the system's functionality by adding new techniques and algorithms. To boost performance, we'll use a variety of deep learning models for things like feature extraction and dynamic modeling. When conducting future research, we plan to apply more complex fusion approaches that include EOG, PPG, GSR and EMG signals. EEG and eye movement feature similarity will be used in future study, as will the method of multi-modal deep learning for emotion recognition from EEG and other physiological inputs. Also, EEG can be linked to other physiological parameters to conduct further research. If we can improve EEG signal outcomes and make them more consistent in the future, we'll focus on other noise-removal pre-processing procedures. Finally, we'll look at how each channel's contribution might be improved even more.

Bibliography

- [1] J. D. Morris, “Observations: Sam: The self-assessment manikin an efficient cross-cultural measurement of emotional response 1,” *Journal of Advertising Research*, 1995.
- [2] J. Jin, X. Wang, and B. Wang, “Classification of direction perception eeg based on pca-svm,” in *Third International Conference on Natural Computation (ICNC 2007)*, vol. 2, 2007, pp. 116–120. DOI: 10.1109/ICNC.2007.298.
- [3] X. Cheng, C. Pei Ying, and L. Zhao, “A study on emotional feature analysis and recognition in speech signal,” *Measuring Technology and Mechatronics Automation, International Conference on*, vol. 1, pp. 418–420, Apr. 2009. DOI: 10.1109/ICMTMA.2009.89.
- [4] M. Sifuzzaman, M. Islam, and M. Ali, “Application of wavelet transform and its advantages compared to fourier transform,” *J. Phys. Sci*, vol. 13, Jan. 2009.
- [5] C. Huang, Y. Jin, Q. Wang, L. Zhao, and C. Zou, “Multimodal emotion recognition based on speech and eeg signals,” vol. 40, pp. 895–900, Sep. 2010. DOI: 10.3969/j.issn.1001-0505.2010.05.003.
- [6] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, “Scikit-learn: Machine learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [7] M. Wali, M. M, R. Ahmad, and S. Z. Bong, “Development of discrete wavelet transform (dwt) toolbox for signal processing applications,” Feb. 2012. DOI: 10.1109/ICoBE.2012.6179007.
- [8] A. Sano and R. Picard, “Stress recognition using wearable sensors and mobile phones,” Sep. 2013, pp. 671–676. DOI: 10.1109/ACII.2013.117.
- [9] D. Wang and Y. Shang, “Modeling physiological data with deep belief networks,” *International journal of information and education technology (IJIET)*, vol. 3, pp. 505–511, Jan. 2013. DOI: 10.7763/IJIET.2013.V3.326.
- [10] Y. Wang, X. Yang, and J. Zou, “Research of emotion recognition based on speech and facial expression,” *TELKOMNIKA Indonesian Journal of Electrical Engineering*, vol. 11, Jan. 2013. DOI: 10.11591/telkomnika.v11i1.1873.
- [11] Y. Y. Lee and S. Hsieh, “Classifying different emotional states by means of eeg-based functional connectivity patterns,” *PloS one*, vol. 9, e95415, Apr. 2014. DOI: 10.1371/journal.pone.0095415.

- [12] M. Wyczesany and T. Ligeza, “Towards a constructionist approach to emotions: Verification of the three-dimensional model of affect with eeg-independent component analysis,” *Experimental brain research*, vol. 233, Nov. 2014. DOI: 10.1007/s00221-014-4149-9.
- [13] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, “Learning representations from eeg with deep recurrent-convolutional neural networks,” Nov. 2015.
- [14] M. N. Fakhruzzaman, E. Riksakomara, and H. Suryotrisongko, “Eeg wave identification in human brain with emotiv epoc for motor imagery,” *Procedia Computer Science*, vol. 72, pp. 269–276, 2015, The Third Information Systems International Conference 2015, ISSN: 1877-0509. DOI: <https://doi.org/10.1016/j.procs.2015.12.140>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877050915036017>.
- [15] X. Li, P. Zhang, D. Song, G. Yu, Y. Hou, and B. Hu, “Eeg based emotion identification using unsupervised deep feature learning,” 2015.
- [16] F. Schroff, D. Kalenichenko, and J. Philbin, “Facenet: A unified embedding for face recognition and clustering,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 815–823. DOI: 10.1109/CVPR.2015.7298682.
- [17] I. Belakhdar, W. Kaaniche, R. Djmel, and B. Ouni, “A comparison between ann and svm classifier for drowsiness detection based on single eeg channel,” *2016 2nd International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp. 443–446, 2016.
- [18] W. Liu, W.-L. Zheng, and B.-L. Lu, “Emotion recognition using multimodal deep learning,” vol. 9948, Oct. 2016, ISBN: 978-3-319-46671-2. DOI: 10.1007/978-3-319-46672-9_58.
- [19] V. Vanitha and P. Krishnan, “Real time stress detection system based on eeg signals,” vol. 2016, S271–S275, Jan. 2016.
- [20] S. Alhagry, A. Aly, and R. El-Khoribi, “Emotion recognition based on eeg using lstm recurrent neural network,” *International Journal of Advanced Computer Science and Applications*, vol. 8, Oct. 2017. DOI: 10.14569/IJACSA.2017.081046.
- [21] A. Ang and Y. Yeong, “Emotion classification from eeg signals using time-frequency-dwt features and ann,” *Journal of Computer and Communications*, vol. 05, pp. 75–79, Jan. 2017. DOI: 10.4236/jcc.2017.53009.
- [22] X. Li, J.-Z. Yan, and J.-H. Chen, “Channel division based multiple classifiers fusion for emotion recognition using eeg signals,” *ITM Web of Conferences*, vol. 11, p. 07 006, Jan. 2017. DOI: 10.1051/itmconf/20171107006.
- [23] M. Menezes, A. Samara, L. Galway, A. Sant’Anna, A. Verikas, F. Alonso-Fernandez, H. Wang, and R. Bond, “Towards emotion recognition for virtual environments: An evaluation of eeg features on benchmark dataset,” *Personal and Ubiquitous Computing*, vol. 21, Dec. 2017. DOI: 10.1007/s00779-017-1072-7.
- [24] Z. Mohammadi, J. Frounchi, and M. Amiri, “Wavelet-based emotion recognition system using eeg signal,” *Neural Computing and Applications*, vol. 28, Aug. 2017. DOI: 10.1007/s00521-015-2149-8.

- [25] K. Potdar, T. Pardawala, and C. Pai, “A comparative study of categorical variable encoding techniques for neural network classifiers,” *International Journal of Computer Applications*, vol. 175, pp. 7–9, Oct. 2017. DOI: 10.5120/ijca2017915495.
- [26] M. I. Singh and M. Singh, “Development of a real time emotion classifier based on evoked eeg,” *Biocybernetics and Biomedical Engineering*, vol. 37, no. 3, pp. 498–509, 2017, ISSN: 0208-5216. DOI: <https://doi.org/10.1016/j.bbe.2017.05.004>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0208521616303035>.
- [27] V. Bajaj, A. Krishna, a. sri aravapalli, K. Priyanka, and S. Taran, “Emotion classification using eeg signals based on tunable-q wavelet transform,” *IET Science, Measurement Technology*, vol. 13, Dec. 2018. DOI: 10.1049/iet-smt.2018.5237.
- [28] S. Chambon, V. Thorey, P. J. Arnal, E. Mignot, and A. Gramfort, “A deep learning architecture to detect events in EEG signals during sleep,” in *MLSP 2018 - IEEE International Workshop on Machine Learning for Signal Processing*, Aalborg, Denmark, Sep. 2018. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01917529>.
- [29] H. Chao, H. Zhi, D. Liang, and Y. Liu, “Recognition of emotions using multi-channel eeg data and dbn-gc-based ensemble deep learning framework,” *Computational Intelligence and Neuroscience*, vol. 2018, pp. 1–11, Dec. 2018. DOI: 10.1155/2018/9750904.
- [30] M. Ghofrani Jahromi, H. Parsaei, A. Zamani, and D. W. Stashuk, “Cross comparison of motor unit potential features used in emg signal decomposition,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 5, pp. 1017–1025, 2018. DOI: 10.1109/TNSRE.2018.2817498.
- [31] M. Li, H. Xu, X. Liu, and S. Lu, “Emotion recognition from multichannel eeg signals using k-nearest neighbor classification,” *Technology and Health Care*, vol. 26, pp. 1–11, Apr. 2018. DOI: 10.3233/THC-174836.
- [32] J. Liu, H. Meng, M. Li, F. Zhang, R. Qin, and A. Nandi, “Emotion detection from eeg recordings based on supervised and unsupervised dimension reduction,” *Concurrency and Computation: Practice and Experience*, vol. 30, e4446, Mar. 2018. DOI: 10.1002/cpe.4446.
- [33] J. Thomas, L. Comoretto, J. Jin, J. Dauwels, S. Cash, and M. Westover, “Eeg classification via convolutional neural network-based interictal epileptiform event detection,” *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3148–3151, 2018.
- [34] H. Zamanian and H. Farsi, “A new feature extraction method to improve emotion detection using eeg signals,” *ELCVIA Electronic Letters on Computer Vision and Image Analysis*, vol. 17, p. 29, Nov. 2018. DOI: 10.5565/rev/elcvia.1045.

- [35] M. A. Asghar, M. J. Khan, Fawad, Y. Amin, M. Rizwan, M. Rahman, S. Badnava, S. S. Mirjavadi, and S. S. Mirjavadi, "Eeg-based multi-modal emotion recognition using bag of deep features: An optimal feature selection approach," *Sensors (Basel, Switzerland)*, vol. 19, no. 23, Nov. 2019, ISSN: 1424-8220. DOI: 10.3390/s19235218. [Online]. Available: <https://europepmc.org/articles/PMC6928944>.
- [36] D. Barahona-Pereira, "Evaluation of feature extraction techniques for an internet of things electroencephalogram," 2019.
- [37] Y. Hou and S. Chen, "Distinguishing different emotions evoked by music via electroencephalographic signals," *Computational Intelligence and Neuroscience*, vol. 2019, pp. 1–18, Mar. 2019. DOI: 10.1155/2019/3191903.
- [38] W. Ng, A. Saidatul, Y. Chong, and Z. Ibrahim, "Psd-based features extraction for eeg signal during typing task," *IOP Conference Series: Materials Science and Engineering*, vol. 557, p. 012032, Jun. 2019. DOI: 10.1088/1757-899X/557/1/012032.
- [39] X. Xing, Z. Li, T. Xu, L. Shu, B. Hu, and X. Xu, "Sae+lstm: A new framework for emotion recognition from multi-channel eeg," *Frontiers in Neurobotics*, vol. 13, p. 37, 2019, ISSN: 1662-5218. DOI: 10.3389/fnbot.2019.00037. [Online]. Available: <https://www.frontiersin.org/article/10.3389/fnbot.2019.00037>.
- [40] Y. Çimtay and E. Ekmekcioglu, "Investigating the use of pretrained convolutional neural network on cross-subject and cross-dataset eeg emotion recognition," *Sensors*, vol. 20, Apr. 2020. DOI: 10.3390/s20072034.
- [41] S. A. Hussain and A. S. A. A. Balushi, "A real time face emotion classification and recognition using deep learning model," *Journal of Physics: Conference Series*, vol. 1432, p. 012087, Jan. 2020. DOI: 10.1088/1742-6596/1432/1/012087. [Online]. Available: <https://doi.org/10.1088/1742-6596/1432/1/012087>.
- [42] T. Kusumaningrum, A. Faqih, and B. Kusumoputro, "Emotion recognition based on deap database using eeg time-frequency features and machine learning methods," *Journal of Physics: Conference Series*, vol. 1501, p. 012020, Mar. 2020. DOI: 10.1088/1742-6596/1501/1/012020.
- [43] Q. Xiong, X. Zhang, W.-F. Wang, and Y. Gu, "A parallel algorithm framework for feature extraction of eeg signals on mpi," *Computational and Mathematical Methods in Medicine*, vol. 2020, pp. 1–10, May 2020. DOI: 10.1155/2020/9812019.
- [44] A. Aydin, H. Öğmen, and H. Kafaligonul, "Neural correlates of metacontrast masking across different contrast polarities," *Brain Structure and Function*, Mar. 2021. DOI: 10.1007/s00429-021-02260-5.
- [45] Y. Chen, R. Chang, and J. Guo, "Emotion recognition of eeg signals based on the ensemble learning method: Adaboost," *Mathematical Problems in Engineering*, vol. 2021, pp. 1–12, Jan. 2021. DOI: 10.1155/2021/8896062.
- [46] Y. Liu and G. Fu, "Emotion recognition by deeply learned multi-channel textual and EEG features," *Future Gener. Comput. Syst.*, vol. 119, pp. 1–6, 2021. DOI: 10.1016/j.future.2021.01.010. [Online]. Available: <https://doi.org/10.1016/j.future.2021.01.010>.

- [47] S. D. Rama Chaudhary Ram Avtar Jaswal, “Emotion recognition based on eeg using deap dataset,” *European Journal of Molecular amp; Clinical Medicine*, vol. 8, no. 3, pp. 3509–3517, 2021, ISSN: 2515-8260.
- [48] N. Donges. (). “A guide to rnn: Understanding recurrent neural networks and lstm networks,” [Online]. Available: <https://builtin.com/data-science/recurrent-neural-networks-and-lstm>. (accessed: 24.09.2021).
- [49] S. Koelstra. (). “Deapdataset a dataset for emotion analysis using eeg, physiological and video signals,” [Online]. Available: <https://www.eecs.qmul.ac.uk/mmv/datasets/deap/>. (accessed: 12.07.2021).
- [50] D. S-l. (). “Sklearn.preprocessing.onehotencoder –scikit-learn 0.21.3 documentation,” [Online]. Available: <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OneHotEncoder.html>.. (accessed: 29.07.2019).
- [51] (). “What is the architecture behind the keras lstm cell?” [Online]. Available: <https://stackoverflow.com/questions/50488427/what-is-the-architecture-behind-the-keras-lstm-cell>. (accessed: 01.05.2017).