Multi-layer Stacking-based Emotion Recognition using Data Fusion Strategy

Saba Tahseen[1]
PhD Scholar
Dept. of Computer Science and Engineering
Christ University, Bangalore, India

Ajit Danti[2]
Professor
Dept. of Computer Science and Engineering
Christ University, Bangalore, India

***Abstract*—Electroencephalography (EEG), or brain waves, is a commonly utilized bio signal in emotion detection because it has been discovered that the data recorded from the brain seems to have a connection between motions and physiological effects. This paper is based on the feature selection strategy by using the data fusion technique from the same source of EEG Brainwave Dataset for Classification. The multi-layer Stacking Classifier with two different layers of machine learning techniques was introduced in this approach to concurrently learn the feature and distinguish the emotion of pure EEG signals states in positive, neutral and negative states. First layer of stacking includes the support vector classifier and Random Forest, and the second layer of stacking includes multilayer perceptron and Nu-support vector classifiers. Features are selected based on a Linear Regression based correlation coefficient (LR-CC) score with a different range like n_1, n_2,n_3,n_4 a, for d_1 used n_1 and n_2 dataset ,for d_2 dataset, combined dataset of n_3 and n_4 are used and developed a new dataset d_3 which is the combination of d_1 and d_2 by using the feature selection strategy which results in 997 features out of 2548 features of the EEG Brainwave dataset with a classification accuracy of emotion recognition 98.75%, which is comparable to many state-of-the-art techniques. It has been established some scientific groundwork for using data fusion strategy in emotion recognition.**

Keywords*—Electroencephalograph (EEG); linear regression based correlation coefficient; feature selection; multi-layer stacking model; machine learning techniques; emotion recognition

I. INTRODUCTION

With the rapid advancement of computers and human contact technology, there is a significant demand in the area of human interaction for a more intelligent and humanized human-machine interface (HMI). A (BCI) brain-computer interface transforms a way of transforming brain processes that are connected directly to the brain of a living organism, such as a human or an animal. BCI serves as a bridge for communication between both the human brain and as a tool to detect different applications, such as emotion identification and different applications [1].

Human communication, daily life, and work all rely heavily on emotional expressiveness. It can be characterized as positive, neutral, or negative experiences arising from a variety of physiological activities, and it includes a wide range of emotions such as sadness, happiness, surprise, anger, and disgust [2][3][4].

Emotion recognition research has become more common among researchers with the development of sensor-based technology and processes and accessibility have improved. Emotion recognition can have important applications, whether professional, personal, or personal [5] such as in the fields of medicine [6], education, psychology, computer games, driving, security, entertainment [7], and workload evaluation, and many others [8].

Emotions can be detected in a variety of ways, including brain waves and facial expressions. Brain waves are a means of obtaining emotion that can be both intrusive and non-invasive. In wires and an intrusive Brain-computer-interface (BCI) are surgically implanted on the accessible brain surface. The non-invasive method is known as BCI, and it provides a simple, fast, and beneficial method for collecting the brainwaves, which comprises functional magnetic resonance imaging (fMRI), Magnetoencephalogram (MEG), Electroencephalogram (EEG), and numerous signaling have already been approved, recognized, and classified as non-physiological and physiological signals, respectively. In the practical application of emotion recognition, Gesture, Text, movement, speech, voice intonation, and facial expression, among other non-physiological signals, are indeed the original concept is a term that has been used a lot in the past. More studies have recently been conducted using physiological signals such as electrocardiogram (ECG) and electroencephalogram (EEG) [9].

This study proposes an electroencephalography (EEG) signal analysis technique for recognizing and classifying emotional states, as well as a correlation-based data reduction strategy coefficient score between features with a different range and developed a new dataset. Machine learning models are grouped into three types: supervised, unsupervised, and reinforcement learning, as well as a specific form termed ensembles learning, based on the methodology utilized [10].

An ensemble machine learning algorithm is often known as stacking. Stacking is the process of learning how to aggregate the prediction of participating ensemble components using a machine learning model and minimize the variation. In this paper extension of Stacking Classifier with two layers of learning models i.e., Multi-layer Stacking Classifier is developed.

In this research article there are four contributions: 1) Calculated the correlation coefficient score of features based on linear regression model. 2) Prepared a different dataset with a particular range and applied a data fusion strategy. 3) Developed a method which has already been tested through a developed dataset. 4) Calculated the time complexity of each classifier.

This manuscript's structure is as follows: Section II describes the related work on emotion recognition, data reduction, and on machine learning techniques. Section III describes about possible approach for reducing data in an EEG analysis by generating restricted electrode correlations zones. Section IV explains the configuration for the classification procedure as well as the methods used to carry out the experiments. The results of executing the proposed solutions are discussed in Section V, and the work's conclusions are discussed in Section VI.

II. Related Work

A method for classifying emotional states using wavelet compression, EEG data, and sensor classification procedure, links electrodes in a 10/20 model to Brodmann n areas and reduces computational load. The emotions modelling is based on hold value of an adjusted space from the Russell Arousal Valence Space and the Geneva model, and the classification procedure was accomplished using an SVM Classification process, which achieved an 81.46%, classification performance for a multi-class problem [11].

Researchers used the zero-time window-based using the numerator group-delay to derive immediate frequency features function to correctly detect the periods in each emotional situation. Using QDC and RNN, as well as the DEAP database, separate class systems were constructed used to test them [12].

There are two types of network entropy metrics calculated: nodal degree entropy and clustering coefficient entropy. The effective characteristics are fed into the SVM classifier using the AUC method to accomplish emotion recognition across participants. The findings of the experiment revealed that the properties of 18 channels selected by AUC were significant (p 0.005) for the EEG signals of 62 channels [13].

The program collects features from EEG data and uses machine learning techniques to classify emotions, with different segments of a trial being utilized to train the proposed model and examine its impact on emotion detection outcomes. Second, using the classification performance and two emotion coefficients, namely, the correlations and entropy coefficients, a unique activation curve for emotions is created. The activation curve can not only define emotions, but it may also indicate the emotional activating mechanism to some extent. Next, the two factors are combined to provide a weight coefficient, which improves emotion recognition accuracy. Experiments on the DEAP and SEED datasets were conducted to validate the suggested technique [14]. Based on the SJTU emotion EEG datasets (SEED) and the ResNet50 and Adam optimizer, the CNN model is being used to train the features and recognize the emotions of positive, negative and neutral states of real EEG signals in a single model [15]. This study presented a new model called the "hybrid model" that merged three ensemble models. For classification challenges, a set of features is retrieved from raw IoT datasets from various IoT domains utilizing linear discriminant analysis (LDA), Principal component analysis (PCA), and Isomap. The classifiers' accuracy, area under the curve (AUC), and F1 score are used to compare their performances [16].

A feature extraction subsystem and a classifier subsystem are created in this paper for an EEG-based emotion recognition system. 9 features extracted from the time and frequency domain from the EEG sign were used because the greater performance of the feature extraction module may result in higher recognition accuracy [17]. Authors have created a machine learning algorithm based on ensembles. The data was cleaned using a pre-processing technique, and feature selection was done using wrapper-based methods; additionally, a stacking-based ensemble learning model was used to identify the MDD participants in the final stage [18]. The proposed research makes a significant addition by presenting an enhanced version of the agent-based data reduction algorithm that incorporates the stacking generalization mechanism for data reduction. Enhancing the performance of the categorization results in the model discussed in [19]. Considering the aforementioned factors and applying the skills in this sector, it is inspired to write this research to address the following problem.

- How accuracy and prediction performance can be improved?
- How to overcome from overfitting problem?
- How precise can the EEG waves be classified?
- What are other essential features can derive using EEG data?

III. Dataset Details

Performed thorough analysis on the EEG Brainwave Dataset, which is publicly available, in each state: happy, neutral, and negative - data was collected for 3 minutes from two subjects (1 male, 1 female). To record the TP9, AF7, AF8, and TP10 EEG placements, the author employed a Muse EEG headgear with dry electrodes.

IV. Proposed Methodology

In this section, discussed about a possible approach for reducing data in an EEG analysis by generating restricted electrode correlations zone and explains about the configuration for the classification procedure as well as the methods used to carry out the experiments.

A. Data Preprocessing

In Machine Learning, data preprocessing relates to the procedure of organizing and managing basic information to make it appropriate for creating and training Machine Learning models. First libraries are required for data preprocessing, identifying missing values, so, the EEG brainwave dataset has no missing values, and provides the label for each category like positive, negative, and neutral.

B. Feature Selection

For features, correlation coefficient score is determined with a linear regression model. For each feature set, selected the subset of features based on the range and created a new dataset.

C. Correlation Coefficient Score

To calculate the correlation coefficient, the linear regression model is used. The linear relationship between multiple Correlations is used to assess variables. Correlation is used to predict one variable from the other. Because the good variables are so closely related to the aim, using correlation to pick features stands to reason. Furthermore, variables should really be relevant to the goal yet unrelated to each other. It anticipates one from the other if the two variables are correlated. As a result, if features are correlated, the model only needs one among them because the other provides no additional information. The linear correlation coefficient is a mathematical expression that measures the degree and direction of a relationship between two variables:

$$r = \frac{\sum \frac{(x_i - \bar{x})}{s_x} \frac{(y_i - \bar{y})}{s_y}}{n-1} \tag{1}$$

Where $\bar{x}$ and s_x are represented, the sample mean and sample deviation of the x's. And $\bar{y}$ and s_y are represented the mean and standard deviation of the y's.

A different method of calculating the correlation coefficient is as follows:

$$r = \frac{s_{xy}}{\sqrt{s_{xx} s_{yy}}} \tag{2}$$

Where,

$$s_{xx} = \sum x^2 - \frac{(\sum x)^2}{n}$$

$$s_{xy} = \sum xy - \frac{(\sum x)(\sum y)}{n}$$

$$s_{yy} = \sum y^2 - \frac{(\sum y)^2}{n}$$

The properties of "r" are as follows:

- It is always in the range of -1 to +1.
- Because it is a dimensionless quantity measure, "r" would have been the same that whether two variables were measured in pounds and inches or grams and centimeters.
- Favorable "r" levels are linked to positive connections.
- Bad "r" values are linked to negative relationships.

First, Linear Regression model is used to find the feature importance of each feature. Then correlation coefficient Scoreis determined for each feature. Following ranges are considered for each dataset.

$n_{1=}0.0\ to\ 0.3;$

$n_2 = 0.31\ to\ 0.5;$

$n_3 = 0.51\ to\ 0.75;$

$n_4 = 0.76\ to\ 1.0.$

Based on the Correlation Coefficient score value, it is found that 610 features importance score are within 0.0-0.3, 179 features are within 0.31-0.5, 136 features are within the range 0.51-0.75 and 77 features are within 0.76 to 1.0. After applying the data fusion technique for the EEG brainwave dataset, 997 features are selected. At this stage developed a four set of datasets with different feature values. Then selected all the unique features from n_1 and n_2 and generated the new dataset called d_1 which has 787 columns excluding the label. The same procedure has been applied on n_3 and n_4 and after combining the features and filtering the unique features, around 211 are considered, i.e., d_2. And finally, by considering d_2 and d_3 determined 997 unique features and that is final dataset to work on.

$d_{1,} d_{2,}$and d_3 are the final combined dataset of $d_{1,}$

and $d_{2,}$ which details as follows:

$n_1 = (2132\ trails \times\ 610\ features)$

$n_2 = (2132\ trails \times\ 179\ features)$

$d_1 = n_1 + n_2 = (2132\ trails \times\ 787\ features)$

$n_3 = (2132\ trails \times\ 136\ features)$

$n_4 = (2132\ trails \times\ 77\ features)$

$d_2 = n_3\ +\ n_4 = (2132\ trails \times\ 210\ features)$

$d_3 = d_1 + d_2 =\ (2132\ trails \times\ 997\ features).$

D. Machine Learning Classifiers

The method of feature selection, the LR-CC algorithm chooses the most essential features for predicting emotional states. The next step is to classify the emotions by using a dataset using a machine learning technique. Multilayer stacking Classifier is developed based on Stacking Classifier 1 and Stacking Classifier 2 method on Support vector classifier, Random Forest for a Stacking Classifier 1 (layer1), and Nu-Support vector classifier, r, and Multi-Layer Perceptron for Stacking Classifiers 2 (layer 2) as low-level base learners and Random Forest as meta learner algorithm. This section examines step-by step working of the suggested model and overall design. Fig. 1 establishes the framework. First, there are the EEG signals that have been pre-processed from the dataset based on correlation coefficient score; the feature selection stage selects the features based on the data fusion technique and makes a new dataset, i.e., d_3

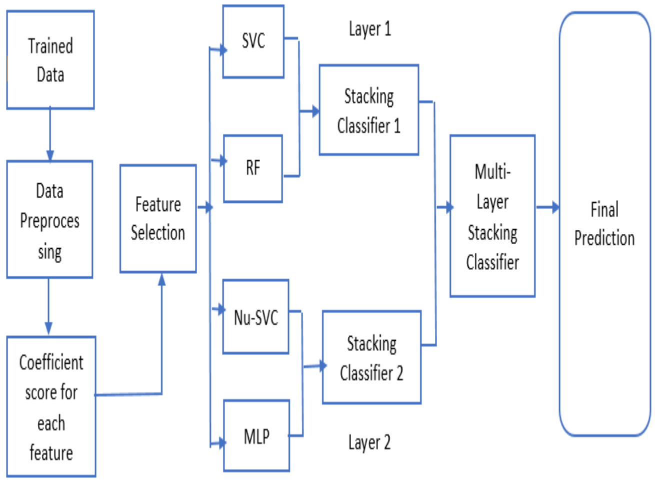


Fig. 1. Architecture Diagram of Multi-layer Stacking Model based on different Machine Learning Classifier Called as Learners and Meta Classifier.

Two layers of Classifiers are implemented separately first and created as a Stacking Classifiers 1 and 2 with a random forest as a meta learner, after implementation of these stacking 1 and 2, developed a multi-layer Stacking Classifier which is based on the predictions of Stacking Classifier 1, and Stacking Classifier 2 as an input and generate a final prediction. A feedforward artificial neural network is a multilayered perceptron in which the functions and structure of the ANN are like the brain's activities and structure [20]. Neurons are basic components that are connected and work in tandem. These neurons are linked by a vital link that contains the information to solve difficulties [21]. For regression and classification issues, a support vector machine (SVM) is a supervised machine learning technique. Each property is represented by a point in n-dimensional space (n = number of features). A hyperplane is created to do linear classification, which effectively separates the two classes. Kernel approaches such as Gaussian kernel, Laplace kernel, and Polynomial kernel are utilized for non-linear classification [22]. Random forest is a decision tree-based learning technique that combines numerous decision trees. It forecasts by aggregating decision tree forecasts [23]. The resilience of SVC and Nu-SVC isn't always the best option, and random forest is suitable in situations [24]. It is difficult to acquire using a Multilayer Perceptron to find the best parameters. Developed a multi-layer stacking ensemble model to address these concerns and increase forecast accuracy. Support Vector Classifier (SVC), Nu-Support Vector Classifier (NuSVC), Multi-layer Perceptron (MLP), and Random Forest classifier,layer 1, and layer 2 will be trained and tested individually. These four models will compensate for shortcomings and provide superior outcomes when stacking. The following shows the algorithm for multilayer stacking with 2 layers. SVC, RF classifier for Stacking 1 or layer 1 as L_1 Classifier and for Stacking 2 or layer 2 as L_2 has Nu-SVC and MLP. After implementation of these classifiers, developed a new model i.e., multi-layer stacking which is based on L_1 and L_2 and for meta classifier *M* used Random Forest with cross validation K and generate a prediction *P*. Confusion matrix will be used to evaluate each classifier's performance. Finally, results are compared by stacking the predictions of different classifiers.

Algorithm: Multilayer Stacking Classification

Input:
Training Datasets *D*, L_1, L_2, L_3, *M, SVC, RF,MLP,Nu-SVC*
Output: An Ensemble Classifiers, L_3,*P*
Step 1: Load SVC, RF Model
load (SVC, RF)
Step 2: Train First Level Classifiers L_1
Apply First level Classifier on dataset *D*
L_1 <- *(SVC, RF)*

Step 3: Load MLP, Nu-SVC Model
Load (MLP, Nu-SVC)
Step 4: Train a Second Level Classifier L_2
Apply Second level Classifier on dataset *D*
L2<- (MLP, Nu-SVC)

Step 5: Construct a new training model based on L_1 and L_2
Adopt a Cross-validation approach *K* in preparing a new training set for Meta classifier *M*
Step 6: Learn a meta–Classifiers *M*
Return Multi-Layer Stacking model L_3
Return *P*

Where,

D: Dataset

L_1: Layer 1 Classifiers

L_2: Layer 2 Classifiers

L_3: Multi-layer Stacking Model

M: Meta Classifiers

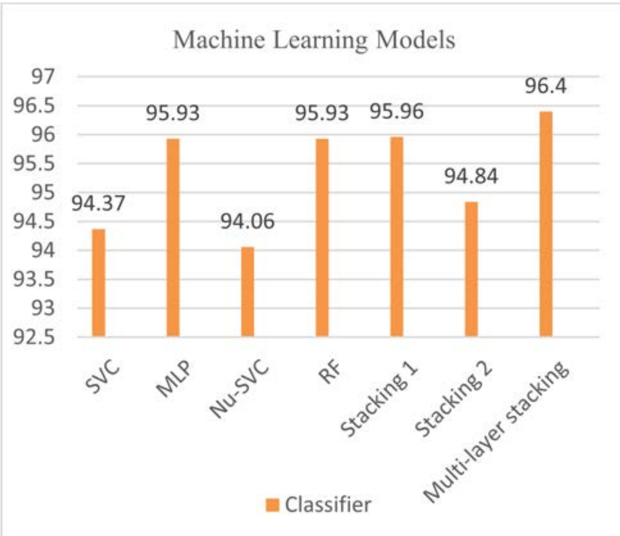
P: Prediction

V. Experimental Results

The main source of dataset consists of 2132 records and 2548 features, after data pre-processing, feature selection is made based on the Linear Regression-Correlation Coefficient (LR-CC) with a Correlation coefficient score of features, 997 features have been selected by applying the data fusion technique, The data fusion method focuses on a group of features that need to be improved, refined, or obtained. on the original dataset. Data fusion techniques are applied with feature level for selecting the features from the same source, i.e., the EEG Brainwave dataset to develop a single dataset.

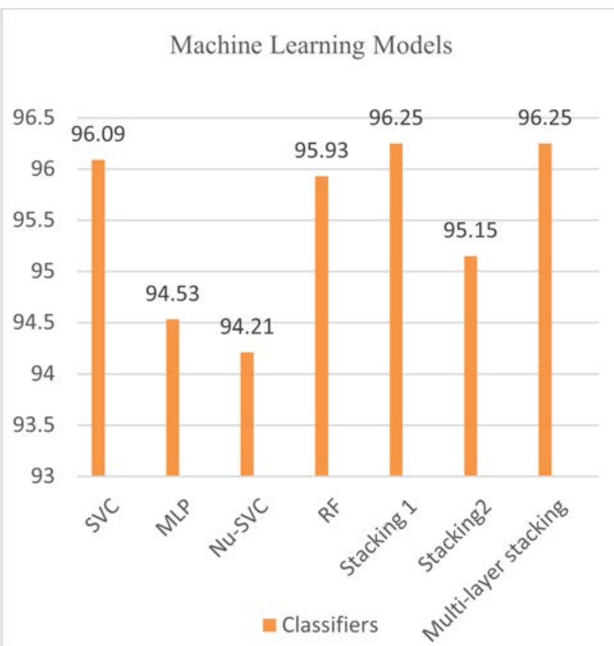
The classification is divided into 3 parts, in the first part i.e., base learners are learned, and findings are predicted after selecting the features separately in the second part layer 1 and layer 2 are learned and findings are predicted separately, and in the third part, to solve the problems with the individual implements Stacking Classifiers 1 and 2 are developed and made a new prediction as input for the multilayer stacking classifier. In the multi-layer stacking, classifiers are trained layer 1(SVC+RF) and layer 2(Nu-svc+MLP) and the base learners predicted output to the multi-layer stacking as an input.

Developed different datasets and combined them into one dataset by using data fusion techniques. Proposed models are tested are tested on different datasets and accuracy of the algorithms are shown in Fig. 2. It is observed the accuracy of the existing algorithm and Proposed algorithm i.e., multi-layer stacking with n_1 dataset which has 610 features.



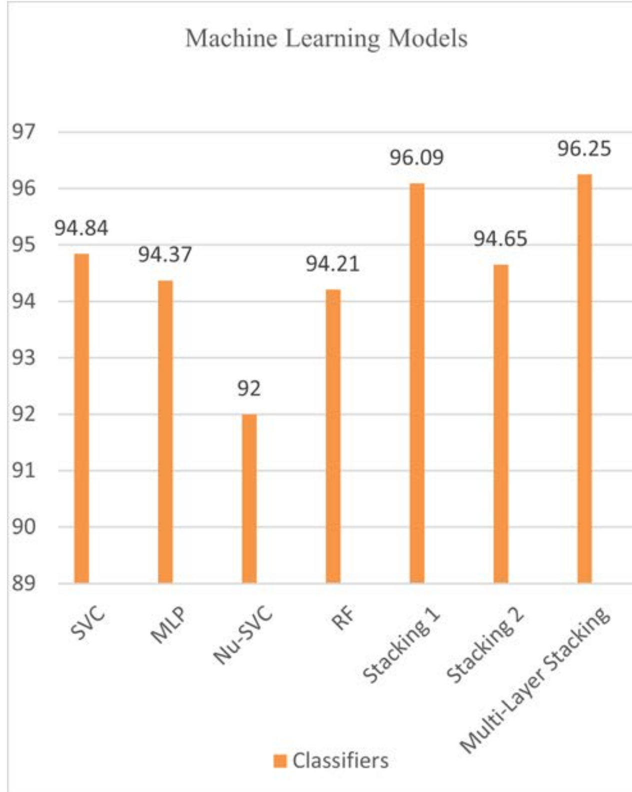
n_1 Dataset=610 features

Fig. 2. Machine Learning Model Performance of Dataset n_1=*610* Features.



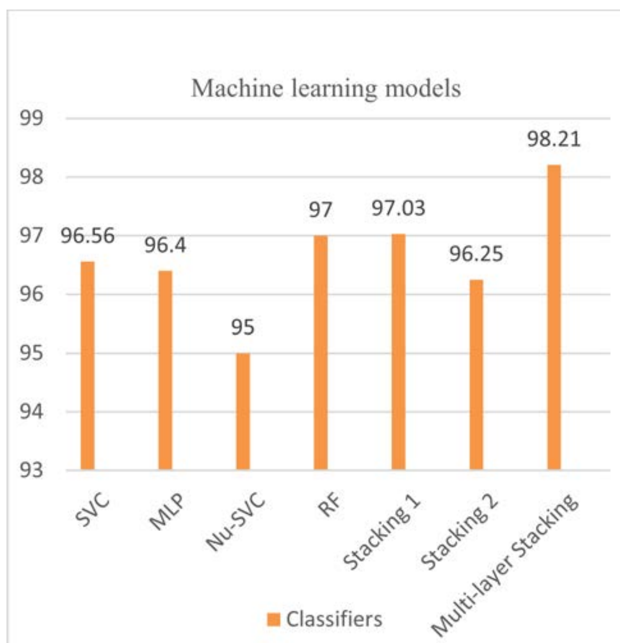
n_2 Dataset=179 Features

Fig. 3. Machine Learning Models Performance of Dataset n_2 = 179 Features.



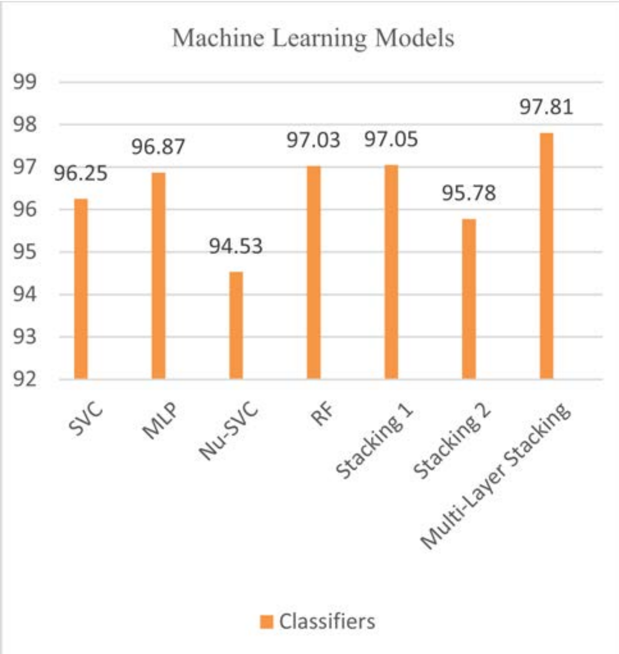
n_4 Dataset=77 features

Fig. 4. Machine Learning Models Performance of Dataset n_3 **=136** Features.



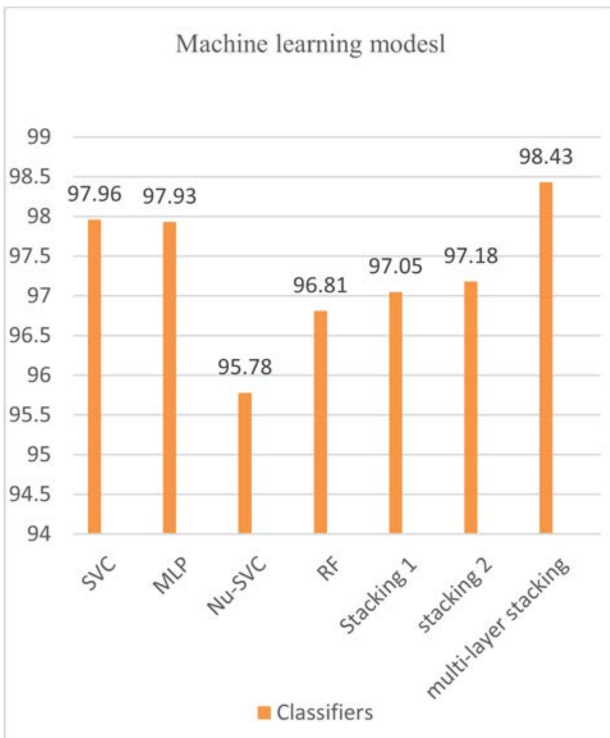
n_3 Dataset=136 Features

Fig. 5. Machine Learning Models Performance of Dataset n_4=77 Features.



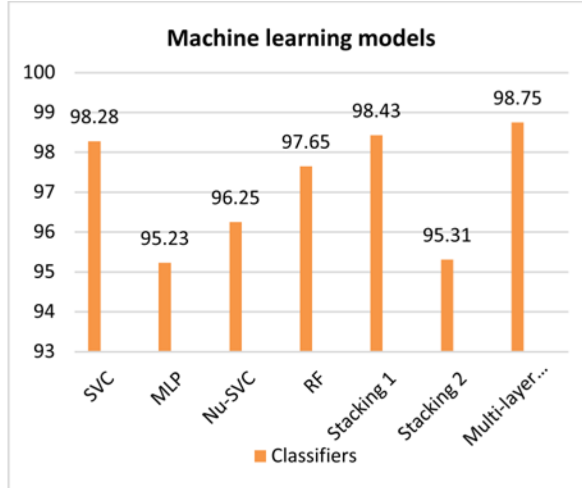
d_1 Dataset=n_1+n_2=787 Features

Fig. 6. Machine Learning Models Performance of Dataset d_1=*787* Features.



d_2 dataset=n_3+n_4=211 features

Fig. 7. Machine Learning Models Performance of Dataset d_2=211 Features.



d_3=dataset=d_1+d_2=997 features

Fig. 8. Machine Learning Models Performance of Dataset d_3 = *997* Features.

Represented the summary of Fig. 2 to Fig. 8 in the form of Table I, in this table, in which proposed model achieved highest accuracy with the unique combination of datasets.

A. Performance Evaluation

To evaluate the developed models, confusion matrix is determined. A confusion matrix is a table that shows how well a classification model (or "classifier") performs on a set of test data for which the true values are known. This is a list of rates that are frequently generated using a binary classifier's confusion matrix:

- Accuracy: What percentage of the time does the classifier get it right?

$$\frac{TP+TN}{N} \tag{3}$$

- Precision: When it predicts yes, how often does it get it right?

$$\frac{TP}{TP+FP} \tag{4}$$

- Recall: The number of true positives divided by the total number of true positives and false negatives equals recall.

$$\frac{TP}{TP+FN} \tag{5}$$

- F1-score: The genuine general positive rate (recall) and precision are weighted averages.

B. Confusion Matrix for Each Classifier

The confusion matrix, which includes metrics such as Sensitivity, Accuracy, Precision, Specificity, and measure, is used to evaluate the algorithm's efficiency after it has been implemented. Fig. 9 to 14 are the confusion matrix of SVC,MLP, Nu-SVC, RF, stacking 1, stacking 2 and Multi-Layer stacking are shown in Fig. 15. Similarly, ensembles models.

TABLE I. ACCURACY OF CLASSIFIERS

Models \ Dataset	n_1	n_2	n_3	n_4	d_1	d_2	d_3
SVC	94.37	96.09	96.56	94.84	96.25	97.96	98.28
RF	95.93	94.53	96.40	94.37	96.87	97.03	95.23
Nu-SVC	94.06	94.21	95	92	94.53	95.78	96.25
MLP	95.93	95.93	97	94.21	97.03	96.81	97.65
Stacking 1	95.96	96.25	97.03	96.09	97.05	97.05	98.43
Stacking 2	94.84	95.15	96.25	94.65	95.78	97.81	95.31
Multi-layer stacking	96.4	96.25	98.24	96.25	97.81	98.43	98.75

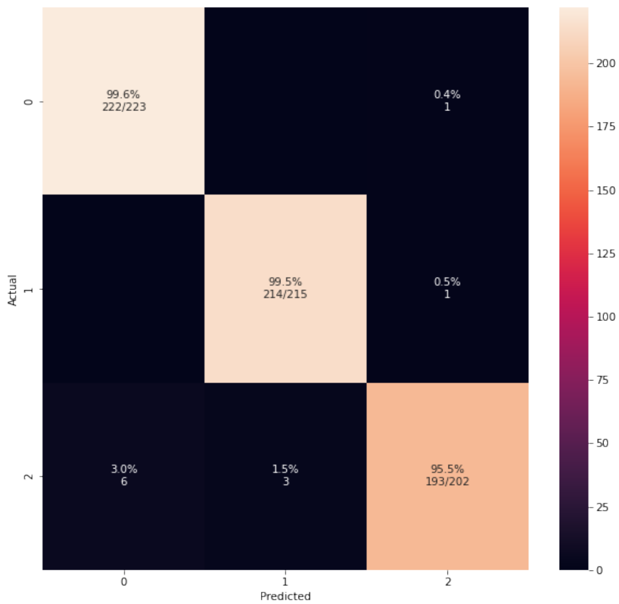


Fig. 9. Confusion Matrix for Support Vector Classifier Model.

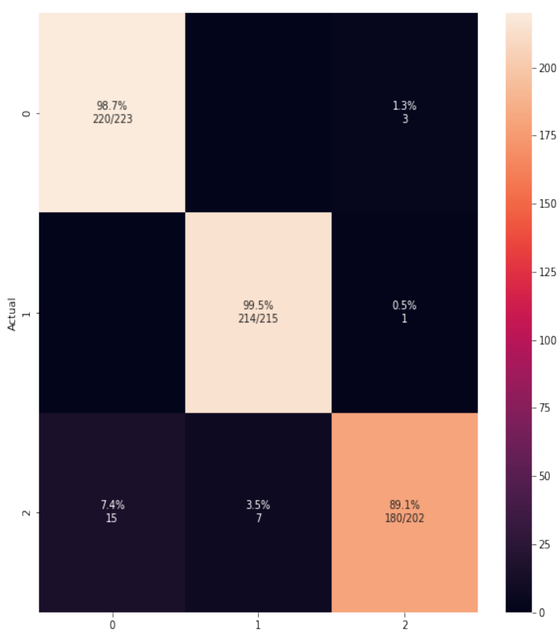


Fig. 10. Confusion Matrix of Multilayer Perceptron Model.

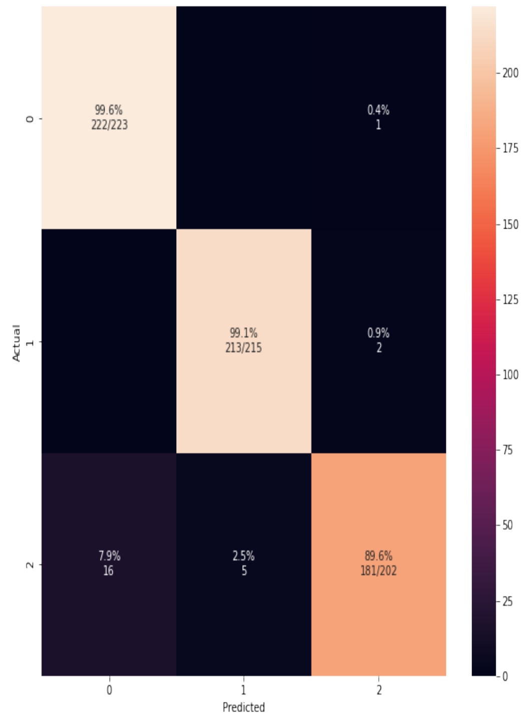


Fig. 11. Confusion Matrix of Nu-SVC Model.

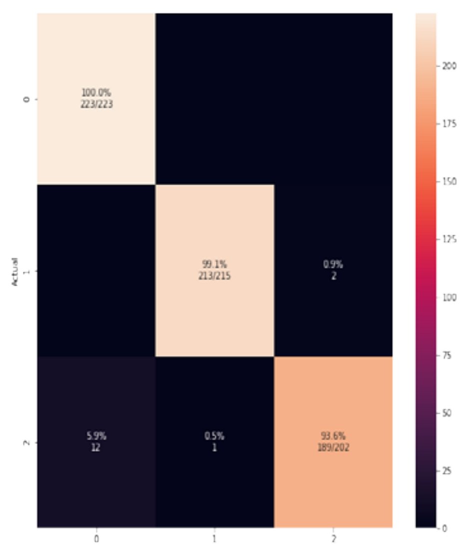


Fig. 12. Confusion Matrix of Random Forest Model.

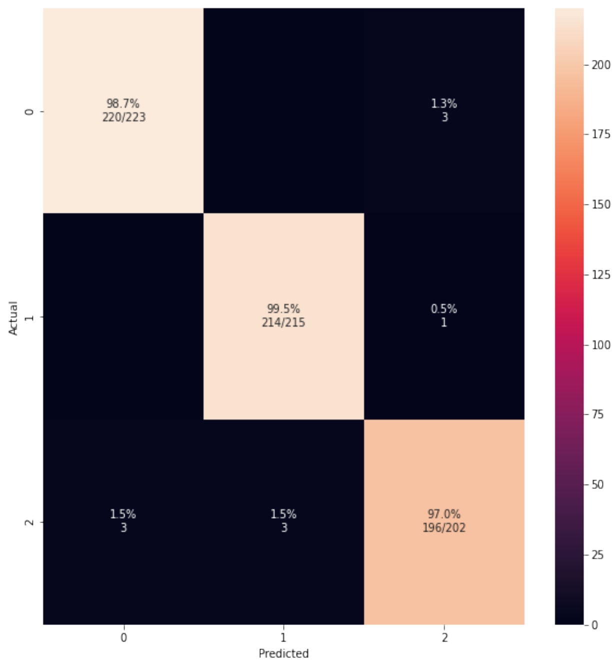


Fig. 13. Confusion Matrix of Stacking 1 Model.

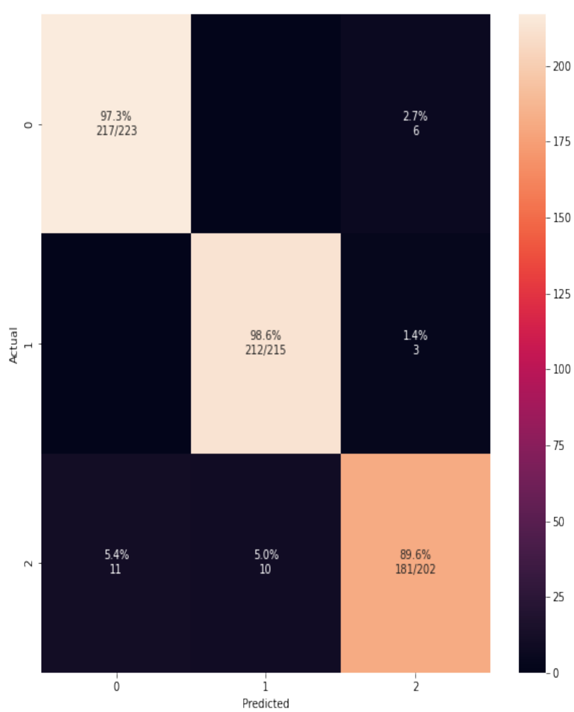


Fig. 14. Confusion Matrix of Stacking 2 Model.

As a result of the increased focus on boosting prediction performance, multi-layer stacking ensembles model depicted to improved predictive performance in this investigation. Various classifications are used such as SVC, MLP, Nu-SVC, RF as a base learner and divided these into two layers stacking 1 and stacking 2, and Random Forest as a meta learner. Eventually, merged these four different models to trade-off various constraints and which provided higher performance. Fig. 8 is a summary of the findings. Evaluated the developed multilayer stacking model to establish state strategies and discovered that suggested approach outperforms them by a wide margin. In contrast, tested proposed models on different datasets and found that proposed method works better compared to other state-of-the-art methods upon these datasets, it's helpful to improve forecast accuracy. Table II shows the comparison state of the art methods for the recognition of neutral, negative, and positive emotions. Accuracy comparison of this research method with other methods of data reduction strategies or feature selection is the recognition of neutral, negative, and positive emotions.

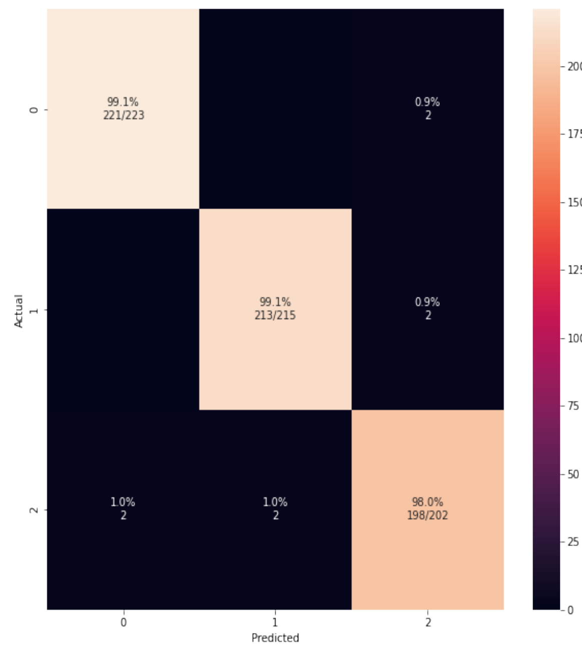


Fig. 15. Confusion Matrix for Multi-layer Stacking.

TABLE II. COMPARISON OF THE PROPOSED MODEL

Study	**Classifiers**	**Dataset**	**Feature Selection**	**Accuracy**
Proposed method	Multi-Layer Stacking	EEG brainwave Data	997 features selected based on LR-CC	98.75
[26]	Deep neural network	EEG brainwave dataset	63 features with information gain	94.89
[17]	Random Forest	DEAP dataset	Features Selected from the time and frequency domains	62.85
[25]	KNN	EEG Brainwave dataset	PCA used for feature selection	96.22
[13]	SVM	DEAP Dataset	Select features based on network entropy measures	68.44

TABLE III. ACCURACIES OF EMOTIONS

Emotions	accuracy	Emotional State (0,1,2)
Happy	99.66	Positive-2
Fear	68.00	Positive-2
Surprise	77.00	positive-2
Sad	98.00	Negative-0
Angry	71.00	Negative-0
Disgust	61.00	Negative-0
Neutral	100	Neutral-1

Table III represents the average accuracy of recognition of emotions in positive, negative, and neutral state using the multi-layer classification approach. Highest accuracy achieved in neutral state as compared to other emotional state.

C. Time Complexity

The amount of time needed for an algorithm to run as a function of the length of the input is known as temporal complexity. It calculates how long each statement of code in an algorithm takes to execute. Calculated time for each and every model and proposed multi-layer stacking model which shown in Table IV.

TABLE IV. TIME COMPLEXITY OF EACH MODEL

Model	Time
SVC	0.76 sec
MLP	1.83 sec
Nu-SVC	1.27 sec
RF	0.27 sec
Stacking 1	4.69 sec
Stacking 2	6.70 sec
Multi-Layer Stacking	19.09 sec

VI. CONCLUSION

In this study, a multi-layer stacking model for boosting the prediction accuracy of emotion recognition using several machine learning techniques to solve a multi-classification problem even with a small dataset. In the first step, Linear Regression-Correlation Coefficient (LR-CC) method is applied for selecting features based on their content which assists in improving forecast accuracy. An original dataset has 2548 features after applying the data fusion technique final dataset of 997 features are selected. The data has been divided into two categories: training and Testing. The training data is used to train the multi-layer stacking, and the testing data is used to make the predictions. Multi-Layer stacking is implemented by two layers, in each layer two base learners or machine learning algorithms are considered like SVM, MLP, Nu-SVC, RF, and set meta learner as a Random Forest. Individual classifiers are also implemented to make comparisons. The 98.75% accuracy is obtained for Multi-Layer Stacking. As compared to the single base learners, the results indicate that the multi-layer stacking model improves the predictive performance. However, proposed method took much time during the training process. In Future, planning to work on these issues to improve the computational cost and focus on multimodal data fusion strategy with higher classification performance.

ACKNOWLEDGMENT

This study received no financing help from public, private, or nonprofit organizations. Our Sincere acknowledgement goes to faculty members of Department of computer science engineering from Christ University.

REFERENCES

[1] Kanjo, L. Al-Husain, and A. Chamberlain, "Emotions in context: examining pervasive affective sensing systems, applications, and analyses," Personal and Ubiquitous Compu 2015, ting, vol. 19, no. 7, pp. 1197-1212.

[2] A. Raheel, M. Majid, M. Alnowami, and S. M. Anwar, "Physiological sensors based emotion recognition while experiencing tactile enhanced multimedia," Sensors, 2020, vol. 20, no. 14, p. 4037.

[3] S. Paul, A. Banerjee, and D. Tibarewala, "Emotional eye movement analysis using electrooculography signal," International Journal of Biomedical Engineering and Technology, 2017, vol. 23, no. 1, pp. 59-70.

[4] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki, "Emotionmeter: A multimodal framework for recognizing human emotions," IEEE transactions on cybernetics,2018, vol. 49, no. 3, pp. 1110-1122.

[5] S. Farashi and R. Khosrowabadi, "EEG based emotion recognition using minimum spanning tree," Physical and Engineering Sciences in Medicine, 2020, vol. 43, no. 3, pp. 985-996.

[6] J. Z. Lim, J. Mountstephens, and J. Teo, "Emotion Recognition Using Eye-Tracking: Taxonomy, Review and Current Challenges," Sensors, 2020, vol. 20, no. 8, p. 2384.

[7] H. Ullah, M. Uzair, A. Mahmood, M. Ullah, S. D. Khan, and F. A. Cheikh, "Internal emotion classification using EEG signal with the sparse discriminative ensemble," IEEE Access, 2019, vol. 7, pp. 40144-40153.

[8] J. Li, Z. Zhang, and H. He, "Hierarchical convolutional neural networks for EEG-based emotion recognition," Cognitive Computation, 2018, vol. 10, no. 2, pp. 368-380.

[9] X. Chai et al., "A fast, efficient domain adaptation technique for cross-domain electroencephalography (EEG)-based emotion recognition," Sensors, 2017, vol. 17, no. 5, p. 1014.

[10] J. Huang, Y.-F. Li, and. Xie, ‘‘An empirical analysis of data preprocessing for machine learning-based software cost estimation,’’ Inf. Softw. Technol., vol. 67, pp. 108–127, Nov. 2015, DOI: 10.1016/j.infsof.2015.07.004.

[11] Adrian Rodriguez Aguinaga, Miguel Angel Lopez Ramir,ez, EZnerd Maria del Rosario Baltazar Flores, “Cl, classification model of arousal and valence mental states by EEG signals analysis and Brodmann correlations” International Journal of Advanced Computer Science and Applications(IJACSA), 6(6), 2015. http://dx.doi.org/10.14569/IJACSA.2015.060633.

[12] Gannouni, S., Aledaily, A., Belwafi, K. et al. Emotion detection using electroencephalography signals and a zero-time windowing-based epoch estimation and relevant electrode identification. Sci Rep 11, 7071 (2021). https://doi.org/10.1038/s41598-021-86345-5.

[13] L. Yao, M. Wang, Y. Lu, H. Li, and X. Zhang, “EEG-Based Emotion Recognition by Exploiting Fused Network Entropy Measures of Complex Networks across Subjects,” Entropy, vol. 23, no. 8, p. 984, Jul. 2021, DOI: 10.3390/e23080984.

[14] Qing, R. Qiao, X. Xu, and Y. Cheng, "Interpretable Emotion Recognition Using EEG Signals," in IEEE Access, vol. 7, pp. 94160-94170, 2019, DOI: 10.1109/ACCESS.2019.2928691.

[15] Isah Salim Ahmad, Shuai Zhang, Sani Saminu, Lingyue Wang, Abd El Kader Isselmou, Ziliang Cai, Imran Javaid, Souha Kamhi, Ummay Kulsum, "Deep Learning Based on CNN for Emotion Recognition Using EEG Signal," WSEAS Transactions on Signal Processing, vol. 17, pp. 28-40, 2021.

[16] Mr. Vijay, M. Khadse. “A Novel Approach of Ensemble Learning with Feature Reduction for Classification of Binary and Multiclass IoT Data.” (2021).

[17] Kusumaningrum, T & Faqih, Akhmad & Kusumoputro, Benyamin. (2020). Emotion Recognition Based on DEAP Database using EEG Time-Frequency Features and Machine Learning Methods. Journal of Physics: Conference Series. 1501. 012020. 10.1088/1742-6596/1501/1/012020.

[18] N. Mahendran, P. M. Durai Raj Vincent, K. Srinivasan, V. Sharma, and D. K. Jayakody, "Realizing a Stacking Generalization Model to Improve the Prediction Accuracy of Major Depressive Disorder in Adults," in IEEE Access, vol. 8, pp. 49509-49522, 2020, DOI: 10.1109/ACCESS.2020.2977887.

[19] Czarnowski and P. Jędrzejowicz, "An approach to machine classification based on stacking generalization and instance selection," 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2016, pp. 004836-004841, DOI: 10.1109/SMC.2016.7844994.

[20] S. Mabu, M. Obayashi, andT. Kuremoto, ‘‘Ensemble learning of rule-based evolutionary algorithm using multi-layer perceptron for supporting decisions in stock trading problems,’’ Appl. Soft Comput., vol. 36, pp. 357–367, Nov. 2015, DOI: 10.1016/j.asoc.2015.07.020.

[21] J. Huang, Y.-F. Li, and. Xie, ‘‘An empirical analysis of data preprocessing for machine learning-based software cost estimation,’’ Inf. Softw. Technol., vol. 67, pp. 108–127, Nov. 2015, DOI: 10.1016/j.infsof.2015.07.004.

[22] Y. Hou, J. Xu, Y. Huang, and. Ma, ‘‘A big data application to predict depression in the university based on the reading habits,’’ in Proc. 3rd Int. Conf. Syst. Information. (ICSAI), Nov. 2016, pp. 1085–1089, DOI: 10.1109/ICSAI.2016.7811112.

[23] Y. Zhou and. Qiu, ‘‘Random forest for label ranking,’’ Expert Syst. Appl., vol. 112, pp. 99–109, Dec. 2018.

[24] J. Xiao, ‘‘SVM and KNN ensemble learning for traffic incident detection,’’ Phys. A Stat. Mech. Appl., vol. 517, pp. 29–35, Mar. 2019, DOI: 10.1016/j.physa.2018.10.060.

[25] Saxena A., Tripathi K., Khanna A., Gupta D., Sundaram S. (2020) Emotion Detection Through EEG Signals Using FFT and Machine Learning Techniques. In: Khanna A., Gupta D., Bhattacharyya S., Snasel V., Platos J., Hassanien A. (eds) International Conference on Innovative Computing and Communications. Advances in Intelligent Systems and Computing, vol 1087. Springer, Singapore. https://doi.org/10.1007/978-981-15-1286-5_46.

[26] J. J. Bird, D. R. Faria, L. J.Manso,., A. Ekárt, & C. D. Buckingham “Mental Emotional Sentiment Classification with an EEG-based Brain-machine Interface.:2019.