International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 15, No.1 (Feb-24)

https://dx.doi.org/10.12785/ijcds/150141

Emotion Detection using Stack Auto Encoder, Deep Learning and LSTM Model

Disha Shah[1], Rashmi Rane[1] and Shakti Kinger[1]

[1] *Department of Computer engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India*

Received 24 Apr. 2023, Revised 7 Jan. 2024, Accepted 21 Jan. 2024, Published 1 Feb. 2024

Abstract: There is an increasing need for distress emotion acknowledgement and channeling the same is very important. There is an increasing need for machines to understand human and their complex emotions deeply. This research describes a unique framework for emotion detection that helps brain-computer interface/machine (BCI) to understand human emotions and brain complexity and working using multi-channel electroencephalograms (EEG). We have employed two datasets in our work, Database for Emotion Recognition Through EEG and ECG Signals from Wireless Low-Cost Off-the-Shelf Devices (DREAMER) and dataset for emotion analysis using EEG, physiological and video signals (DEAP) for LSTM model implementation and validation respectively. A linear model is considered for EEG signal mixing and an emotion timing model comprise the framework. Leveraging the contextual correlations within EEG feature sequences, our proposed methodology effectively recovers EEG source signals from the obtained EEG signals, thereby improving the accuracy rate of classification. Also, stress bins were set up for individual users to assess their degree of stress and calmness following exposure to external stimuli. DEAP dataset using LSTM framework was enforced for emotion recognition, and mean recognition accuracy using area under curve as evaluation matrix for valance and arousal was 82.02% and 76.52% respectively, validating competence of framework. Novelty of our work is it improved competency in feature extraction, use of context correlations increasing accuracy and use of spatio-temporal features in the proposed model framework.

Keywords: Electroencephalogram (EEG), Distress Emotions, Emotion Detection, OpenBCI, Stack Auto Encoder, LSTM, fully connected network.

1. Introduction

In our contemporary, material-driven society, the importance of emotional intelligence is widely recognized, especially amid the reflective period of the ongoing pandemic. This challenging era has subjected many individuals to various extremes on the emotional spectrum, ranging from heightened anxiety to instances of suicide and an increased prevalence of depression [1]. In recent epochs, there has been a growing realization of the pivotal role that mental or psychological health plays in advancing global developmental aspirations. Evidently, despondency emerges as a substantial contributor to incapacitation.

The prima-facie of this paper is to find cause-effect relationship between eustress and distress emotion using EEG signal and spatio-temporal data to develop a stress bin empirically and scientifically. Objective and significance of our work is further articulated in detail in section 1 Aims and Our Contribution section.

These are the following statistics of mental health issues in recent years for various age groups and sex through graphs and charts.

Figure 1 shows how mental health in kids and teenagers in 5-16 years of age has worsened. The graph shows percentage of children in England with potential mental health concerns by age group. The graph shows percentage by year, of 2017 and 2020, grouped by age 5-10 and 11-16 and sex, distinguished by color blue and pink for respective year. There was 6.4% and 3.8% hike in percentage of deteriorating mental health from 2017 to 2020 in boys and girls respectively in 5–10-year age. While 4.1% and 6.2% in the ages 11-16 years respectively.

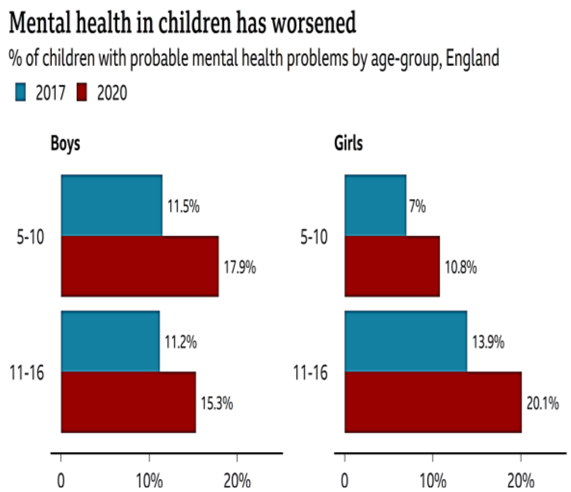

Figure 1. Percentage of children in England with potential mental health concerns by age group [2]

E-mail:itsmedisha.23@gmail.com, rashmi.rane@mitwpu.edu.in, shakti.kinger@mitwpu.edu.in

Figure 2 shows almost 16% of adults exhibit symptoms of depression. In this graph citizens of England and Wales are divided by the age group of 16-39, 40-69,70 and above age. The graph concludes students and adults of age group 16-39 are overly depressed and there was a major push through from 10.9% in 2017 to 31% in 2020. Also, elderly people able 70 have great depression which hiked from 5% to 10.3% in 2020.

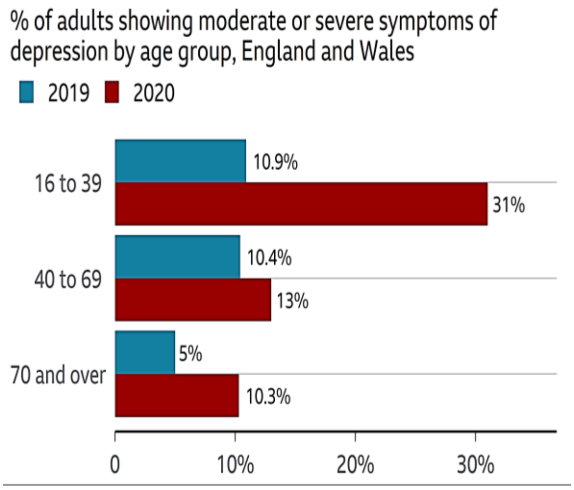

Figure 2. Percentage of adults exhibiting symptoms of depression divided by age group [2]

We can conclude from the graph of Figure 3 that the regular feelings and emotions concerns us the most which goes unnoticed by people, like being alone, feeling burden, stress or anxiety, relationships. People need to be aware and acknowledged of the mental issues which they are facing before it reaches to that threshold that requires professional help, which we think they should definitely do. Our purpose is to detect the stress beforehand and suggest them (the user/participant) to perform activities that causes eustress i.e., happy and calm emotion, even if they are in their own company.

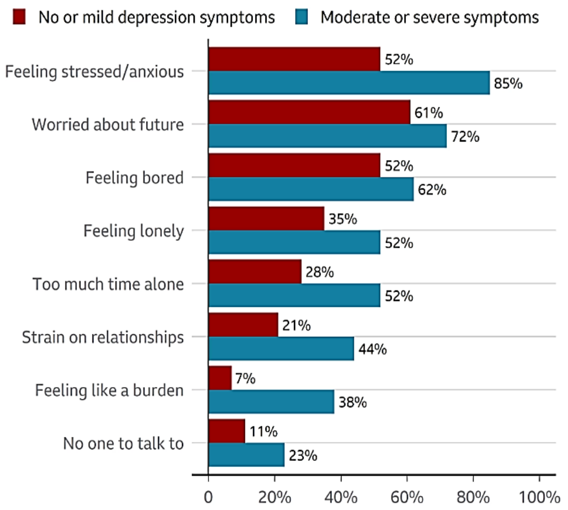

Figure 3. Impact and causes of mental illness [2]

The recent statistics for mental health problems in United States in Figure 4 highlighting depression and are hopeless (42%) in the Gen Z (Z Generation People) and 18% people are suicidal. Other generations and age groups are also dealing with severe mental health problems or know people who are suffering from one. As we compare it with Figure 6 which is data based on year 2019-2020 when COVID-19 hit us, we can conclude from this comparison between Figure 4 and 6 that mental health is not completely depended on COVID-19 or Movement Control Order (MCO).

The domains of BCI and medical diagnostics are becoming more and more prominent due to the increased focus on EEG signal processing [3]. Emotions may reveal details about a person's habits, behavior, interests, and even health. Automation can benefit from human emotion recognition to improve the robustness of brain-computer interfaces while also supporting them with action processing and social cognition. As a result, development on EEG-based automated Emotion recognition is essential for neurological devices such as computers and robots because it allows them to interpret people's interacting motives and emotions using wirelessly received EEG [4].

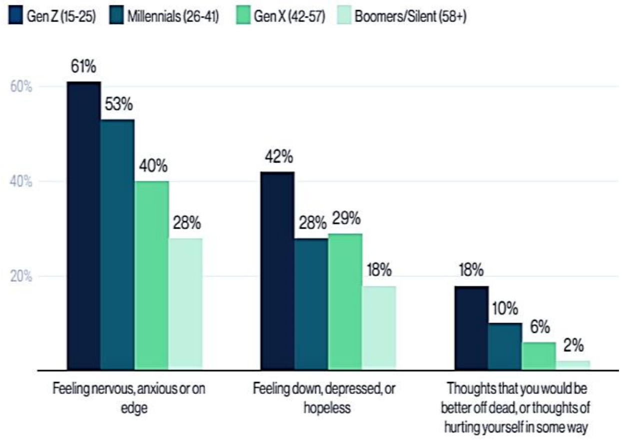

Figure 4. Recent statistics for mental health [5]

EEG signals is composure of source signal and other redundancies which we can call noise. A standard EEG equipment consists of numerous electrodes capable of recording data in both time (temporal) and space (spatial). After successful extraction of required EEG signal, researchers concluded that the spatial and temporal data of signal is also key information. We will look into this in further sections of this paper.

A. Aim

Aim of this study is to detect distress emotion mostly, create stress bins and train our model on that data. Significance of this work is to be able to distinguish between eustress and distress emotions, its prima-facie factors and cause-effect relationship between them. Though with today's weak AI we cannot predict the temporal behavior of humans, but we might be able to find the root cause. The purpose of this study is very crucial and its untapped potential for future research work can really help in understanding human emotions neurologically and empirically; thus, we want to contribute in it.

B. Background Motivation

With 'N' number of thoughts processing in real time, influencing us vividly, we never realise on what point of time we might be ailing from what psychological disorder. People already have low Emotional Quotient (EQ), but often are unable to recognise it, which is necessary, at the same time complex. To cut short, stress has 2 types: eustress (is positive one, can lead to a personal healthy growth) & distress (the negative one, which affects our persona as well). Distress must be dealt adequately, since depression and disorders of mood, anxiety and stress are ever rising in human history. Following Figure 5 depicts the anxiety and/or depressive conditions in the USA in 2019-20.

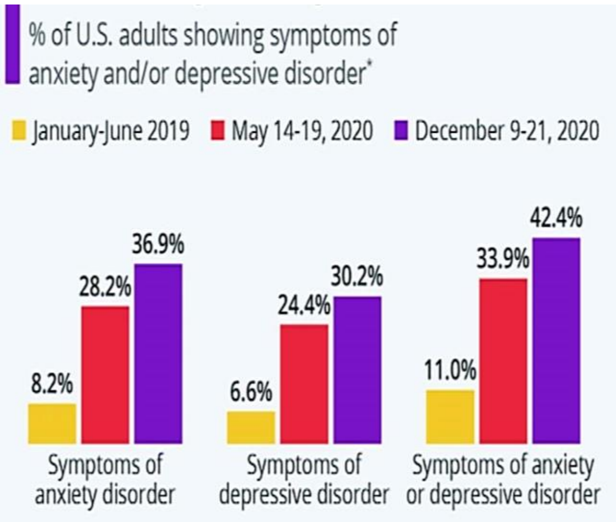

Figure 5. Proportion of anxiety and/or depressive conditions in the United States [2]

C. Our Contribution

To investigate and categorize distress emotions using EEG signals and develop a stress bin for user to detect the stress level or calmness they experience under external stimuli. Also, constructing a neural network model to train on such very extreme high and low valance and arousal levels and detect the targeted emotion. We also keep a goal to use spatio-temporal signal data for model development.

The further work in Section 2 explains inspired works from different authors in our domain. Section 3 includes methodology and flowcharts explaining our work, 4 discusses the results, discussion and limitations around our study and finally 5 have future work and conclusion.

2. RELATED WORKS

Research study addresses multiple research initiatives linked to emotion recognition using different physiological signals such as EEG, Electrocardiogram (ECG), skin temperature (SKT), etc. [6]. The studies utilized machine learning algorithms such as Gaussian Mixture Model (GMM), clustering algorithms like K means, Machine learning classifiers like the support vector machine, and artificial neural networks like ELM's (Extreme learning machine) [7] to classify emotional EEG data into different categories. The research also involved pre-processing techniques such as bandpass filtering, Welch period technique, Digital Signal Processing (DSP), Power Spectral Density, Wavelet Decomposition and transforms, and Hjorth Parameter for extracting feature vectors from filtered EEG data [8].

The datasets used in the studies include MAHNOB-HCI, DEAP, SJTU Emotion EEG Dataset (SEED), DREAMER, and WeDea.

WeDea constitutes a diverse dataset gathered as thirty participants viewed 79 brief video clips, each designed to elicit distinct emotions. The data acquisition involved the utilization of a portable headgear device. Pre-processing techniques were used to reduce artifacts from the data. The studies achieved classification accuracy ranging from 50.7% to 100% using various machine learning algorithms and feature extraction methods.

In Brain-IoT-based inspired study [9], an EEG headset, a mobile application, a ThingSpeak database, a heart rate variability sensor, and a galvanic skin reaction (GSR) sensor were used to detect and alleviate negative emotions [3]. Researchers analyzed EEG, heart rate variability sensor, and a galvanic skin reaction sensor to detect emotions and a mobile application to play a movie to evoke positive emotions. The to compute metrics like arousal and valence they used MathWorks MATLAB for coding using pre-processed data, which are then submitted to a database. The study also found that including inertial sensing data in the data given to classifiers increased classification accuracy, with Distributed Random Forest achieving the highest accuracy of 82.49%. The investigators reach the conclusion that Brain-IoT has the potential to be employed in crafting a system capable of transforming adverse emotions into positive sentiments.

Xiang Li's, 2018 paper on exploring cross-subject emotion recognition, they want to do a more thorough examination into identifying cross-subject emotions using brain imaging data from 18 various types of linear and non-linear EEG data characteristics [10]. They used DEAP and SEED dataset to test the effectiveness of these features. Using Automatic feature selection method and Support Vector Machine (SVM) classifier they achieved highest accuracy of 59.06% and 83.33% with DEAP and

SEED datasets respectively [10]. They also concluded that compared to the other features Hjorth parameter achieved the best mean recognition accuracy and various other comparisons and observations have been made in different features.

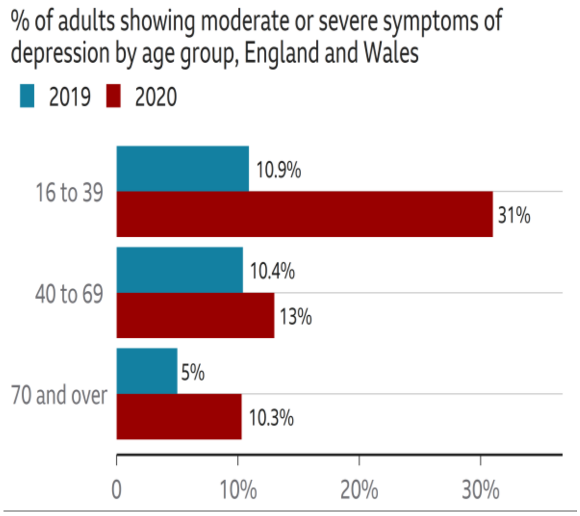

Figure 6. Statistic of mental health [2]

Xiang Li's further research in this domain in 2020 [11] using latent factor decoding hypothesize that in emotion processes, there are subject-specific default brain variables. They propose an unsupervised model such as variational auto encoder (VAE) that extract latent factors and using the decoded latent factor sequences, Long Short-Term memory (LSTM)-sequence-model is used for mining emotion-related information [11]. This model was validated using two multichannel datasets, DEAP and SEED, and the results were compared with the autoencoder approach and independent component analysis. They achieved superior recognition performance on the VAE-decoded latent factors and it can also detect hidden state space of brain and was able to determine depression, Alzheimer, brain impairment, etc.

Z. Khalili and M. H. Moradi's study on emotion recognition using correlation dimension proposed a multimodal fusion for emotion detection using fusion between brain signals and peripheral signals [12]. EEG signals, Galvanic Skin Response (GSR), body warmth, BP and breathing were input and were acquired based on valance and arousal levels to be positive, negative or calm. A most influential non-linear feature, correlation dimension is used to improve the results of extracted feature. Extracted features are divided into 3 different feature set of peripheral signals, EEG signals and a set containing both and all these sets are evaluated using pattern recognition technique the Quadratic Discriminant Classifier. They concluded that EEG signals perform better with 66.66% highest and can be improved to 76.66% using correlation dimension but results of EEG and peripheral signals together are more robust compared to single results.

Guillaume Chanel's assessment [13] of using these signals adapts game difficulty level according to player's emotion through EEG signals to maintain engagement of player. They have extracted signals from players engaging themselves in a game with 3 different complexity levels experiencing variety of emotions but it was observed that they experience boredom when playing a difficulty level several times. They built and trained classifiers to detect the emotion classes. Easy level has low arousal and pleasure which is boredom, medium has higher arousal and pleasure and amusement, and hard associates with high pressure and arousal but low arousal which is anxiety. Figure 7 shows player's automatic adaption to emotional reaction, where 3 different kind of arrow shows change in competence and difficulty level and bold arrow shows automatic adaption. They used fast correlation-based filter (FCBF) and Analysis of variance (ANOVA) for peripheral and EEG signal and got result of 59% and 56 % respectively [23]. But accuracy obtained after fusion of both signals increased result to 63%.

It was also analyzed that EEG features were more robust for short term assessment for emotions.

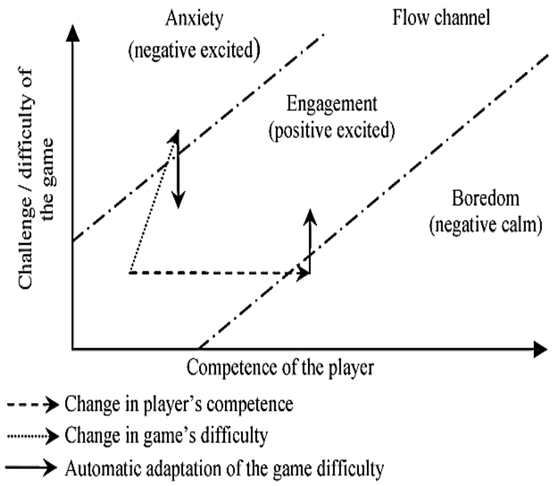

Figure 7. Automatic adaptation to emotional reactions [13]

Raja Majid Mehmood, in 2022, proposed a method to achieve emotion recognition using Hjorth-activity covers research gap of employing a few emotions class or large feature set using objective of Hjorth-feature [14]. This emotion recognition model overcome their research gap and traverses more in valance-arousal domain to provide variety of emotion class. They have employed datasets such as DEAP, SEED-IV, DREAMER, SELEMO, and ASCERTAIN and for feature extraction they have use Hjorth's parameters and extracted features with minimum error rate. Models utilized for comparative analysis are machine and deep learning, voting ensemble, and tree algorithms, like random forest and decision tree. The average accuracy was 69%, 76%, 85%, 59%, and 87% in the respective sequence of datasets.

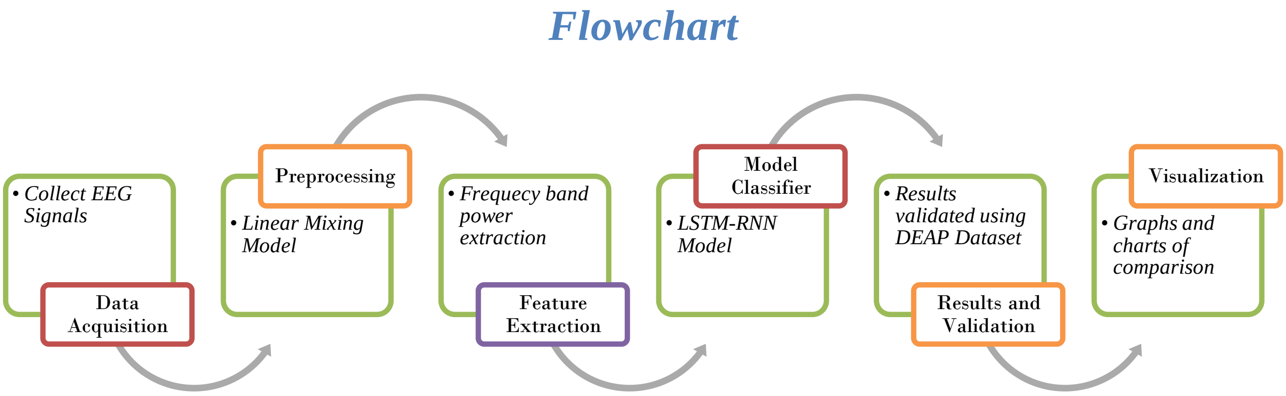

Figure 8. Brief architecture

Hybrid model combining fuzzy cognitive maps (FCM) and SVM Classifier is used in paper [15]. A method for emotional response classification using a combination of FCM and SVM classifiers is employed, sophisticated DSP techniques such as CS, Furrier transform, and Local Binary Pattern (LBP) are used to simplify data and retrieve characteristics for emotional response classification [3]. The characteristics extracted are fed into the SVM classifier and undergo adjustment through a defuzzification technique. The resultant output involves a hidden layer with seven nodes, achieving stability upon convergence. Employing the defuzzification approach, the values of nodes on the "NNP" (Negative Neutral Positive) plane are tuned. This methodology is applied to streamline the data and distill the key features of emotional responses.

Another paper proposes a fully connected neural network (FCNN) called E2ENNet for emotional BCI that can recognize emotions directly from EEG signals [16]. The proposed E2ENNet comprises of three components: feature extraction, channel attention, and emotion classification. The feature extraction component uses convolutional neural network (CNN) to extract spatial and temporal features from the input EEG signals [16]. The channel attention component applies method to choose important features from several EEG channels to improve the model's interpretability and generalization capability [16]. Finally, to categorize emotions based on the given features, the emotion classification component employs a FCNN [17]. The E2ENNet is tested on the DEAP data and compared to other techniques [18]. The outcomes reveal that E2ENNet surpasses previous approaches in evaluation matrix in regards of accuracy, F1 score, and area under the curve [19]. The proposed channel attention mechanism also provides insights into the importance of different EEG channels in recognizing emotions. The study highlights the potential of the proposed E2ENNet for real-time emotional BCI applications. The end-to-end nature of the network allows efficient processing of EEG signals, eliminating the need for manual feature engineering. The attention mechanism also provides interpretable insights into the decision-making process of the model. Future research can explore the generalizability of the proposed approach across different emotional states and individuals with different EEG signal characteristics.

The recent 2023 paper, summarized the characteristics of emotional tendencies among different demographic groups of museum visitors under digital/technological external environment [20]. For digital emotional experience they used EEG data and PAD model (Psychological model based on Pleasure, Arousal and Dominance). Classification of model was done using linear, polynomial kernel and Gaussian kernel SVM and got success rates of 73%, 72.4%, 72.4% respectively in quasi-experimental setting [20]. The emotional features of museum visitors are analyzed using digital emotional experience assessment, and the results show that negative feelings are typically displayed in reaction to the digital materials' subject. Future research suggested increasing sample size, exploring different emotions and additional brain signal for employing other models than SVM.

Significance of all the studied works using different models and feature extraction method of machine learning and deep learning domains, using different physiological signals in few papers and mapping them with ML models, studying different application uses such as game play, museum visiting give us a spectrum of use cases and improvement opportunities. Limitation of few of the works were using basic ML methodology giving comparatively higher results than variant technologies experiment in rest other work using automatic feature selection, quadratic discriminant classifier, Hjorth-feature still giving accuracy of 59.06%, 66.66% , 69% accuracy respectively. Using different methods like fast correlation-based filter (FCBF) and Analysis of variance (ANOVA) for peripheral and EEG signal and got result of 59% and 56 % respectively, fusion of signals got 63%. There are many limitations to their work mainly being not enough sample space and not utilizing enough physiological methods. Using VAE got similar approach to us that extract latent factors. They used 2 approaches independent and VAE approach and VAE can also detect hidden state of brain which is limitation from our end. Significance of our work is we are using similar encoder named stacked auto encoder is deriving EEG channel correlations from source, using linear mixing model to emulate the timing of emotions and improved computational efficiency for better result.

A. *Research Gap*

Augment the quantity of EEG signals pertaining to emotions and explore more effective feature extraction methods and optimize the current methods. The studied researches were unable to categorize all distress emotions with precision and was just dealing majorly with basic emotion palette. No use of Spatial and temporal data of EEG features was seen in literature review.

3. METHODOLOGY

Figure 8 shows the brief architecture of our model while Figure 10 is framework design of our work comprising of linear mixing model for signal decomposition, feature extraction and digital processing, and LSTM model for emotion classification.

In this section the main contribution is employing LMM with stacked auto encoder, timing emotions precisely using LSTM model resulting in improved computational efficiency in emotion detection. Data used in this study is DREAMER dataset and for further validation we used DEAP dataset which are available for research purposes. Figure 10 further helps in better understanding of our framework. Briefly explaining the framework in Figure 10, concepts related to emotion detection in neural networks and signal processing are visualized in this figure. The auto encoder helps with dimensionality reduction and unsupervised learning, while the linear mixing model merges sources for a mixed signal. Frequency Bands classify signals, and the Hanning Window minimizes spectrum leakage. Signal intensity is quantified using feature band power, and the power content to frequency ratio is represented by power spectrum density. Power Spectral Density is calculated using the Welch Method, and LSTM is a recurrent neural network that handles time-dependent data in emotion detection models.

The details of how it is achieved is described step by step further in this section.

A. *Data Acquisition*

- *Collect EEG Signal*

We acquired the multi-modal DREAMER dataset from zendo [21] in aspects of valence, arousal, and dominance for physiological signals recorded with audio-visual stimuli for 25 participants using electrodes [22]. For EEG and ECG, the Emotiv EPOC wireless EEG headset and the Shimmer2 ECG sensor were used [22]. These wireless wearables are easily adaptable to regular use. Figure 9 shows details considered for collecting data for DREAMER dataset.

Audio-visual stimuli	
Number of videos	18
Video content	Audio-Video
Video duration	65 - 393 s (*M*=199 s)
Experiment information	
Number of participants	25 (23)
Number of males	14 (14)
Number of females	11 (9)
Age of participants	22 - 33 (*M*=26.6, *SD*=2.7)
Rating scales	Arousal, Valence, Dominance
Rating values	1 - 5
Recorded signals	14-channel 128Hz EEG, 256Hz ECG

Figure 9. DREAMER dataset details

B. *Pre-processing*

- *Linear Mixing Model (LMM)*

The EEG signals obtained are a blend of source signals. Other studies and research presented the LMM, which is extensively used in medical fields, to stimulate the mixing process. The aim is to design an encoder that will recover the source signals from the acquired data. For optimization, an auto-encoder with numerous layers, known as a stacked autoencoder, is utilized. The purpose of building LMM based on auto encoder is because LMM and auto encoder have comparable expressions for linear activation function. Mini-batch-gradient-descent (MB-GD) was used to step up this technique and boost it while training LMM [23]. MB-GD randomly selected small chunks of data records to compute the loss gradient at every step, resulting in an impressive convergence speed and processing efficiency. The loss function used was mean square error (MSE). The model is then validated by feeding the training data into it and determining the corrected R-squared between the test and generated data. When the modified R-square surpassed 0.9, it signified that our model's encoder nearly maintained all of the information from the raw EEG signals. It accurately depicted the EEG source signals, demonstrating the effectiveness of the decomposition process.

C. *Feature Equation*

- *Frequency Band Power Extraction*

To streamline signal processing on these signals, they are consistently partitioned into brief time frames, with an underlying assumption of signal stationarity [24]. Where the EEG source signals are divided into 125 frames of data

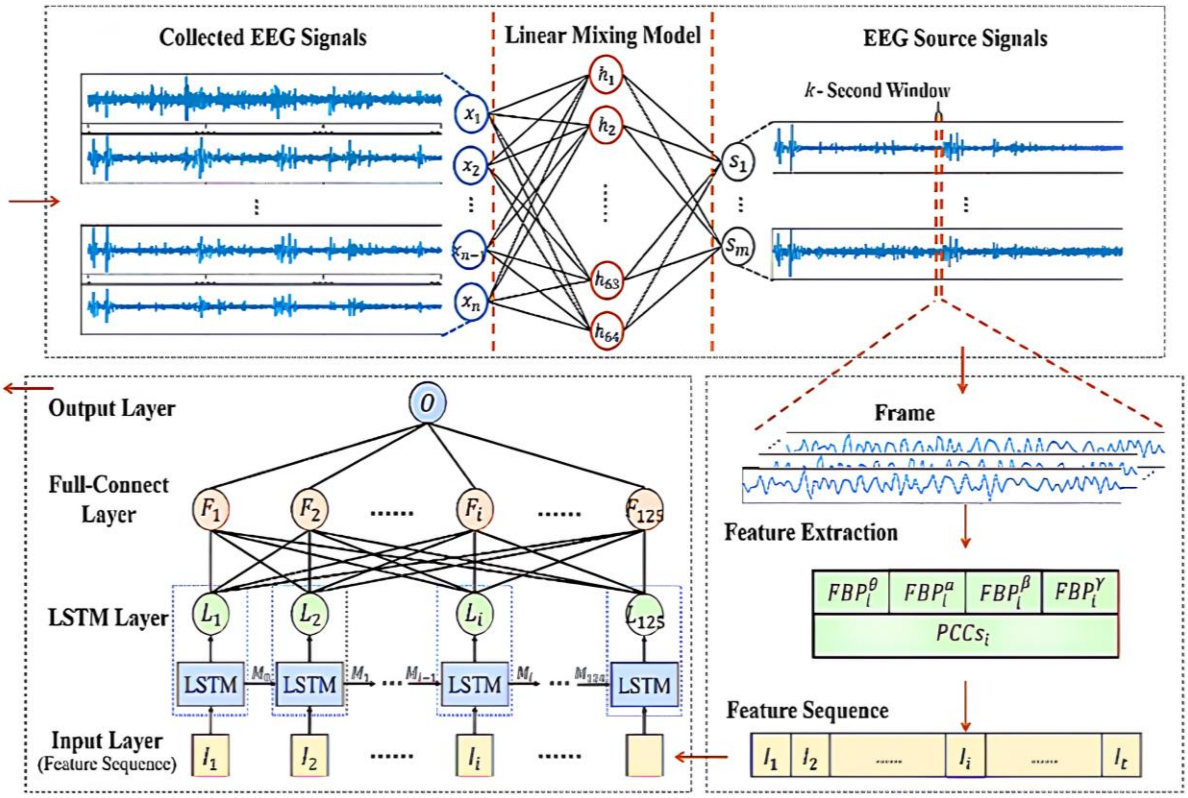

Figure 10. Framework network

by using a 1-second window is split into two parts with 50% of the window being overlapped [23]. Although we tried different window sizes, we selected the specified one because we can collect more data by overlapping the 1s frames rather than 2s, 3s or 5s since neural network requires more data for better training. Following the separation of the signal into frames, the EEG-based features from each frame are identified and arranged into an array. Eventually, EEG feature array containing 125-frames are acquired.

EEG signals are classified into five frequency bands based on their frequency range from low to high: Delta (δ), theta (θ), alpha (α), beta (β), and gamma (γ) are the frequency bands ranging from 0.5 to 50 Hz respectively [23]. For reducing ripple each EEG channel was given the Hanning window to determine frequency spectrum of original signal and welch period was used for the power spectral density (PSD) [17]. In our investigation, four frequency band powers (FBP) of these EEG signals were chosen. The lowest band power of frequency range is ignored our experiment because delta band is associated with deep sleep or unconsciousness. Data segmentation and FBP extraction was also carried out using Hanning window, resulting in low spectrum leakage [23].

For inter-channel correlations, the Pearson correlation coefficient (PCC) is used linearly to measure correlation between the two signals. One of the frame signals is chosen as the reference signal, and the PCC between reference and frame signals may be determined using (1) [25] as shown in Figure 11.

$$PCC = \frac{\sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N} (x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{N} (y_i - \bar{y})^2}} \quad (1)$$

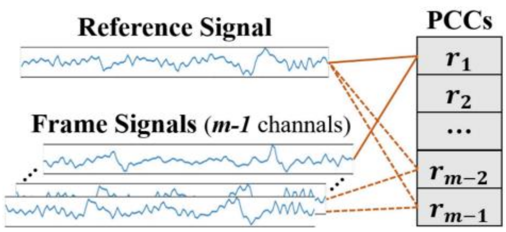

Figure 11. PCC for inter channel correlations

In this step, stress bins were set up for individual users to assess their degree of stress and calmness following exposure to external stimuli. We created new data frames for participant affecting factors like sex, age, valence level, arousal and dominance level, etc. [16] for each participant and classified the targeted emotions in classes as 1,2 and 3 as 'surprise': 0, 'excitement': 0, 'amusement': 0, 'happiness': 0, 'fear': 1, 'anger': 1, 'sadness':2, 'disgust':2, 'calmness': 3 [3]. Stress bins coordinates targeted emotions to respective classes.

D. Classifier

- *LSTM – RNN Model*

The emotion categorization is performed using a LSTM, which is improved and reinforced version of Recurrent Neural Network (RNN). RNN suffers from long-term dependencies, making it unsuitable for time series analysis, but LSTM can overcome this using module repetition. The LSTM layer is the initial layer of the deep neural network's architecture, and it is used to evaluate the connections between the EEG characteristics in the sequence of signal inputs. The next layer, 2nd layer, is the classifier's primary function which is used to compile data. Each sequence's 125 frame attributes are allocated to respective LSTM cells.

Connection units are assigned the same number in the full-connect layer. Sigmoid was used at output layer as activation function. The classifier training was conducted using the MB-GD optimizer and the MSE (Mean Squared Error) was applied as loss function, to minimize over-fitting, "dropout" was introduced to the LSTM and full-connect layers [23]. The training of the model was conducted over a few thousand epochs. For training acceleration, learning rate was initially set to a high value, then reducing it to produce more reliable results. The training was finished when the AUC attained the predetermined objective.

The evaluation matrix used for methodology is AUC (Area Under Curve) and Average recognition accuracy is determined using equation (2) where N_{test} is test samples and $N_{correct}$ is correctly classified samples.

$$AUC_{mean} = \frac{1}{10}\sum_{k=1}^{10}\frac{N^k_{correct}}{N_{test}} \qquad (2)$$

4. Results And Discussion

1) Result validation using DEAP Dataset

a) Validation Dataset

We validated our methodology with EEG data from the DEAP dataset. DEAP is a database that analyses human emotional states by utilizing many types of physiological inputs. It includes 32 participants, multichannel (32) EEG and 8 input peripheral channels [26]. Each participant had to watch one-minute video clips while their data were monitored. Participants assigned ratings of 1 to 9 to each movie based on valence, arousal, dominance, and pleasure [27]. The DEAP database's EEG signals were down sampled to 128 Hertz and using filter called band-pass-filter frequency is kept in range of 4 to 45 Hertz [28]. Since Electrooculography (EOG) can corrupt and mix with EEG signals, Independent Component Analysis (ICA) removed the EOG noise in the DEAP dataset [23]. We separated the dataset into two classes based on the valence (or arousal) value, labelling them "High" if >5.5 and "Low" if < 4.5. We have attribute values of valance, arousal, dominance, and calculated class value. We calculated high valance and arousal according to high>5.5 and we have mapped those emotion classes and feature into stress bins. Valence-arousal-dataset was formed and down sampling was done to balance data forming new attributes 'valance-high' and 'arousal-high' [23].

b) Visualization

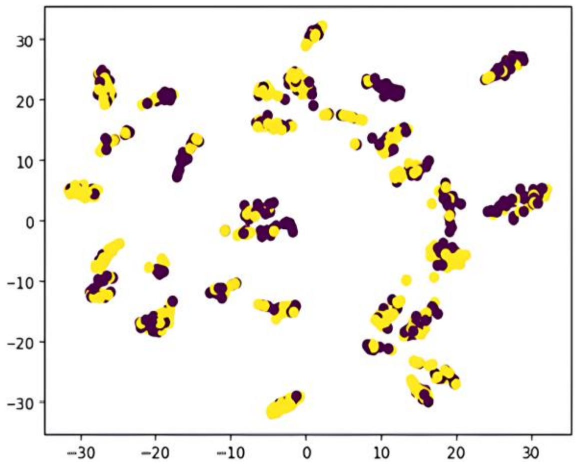

Figure 12: (a)

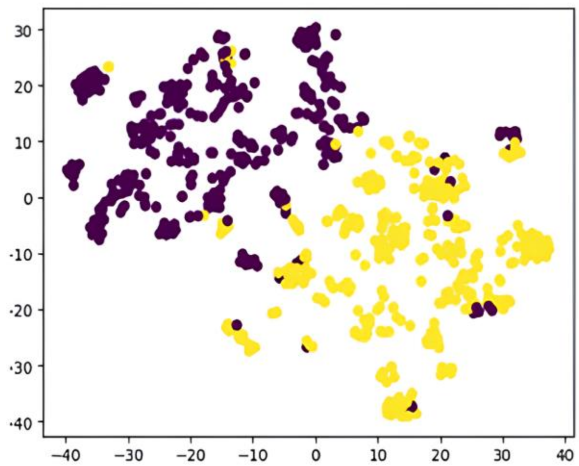

Figure 12: (b)

Figure 12: Feature clustering before and after completion of training
(a) Before Training (b) After Training

Eustress and distress have been grouped into two distinct categories, each denoted by a different color, as illustrated in Figures 12(a) and 12(b). The positive emotion category is visually represented by the color yellow, while the negative emotion category is denoted by the color purple.

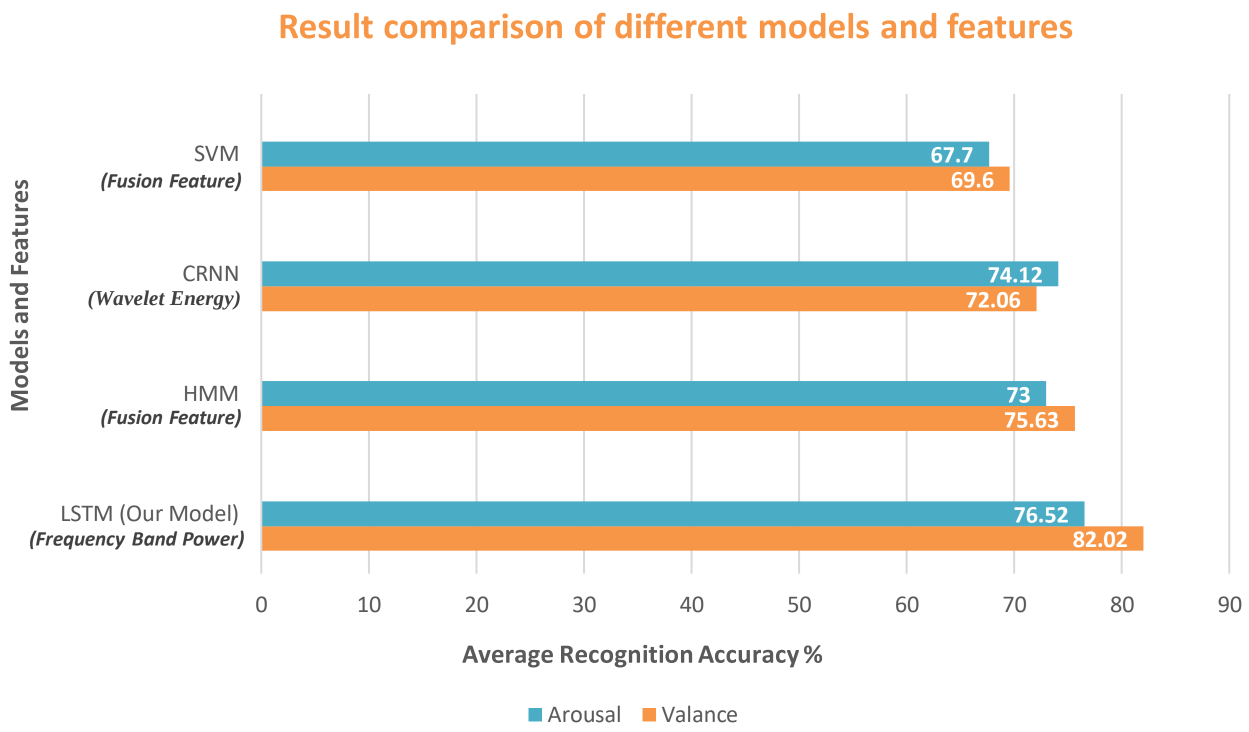

Figure 13. Model Comparison

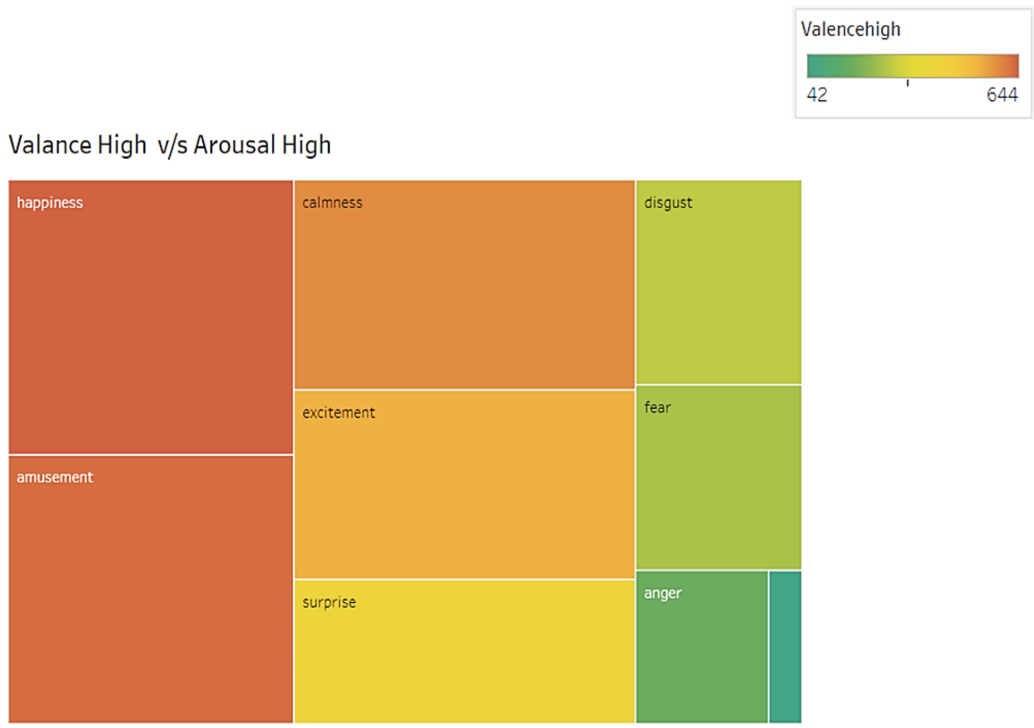

Figure 14: Tree map showing valance high and arousal high by target emotions

International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 15, No.1 (Feb-24)

Figure 15: Charts showing valance high and arousal high by gender and target emotions

2) Results

We validated our classifier outcomes using the 10-fold cross validation approach. Our model framework has been found to be the most effective in recognizing emotions based on valence (82.03%) and arousal (76.52%) when compared to other approaches as depicted in above graph in Figure 13 where the first axis title on Y axis is model and second title below it in bold italic is feature method used. Figure 13 shows comparison of different models and features. It compares SVM, convolutional recurrent neural network (CRNN), hidden Markov model (HMM) and LSTM (our model) with the respective model features.

We also used Tableau for further visual analysis. Figure 14 shows 'valance-high' and 'arousal-high' when they are the highest across all target emotions. The derivation of new attributes, namely 'valance-high' and 'arousal-high' is detailed in section 4. I. (a). The categorization of the the target emotions is done using a tree map and marks card 'color' is added for 'valance-high'. The same graph can work as heat map for valance and tree map for arousal. Color scale for heat map is given on top. As concluded from the graph, emotion with boxes having darker shade of red has higher 'valance-high' transitioning to greener side indicating darker pigment of green lower the 'valance-high' value. Similar to this, for arousal the size of box decides the value of 'arousal-high'. Larger the box higher the 'arousal-high' value. Hence, happiness have high 'valance-high' and high 'arousal-high' while surprised and excitement have lower value of 'valance-high' and higher value of 'arousal-high' in comparison, lastly anger has lowest 'valance-high' and 'arousal-high'.

A comparison of attributes 'valance-high' and 'arousal-high' by gender and target emotions is depicted in Figure 15. Initially, the bar graph shows reading of male and female for 'valance-high' and 'arousal-high' to get a generic idea of participants valance-high and arousal-high values based on gender. Then pie chart was made, since we are more focused on distress emotions we put a filter for target emotion in tableau and showcase distress emotion 'valance-high' and 'arousal-high' value comparing for both genders. As the filter chart on the most right shows, bigger the size of circle, higher the value of 'valance-high'. Similarly, for arousal aspect, pie chart shows that bigger the quadrant/angle of pie higher the 'arousal high'. The number written along the pie are value of 'arousal-high'. Gender is visualised by color. Blue is female and orange is male. The reading of the chart highlighted the values of target emotion 'saddess' in the box in middle.

This might be attributed to the following factors:

a) HMM- Since each phase output of HMM was only connected to a subset of prior states, the model classifier could not learn automatically, like LSTM.

b) CRNN- CNN needs an immense training to learn and extract, which the DEAP dataset does not provide.

c) SVM- The SVM model's performance was limited since it was unable to analyze the contextual relationships of the EEG feature array.

Our valance categorization model accuracy (82.03%) was higher than that of another research.

The significant aspect of this framework was that a linear EEG mixing model efficiently used spatio-temporal data of EEG signals. An emotion timing model (ETM) based on LSTM considerably gains valence classification while having no effect on arousal categorization.

3) Discussion

This study holds significance in achieving elevated EEG-based emotion recognition rates, specifically attaining 82.03% for valence and 76.52% for arousal. Though with today's weak AI we cannot predict the temporal behavior of humans but we might be able to find the root cause, hence root cause exploration adds more significance to work. In addition to this, we have more improvement opportunities in hand as 76.52% accuracy for arousal can be enhanced and results for valance can be improved by using better and complex neural network model. However, major limitation in our study is dataset size constraint, use larger data for better training of model as the result is insufficient for real life model application. Also, one more limitation to our work is our objective was to use just EEG data but for future work we can use other physiological measures as we are currently working on single measure i.e., EEG. Future goal might also include working with specialized demographic of the participant which will expand the schema and was not considered in our general study on this topic, but is an important factor to look into as it might affect results.

5. Conclusion And Future Work

We introduce a unique and novel method of emotion recognition and the framework network system based on a LMM and an ETM in this research [23]. Our key contribution in this study is LMM using auto encoder and LSTM model. To dissect EEG source inputs and uncover EEG channel correlations, one can employ the linear EEG mixing model. This approach not only facilitates the breakdown of EEG sources but also enhances computational efficiency in feature extraction, thereby improving emotion detection performance [29]. A LSTM system is employed to emulate the timing of emotions, thereby increasing the precision of EEG feature sequence context recognition. Our framework's efficiency was confirmed by the comparative outcomes in our experiment. Our model framework has been found to be the most effective in recognizing emotions based on valence (82.03%) and arousal (76.52%) when compared to other approaches. Yet, our model has few limitations as explained in discussion section. We hope to work on those in future work.

As a future goal, we can add specialized demographics, physiological measures, larger records and mentioned above in discussion section. Furthermore, our theory is that since there exists no prototype as such yet, a new wearable sensory device should be devised, which can collect and comprehend the user's data in real time, even for leisure activities like hobbies, all around the day-during which it can identify and maintain the negative emotions by actively recommending activities customized for the user; thus, countering detected distress with positive emotions. There is no proof to this but we can research further in this to get to know this better because if successfully implemented it, there will be a major breakthrough in this domain.

Acknowledgment

The authors express gratitude to Dr. Vishwanath Karad of MIT World Peace University, Pune, for generously providing the resources and support essential for the completion of this research. Special appreciation is extended to Zendo and Queen Mary University of London for granting access to the DREAMER and DEAP datasets, respectively. These datasets proved instrumental in conducting the research, significantly enhancing the quality of the work. Additionally, the authors would like to convey their thanks to the reviewers and editors for their valuable feedback, which contributed to the enhancement of the paper's overall quality.

References

[1] M. Soleymani, S. Asghari-Esfeden, M. Pantic and Y. Fu, "Continuous emotion detection using EEG signals and facial expressions," 2014 IEEE International Conference on Multimedia and Expo (ICME), Chengdu, China, 2014, pp. 1-6, doi: 10.1109/ICME.2014.6890301

[2] "How has coronavirus affected mental health?," *BBC News*. https://www.bbc.com/news/explainers-55957105

[3] Shah, D., Rane, R. (2023). Emotion Recognition Through Physiological Signals and Brain Sensing. In: Ranganathan, G., Papakostas, G.A., Rocha, Á. (eds) Inventive Communication and Computational Technologies. ICICCT 2023. Lecture Notes in Networks and Systems, vol 757. Springer, Singapore. https://doi.org/10.1007/978-981-99-5166-6_55

[4] Xing X, Li Z, Xu T, Shu L, Hu B and Xu X (2019) SAE+LSTM: A New Framework for Emotion Recognition From Multi-Channel EEG. Front. Neurorobot. 13:37. doi: 10.3389/fnbot.2019.00037[1]S. Koelstra, "DEAP: A Dataset for Emotion Analysis using Physiological and Audiovisual Signals," *DEAP: A Dataset for Emotion Analysis using Physiological and Audiovisual Signals*. http://www.eecs.qmul.ac.uk/mmv/datasets/deap/index.html

[5] https://www.yahoo.com/news/opinion-slow-motion-crisis-gen-190100633.html

[6] Pichara, Karim & Soto, Alvaro. (2011). Active learning and subspace clustering for anomaly detection. Intell. Data Anal.. 15. 151-171. 10.3233/IDA-2010-0461.

[7] Abdulrahman, A., Baykara, M. (2021). A comprehensive review for emotion detection based on EEG signals: Challenges, applications, and open issues. Traitement du Signal, Vol. 38, No. 4, pp. 1189-1200. https://doi.org/10.18280/ts.380430

[8] I. Mazumder, "An Analytical Approach of EEG Analysis for Emotion Recognition," 2019 Devices for Integrated Circuit (DevIC), 2019, pp. 256-260, doi: 10.1109/DEVIC.2019.8783331.

[9] B. Shirke, J. Wong, J. C. Libut, K. George and S. J. Oh, "Brain-IoT based Emotion Recognition System," 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), 2020, pp. 0991-0995, doi: 10.1109/CCWC47524.2020.9031124.

[10] Z. Khalili and M. H. Moradi, "Emotion recognition system using brain and peripheral signals: Using correlation dimension to improve the results of EEG," 2009 International Joint Conference on Neural Networks, Atlanta, GA, USA, 2009, pp. 1571-1575, doi: 10.1109/IJCNN.2009.5178854.

[11] Raja Majid Mehmood, Muhammad Bilal, S. Vimal, Seong-Whan Lee, "EEG-based affective state recognition from human brain signals by using Hjorth-activity", Measurement, Volume 202, 2022, 111738, ISSN 0263-2241, doi: 10.1016/j.measurement.2022.111738.

[12] G. Chanel, C. Rebetez, M. Bétrancourt and T. Pun, "Emotion Assessment From Physiological Signals for Adaptation of Game Difficulty," in IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, vol. 41, no. 6, pp. 1052-1063, Nov. 2011, doi: 10.1109/TSMCA.2011.2116000

[13] Li X, Zhao Z, Song D, Zhang Y, Pan J, Wu L, Huo J, Niu C and Wang D (2020) Latent Factor Decoding of Multi-Channel EEG for Emotion Recognition Through Autoencoder-Like Neural Networks. Front. Neurosci. 14:87. doi: 10.3389/fnins.2020.00087

[14] Bagherzadeh S, Maghooli K, Shalbaf A, Maghsoudi A. Emotion Recognition Using Continuous Wavelet Transform and Ensemble of Convolutional Neural Networks through Transfer Learning from Electroencephalogram Signal. Frontiers Biomed Technol. 2022;10(1):47-56.

[15] K. Guo, H. Yu, R. Chai, H. Nguyen and S. W. Su, "A Hybrid Physiological Approach of Emotional Reaction Detection Using Combined FCM and SVM Classifier," 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 7088-7091, doi: 10.1109/EMBC.2019.8857698.

[16] Han Z, Chang H, Zhou X, Wang J, Wang L and Shao Y (2022) E2ENNet: An end-to-end neural network for emotional brain-computer interface. Front. Comput. Neurosci. 16:942979. doi: 10.3389/fncom.2022.942979

[17] https://link.springer.com/book/10.1007/978-3-030-10374-3#toc

[18] Zhen Liang, Xihao Zhang, Rushuang Zhou, Li Zhang, Linling Li, Gan Huang, and Zhiguo Zhang. 2022. Cross-individual affective detection using EEG signals with audio-visual embedding. Neurocomput. 510, C (Oct 2022), 107–121. https://doi.org/10.1016/j.neucom.2022.09.078

[19] https://tesis.pucp.edu.pe/repositorio/

[20] Zhang T, Qi Z, Guan W, Zhang C and Jin D (2023) Research on the measurement and characteristics of museum visitors' emotions under digital technology environment. Front. Hum. Neurosci. 17:1251241. doi: 10.3389/fnhum.2023.1251241

[21] DREAMER dataset https://zenodo.org/record/546113#.Y3_nC3ZBy3A

[22] DREAMER S. Katsigiannis, N. Ramzan, "DREAMER: A Database for Emotion Recognition Through EEG and ECG Signals from Wireless Low-cost Off-the-Shelf Devices," IEEE Journal of Biomedical and Health Informatics, 2017. In press. doi: 10.1109/JBHI.2017.2688239

[23] Xing X, Li Z, Xu T, Shu L, Hu B and Xu X (2019) SAE+LSTM: A New Framework for Emotion Recognition From Multi-Channel EEG. Front. Neurorobot. 13:37. doi: 10.3389/fnbot.2019.00037[1]S. Koelstra, "DEAP: A Dataset for Emotion Analysis using Physiological and Audiovisual Signals," *DEAP: A Dataset for Emotion Analysis using Physiological and Audiovisual Signals*. http://www.eecs.qmul.ac.uk/mmv/datasets/deap/index.html

[24] Bonita, J. D., Ambolode, L. C. C. II., Rosenberg, B. M., Cellucci, C. J., Watanabe, T. A. A., Rapp, P. E., et al. (2014). Time domain measures of inter-channel eeg correlations: a comparison of linear, nonparametric and nonlinear measures. Cogn. Neurodyn. 8, 1–15. doi: 10.1007/s11571-013-9267-8[1]S. Narkhede, "Understanding AUC - ROC Curve," Medium, Jun. 15, 2021. https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5

[25] Jang, Eun-Hye, Byoung-Jun Park, Sang-Hyebo Kim, Myoung-Ae Chung, Mi-Sook Park and Jin-Hun Sohn. "Classification of Three Negative Emotions based on Physiological Signals." International Conference on Interaction Sciences (2013).

[26] "DEAP: A Database for Emotion Analysis using Physiological Signals", S. Koelstra, C. Muehl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras, IEEE Transactions on Affective Computing, Special Issue on Naturalistic Affect Resources for System Building and Evaluation, in press.

[27] Li X, Song D, Zhang P, Zhang Y, Hou Y and Hu B (2018) Exploring EEG Features in Cross-Subject Emotion Recognition. Front. Neurosci. 12:162. doi: 10.3389/fnins.2018.00162

[28] Shu L, Xie J, Yang M, Li Z, Li Z, Liao D, Xu X, Yang X. A Review of Emotion Recognition Using Physiological Signals. Sensors (Basel). 2018 Jun 28;18(7):2074. doi: 10.3390/s18072074. PMID: 29958457; PMCID: PMC6069143.

[29] Cai J, Xiao R, Cui W, Zhang S and Liu G (2021) Application of Electroencephalography-Based Machine Learning in Emotion Recognition: A Review. Front. Syst. Neurosci. 15:729707. doi: 10.3389/fnsys.2021.729707

Disha Shah is a Master's student at Dr. Vishwanath Karad MIT World Peace University in Pune, pursuing M. Tech. in Data Science and Analytics. She is the study's corresponding author. She has another review paper in the same domain that is in publication, and she has completed several projects in data analysis and machine learning working with multiple tool skills and certifications.

Rashmi Rane is postgraduate in Computer Science and Engineering, 2009 and now is pursuing a Ph.D. in Machine Learning at Sathyabama University in Chennai. She is an ISTE life member. She has 22 years of teaching experience. She has various national and international publications to her name. Her research interests are Social Data Mining, Deep Learning.

Shakti Kinger is a post graduate from Dr. Vishwanath Karad MIT World Peace University, 2012. She presently works as an Assistant Professor and an active member of the Industry Collaboration Committee and the SPOCs of several IT businesses. Her research interests include artificial intelligence, deep learning, and network security. She has authored national and international articles