EEG BASED EMOTION DETECTION USING DEEP NEURAL NETWORKS

A Project Report

Submitted by

Ms. GANGA V SAJI

REG NO : TKM20MEAI07

SEMESTER : IV

In partial fulfillment for the award of the degree of

MASTER OF TECHNOLOGY
IN
Mechanical Engineering (Artificial Intelligence)

Under the guidance of
Dr. SABEENA BEEVI K

Thangal Kunju Musaliar College of Engineering
Kerala

JULY 2022

DECLARATION

I undersigned hereby declare that the project report **"EEG Based Emotion Detection Using Deep Neural Networks"**, submitted for partial fulfillment of the requirements for the award of degree of Master of Technology of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by me under supervision of **Dr.Sabeena Beevi K**. This submission represents my ideas in my own words and where ideas or words of others have been included, I have adequately and accurately cited and referenced the original sources. I also declare that I have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Place: Kollam
Date: GANGA V SAJI

Thangal Kunju Musaliar College of Engineering
Centre for Artificial Intelligence

C E R T I F I C A T E

This is to certify that, this report titled ***EEG BASED EMOTION DETECTION USING DEEP NEURAL NETWORKS*** is a bonafide record of the **Project** presented by **GANGA V SAJI (TKM20MEAI07)**, under our guidance and supervision, in partial fulfillment of the requirements for the award of the degree, **M.Tech in Mechanical Engineering (Artificial Intelligence)** in **APJ Abdul Kalam Technological University** .

Internal Supervisor

Dr. Sabeena Beevi K
Assoc. Professor & Head
Dept. of EEE,
TKMCE

Project Coordinator

Prof. Sumod Sundar
Assistant Professor
Centre for Artificial Intelligence,
TKMCE

Head of the Department

Dr. Imthias Ahamed
Professor & Head
Centre for Artificial Intelligence,
TKMCE

Internal Examiner

External Examiner

ACKNOWLEDGEMENT

A successful project is a fruitful culmination of efforts by many people, some directly involved and some others indirectly, by providing support and encouragement. Firstly I would like to thank the almighty for giving me the wisdom and grace for making my project a memorable one. I thank him for steering me to the shore of fulfillment under his protective wings.

I express my sincere gratitude to **Dr. T A Shahul Hameed**, Principal of T.K.M College of Engineering for giving me an opportunity to present my project. I would like to thank **Dr. Imthias Ahamed**, Professor and Head of the Department, Centre for Artificial Intelligence, TKMCE and Project coordinator **Prof. Sumod Sundar**, Assistant Professor, Centre for Artificial Intelligence, TKMCE, for their constant support and encouragement throughout the project work.

With a profound sense of gratitude, I would like to express my heartfelt thanks to my guide **Dr. Sabeena Beevi K**, Associate Professor and HOD, Department of EEE, TKMCE, for her expert guidance, cooperation and immense encouragement. I also extend my thanks to the entire faculty and staff of the Centre for AI, TKMCE, who has encouraged me throughout this work.

I also express my thanks to my loving parents, brother and friends, for their support and encouragement in the successful completion of this project work.

Ganga V Saji

ABSTRACT

The EEG signals captured from central nervous system, have the property to directly reflect the brain activity and thus have an intrinsic relationship with the emotional states of humans. Researchers are paying more attention to emotion recognition from EEG signals with the development of Brain Computer Interface (BCI) technology. BCI is a direct communication or mapping between the electrical activity of the brain and external devices. Emotion recognition gained more attention in the past decade as it is directly connected to various fields like psychology, healing, physiology, marketing and studies. The gaming market has expanded to be one of the most significant entertainment markets because of recent advanced technologies. Understanding the automatically inferred player's emotions during game play can be used to enhance the quality of the game. This work aims to identify the emotions using a reliable LSTM network from the selected features of EEG with a better recognition rate. Interactive software like video games can elicit a range of feelings in their users. Here in this work, GAMEEMO dataset is used where the EEGs are in response to 4 different game plays. A video game can affect players' thoughts and feelings through game play, storytelling, and the gaming environment. At first the extraction of maximum features using a suitable feature extraction method from the dataset is done. Then emotion recognition using different machine learning models and recurrent neural networks are done followed by their performance comparison. The LSTM model emerged as the one with better performance . The model is then used for emotion detection from selected features of the dataset to study the effect of feature selection on the performance. The results shows performance improvement of the model with selected features. Finally, the model was used on various experimental datasets like EEG Brainwave Dataset and an experimental sample dataset from DEAP to study its applicability and reliability.

Contents

List of Figures

List of Tables

ABBREVIATIONS

BCI	Brain Computer Interface
BiLSTM	Bidirectional Long Short Term Memory
CNN	Convolutional Neural Network
CNS	Central Nervous System
DEAP	Dataset for Emotion Analysis using Physiological & Audiovisual Signals
EEG	Electroencephalogram
GCNN	Graph Convolutional Neural Networks
GRU	Gated Recurrent Units
HALA	High Arousal-Low Arousal
HANV	High Arousal Negative Valence
HAPV	High Arousal Positive Valence
HVLV	High Valence-Low Valence
LANV	Low Arousal Negative Valence
LAPV	Low Arousal Positive Valence
LF-DfE	Linear Formulation of Differential Entropy
LSTM	Long Short Term Memory
RNN	Recurrent Neural Network
SEED	SJTU Emotion EEG Dataset
SRU	Simple Recurrent Unit
SVM	Support Vector Machine
XG Boost	eXtreme Gradient Boosting

Chapter 1

INTRODUCTION

The Electroencephalography (EEG) signals captured from central nervous system (CNS), have the property to directly reflect the brain activity and thus have an intrinsic relationship with the emotional states of humans. A brain-computer interface, also called a brain-machine interface, is a direct mapping or communication between the electrical activity of the brain and external devices. Researchers are giving more importance to emotion recognition from EEG signals with the development of this BCI technology. Emotions are not just the physiological states of the different feelings, behaviors, and thoughts of human beings but also psychological along with the physiological reactions brought about by different external stimuli. In a broader sense, emotion recognition can be achieved from speech, limb movements and facial expressions. Emotion recognition gained more attention in the past decade as it is directly connected to various fields like psychology, healing, physiology, marketing and studies. The gaming market has expanded to be one of the most significant entertainment markets because to recent advanced technologies Both the scientific community and other for businesses have worked very hard on both the technological and player-player interaction components of video games. It is important to understand how the automatically inferred players' emotions during game play can be used to enhance the quality of the game. The game designer typically strives to develop a set of game elements that can entertain, attract, engage, and/or educate the users. A video game can affect players' thoughts and feelings through gameplay, storytelling, and the gaming environment. Automatic emotion recognition, it is an integral and challenging task. Traditional machine learning and deep learning are the two aspects of EEG-based emotion recognition methods. In traditional machine learning methods, including Naive Bayes, SVM and other classifiers are used to classify and recognize. The deep learning methods automatically learn deep features and recognize emotions through models such as RNN, LSTM and BiLSTM. Several experiments are done on various public EEG datasets such as DEAP, SEED, DREAMER and GAMEEMO.

1.1 Brain Human Interface (BCI)

Artificial intelligence has a sub-field called "Affective computing" that can recognise, analyse, process and simulate human emotions. It is a multidisciplinary field that is constantly expanding and looks at how affect can influence human-technology interactions, how systems can be created to use affect to improve capabilities, how technology can help to understand human affect, and also how sensing and affective techniques can change how people interact with computers. In BCI the brain signals are gathered, analyzed, and interpreted into commands that are transmitted to output devices that carry out the intended tasks. BCIs would not employ standard neuromuscular output channels. It is based on EEG, it is able to identify the emotional states of people.

1.1.1 EEG-Based BCI in Emotion Recognition

Many investigations have discovered a connection between emotions and the electrical activity the central nervous system produces. By identifying electrical impulses in the brain and tracking changes, locations, and functional links among them, EEG equipment can detect brain activity. High temporal resolution EEG recordings can be utilised to directly monitor brain activity. These signals are a dependable source of information because they can't be altered or duplicated to produce a fake emotional state. The objective is to translate this data and relate to other emotions. One simple and practical method to detect EEG signals is with EEG-based BCI devices that are affordable, non-invasive and even wearable like helmets and headbands. Hence this technology, will soon be more practical. They could therefore soon be used regularly for emotion detection in a wide range of contexts, such as optimising work performance, emotion monitoring in healthcare facilities, different teaching-learning scenarios, gaming and other forms of entertainment[9].

1.2 Electroencephalogram (EEG)

The human brain starts to have neuronal activity between the 17th and the 23rd week of prenatal development. The brain's electrical activity is assumed to reflect not just how well the brain is functioning but also how the body is doing overall, beginning right from infancy and continuing throughout life. This theory drives the employment of advanced digital signal processing methods on EEG signals captured from a human subject's brain. For those working with these signals for BCI applications as well as the detection and treatment of brain disorders and diseases, a thorough understanding of neuronal processes, neuro physiological characteristics of the brain, as well as the mechanism behind the signal generation and recording is important.

1.2.1 EEG Generation:

An EEG signal is used to measure the currents that flow when the pyramidal neurons' dendrites are synaptically excited in the cerebral cortex. Activated brain cells (neurons) generate synaptic currents within their dendrites. The magnetic and secondary electrical fields produced by this current over the scalp is detected by electromyogram (EMG) and electroencephalogram (EEG) instruments respectively. Electrical potential differences arise from electrical dipoles formed by combined postsynaptic graded potentials from pyramidal cells between the soma and apical dendrites, which branch from neurons (Figure 2.1). The negative ion of chlorine and the positive ions of sodium, magnesium, calcium, and potassium are pushed to produce the current into the membranes of the neurons in the direction defined by the membrane potential.

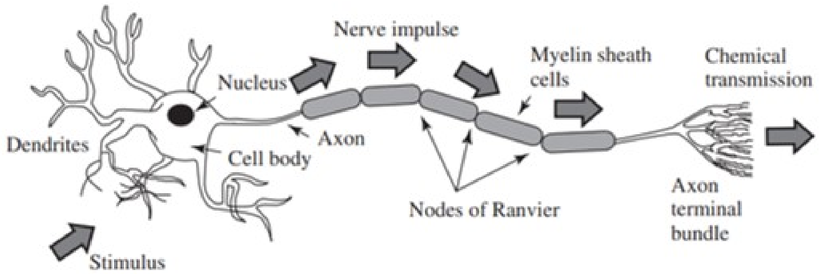

Figure 1.1: Neuron Structure.

The human head is comprised of the scalp, the skull, the brain, and several other thin layers. The structure and properties of the skull attenuates signals 100 times more than the other soft tissue does. Contrarily, most noise is produced either inside the brain called internal noise or over the scalp called external noise or system noise. Thus, only extremely large groups of activated neurons can create sufficient voltage to record using scalp electrodes[8].

1.2.2 Brain Rhythms

The five major brain waves are distinguished by their distinct frequency ranges. Alpha, theta, beta, delta, and gamma are the names of these low-to-high frequency ranges. The alpha and beta waves were initially proposed by Berger in 1929. Jasper and Andrews in 1938 first used the word "gamma" to refer to waves having a frequency higher than 30 Hz. Those frequencies below the alpha range were referred to as "delta rhythm" by Walter in 1936. He added that theta waves have frequencies in the range of 4 to 7.5 Hz. The idea of a theta wave was proposed in 1944 by Wolter and Dovey.

The frequency of delta waves ranges between 0.5–4 Hz. The real delta response may be confused with the artifact signal received from the large muscles in the mouth and neck. The reason for this is that, in contrast to the signal of interest, which arises deep inside the brain and is significantly attenuated as it goes through the skull, muscles produce large signals close to the skin's surface. However, it is relatively easy to identify whether a response is brought on by excessive movement by applying standard signal analysis techniques to the EEG.

The frequency range for theta waves is 4-7.5 Hz. The choice of theta might have been made to suggest that it originated in the thalamus. Theta waves have close relation with deep meditation, access to unconscious thoughts, and creative inspiration. Theta waves are commonly accompanied by several other frequencies and seem to be related to arousal levels. Alpha waves are said to eventually diminish in frequency in the brains of healers and skilled mediators.

The occipital area of the brain is where alpha waves occur frequently. They can be discovered across the posterior lobes of the brain. The Alpha waves often appear as a signal with a circular or sinusoidal shape and a frequency of 8 to 13 Hz. However, it might occasionally appear as sharp waves. In such cases, the positive component appears rounded, similar to the wave shape of the rolandic mu rhythm, while the negative component appears sharp. The alpha wave, which can occur across a wider range, is the most observable rhythm in every brain activity.

A beta wave is a term used to describe the brain's electrical activity, which ranges between 14 and 26 Hz. A beta wave is the brain's typical waking pattern in healthy individuals and is connected to active thinking, focusing, resolving concrete and/or specific problems or active attention on the outside world. There is a strong beta wave that can be seen when someone panics. The brain has significant rhythmical beta activity, particularly in the frontal and central areas. The rolandic mu rhythm is connected to a central beta rhythm that is susceptible to disruption from tactile or motor stimulation.

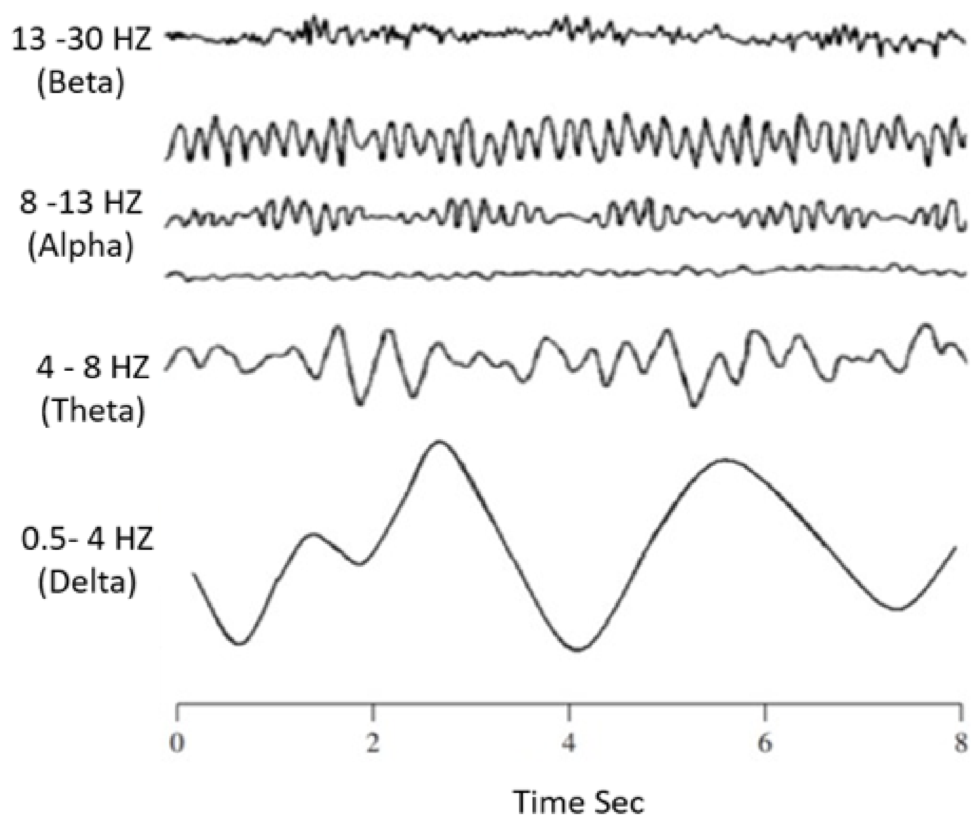

Figure 1.2: EEG Frequency Bands.

1.2.3 EEG Recording and Measurement

Several applications, including BCI and studies of mental activity, frequently include the stimulation of some electrodes near movement-related regions. The amplitudes of the unprocessed EEG signals are in the microvolt range, and their highest frequency components reach 300 Hz. Before the ADC, the signals must be amplified, and then either before or after the ADC, filters must be applied in order to clean up the signals and get them ready for processing and presentation. The signals cannot be altered or distorted because of the way the filters are made.

EEG signals are traces of neural activity. They can be captured in a variety of formats by multiple-electrode EEG equipment from the brain, the cortex under the skull, or certain areas on the scalp. The data are usually in the time domain, but many modern EEG equipment have imaging capabilities to display EEG topographies and can do frequency analysis using basic signal processing methods like the Fourier transform. There are numerous methods currently in development for analyzing EEG signals. The processes include multiway processing, time-domain processing, frequency-domain processing, and spatial-domain processing. Numerous methods have also been developed for seeing brain activity purely from reconstructed images[8].

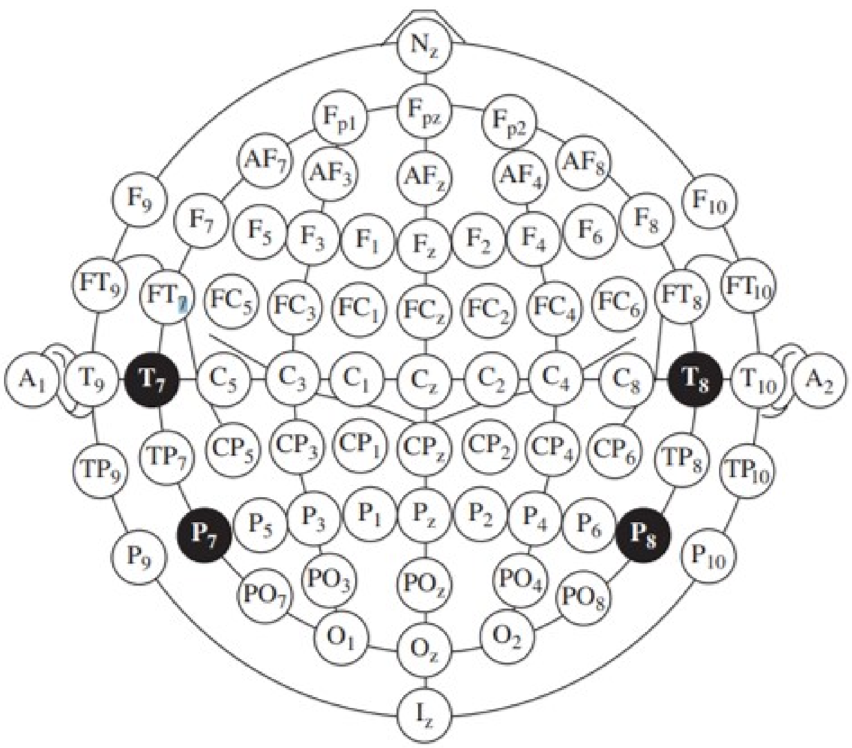

Figure 1.3: electrode channel placement on scalp

1.3 Objectives

- To propose and implement suitable feature extraction of GAMEEMO EEG dataset for emotion recognition.
- Identify the emotions from the selected features using the proposed LSTM model with better recognition rate.
- Compare and study the performance of the proposed model with machine learning models.

- Understand the effects of feature selection on the performance
- Compare and study the proposed model for different data sets to study the reliability of the model.

Chapter 2

LITERATURE SURVEY

2.1 Existing Works

Xing X et al.[1] worked on emotion detection using recurrent neural networks focused on the development of brain-inspired robots, specifically in the branch of human-robot interaction. It presented a framework for multi-channel EEG emotion recognition. A linear EEG mixing model and an emotion timing model make up the framework. While the emotion timing model is based on the LSTM Recurrent Neural Network, the linear EEG mixing model was created and solved using Stack AutoEncoder (SAE) (LSTM-RNN). The framework's efficacy was demonstrated by an emotion recognition experiment using the model on the DEAP dataset, which produced mean emotion recognition accuracy values of 81.10 percent for valence and 74.38 percent for arousal.

Wei C et al.[2] aimed to create an EEG-based emotion detection system that can distinguish between three different emotions: positive, neutral, and negative. Using a deep Simple Recurrent Units (SRU) network, they were able to address the issue of long-term dependencies that might appear in a normal Recurrent Neural Network in addition to processing sequence data (RNN). To improve classification outcomes, three ensemble processes were used with base SRU models. The performances of these shallow models, deep models, and ensemble models were then evaluated and compared. The model may distinguish physiological elements of EEG from many domains and multiple frequency bands while analyzing the stability of the emotion detection technique across time.

Yin Y et al.[3] developed an EEG emotion recognition using a fusion model of graph convolutional neural networks and LSTM. They proposed a novel method for emotion detection that is based on a special deep learning model. EEG data is calibrated with 3s reference standards before being divided into sections with a 6s time window to create a feature cube. Then each segment is extracted for its differential entropy. Second, a new deep learning model that mixes Graph Convolutional Neural Networks (GCNN) and LSTM neural networks is given each segment's feature cube. The fusion model employs several GCNNs to capture graph domain features, LSTM cells to extract temporal features and memorize the change in the relationship between two channels over time, and a Dense layer to produce emotion classification results. They achieved better recognition accuracy for both subject-dependent and independent cases.They achieved better recognition accuracy for both subject dependent and independent cases.

Ramzan M et al.[4]in their work, High-Arousal-Low-Arousal (HALA), High-Valence-Low-

Valence (HVLV), familiarity, dominance, and liking emotions have all been classified using the DEAP dataset. For the study of emotions using EEG signals, the merging of deep learning models, specifically CNN and LSTMRNN, tends to perform better. For HALA, HVLV, familiarity, dominance, and liking, the average accuracies examined by the fused deep learning classification model for DEAP are 97.39 percent, 97.41 percent, 98.21 percent, 97.68 percent, and 97.89 percent, respectively. The model was further tested on the SJTU Emotion EEG Dataset (SEED) for the detection of positive and negative emotions, yielding a 93.74 percent average accuracy. The findings reveal that the created model is capable of classifying the inner feelings of various EEG-based emotion databases.

Joshi VM et al.[5] proposed an emotion detection utilising Electroencephalography (EEG) signals based on the Linear Formulation of Differential Entropy feature extractor and the BiLSTM network classifier. The nonlinearity and non-Gaussianity of the EEG signal are successfully detected by LF-DfE. The BiLSTM network learns spatial information from several brain regions and captures long-term reliance of the EEG signal. On the SEED database, the proposed model is employed to detect positive, negative, and neutral emotions, along with valence and arousal on the DEAP database. On the SEED database, contingent, non-contingent, and inter-dependent (cross-session) experiments are done to evaluate the suggested model subject. On the SEED database, the resultant average accuracy of emotion identification for the subject contingent approach improved by about 4.12 percentage, for non-contingent approach by 4.5 percentage and for inter-dependent approach it is improved by 1.3 percentage.On the DEAP database, the average accuracy of subject non-contingent experiments has improved by 7.04 percent.

2.2 Inferences

Recurrent neural networks are dependable to mine the features of EEG to detect emotions. LSTM overcomes the problem of long-term dependencies and vanishing gradient in RNNs. It is reliable for emotion detection from EEG dataset as it gives more features and include low level to high level feature extraction from large data. The temporal characteristics of EEG waves are mined and studied for emotion recognition.

Title	Technique Used	Advantages	Disadvantages
SAE+ LSTM: A New framework for emotion recognition from multi-channel EEG [1] (2019)	Stack Auto Encoder and LSTM	81.10 percentage accuracy for valence and 74.38 percentage for arousal from DEAP dataset.	Framework exhibited low recognition accuracy in arousal
EEG-based emotion recognition using simple recurrent units network and ensemble learning [2] (2020)	Simple Recurrent Units (SRU) network and ensemble learning	Exploring the stability of the emotion recognition system over time, by extracting physiological EEG data from multi domains and frequency bands.Three different approaches were employed by the ensemble learning methods to combine numerous SRU models for better outcomes. .	The parameter selection had no systematic approach. The model is subject specific which limits its applicability. In addition it have high computational cost and training time.
EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM. [3] (2020)	C-Fusion model of Graph Convolutional Neural Networks and LSTM	Both subject independent and subject dependent experiments on DEAP achieved better recognition accuracy.	Only explored the effectiveness of GCNN on the binary classification of emotions (low-high arousal or negative-positive valence). High computational complexity
Fused CNN-LSTM Deep learning emotion recognition model using Electroencephalography signals [4] (2021)	Fused CNN-LSTM	Could classify without much hand-engineered feature extraction and/or selection method(s).	Achieved less accuracy using SEED dataset.
EEG based emotion detection using fourth order spectral moment and deep learning [5] (2021)	BiLSTM .	In comparison to other established approaches, the network had a higher recognition rate.	Better discrimination was for positive emotions than negative emotions.

Table 2.1: Existing Works

Chapter 3

DATASETS USED

3.1 GAMEEMO

This is a game based EEG dataset and was downloaded from Mendeley Data. The EEG readings from 28 separate people were collected using the 14-channel EMOTIV EPOC+, a wearable and portable EEG equipment. The 20 minutes of available EEG data for each subject were collected while the subjects played four distinct video games that recorded emotions for five minutes each. By using the self assessment form, the individuals evaluated each video game using the arousal and valence scales. The emotions were boring, calm, horror and funny.EEG electrodes are located on 16 different scalp zones (AF3, F7, F3,FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, P3 and P4) as shown in Figure 3.1. EEG device

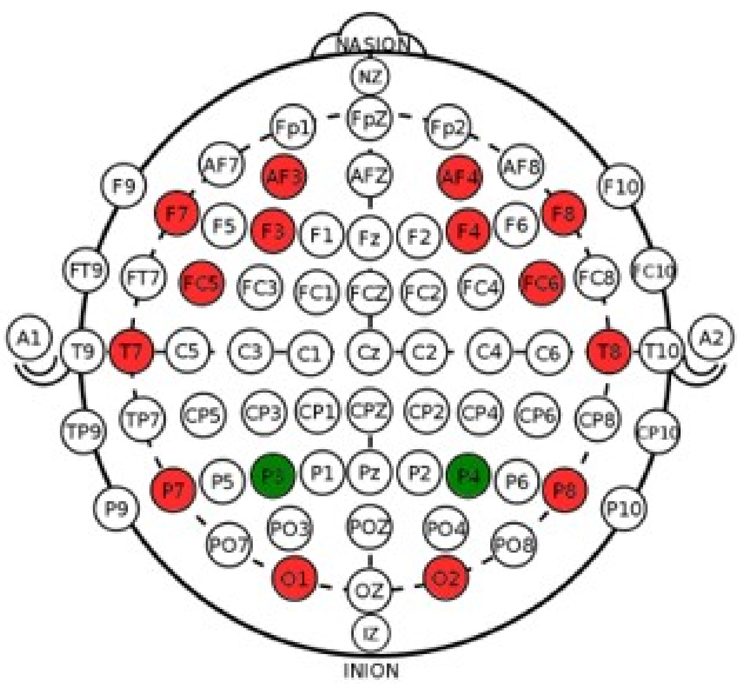

Figure 3.1: Electrodes on 16 Different Scalp Zones

has predefined 16 electrode locations.Two of them, P3 and P4 are reference electrodes.The device's sampling rate is 2048 Hz, which is down-sampled to 128 Hz. The 5th order sinc filter, a built-in filter of the device, is used to pre-process the signals. The fundamental goal of employing a sinc filter is to eliminate all frequency components that are higher than a specific cutoff frequency while leaving the lower frequencies unaffected. This filter eliminates motion artifacts brought about by arm, head, and hand movements. There are 1568 ie. 4x14x28 EEG data with a 20-minute duration and thus 38,252 total samples are available. 4 denotes the total number of games used, 14 represents the number of EEG channels, and 28 is the total number of subjects[10].

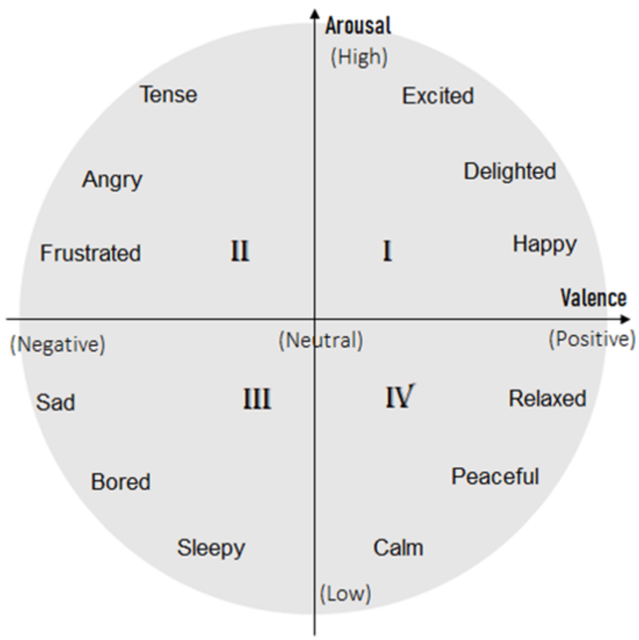

Figure 3.2: Emotion Model

- Game 1: **Train Sim World**: Boring : III- Low Arousal Negative Valence (LANV).
- Game 2: **Unravel**: Calm: IV - Low Arousal Positive Valence (LAPV).
- Game 3: **Slender-The Arrival**: Horror: II - High Arousal Negative Valence (HANV).
- Game 4: **Goat Simulator**: Funny: I - High Arousal Positive Valence (HAPV).

3.2 EEG Brainwave Dataset: Feeling Emotions

This is a publicly available Kaggle dataset which contain EEG brainwave data that has been processed using a unique extraction method based. The data was gathered from two individuals—a male and a female—for three minutes for each of the three states— neutral,positive and negative.A Muse EEG headgear to capture the EEG locations at TP9, AF7, AF8, and

TP10 using dry electrodes. In addition to that a six minutes of resting neutral data are also recorded[11].The stimuli used to induce the emotions are:

- The Death Scene from movie 'Marley and Me' by Twentieth Century Fox - Negative
- Opening Death Scene from movie 'Up' by Walt Disney Pictures - Negative
- Musical number in opening from 'La La Land' by Summit Entertainment - Positive
- Funeral Scene from 'My Girl' by Imagine Entertainment - Negative
- A Nature timelapse from 'Slow Life' by BioQuest Studios- Positive
- Funny clips from 'Funny Dogs' by MashupZone - Positive.

3.3 A Sample Dataset

A sample experimental dataset is created from the DEAP dataset, observations of 20 individuals are considered. The DEAP dataset is divided into two sections, they are the results of an online self-evaluation in which 14–16 volunteers evaluated each of 120 one-minute music video samples on the basis of arousal, valence, and dominance.32 volunteers watched a selection of 40 of the mentioned music videos for an experiment, and ratings, physiological readings, and face videos were taken. Each participant also rated the videos as above, and EEG and physiological signs were also captured. Each channel's data is extracted into a separate file in this assignment, along with the labels. For each trail, each person's data is saved in a column by time instead of row wise for each channel.

Chapter 4

FEATURE EXTRACTION AND SELECTION

There are four .csv files each of 28 subjects with signal values of 14 channels each in GAMEEMO dataset.Time based and frequency based features are extracted using nitime, spectrum etc. by unifying these files as shown in Figure 5.1. A total of 66 features are extracted from the EEG datafiles.

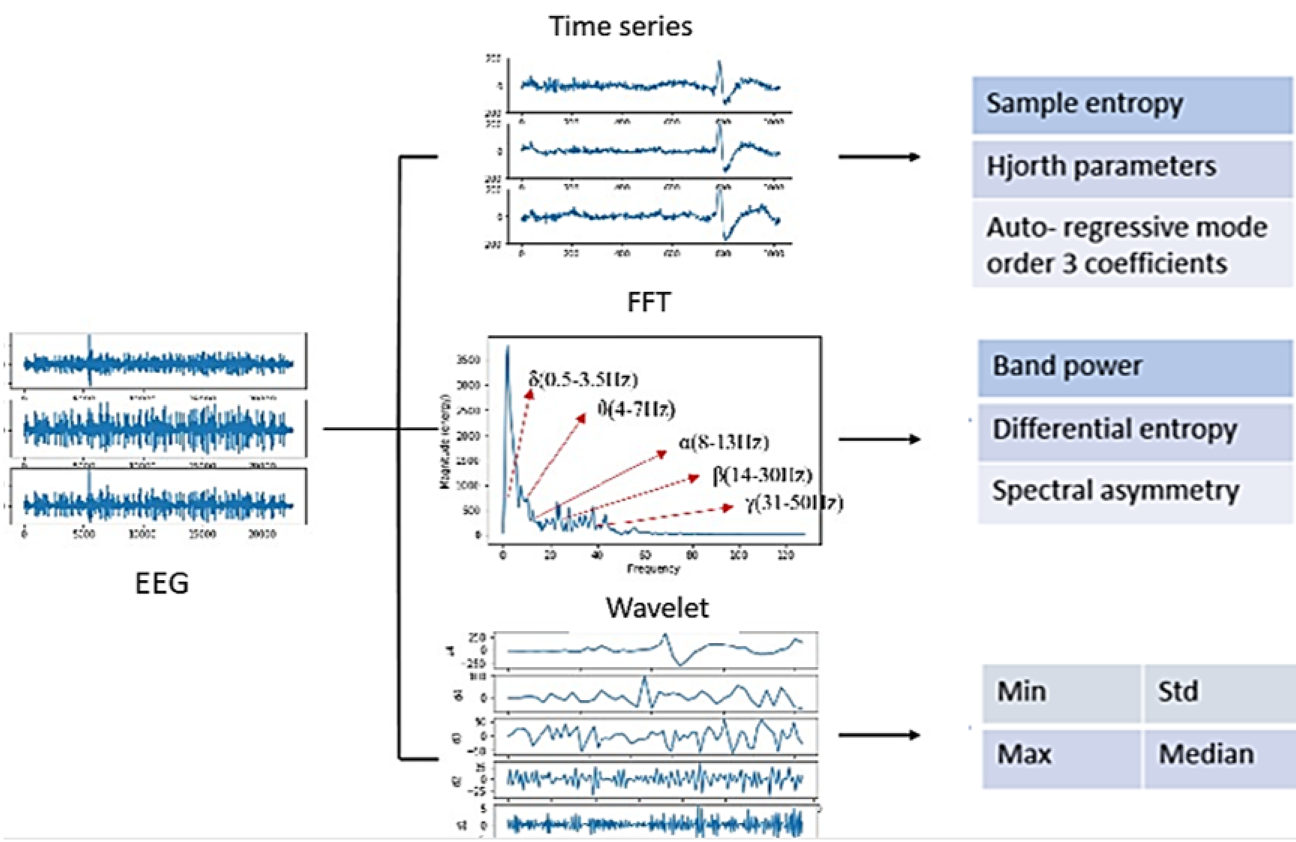

Figure 4.1: Feature Extraction from EEG

4.1 Extracted Features

4.1.1 Hjorth Parameters

They are characteristics of the time domain called Activity, Mobility, and Complexity.

- Activity: Represents a time function's variance or signal power.
- Mobility: Represents the power spectrum's mean frequency or standard deviation percentage.
- Complexity: a metric that depicts the frequency change.

Hjorth parameters are typically used as markers of statistical features in time-domain signal processing.

4.1.2 Higher Order Moments

Time domain characteristics such as skewness and kurtosis are higher order moments. They are statistical measures of the distribution of the signal elements and the complexity of the EEG signals.

- Skewness: A measure of an EEG signal data set's lack of symmetry or asymmetries is called skewness.
- Kurtosis: This characteristic defines whether the EEG signal is flat at the mid point of the waveform or has a peak.

4.1.3 Autoregressive Mode Order 3 Coefficients

When a value from a time series is regressed on earlier values from the same time series, an autoregressive model is constructed. In a regression model, the response variable out from previous time period has been converted into a predictor. The order of an autoregressive model is determined by how many values in the series immediately before the current value were used to predict it.

4.1.4 Wavelet Coefficients

A wavelet transform breaks down a signal into a weighted linear mixture of scaling and wavelet functions. Low frequency data are captured by approximation coefficients, which are scaling function-related coefficients (weights). To store high-frequency data, detail coefficients connected to the wavelets are used. In contrast to the Fourier transform, the wavelet transform provides local information on a particular signal. Wavelet entropy is a technique that can be used to examine transient elements of non-stationary data (WE). This metric combines wavelet decomposition and entropy to analyze the extent of order or disorder in a signal with a high time-frequency resolution. It evaluates traits such as mean, standard deviation, and wavelet energy.

4.1.5 FFT Band Powers

Decomposing the signal into five functionally distinct frequency bands—Delta (0.5–4 Hz), Theta (0.8–12 Hz), Alpha (8–12 Hz), Beta (12–30 Hz), and Gamma (30–100 Hz)—is one of the most used techniques for analysing EEG data. This suggests that the EEG signal has been broken down into its component frequencies, which is often done using Fourier transformations.

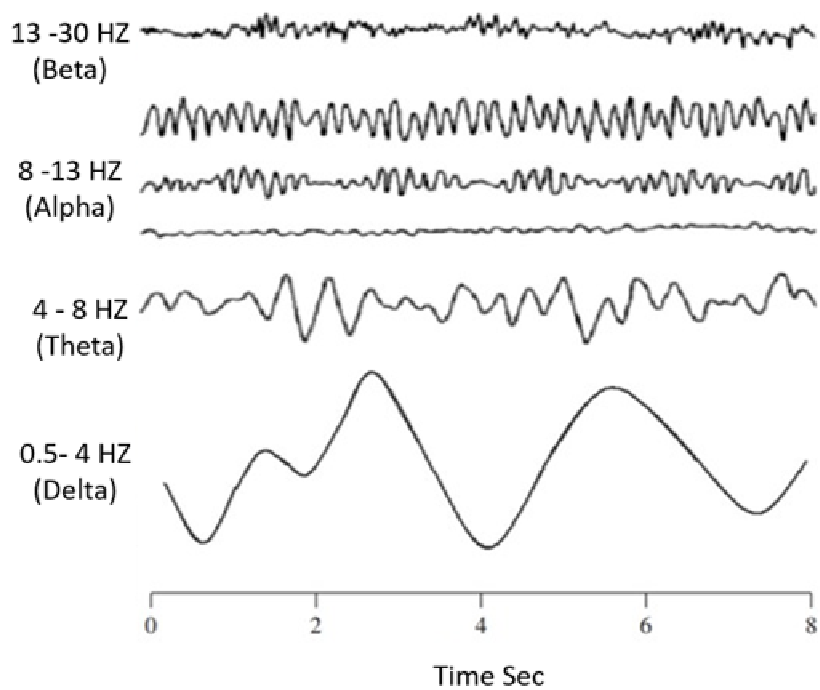

Figure 4.2: EEG Powerband

4.2 Feature Selection

When creating a predictive model, the process of feature selection involves limiting the number of input variables as they decreases the number of features that have to be measured and processed.In some circumstances, reducing the number of input variables may enhance the efficiency of the model along with lowering the computing cost of modelling. Feature reduction helps improves computational speed in lower dimensional feature spaces.

Feature selection is broadly classified as supervised and unsupervised. The supervised method is divided into intrinsic, wrapper and filter methods. Here a filter method and a intrinsic methods are used.

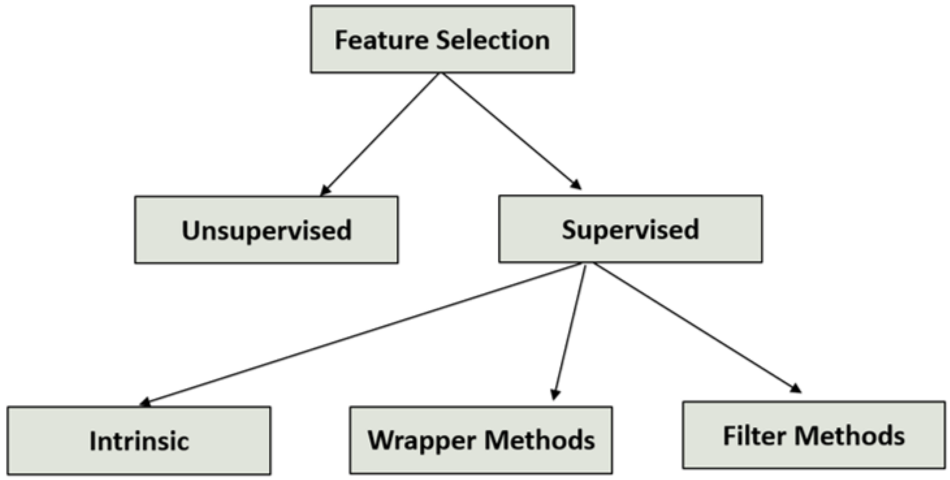

Figure 4.3: Feature Selection Techniques used

4.2.1 Constant and Quasi-Constant Feature Elimination

This is a filter based method. The constant features have same value throughout all its records. Quasi Constant features have one of the values with a 99.9% dominance rate. Both these type of features are dropped from the dataset.

4.2.2 Extreme Gradient Boosting (XGBoost)

Estimates of feature importance from a trained predictive mode can be automatically provided by ensembles of decision tree algorithms, such as gradient boosting. It is a method for building boosted trees where the relevance scores for each feature can be easily retrieved. In general, importance assigns a score based on the usefulness or value of each feature in building the boosted decision trees within the model. The relative relevance of an attribute increases when it is utilised more frequently to make important decisions with decision trees. Each feature in the dataset has its relevance evaluated directly, allowing for ranking and comparison of the attributes.

Chapter 5

MACHINE LEARNING MODELS

5.1 K-Nearest Neighbors (KNN)

K-nearest neighbors (KNN) algorithm is a supervised machine learning technique that is non parametric as it does not describe the classes in any predefined methods. It uses feature similarity for classification. Depending on how closely the new data point resembles the points in the training set, a value will be assigned to it. It is simple to understand and implement. The KNN classifier works well with small datasets than large datasets.

5.2 Gaussian Naive Bayes

A Gaussian N B algorithm is a special type of classification NB algorithm which is probabilistic. Gaussian NB applies Bayes theorem. It is applied with strong independence assumption. When the features contain continuous values, it is specially employed. Additionally, it is expected that each feature has a normal distribution, or a Gaussian distribution

5.3 Support Vector Machine (SVM)

SVM is a supervised machine learning method used for classification and regression on both linear and nonlinear data by using less memory space. It can handle the outliers in an efficient way. It doesn't intuitively enable multi class classification in its most basic form. It facilitates categorising data points into two classes and using binary classification. After dividing the multi class classification problem into numerous binary classification problems, the same method is applied to multi class classification.

5.4 XG Boost

Extreme Gradient Boosting referred to as XG Boost is an ensemble machine learning model. Decision trees are generated sequentially in this approach. Weights are significant in XG Boost as each independent variable is given a weight before being fed into the decision tree that predicts outcomes. Variables that the tree incorrectly predicted are given more weight before being placed into the second decision tree. XG Boost is a scalable application of gradient boosting which reduces computational power loss.

5.5 Decision Tree

A decision tree is mainly used for handling numerical and categorical non-linear data in an effective manner. It is a tree structure that resembles a flowchart. Any test on an attribute is indicated by each internal node. Each leaf node contains a class label, and each branch corresponds to a test result. It is simple to understand and implement.

5.6 GB Decision Tree

To create a single strong learner in gradient boosting decision trees, we merge numerous weak learners or individual decision trees. Each tree attempts to reduce the error of the one before it, and all of the trees are connected in series. Boosting algorithms are typically difficult to learn but extremely precise because of this sequential relationship and frequently has remarkable forecasting accuracy. It offers a number of hyper parameter tuning choices and it may optimise on various loss functions, making the function fit very adaptable. discrete bins are used in place of continuous data to reduce memory usage.

Chapter 6

RECURRENT NEURAL NETWORKS

The connections between nodes in a recurrent neural network (RNN) gradually form a directed or undirected graph over time. It can now act in a temporally dynamic way as a result. Utilizing their internal state, RNNs, which are evolved from feed-forward neural network models, can process input sequences of differing lengths.

6.1 Gated Recurrent Units (GRU)

Recurrent neural networks can have long-term dependencies, which can cause gradients to vanish, like in the case of RNN networks. Gated Recurrent Units are a kind of RNN that address this issue. In order to help the network predict outcomes in the future, GRUs solve this problem by storing memory from earlier time points.

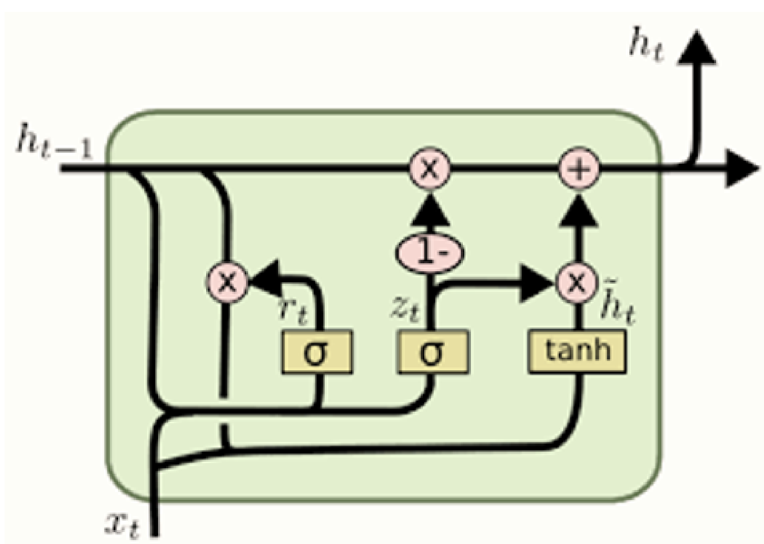

Figure 6.1: A Single GRU Cell

It has three gates and doesn't keep track of the internal state of the cell as shown in

Figure 6.1. The GRU used here have 256 layers, with Softmax activation and output dense layer.

6.2 Long Short-Term Memory(LSTM)

The LSTM network is used as a modification or enhancement to the Recurrent Neural Network. RNN suffers with long-term dependencies, leaving time series analysis impractical. LSTM can therefore resolve the issue due to the architecture of its repeating module. Each module includes an input, an output, and a forget gate that can add or remove data from the cell state. The forget gate employs a sigmoid function to determine what information from the prior cell state must be forgotten. Using point-wise multiplication of "sigmoid" and "tanh," the input gate regulates the information flow to the present cell state. The output gate decides which information should be transferred to the hidden state.

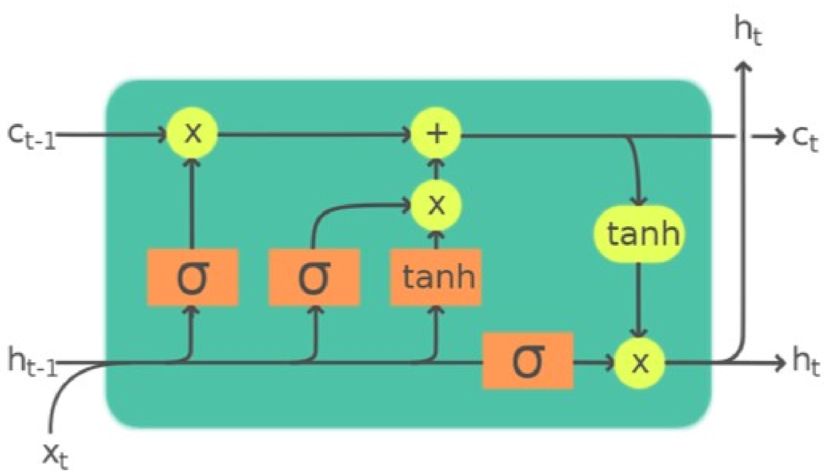

Figure 6.2: A Single LSTM Cell

Sl.No.	Layer	Type	Activation	Parameters
1	lstm	LSTM	Relu	66800
2	dropout	Dropout		0
3	lstm_1	LSTM	Sigmoid	30200
4	dropout_1	Dropout		0
5	dense	Dense		255

Table 6.1: LSTM Model Used

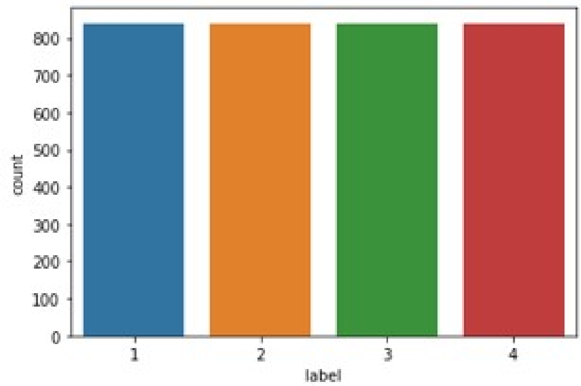

Figure 6.3: Data Count under Each Label

The extracted features are fed into an LSTM model as input data for emotion detection as shown in the Figure 6.3. The classes are: label 1: LANV, label 2: LAPV, label 3: HANV and label 4: HAPV. The LSTM model employed have 2 LSTM layers as shown in the Table 6.1, with 100 and 50 Units and activations Relu and Sigmoid respectively with 0.2 of Dropouts each. Finally the output layer is a Dense layer. The optimizer used is Adam optimizer.

Chapter 7

RESULTS AND DISCUSSIONS

In this work the extracted features from GAMEEMO dataset are classified using traditional machine learning as well as deep learning methods.

7.1 Experimentation Using Machine Learning Models

7.1.1 K-Nearest Neighbors (KNN)

The KNN algorithm is the first model experimented and it did not seem to perform well with data set. Table 7.1 shows the results. The total model accuracy obtained is 29 percentage.

Emotion Class	**Precision**	**Recall**	**f1-score**
LANV	0.29	0.31	0.30
LAPV	0.31	0.31	0.31
HANV	0.25	0.25	0.25
HAPV	0.30	0.29	0.30

Table 7.1: Experimentation Result Using KNN

7.1.2 Gaussian Naive Bayes

The next model Gaussian NB as shown in the Table 7.2 could not identify the emotion with much precision and accuracy, out of the four class labels, label 3 HANV show maximum recall and f1 score. The resultant model accuracy is 24 percentage.

7.1.3 Support Vector Machine (SVM)

The SVM also had a very poor performance on emotion recognition from the data set. The Table 7.3 shows the resultant values of precision, recall and f1 scores. 24, 22.9 and 23.9 percentages for overall accuracy, accuracy using ploy kernel and accuracy using rbf kernel respectively were also obtained.

Emotion Class	Precision	Recall	f1-score
LANV	1.00	0.00	0.01
LAPV	0.29	0.24	0.26
HANV	0.23	0.78	0.35
HAPV	0.20	0.00	0.01

Table 7.2: Experimentation Result Using Gaussian NB

Emotion Class	Precision	Recall	f1-score
LANV	0.29	0.31	0.30
LAPV	0.31	0.31	0.31
HANV	0.25	0.25	0.25
HAPV	0.30	0.29	0.30

Table 7.3: Experimentation Result Using SVM

7.1.4 Decision Tree

The Decision Tree classifier performed better than the previously experimented models. The results are as in Table 7.4, out of the four classes LAPV is classified with better precision and HANV with high recall and f1 score. The accuracy of model is about 50 percentage.

Emotion Class	Precision	Recall	f1-score
LANV	0.57	0.50	0.53
LAPV	0.58	0.33	0.42
HANV	0.46	0.71	0.56
HAPV	0.46	0.48	0.47

Table 7.4: Experimentation Result Using Decision Tree

7.1.5 XG Boost Classifier

The XG Boost Classifier resulted an accuracy of 54 percentage, the results are as shown in the Table 7.5. classified the class LANV is with maximum precision and HANV with high recall and f1 score.

Emotion Class	Precision	Recall	f1-score
LANV	0.64	0.51	0.57
LAPV	0.49	0.45	0.47
HANV	0.54	0.70	0.61
HAPV	0.48	0.50	0.49

Table 7.5: Experimentation Result Using XG Boost classifier

7.1.6 GB decision tree

The GB Decision tree could recognise the emotions with a better performance due to the gradient boosting applied. Here as shown in Table 7.6, class label 1 LANV was detected with maximum precision and class label 3 HANV with high recall and f1 score. The obtained overall accuracy of the model is 75 percentage.

Emotion Class	Precision	Recall	f1-score
LANV	0.83	0.77	0.80
LAPV	0.73	0.74	0.73
HANV	0.74	0.78	0.76
HAPV	0.72	0.72	0.72

Table 7.6: Experimentation Result Using GB Decision Tree

Of the employed machine learning models, Gradient boosting decision tree achieved an accuracy of 75 percent which was found the best.

7.2 Recurrent Neural Network

7.2.1 Gated Recurrent Units (GRU)

The GRU model had 256 units and performed with less training time and being a recurrent neural network it was expected to achieve a better result than ML models, However, it could not show a very gud performance with high accuracy. The model achieved an accuracy 64.583 percentage. LANV is the class with maximum precision and f1 score (Table 7.7).

Emotion Class	**Precision**	**Recall**	**f1-score**
LANV	0.71	0.65	0.68
LAPV	0.64	0.34	0.38
HANV	0.60	0.66	0.63
HAPV	0.64	0.64	0.64

Table 7.7: Experimentation Result Using GRU

7.2.2 Long Short-Term Memory (LSTM)

1The LSTM model which is known to mine the temporal features of a time series data could achieve an accuracy of 91.7 percentage during the experiment. The model accuracy, performance and loss are depicted in Figure 7.1, 7.2, 7.3 respectively.

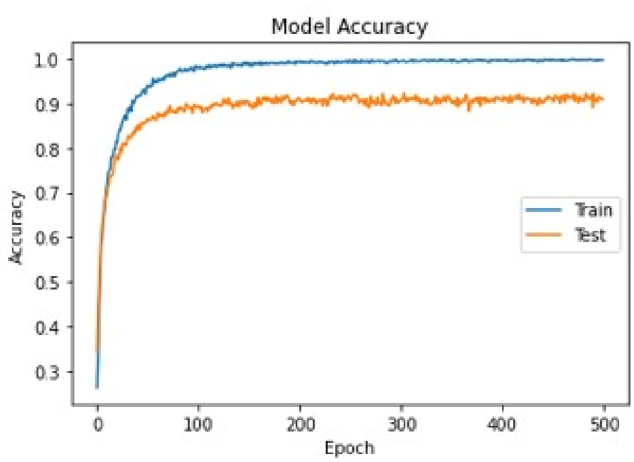

Figure 7.1: Accuracy of Experimentation Using LSTM

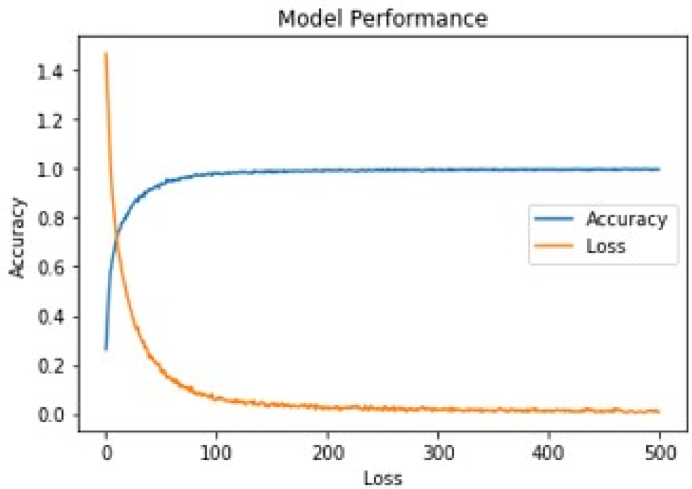

Figure 7.2: Performance of Experimentation Using LSTM

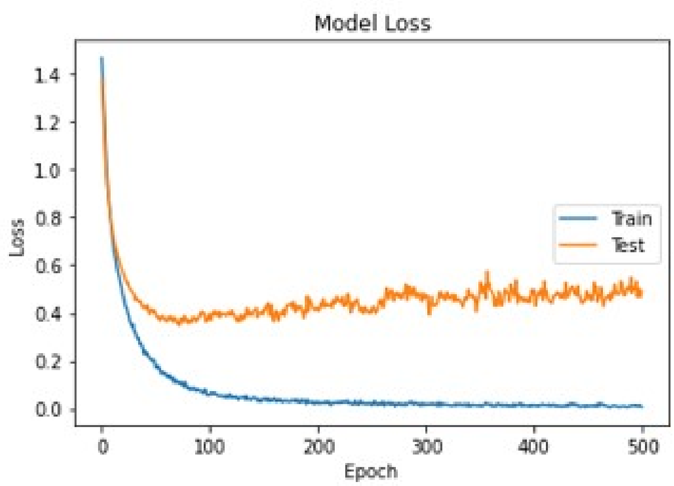

Figure 7.3: Loss of Experimentation Using LSTM

7.3 Feature Selection

The performance of the model could be improved using better systematic feature selection method, this was experimented using two methods,filter based feature elimination and feature importance based selection method.

7.3.1 Constant and Quasi-Constant Feature Elimination

This is a Filter based feature selection method to eliminate the constant and quasi constant features. The constant features only has one value for each record in the dataset. Quasi constant features are almost constant and have a single value in record that have 99% dominance rate. They don't offer any information that is helpful in the classification. Therefore, they are eliminated from the dataset. As a result, 66 features were reduced to 64 features. There was an observable improvement in the performance of the LSTM model. The accuracy achieved is 92.113 percentage which is higher than the results shown without using feature elimination.

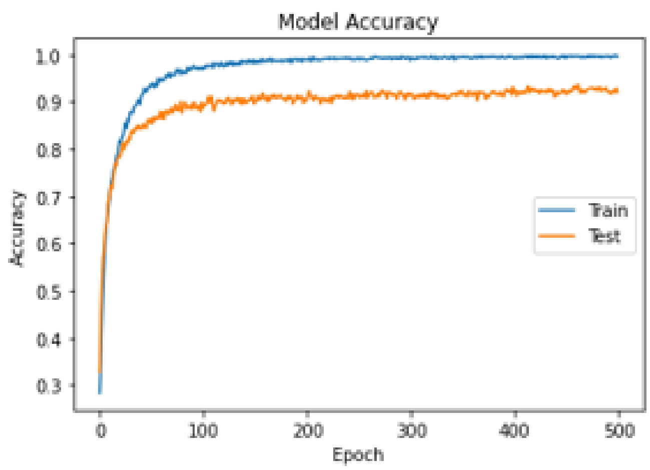

Figure 7.4: Accuracy of Experimentation Using Feature Elimination

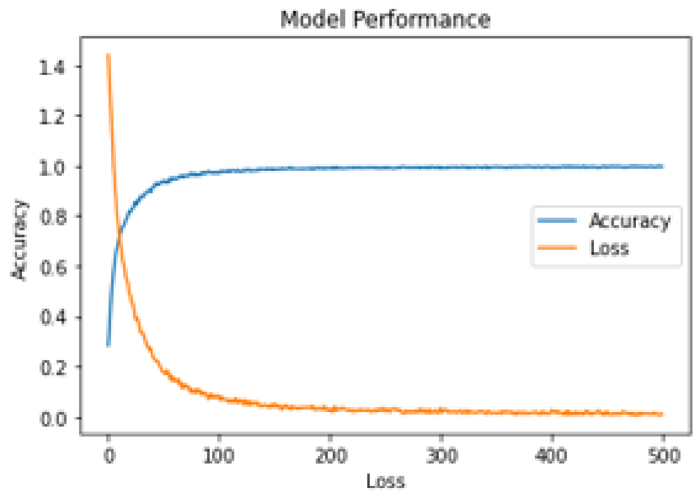

Figure 7.5: Performance of Experimentation Using Feature Elimination

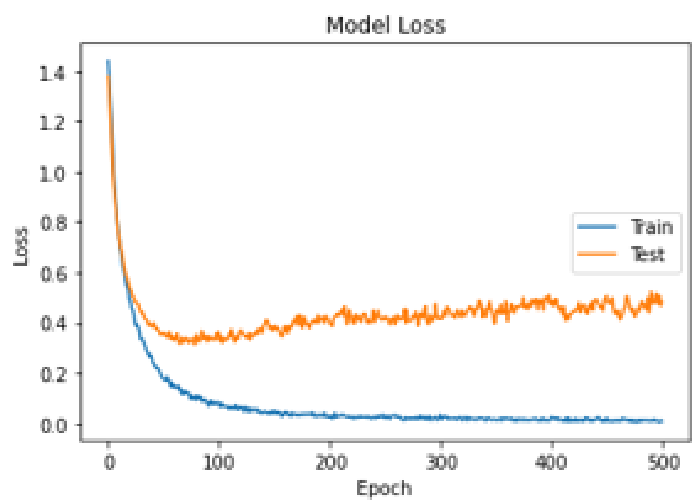

Figure 7.6: Loss of Experimentation Using Feature Elimination

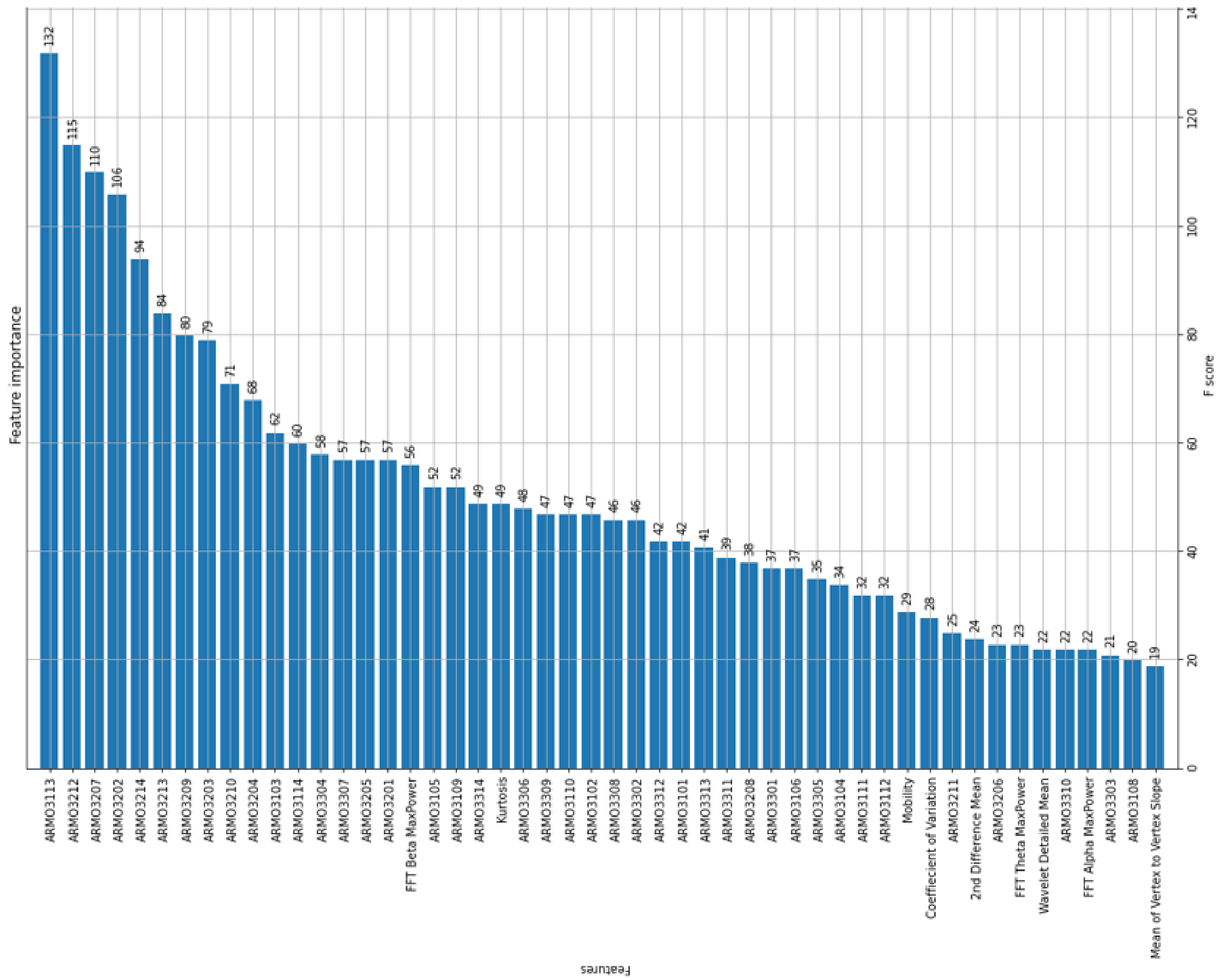

Figure 7.7: Ranking based on Feature Importance

7.3.2 Extreme Gradient Boosting (XGBoost)

A score based on feature importance indicates how significant or useful each feature was when building the boosted decision trees inside the model. Each attribute in the dataset has its importance assessed directly, enabling ranking and comparison of the attributes. The feature importance provides a score as shown in Figure 7.7. This importance is calculated explicitly for each feature in the dataset, allowing features to be ranked with an F score ranging from 1 - 140. The intensity of each feature in the bar-plot indicates its importance. Here 132 is the score for most important feature while 19 is the score for least important feature. After this, the features are selected based on a threshold dropping the less important features from the dataset. The accuracy achieved a 93.55 percentage.

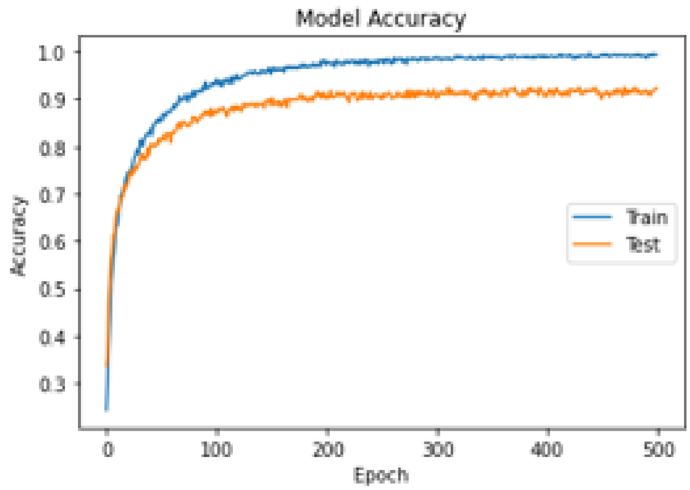

Figure 7.8: Model Accuracy of Extreme Gradient Boosting

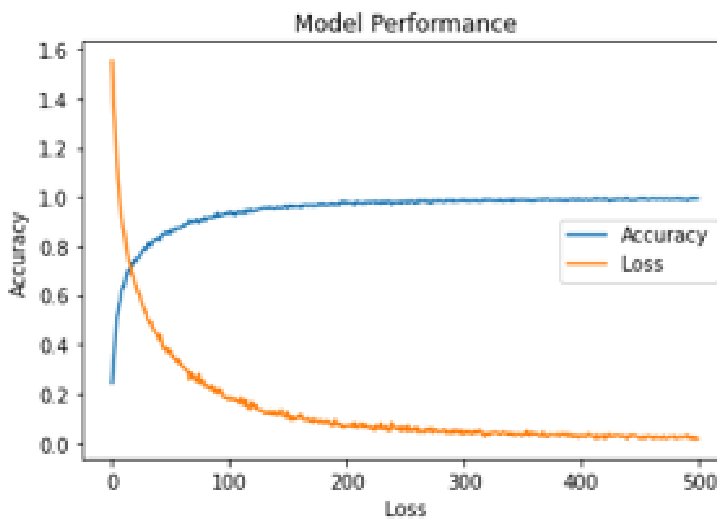

Figure 7.9: Performance of Experimentation Using Extreme Gradient Boosting

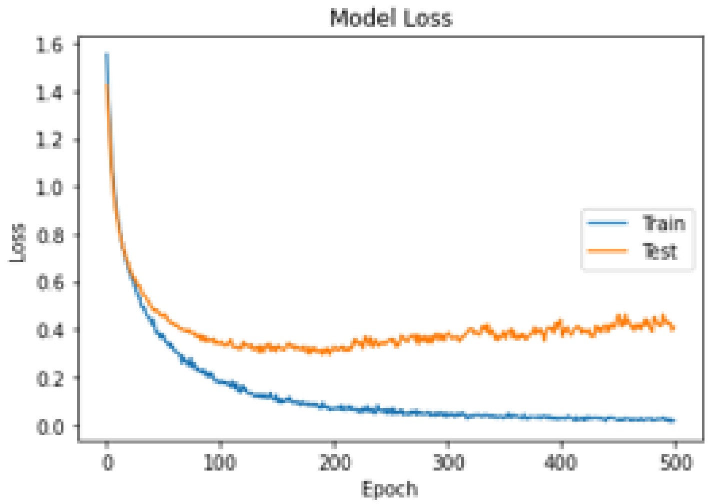

Figure 7.10: Loss of Experimentation Using Extreme Gradient Boosting

7.4 LSTM on different EEG Emotion Dataset

7.4.1 Sample Dataset

The features were extracted from Sample Dataset as it was done on the GAMEEMO dataset. The model performance on the dataset is as shown in Figure 7.12. An accuracy of 84.1 percentage was achieved as result.

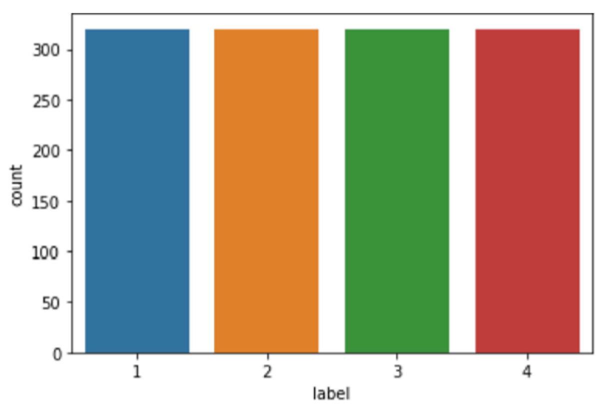

Figure 7.11: Sample Data Count

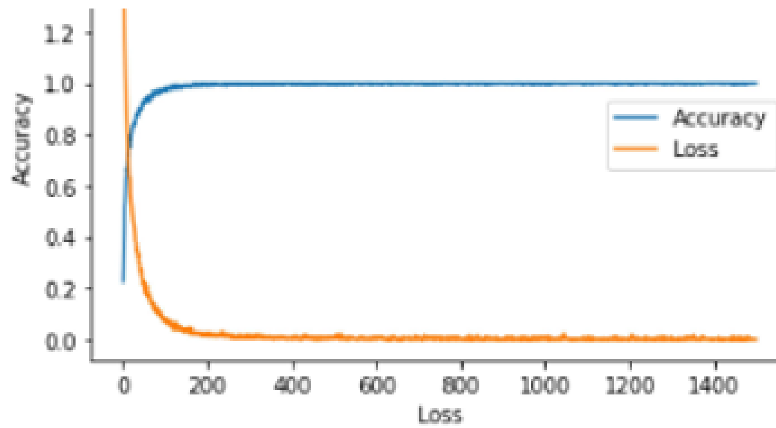

Figure 7.12: LSTM Model Performance on Sample dataset

7.4.2 EEG Brainwave Dataset: Feeling Emotion

From the EEG Brainwave dataset an accuracy of 98.59 percentage was the result using the model. Classifies data into three categories: negative, neutral, and positive with an accuracy of 98.59 percentage. Similarly the Positive-Negative Valence recognition of the Gameemo dataset achieved 94 percentage accuracy.

7.5 Performance analysis and Comparison

7.5.1 Machine Learning Models vs Recurrent Neural Networks

Deep learning methods automatically learns deep features and recognizes emotions through such models as Long Short-Term Memory, Gated Neural Network and other Recurrent Neural Networks. In this work both aspects are explored as machine learning models like KNN, SVM , Gaussian NB , Decision Tree , Xg Boost Classifier, GB decision tree and deep learning models like LSTM and GRU are employed. For this work a dataset named GAMEEMO is used. There are 1568 (4×14×28) EEG data availablethat are 20 min long, and the total number of samples is 38,252. 62 features were extracted from this EEG dataset and then given to various classifier models. The machine learning classifier models could provide a good testing accuracy.Of the employed machine learning models, Gradient boosting decision tree achived an accuracy of 75 percent which was found the best among them. LSTM model achieved an accuracy of 91.3 percent, which was found best among all the models experimented. When comparing the performances of the models experimented as shown in the Table 7.8, LSTM model had the best performance. GRU took less training time than LSTM, but the latter was more accurate.

Sl.No	Model	Accuracy(in percentage)
1	KNN	29
2	Gaussian NB	24
3	SVM	25
4	Decision Tree	50
5	XG Boost Classifier	54
6	Gradient Boosting Decision Tree	75
7	GRU	65
8	LSTM	**91.7**

Table 7.8: Performance Comparison of Different Experimented Models

7.5.2 Importance of Feature Selection

The feature selection method improved the performance of the model to an extend. Out of the two method applied the XG Boost based feature selection method which used the feature importance helped the model to achieve an accuracy of 93.55 percentage. Finding the best features for a dataset is thus helpful in improving the performance of the model.

Sl.No.	Feature selection method	Training Accuracy(in percentage)
1	No methods applied.	91.7
2	Constant and Quasi constant feature elimination	92.113
3	XG boost	**93.55**

Table 7.9: Performance comparison of the model with feature selection

Chapter 8

CONCLUSION

Emotions are not just the physiological states of the different feelings, behaviors, and thoughts of human beings but also psychological along with physiological reactions brought about by a number of external stimuli. The EEG signals captured from CNS, have the property to directly reflect the brain activity and thus have an intrinsic relationship with the emotional states of humans. Emotion recognition gained more attention in the past decade as it is directly connected to various fields like psychology, healing, physiology, marketing and studies. The gaming market has expanded to be one of the most significant entertainment markets because of recent advanced technologies. Understanding how the player emotions automatically inferred during game play can be used to enhance the quality of the game is important. Numerous ideas regarding game design have been put up to outline some standards for creating games that might elicit particular emotions in players. They mostly rely on interviews or observations to draw conclusions about the effectiveness of their strategy from qualitative data. There are various empirically based works that use quantitative data to directly analyse the emotional states of players. These studies, however, typically see data analysis as a categorization issue involving, primarily, game events. Our purpose is to understand how players' emotions inferred can be used to enhance the quality of the game. The two primary components of EEG-based emotion identification techniques are deep learning and conventional machine learning. Deep learning techniques use models like LSTM, Gated Neural Networks, and other Recurrent Neural Networks to automatically learn deep features and distinguish emotions. Both aspects are investigated in this work, which makes use of both deep learning models like LSTM and GRU as well as machine learning models including KNN, SVM, Gaussian NB, Decision Tree, Xg Boost Classifier, and GB decision tree.This study makes use of the GAMEEMO dataset. There are 38,252 total samples and 1568 (41428) EEG data with a 20-minute duration available. From this EEG dataset, 67 characteristics were taken out and applied to multiple classifier models.

The classifier models for machine learning could offer good testing accuracy. The Gradient Boosting Decision Tree was determined to have the highest accuracy among the used machine learning models, with a 75 percentage accuracy rate. Among all the models tested, the LSTM model's accuracy of 91.7 percentage was found to be the highest. The LSTM model performed the best when comparing the experimental models' performances. LSTM was more accurate, but GRU required less training time. In future works, a better systematic feature selection strategy may be used to enhance the model's performance.feature selection methods like were applied to increase the accuracy of which XB boost method give an accuracy of 93.55 percentage. The model was tested for its reliability on a Sample dataset which followed the same feature extraction prior to the detection and achieved an accuracy of 84 percentage. The emotions dataset was used for positive-negative valance emotion detection which resulted in 98.59 percentage accuracy.The Real time Emotion recognition and analysis methods can enhance the process to a much better level. The suitable feature extraction and selection can improve the performance in further studies and works.

References

[1] Xing X, Li Z, Xu T, Shu L, Hu B, Xu X. SAE+ LSTM: A New framework for emotion recognition from multi-channel EEG. Frontiers in neurorobotics. 2019 Jun 12;13:37

[2] Wei C, Chen LL, Song ZZ, Lou XG, Li DD. EEG-based emotion recognition using simple recurrent units network and ensemble learning. Biomedical Signal Processing and Control. 2020 Apr 1;58:101756.

[3] Yin Y, Zheng X, Hu B, Zhang Y, Cui X. EEG emotion recognition using fusion model of graph convolutional neural networks and LSTM. Applied Soft Computing. 2021 Mar 1;100:106954.

[4] Ramzan M, Dawn S. Fused CNN-LSTM Deep learning emotion recognition model using Electroencephalography signals. International Journal of Neuroscience. 2021 Jun 13(just- accepted):1-0.

[5] Joshi VM, Ghongade RB. EEG based emotion detection using fourth order spectral moment and deep learning. Biomedical Signal Processing and Control. 2021 Jul 1;68:102755.

[6] Zhuang N, Zeng Y, Tong L, Zhang C, Zhang H, Yan B. Emotion recognition from EEG signals using multidimensional information in EMD domain. BioMed research international. 2017 Aug 16;2017.

[7] Liu H, Zhang Y, Li Y, Kong X. Review on Emotion Recognition Based on Electroencephalography.Frontiers in Computational Neuroscience. 2021:84.

[8] Sanei S, Chambers JA. EEG signal processing. John Wiley Sons; 2013 May 28.

[9] Torres EP, Torres EA, Hernández-Álvarez M, Yoo SG. EEG-based BCI emotion recognition: A survey. Sensors. 2020 Jan;20(18):5083.

[10] Alakus TB, Gonen M, Turkoglu I. Database for an emotion recognition system based on EEG signals and various computer games–GAMEEMO. Biomedical Signal Processing and Control. 2020 Jul 1;60:101951.

[11] Bird JJ, Manso LJ, Ribeiro EP, Ekárt A, Faria DR. A study on mental state classification using eeg-based brain-machine interface. In2018 international conference on intelligent systems (IS) 2018 Sep 25 (pp. 795-800). IEEE.

[12] Bird JJ, Ekart A, Buckingham CD, Faria DR. Mental emotional sentiment classification with an eeg-based brain-machine interface. InProceedings of theInternational Conference on Digital Image and Signal Processing (DISP'19) 2019 Apr.

[13] Granato M. Emotions Recognition in Video Game Players Using Physiological Information.

[14] Kołakowska A, Landowska A, Szwoch M, Szwoch W, Wrobel MR. Emotion recognition and its applications. InHuman-Computer Systems Interaction: Backgrounds and Applications 3 2014 (pp. 51-62). Springer, Cham.

[15] Koelstra S, Muhl C, Soleymani M, Lee JS, Yazdani A, Ebrahimi T, Pun T, Nijholt A, Patras I. Deap: A database for emotion analysis; using physiological signals. IEEE transactions on affective computing. 2011 Jun 9;3(1):18-31.