A Long Short Term Memory Deep Learning Network for the Classification of Negative Emotions Using EEG Signals

1st Divya Acharya
Computer Science Engineering
Bennett University
Greater Noida, India
divyaacharya099@gmail.com

2nd Shivani Goel
Computer Science Engineering
Bennett University
Greater Noida, India
shivani.goel@bennett.edu.in

3rd Harshit Bhardwaj
Computer Science Engineering
Gautam Buddha University
Greater Noida, India
harshitbhardwaj15@gmail.com

4th Aditi Sakalle
Computer Science Engineering
Gautam Buddha University
Greater Noida, India
aditi.sakalle@gmail.com

5th Arpit Bhardwaj
Computer Science Engineering
Bennett University
Greater Noida, India
arpit.bhardwaj@bennett.edu.in

***Abstract*—In cognitive science and human-computer interaction, automatic human emotion recognition using physiological stimuli is a key technology. This research considers classification of negative emotions using EEG signals in response to emotional clips. This paper introduces a long short term memory deep learning (LSTM) network to recognize emotions using EEG signals. The primary goal of this approach is to assess the classification performance of the LSTM model for classifying emotions. The secondary goal is to assess the human behavior of different age groups and genders. We have compared the performance of Multilayer Perceptron (MLP), K-nearest neighbors (KNN), Support Vector Machine (SVM), Deep Belief Network based SVM (DBN-SVM), and LSTM based deep learning model for classification of negative emotions using brain signals. The analysis shows that for four class of negative emotion recognition LSTM based deep learning model provides classification accuracy as 81.63%, 84.64%, 89.73%, and 92.84% for 50-50, 60-40, 70-30, and 10-fold cross-validation. Generalizability and reliability of this approach is evaluated by applying our approach to publicly available EEG dataset DEAP and SEED. In compliance with the self-reported feelings, brain signals of 26-35 years of age group provided the highest emotional identification. Among genders, females are more emotionally active as compared to males.**

***Index Terms*—Emotion Recognition, Deep learning, EEG, Fast Fourier Transformation, LSTM.**

I. Introduction

Emotion recognition is growing as an evolving research area. Emotions have a significant impact on human life as it impinges their psychological and physiological status [17]. Emotion is defined as the response to stimulation lasting for seconds or minutes as a result of that person experiences certain feelings during that time. Emotions are integral parts of human behavior, which are divided into two broad categories [2], i.e., positive emotions and negative emotions. Positive emotions help in improving the quality and health of human life, while negative emotions directly impact the health and reasoning capability of humans. Negative emotions are also an influential factor in causing many mental health problems. Mental health issues like depression, stress, and anxiety are the result of the amassing of negative emotions for an extended period, which can even lead to self-destruction in many cases. Approximately 89% of the inhabitants in India report emotional instability as a contrast to the worldwide average of 86% [10], [18], [15]. Thus, this brings forth the need to develop a new accurate emotion recognition approach that recognizes negative emotions as the recognition of emotion is critical and complex, which helps in improving the quality of human life. In this research work, we are proposing an LSTM based deep learning network for the classification of negative emotion to recognize human emotions using EEG signals.

To show the supremacy of our LSTM based deep learning network, we also analyzed it over the two publicly available benchmark dataset DEAP [12] and SEED [19] and compare our work with other state-of-the-art methods as shown in Table VII. Fast Fourier Transformation is used for feature extraction. However, considering the restricting elements of this study, it is essential to notice that only the FFT feature extraction technique is considered. There are other feature extraction techniques like Empirical Mode Decomposition (EMD), Ant Colony Optimization, and Particle Swarm Optimization, which can be considered. To illustrate the dominance of our method, the partition scheme used is 50-50, 60-40, 70-30, and 10-fold cross-validation methods for the confusion matrix. In terms of the social behaviour analysis age group based response to emotions, and gender-based responsiveness is analyzed. The contribution of this paper can be summarized as below:

- A long short term memory deep learning network (LSTM) to recognize negative emotions using EEG sig-

978-1-7281-6926-2/20/$31.00 ©2020 IEEE

TABLE I
SAMPLE DESCRIPTION OF STIMULI FILM CLIPS

Genre	Film Name	Length (sec)	Clip Content
Sadness	Irreplaceable You	72	A girl was diagnosed with cancer
Fear	The Conjuring	60	Unusual events happening in night
Anger	Aftermath	92	A morgue worker horrible behavior
Surprise	Final Destination	80	Teenager having future vision of dying

nals.

- A new EEG signal dataset for emotional clips is created with a portable,single-channel EEG headset (NeuroSky MindWave 2).
- An gender and age group based human behavior analysis is performed for responsiveness of emotions.

II. RELATED WORK

Different methods are used in literature to recognize emotions. In general, emotion recognition methods are broadly classified into three categories. First, clinical method or subjective measures that contain self-assessment manikin (SAM) forms [9] in which emotion is recognized using a questionnaire in which one has to rate the level of emotion like very, moderately to extremely. All conventional emotion recognition models are explicit time taking process. Second, the method recognizes emotions based on physical measures. Physical measure includes images [16], speech and audio [3]. Although, this approach suffers from low accuracy of recognition. The third method is based on using physiological signals to recognize emotion.

In this category, signals are recorded from brain using electroencephalograph (EEG) [6], [14]. Electroencephalogram (EEG) signals are used to record the electrical activity of brain and also, to analyze the functional changes in the brain [5]. EEG is believed to be the finest way of recording data in various modalities owing to its distinctive elements when interacting with emotional conditions [7].

The main challenge in the field of emotion recognition is controlling of physical response and tampering of subjective measures recorded to analyze emotions. The process used to analyze emotions in the literature suffers from low accuracy and data reliability issues. Also, the clinical method used to recognize emotion is a very time consuming explicit process, which leads to the need for automating the process of emotion recognition to improve the quality of human life. To overcome this issue, in this research work, EEG based analysis is performed to study the response of the brain to emotional stimuli [1]. This also helps in achieving the right level of accuracy in the classification of emotions in this study as it is impossible to falsify brain activity.

In this research work, movie clips are used as stimuli. Movie clips targeting specific emotions have the ability to sway emotional and behavioral changes, which can help in the treatment of any mental health issues in human beings using emotion elicitation therapy. Thus, in this research work, movie clips are used as an emotional stimulus which is easy to implement, economical, and non-invasive tool in recognizing negative emotions which can lead to mental health issues.

TABLE II
PARAMETER VALUES USED TO CREATE THE LSTM MODEL

Parameter	Value
Optimizer	Adam
Learning Rate	0.001
Loss Function	Categorical cross entropy
Metrics	Accuracy
Batch size	10
Epochs	100

TABLE III
PERFORMANCE COMPARISON OF THREE DIFFERENT LSTM ARCHITECTURES FOR PROPOSED FRAMEWORK

	Average Accuracy(%)		
Validation Techniques	**LSTM_1**	**LSTM_2**	**LSTM_3**
50-50	37.29	59.41	78.67
60-40	40.92	61.84	82.85
70-30	45.32	64.77	86.30
10- fold CV	49.64	66.86	89.83

III. METHODOLOGY

This section describes the statutory background for negative emotion classification using LSTM based deep learning network. All the blocks for negative emotion classification are described next.

A. Stimuli

In this research work, we have used 40 English language movie clips as elicitation material that lasted 1 to 2 minutes and contained independent and integrated content to elicit target emotion. The stimuli database is targeting four class of emotions, i.e., sadness, fear, anger, and surprise emotions taken from IMDb. Sample description of clips in each emotion category is shown in Table I.

B. Subjects

Forty five healthy subjects pool was considered for this research work. Participants has participated voluntarily in this research study and also signed a informed consent form. All participants are from different culture and educational background divided into three age group 18-25 years, 26-35 years, and 36-55 years. The principle of Helsinki declaration is followed.

C. Dataset Description

In this research work, we have used 2 publicly available EEG dataset DEAP [12], SEED [19], and the third one is our generated dataset. The description of our dataset is given below:

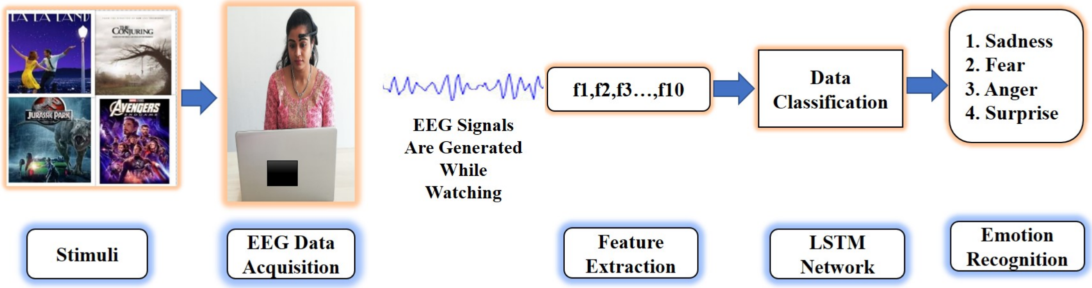

Fig. 1. EEG based emotion recognition framework using LSTM network

Algorithm 1 Algorithm for Negative Emotion Classification Using LSTM based Deep Learning Network

Input Brain signal based EEG training dataset and LSTM network parameter see Table II.
Output Emotion Recognition for four class of negative emotions.
Begin
Initialization Inform consent by subjects
for all watch emotions targeting stimuli **do**
  EEG device setup
  while subject watch emotional clips **do**
    brain signal based EEG data acquisition
  end while
end for
for all feature extraction **do**
  Apply FFT technique on brain signal based EEG dataset
  Get resultant features of dataset as f1,f2,f3,....f10
  Pass extracted features to LSTM network classifier
end for
Return: Recognize four class of negative emotion.

TABLE IV
CONFUSION MATRIX FOR LSTM_3 MODEL

Validation Technique	Classifier-LSTM_3				
		Sadness	Fear	Anger	Surprise
50-50	Sadness	2300	200	250	250
	Fear	200	2450	300	50
	Anger	50	300	1050	100
	Surprise	50	100	70	1280
		Sadness	Fear	Anger	Surprise
60-40	Sadness	2310	200	240	250
	Fear	70	1625	75	30
	Anger	50	60	950	40
	Surprise	20	80	120	1080
		Sadness	Fear	Anger	Surprise
70-30	Sadness	1600	130	110	60
	Fear	40	1430	50	80
	Anger	50	60	930	60
	Surprise	50	35	15	700
		Sadness	Fear	Anger	Surprise
10-fold CV	Sadness	675	10	05	10
	Fear	20	395	30	55
	Anger	10	05	378	07
	Surprise	11	09	10	170

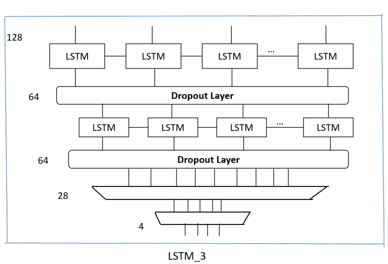

Fig. 2. LSTM_3 network architecture

1) Generated Dataset: The experiment was performed in a soundproof lab, and the subject has to switch off all Bluetooth and wireless devices to prevent possible intrusion. The dataset was prepared by taking brainwave samples of participants. Every subject was first made comfortable with wearing the NeuroSky MindWave 2 headset. Physical components include flexible rubber sensor arms, rounded forehead sensor tip, T-shaped headband, and ear clip contacts. The headset's reference and ground electrodes are on the ear clip, and the EEG electrode is on the sensor arm, resting on the forehead above the eye (FP1 position). It uses the TGAM1 module. For pairing purposes, it has a Static Headset ID. It outputs 12-bit Raw-Brainwaves (3Hz-100Hz) with a sampling rate of 512 Hz and outputs EEG power spectrum in different frequency and morphology bands. Total of 50 sessions were experimented in one set of the experiment for subjects. For one participant, this process is iterated four times. To start a session, the hint of start is given 5s before each movie clip. After these subjects were then shown four clips from the list of stimuli dataset that contains sadness, fear, anger, and surprise emotion targeting clips as shown in Table I. All the film clips were adjusted to the same resolution (720 x 576). They were presented randomly on a 15-inch laptop screen. The volume was adapted to a suitable rate to make the participants comfortable. They had eyes about 0.6 meters away from the middle of the display.

After viewing every movie clip, we gathered the participant's emotions and reactions in the form of self-assessment form which was on a 10-point scale (1 = "not at all" 10 = "extremely") to assess the intensity of each self-reported emotions and 40s time is devoted to it. A rest of 10s is given after every stimuli. The proposed system is using the standardized database of movie-induced emotions as the ground truth to label the EEG signals. Since our model used the standardized database of emotions induced by movies as the ground truth for labeling EEG signals we compared the self-assessments of each participant with the standardized database and discarded all cases in which the target emotion was not at least one point greater than other non-target emotions. The EEG based emotion recognition framework using LSTM based deep learning network model framework is shown in Figure 1.

The participants were encouraged to answer all the questions based on their true feelings when watching each film excerpt, instead of their familiar feelings or general mood. The test data was collected by showing a single clip targeting one emotion to the participants such that we can measure the accuracy of our model. Every clip shown generates raw EEG signal outputs. The raw signals collected are the 10 features that are subjected to FFT by collecting data from NeuroSky Mindwave 2.

D. Feature Extraction-Fast Fourier Transformation

Extracting the more prominent statistical features from the EEG signal is the second step after signals are pre-processed for efficiently classifying the emotions. EEG signals are highly complex and non-linear. The most widely used method to analyze EEG data is to decompose the signal into functionally distinct frequency bands. The on board digital signal processing module is embedded with algorithms to perform signal decomposition using Fast Fourier Transform (FFT). If N samples are present, FFT takes only N*log(N) operations. Hence FFT is a simpler and faster method of implementing DFT. This is very useful when the value of N is large.

1) Calculation of feature extraction: Fourier analysis is a popular signal processing approach to go from time-domain signals to frequency domain signals or vice versa. In this work, the decomposition of the EEG signal into frequency components is achieved through Fourier transforms. The most widely used algorithm to compute the Fourier transform is the Fast Fourier Transform (FFT) that computes the Discrete Fourier Transform (DFT) of a sequence, or its inverse.

$$X_k = \sum_{i=0}^{N-1} x_i(n) e^{\frac{-j2\pi ik}{N}} \qquad \text{for k = 0, 1 ,2 ... N-1} \tag{1}$$

where X(k) is expressed as the discrete Fourier coefficient, N is the length of available data and $x_i(n)$ is the input signal in the time domain.

Since, the aim is to extract key features from our collected data and pass it to LSTM based deep learning network model for classification of negative emotions. We obtain a single number to summarize a certain aspect of the data.

Among the techniques that use FFT for feature extraction, the periodogram and Welch's method is two of the most popular and commonly exploited. The easiest approach to compute PSD is the periodogram. It consists of a frequency

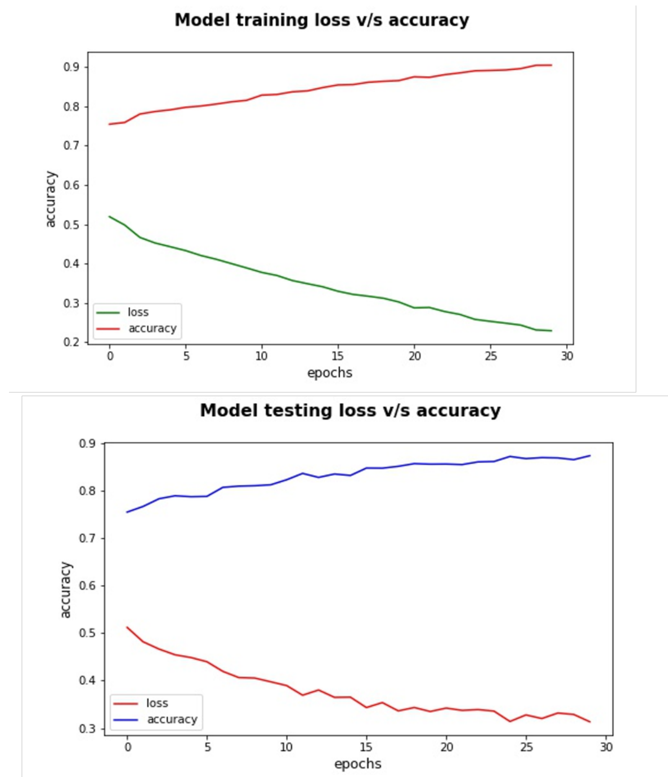

Fig. 3. LSTM_3 Model loss versus accuracy plot for 10-fold cross validation

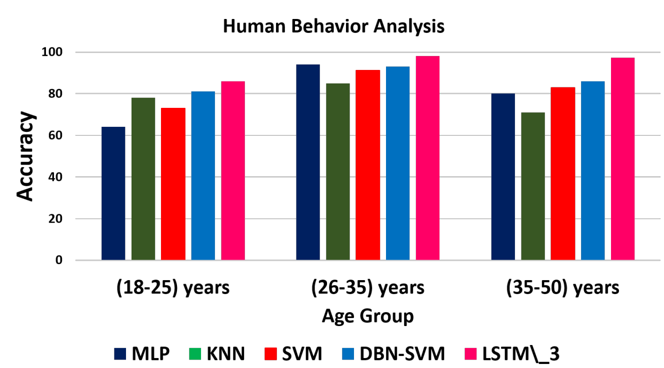

Fig. 4. Age groups based classification comparison of negative emotions

decomposition and is given by the modulus squared of the fourier transform of the signal:

$$\tilde{P}_{xx}(f) = \frac{\Delta t}{N}\left| \sum_{i=0}^{N-1} x_i(n)\, e^{\frac{-j2\pi ik}{N}} \right|^2 \tag{2}$$

Where $\tilde{P}_{xx}(f)$ is PSD of $x_i(n)$, Δt is the space between samples, N is the length of available data and $x_i(n)$ is the input signal in the time domain.

The second method for PSD estimation is the Welch's method that improves the accuracy of the classic periodogram. It is based on the use of overlapping windows to the signal in which a periodogram is calculated for each window and then those periodograms are averaged between them to compute PSD.

We assumed signals x(n) are of finite length, it's power spectral density estimate has the following relationship:

$$x_i(n) = x(n + iD), \qquad n = 0,\ 1,\ 2 \ldots,\ M\text{-}1 \tag{3}$$

$$\textit{while } i = 0,\ 1,\ 2 \ldots,\ L\text{-}1$$

Take *'iD'* to be the point of start of the i^{th} sequence. Then L of length $2M$ represents data segments that are formed. The resulting output periodograms give:

$$\tilde{\tilde{P}}^{(i)}_{xx}(f) = \frac{1}{MU} \left| \sum_{n=0}^{M-1} x_i(n)\, w(n)\, e^{-j2\pi fn} \right|^2$$

Here, in the window function, U gives normalization factor of the power and is chosen such that

$$U = \frac{1}{M} \sum_{n=0}^{M-1} w^2(n), \tag{4}$$

where $w(n)$ is the window function. The average of these modified periodogram gives Welch's power spectrum as follows:

$$P^W_{xx} = \frac{1}{L} \sum_{i=0}^{L-1} \tilde{\tilde{P}}^{(i)}_{xx}(f) \tag{5}$$

An optimal window duration must be decided that should be sufficiently long to cover two full cycles of the lowest frequency in our interest. So depending on the band, for example, to extract the delta band our frequency of interest is 0.5Hz and use a 2/0.5 or 4-second sliding window, this gives a frequency resolution of 4 bins per second, and each step is able to capture 0.25Hz.

The headset gives us access to 10 features named as f1,f2,...f10. In this study, our dataset is extended to accommodate these features as well, as part of our feature set we extracted a total of 10 features from five frequency bands the 10 features are named as Attention, Meditation, Delta (1-3Hz), Theta (4-7Hz), LowAlpha (8-9Hz), HighAlpha (10-12Hz), LowBeta (13-17Hz), HighBeta (18-30Hz), LowGamma (31-40Hz), and HighGamma (41-50Hz) in this research work.

E. Classification

In literature different machine learning algorithms are used to recognize and classify emotions. In this research work long short term memory (LSTM) model based deep learning network are used to classify negative emotions using EEG signals. LSTM model architecture is described next.

1) LSTM Architecture: The recurrent neural network has an LSTM model as the eminent variation, which has shown its efficiency in the extraction of temporal information from long physiological signals. In this section testing of three different LSTM architecture is explained as shown in Figure 2.

The EEG signals are long sequence which is difficult for a recurrent neural network to learn because of the factor that they are trained by back-propagation through time called as BPTT which leads to the problem of vanishing/exploding gradient to overcome this issue, the RNN cell is replaced by gate cell called as LSTM cell. This advantage of LSTM makes it a preferred choice to adopt it for negative emotion classification using EEG signals.

These three LSTM architectures are containing simple to complex internal architecture named as LSTM_1, LSTM_2, LSTM_3 architectures and parameter used for them is shown in Table II. These LSTM networks were built using Keras 2.0.9 upon Tensorflow 1.4.0 backend in Python 3.6. Simplest approach is implemented in LSTM_1 architecture having 56 memory unit in single layer. LSTM_2 architecture is also implemented using single layer approach but the number of

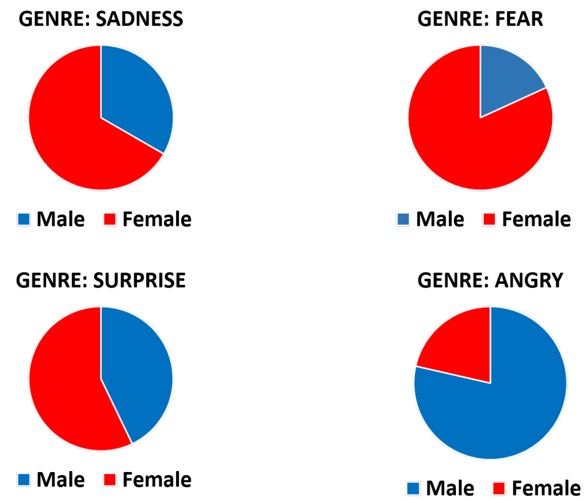

Fig. 5. Gender based responsiveness on targeted genres

TABLE V
COMPARISON OF PRECISION, RECALL, F1-SCORE, AND CLASS ACCURACY OF LSTM_3 MODEL

Validation Technique	Classifier-LSTM_3				
50-50	**Class**	**Precision**	**Recall**	**F1-score**	**Accuracy**
	Sadness	0.77	0.88	0.82	88.89
	Fear	0.82	0.80	0.81	87.22
	Anger	0.70	0.63	0.66	88.11
	Surprise	0.85	0.76	0.81	93.11
60-40	**Class**	**Precision**	**Recall**	**F1-score**	**Accuracy**
	Sadness	0.77	0.94	0.85	88.47
	Fear	0.90	0.83	0.86	92.85
	Anger	0.86	0.68	0.76	91.74
	Surprise	0.83	0.77	0.80	92.50
70-30	**Class**	**Precision**	**Recall**	**F1-score**	**Accuracy**
	Sadness	0.84	0.92	0.88	91.85
	Fear	0.89	0.86	0.88	92.69
	Anger	0.85	0.84	0.84	93.61
	Surprise	0.88	0.78	0.82	94.44
10-fold CV	**Class**	**Precision**	**Recall**	**F1-score**	**Accuracy**
	Sadness	0.96	0.94	0.95	96.33
	Fear	0.79	0.94	0.86	92.83
	Anger	0.94	0.89	0.92	96.28
	Surprise	0.85	0.70	0.77	94.33

TABLE VI
COMPARISON OF SPECIFICITY OF LSTM_3 MODEL FOR 10 FOLD CROSS-VALIDATION PARTITION SCHEME

Validation Technique	**Classifier-LSTM_3**	
	Class	**Specificity (%)**
		Mean ± Std
10-fold cross validation	Sadness	94.76 ± 2.96
	Fear	95.62 ± 3.66
	Anger	95.29 ± 4.86
	Surprise	92.91 ± 2.36

TABLE VII
CLASSIFICATION ACCURACY COMPARISON FOR FOUR CLASS OF EMOTIONS

Method	**Validation Technique**	**Accuracy**		
		Max.	Avg.	Min.
MLP[8]	50-50	70.11	69.33	68.56
KNN[4]	50-50	65.10	63.41	61.73
SVM[11]	50-50	72.44	71.11	69.78
DBN-SVM[13]	50-50	78.88	76.51	75.26
LSTM_3 Classifier	50-50	81.62	78.67	76.45
MLP[8]	60-40	72.48	71.80	70.12
KNN[4]	60-40	68.62	66.43	64.24
SVM[11]	60-40	74.86	73.33	71.89
DBN-SVM[13]	60-40	80.44	78.92	76.73
LSTM_3 Classifier	60-40	84.64	82.85	80.76
MLP[8]	70-30	75.94	73.77	71.84
KNN[4]	70-30	70.09	68.74	67.26
SVM[11]	70-30	77.48	75.56	73.33
DBN-SVM[13]	70-30	82.59	80.44	78.43
LSTM_3 Classifier	70-30	89.73	86.30	84.87
MLP[8]	10-Fold CV	76.27	72.01	70.59
KNN[4]	10-Fold CV	73.88	70.16	71.25
SVM[11]	10-Fold CV	81.83	76.23	79.15
DBN-SVM[13]	10-Fold CV	83.61	81.68	79.86
LSTM_3 Classifier	10-Fold CV	92.84	89.83	87.76

memory unit is increased to 64. In LSTM_3 the number of layers are increased to 2 while the number of dimension is same in the second layer too. The number of memory units in LSTM_3 are 64. LSTM_1, LSTM_2, LSTM_3 networks architecture is fully connected having output of 28 unit using "$tanh$" as activation function.

$$tanh(x) = \frac{2}{1+e^{\text{-2x}}} - 1 \tag{6}$$

Last and final dense layer is using "$Softmax$" as activation function and generating output for multi-class classification i.e., for four negative emotion.

$$\text{Softmax}(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \tag{7}$$

To learn features from EEG signals LSTM cells and dropout layers are used. In our three LSTM architectures dropout layer was used to reduce the over-fitting by limiting units from co-adapting too much. For classification of negative emotions dense layer is used. Loss function used in these networks architecture is categorical cross entropy and batch size is 10. Adaptive Moment Estimation (Adam) optimizer is used with a learning rate of 0.001. Initially, as Input 10 features are given to each network architecture named as f1,f2,...,f10 these inputs are features extracted using FFT technique from raw EEG network. Each of the above mentioned LSTM architecture was evaluated on three dataset i.e., DEAP, SEED, and the dataset that we have created.

IV. RESULTS

This section presents the experimental results of the LSTM based deep learning model for negative emotion classification by the analysis of the EEG signal to validate the effectiveness of our method.

Computing environment have Python (3.6) to implement three LSTM cell architecture i.e., LSTM_1, LSTM_2, LSTM_3, and other state of the art methods i.e., MLP, KNN, SVM, DBN-SVM, and LSTM_3 based deep learning model classifier framework as a classifier on a Windows-based operating system having core processor as Intel i7 7^{th} gen of 3.4 GHz with 32 GB of RAM. The feature extraction technique used for all these classifiers is FFT. For the implementation of MLP, KNN, SVM, and DBN-SVM classifier parameters values are same as defined in [8], [4], [11], [13] respectively. All these algorithms are used to classify human emotions into four class of negative emotions. We have implemented LSTM architectures using the parameter values as explained in section LSTM Architecture and in Table II. The parameters and configurations for these variables were introduced on the same device to ensure that the findings and comparisons presented were unequivocal and univocal.

In this study the performance evaluation is done using 50-50, 60-40, 70-30, and 10-fold cross-validation technique. Therefore, 100 epochs results are assessed for all partition method.

Figure 2 shows the best LSTM model architecture used in this research study for classifying negative emotions using EEG signals. Simple to complex internal architecture of LSTM model as LSTM_1, LSTM_2, LSTM_3 are analyzed in this work. Table III, compares the average classification performance of these different LSTM architectures for negative emotion classification using EEG signals. LSTM_3 model outperformed the classification accuracy by achieving average accuracy as 78.67%, 82.85%, 86.30%, 89.83% for 50-50, 60-40, 70-30, and 10 fold cross-validation training-testing partition respectively.

Table IV, shows the confusion matrix of our proposed LSTM_3 model for the classification of four class of negative emotions. Table IV represents confusion matrix for four class of emotion recognition of LSTM_3 classifier. The TP and TN of our method are much higher than the other state-of-the-art methods, which show that our method is able to classify the correct samples properly.

Table V shows the recall, precision, and F1-score of our proposed method. High recall value of our method shows that it is able to classify minority class samples correctly. The findings demonstrate the evidence of comparability of our model with other techniques. As, this dataset is also, unbalanced, which

TABLE VIII
COMPARISON OF LSTM_3 MODEL PERFORMANCE ON OTHER EEG DATASETS

	LSTM_3 classifier on other EEG datasets					
	Our Dataset Accuracy		**DEAP Dataset Accuracy**		**SEED Dataset Accuracy**	
Partition	Max.	Mean	Max.	Mean	Max.	Mean
50-50	81.62	78.67	80.66	79.02	83.26	81.44
60-40	84.64	82.85	85.42	81.16	86.38	83.29
70-30	89.73	86.30	86.22	84.38	87.29	85.96
10-fold CV	92.84	89.83	91.38	88.42	89.34	87.22

makes recall and F1-score a critical parameter to evaluate the performance of our proposed LSTM classifier. Table V shows that our method LSTM_3 classifier has high recall, precision, and F1-score values for 50-50, 60-40, 70-30, and 10-fold cross-validation which proves that our method is also classifying minority classes more accurately .

Table VI shows the comparison of specificity values of our method for 10-fold cross-validation data partitioning scheme. Specificity evaluates the correctly recognized positive and negative emotions out of all the generated positive and negative emotions. It is noticeable that our method has a higher value of specificity value for each class classification. Apart from classification accuracy gains, LSTM classifier also, has the ability to sustain above 92.65% specificity that quantitatively results in having very low false prediction rate. Which proves that it is able to classify negative emotions more accurately.

Table VII contains the classification accuracy comparison of other state-of-the-art approaches as MLP, KNN, SVM, and DBN-SVM. It contains the maximum, average, and minimum accuracy for training-testing partition. It is clear from the Table VII that among traditional machine learning classifier LSTM_3 model outperformed the classification accuracy by achieving maximum accuracy for four class of negative emotions as 81.62%, 84.64%, 89.73%, 92.84% for 50-50, 60-40, 70-30, and 10 fold cross-validation training-testing partition respectively. The second best classifier after LSTM based deep learning model is DBN-SVM it has achieved maximum accuracy for four class of negative emotions as 78.88%, 80.44%, 82.59%, 83.61% for 50-50, 60-40, 70-30, and 10 fold cross-validation training-testing partition. The results demonstrated that LSTM classifier has provided significant increase in classification performance from + 6.3% and up-to 20.4%. This also, indicate the reliability of LSTM based deep learning classifier. The reason for such high accuracy is the LSTM based deep learning model used for classification of EEG signals having long sequence, which is able to overcome vanishing/exploding gradient problem present in the RNN as the RNN cell is replaced by gate cell called as LSTM cell in this deep learning network.

Figure 3 presents the model training and testing loss versus accuracy plot for 70-30 and 10 fold data partitioning schemes. It is clear from the Figure 3 that the model has minimize the loss upto 12.23% for 70-30 data partitioning and has achieved average accuracy upto 89.83% for 30 epochs in testing phase of 10-fold cross validation scheme. Additionally, empirical testing found that 50 epochs for training of units seemed best but further exploration is required to fine-tune this parameter.

In this study, we also perform experiments on three different age groups (18-25) years, (26-35) years, (36-55) years to analyze their emotions on different video clips. Our dataset includes three age groups. The results are presented in Figure 4. It is clear from the results that age group from (26-35) years has the highest average emotion recognition accuracy, i.e., 92.84% for the five classifiers. For (36-50) years the average accuracy is 82.82% and for (15-25) years it is 75.25%, which is shown in Figure 4. The emotions from clips sensed very well in people aged (26-35) years and (36-55) years. Another demonstrated reality in our studies that the intensity of emotions in the age bracket of 26-35 years is at its peak.

Figure 5 shows a pie graph of the difference in gender-based responsiveness of subjects for targeted genres. From the results, we analyze that females are more responsive as compared to males and express their emotions. When Sad and fear genres are shown to the participants, females responsiveness is more as compared to males. While, when angry genre is shown to the participants responsiveness of males is more as compared to females. In other words, we can say that females are more overt than males in responding to negative emotion targeting visual stimuli.

A. Comparison with other datasets

The approach suggested was also used to demonstrate the generality of our method for other datasets. The dataset used for comparison is DEAP [12], SEED [19] which are larger EEG dataset as compared to our dataset. Table VIII shows the comparison of classification accuracy of our LSTM model over two publicly available datasets and our dataset. On DEAP dataset for 50-50, 60-40, 70-30, and 10-fold cross-validation LSTM model has achieved mean accuracy as 79.02%, 81.16%, 84.38%, and 88.42%. In case of SEED dataset the mean accuracy is 81.44%, 83.29%, 85.96%, and 87.22%. These results show that our model is able to achieve better accuracy for other publicly available benchmark datasets too. These findings demonstrate that our method is generalized and has the potential to handle complex real-life problems also.

V. CONCLUSION

In the duration of this work, we propose brain signals based human emotion recognition system using LSTM based deep learning model, which is analyzing the emotional conditions of participants through brain wave assessment. It is verified after analyzing human behavior that sadness and angry emotions are easiest to classify. Brain signals of age group 26-35 years are recognized to provide the highest precision of emotional identification in compliance with self-reports. For gender-based responsiveness, females tend to respond more as compared to males. In general, the findings of the suggested LSTM based deep learning methodology has confirmed to outperform the outcomes of current state-of-the-art methods.

In future, the data acquisition process can be enhanced with the help of multi-channel EEG devices. Also, techniques like feature engineering can be introduced to enhanced the model performance.

References

[1] Divya Acharya, Anosh Billimoria, Neishka Srivastava, Shivani Goel, and Arpit Bhardwaj. Emotion recognition using fourier transform and genetic programming. *Applied Acoustics*, 164:107260, 2020.

[2] Divya Acharya, Shivani Goel, Rishi Asthana, and Arpit Bhardwaj. A novel fitness function in genetic programming to handle unbalanced emotion recognition data. *Pattern Recognition Letters*, 2020.

[3] Samuel Albanie, Arsha Nagrani, Andrea Vedaldi, and Andrew Zisserman. Emotion recognition in speech using cross-modal transfer in the wild. *arXiv preprint arXiv:1808.05561*, 2018.

[4] Teodiano Freire Bastos-Filho, André Ferreira, Anibal Cotrina Atencio, Sridhar Arjunan, and Dinesh Kumar. Evaluation of feature extraction techniques in emotional state recognition. In *2012 4th International conference on intelligent human computer interaction (IHCI)*, pages 1–6. IEEE, 2012.

[5] Arpit Bhardwaj, Aruna Tiwari, M Vishaal Varma, and M Ramesh Krishna. Classification of eeg signals using a novel genetic programming approach. In *Proceedings of the Companion Publication of the 2014 Annual Conference on Genetic and Evolutionary Computation*, pages 1297–1304, 2014.

[6] Arpit Bhardwaj, Aruna Tiwari, M Vishaal Varma, and M Ramesh Krishna. An analysis of integration of hill climbing in crossover and mutation operation for eeg signal classification. In *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, pages 209–216. ACM, 2015.

[7] Harshit Bhardwaj, Aditi Sakalle, Arpit Bhardwaj, and Aruna Tiwari. Classification of electroencephalogram signal for the detection of epilepsy using innovative genetic programming. *Expert Systems*, 36(1):e12338, 2019.

[8] Adnan Mehmood Bhatti, Muhammad Majid, Syed Muhammad Anwar, and Bilal Khan. Human emotion recognition and analysis in response to audio music using brain signals. *Computers in Human Behavior*, 65:267–275, 2016.

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

[10] Sibnath Deb, Esben Strodl, and Jiandong Sun. Academic stress, parental pressure, anxiety and mental health among indian high school students. *International Journal of Psychology and Behavioral Sciences*, 5(1):26–34, 2015.

[11] Sander Koelstra, Ashkan Yazdani, Christian Soleymani, Mohammad and, Jong-Seok Lee, Anton Nijholt, Thierry Pun, Touradj Ebrahimi, and Ioannis Patras. Single trial classification of eeg and peripheral physiological signals for recognition of emotions induced by music videos. In *International Conference on Brain Informatics*, pages 89–100, 2010.

[12] Sander Koelstra, Christian Muhl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, and Ioannis Patras. Deap: A database for emotion analysis; using physiological signals. *IEEE transactions on affective computing*, 3(1): 18–31, 2011.

[13] Yuan-Pin Lin, Chi-Hong Wang, Tzyy-Ping Jung, Tien-Lin Wu, Shyh-Kang Jeng, Jeng-Ren Duann, and Jyh-Horng Chen. Eeg-based emotion recognition in music listening. *IEEE Transactions on Biomedical Engineering*, 57(7):1798–1806, 2010.

[14] Mufti Mahmud, M Shamim Kaiser, and Amir Hussain. Deep learning in mining biological data, 2020.

[15] AV Medvedev, GI Agoureeva, and AM Murro. A long short-term memory neural network for the detection of epileptiform spikes and high frequency oscillations. *Scientific Reports*, 9(1):1–10, 2019.

[16] Michael James Scott, Sharath Chandra Guntuku, Yang Huan, Weisi Lin, and Gheorghita Ghinea. Modelling human factors in perceptual multimedia quality: On the role of personality and culture. In *Proceedings of the 23rd ACM international conference on Multimedia*, pages 481–490. ACM, 2015.

[17] Olga Sourina, Qiang Wang, Yisi Liu, and Minh Khoa Nguyen. A real-time fractal-based brain state recognition from eeg and its applications. In *Biosignals*, pages 82–90, 2011.

[18] Kostas M Tsiouris, Vasileios C Pezoulas, Michalis Zervakis, Spiros Konitsiotis, Dimitrios D Koutsouris, and Dimitrios I Fotiadis. A long short-term memory deep learning network for the prediction of epileptic seizures using eeg signals. *Computers in biology and medicine*, 99:24–37, 2018.

[19] Wei-Long Zheng and Bao-Liang Lu. Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks. *IEEE Transactions on Autonomous Mental Development*, 7(3):162–175, 2015.