Supplementary

I. Features Discussion

In our study, we have extracted 31 statistical features for the design of the ML system from all 69,552 samples of the EEG signals. The brief discussion about these features with their mathematical formulae is given in Table I. For SEED database,the list of statistical features with their ranking is given in Table II. Table II has three columns, indicating the rank, features, and p-value of each of the 31 features. Feature selection is performed using the analysis-of-variance(ANOVA) method. The ANOVA method identifies dominant features that show significant association with the multiclass ground truth variable (in our case it is nine emotions). The ANOVA method ranks the feature based on the p-value. Out of 31 features, 25 features show a significant association with nine emotions. These 25 dominant features are colored in gray showing p-value less than 0.005 and are used to train the multiclass ML algorithms separately for arousal and valence. dimensions.

II. Hyper-parameters

We trained our proposed ML-based system that contains five EML and five CML algorithms using a concept of "nested cross-validation". The inner k-fold cross-validation is used for hyperparameter optimization, whereas the outer k-fold cross-validation is used for evaluating the performance of ML models. In this proposed study we used 5-fold cross-validation for hyperparameter optimization using the randomsearchCV algorithm and 10-fold cross-validation for evaluating the performance of both EML and CML algorithms. For each ML-based algorithm, the randomsearchCV algorithm forms several combinations of the hyperparameters specified in the corresponding grid. Each combination of hyperparameters is evaluated using 5-fold cross-validation with accuracy as the evaluation metric. Finally, a combination of hyperparameters giving the highest AUC value is selected as the best parameters for the particular algorithm to predict emotions. Note that since we are using a concept of "nested cross-validation", we received the optimized hyperparameters for each of the 10 folds. Table III and Table IV show the optimized parameters used to train the 10 ML-based algorithms for arousal and valence on the DREAMER database.

Table V and Table VI show the optimized parameters used to train the 10 ML-based algorithms for arousal and valence on the SEED database.

Table VII and Table VIII show the optimized parameters used to train the 10 ML-based algorithms for arousal and valence on the INTERFACES database.

Table IX and Table X show the optimized parameters used to train the 10 ML-based algorithms for melody and song on the MUSEC database.

III. Description and mathematical formulations for the performance evaluation metrics

This study used six types of performance evaluation metrics such as accuracy, F1-score, sensitivity, specificity, kappa score, and AUC. All the performance evaluation metrics are computed for each of the nine emotions and at last, the average overall emotions is calculated. The performance evaluation metrics are computed using the concept confusion matrix, which takes the predicted and true labels as input and provides the 2*2 matrix as the output. The confusion matrix contains four parameters such as true positive (TP), true negative (TN), false positive (FP), and false-negative (FN). The TP indicates the number of times the classifier has correctly identified the positive class (i.e., true emotion). Similarly, the TN is defined as the number of times the classifier has correctly identified the negative class. FP is the number of data points that are identified by the classifier into a positive class but actually belongs to the negative class as indicated by the ground truth. Similarly, FN is the number of data points that are identified by the classifier into negative class but actually belongs to the positive class as indicated by ground truth. Using these four parameters, the mathematical expressions of all performance evaluation metrics are given in Table XI

Sensitivity is also called recall or true positive rate. The sensitivity indicates the likelihood of detecting positive class emotion for a subject by an automated ML-based algorithm when the gold standard also indicates the positive class status for the same subject. Similarly, **specificity** is also called **True Negative Rate (TNR)**. It indicates the likelihood of detecting the negative class by the automated ML-based algorithm when the gold standard also indicates the negative class. Ideally, both sensitivity and specificity should be 100%. This indicates that all the positive class and negative class, respectively, are correctly identified by the ML-based system. **F1 Score** is used to measure the model's accuracy. An F1 score ranges from 0 to 1. When the F1 score is equal to 1 means we are correctly identifying true class as it has low false positives and low false negatives. On the other hand, when the F1 score is equal to 0 means incorrectly identifying the true class. **Kappa Score** reflects the relationship amongst the noticed classes and the expected classes while adapting to a relationship that happens by chance. **Accuracy** is defined as the ratio of the total number of correctly identified instances to the total number of instances. The receiver operating characteristic (ROC) curve is a plot between true positive rates (sensitivity) against the false-positive rates (100-specificity) at various cut-off points of the ML classifier probability predictions. Each point on the ROC curve shows the sensitivity/specificity pair at a particular cut-off point. The area under the ROC curve (AUC) is a

TABLE I
FEATURES EXTRACTED FROM THE DECOMPOSED DATA WITH THEIR MATHEMATICAL DESCRIPTION.

SN	Feature	Mathematical Formula	Description
1	Minimum	$Minimum = min_{i=1}^{N}(x(i))$	It is minimum value of EEG signal.
2	Maximum	$Maximum = max_{i=1}^{N}(x(i))$	It is maximum value of EEG signal.
3	Mean	$Mean = \frac{1}{N}\sum_{i=1}^{N} x(i)$	It is average value of EEG signal.
4	Mean Absolute Value (MAV)	$MAV = \frac{1}{N}\sum_{i=1}^{N} \lvert x(i)\rvert$	It is average of absolute value of EEG signal.
5	Enhanced MAV (EMAV) [1]	$EMAV = \frac{1}{N}\sum_{i=1}^{N} \lvert x(i)^p\rvert$	It is extension of MAV by including parameter p to identify the influence of sample in a signal.
6	Modified MAV1 (MMAV1) [1]	$\frac{1}{N}\sum_{n=1}^{N} \omega_n \lvert x_n\rvert$ $\omega_n = 1\ if\ 0.25N \leq n \leq 0.75N$ $= 0.5\ otherwise$	It shows the extension over MAV with an weighting window function ω_n.
7	Modified MAV2 (MMAV2) [1]	$\frac{1}{N}\sum_{n=1}^{N} \omega_n \lvert x_n\rvert$ $\omega_n = 1\ if\ 0.25N \leq n \leq 0.75N$ $= 4n/N\ if\ 0.25N > n$ $= 4(n-N)/N\ if\ 0.75N < n$	It is similar to MAV1 with the change that smoothness of window function ω_n is improved.
8	Mean Absolute Deviation (MAD)	$MAD = \frac{1}{n}\sum_{i=1}^{n} \lvert x_i - m(X)\rvert$ *m(X)* = average value of the data set n=number of data values x_i = data values in the set	It calculate the mean of the absolute deviations from a central point.
9	Root Mean Square (RMS)	$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \lvert x(i)^2\rvert}$	It measures the variations in the amplitude of a signal.
10	Median	$\widetilde{x} = x(\frac{N+1}{2})\ if\ N\ is\ odd$ $= \frac{x(\frac{N}{2})+x(\frac{N}{2})+1}{2}\ if\ N\ is\ even$	Median value in the signal.
11	Median Absolute Deviation (MAD) [2]	$MAD = Median(\lvert X_i - \tilde{X}\rvert)$ X_i=each value $\tilde{X}$=average value	It measure the changeability of a uni-variate sample of quantitative data.
12	Standard Deviation (STD) [2]	$\sigma = \sqrt{\frac{\sum_{i=1}^{N}[E(x(i)-\mu)^2}{N}}$	Standard deviation defines the spreadness of a signal.
13	Difference Absolute STD Value (DASDV) [3]	$DASDV = \sqrt{\frac{\sum_{i=1}^{N-1}(x(i)-x(i-1))^2}{N-1}}$	It is a standard deviation value of the wavelength.
14	Variance (VAR)	$VAR = \frac{1}{N-1}\sum_{i=1}^{N} \lvert x(i)^2\rvert$	It is a measure of statistical dispersion of a random variable.
15	Coefficient Of Variance (COV)	$COV = \frac{\sigma}{\mu}$ σ=population standard deviation μ=population mean	It is a standardized estimate of dispersion of a probability distribution.
16	Skewness (s) [4]	$s = \frac{\sqrt{[E(x(i)-\mu)^3}}{\sigma^3}$	It measures the asymmetry around the mean of a signal.
17	Kurtosis (k) [4]	$k = \frac{\sqrt{[E(x(i)-\mu)^4}}{\sigma^4}$	It is a measure of a degree to which the distribution possesses peaks or 'peakedness'.
18	Waveform Length (WL)	$WL = \sum_{n=1}^{N-1} \lvert x(i+1) - x(i)\rvert$	It is summation of cumulative length of a signal.
19	Enhanced Wavelength (EWL)	$EWL = \frac{1}{N}\sum_{i=1}^{N} \lvert(x(i) - x(i-1))^p\rvert$ $p = 0.75\ if\ i \geq 0.2N\ \&\ i \leq 0.8N$ $= 0.50\ otherwise$	It is extension of WL with inclusion of parameter p to identify the influence of sample in a signal.
20	Interquartile Range (IQR) [2]	IQR=$Q_3 - Q_1$ Q_3 = 1st quartile Q_1 = 3rd quartile	It measures where the bulk of the values lie.
21	Average Amplitude Change (AAC) [5]	$AAC = \frac{1}{N}\sum_{i=1}^{N-1} \lvert x(i+1) - x(i)\rvert$	It is average of cumulative length of the waveform over time.
22	Log Detector [6]	$LG = exp(\frac{1}{N}\sum_{i=1}^{N} log(\lvert x(i)\rvert))$	It gives an estimation of contraction force of muscle.
23	Myopulse % rate (MYOP) [6]	$MYOP = \frac{1}{N}\sum_{i=1}^{N}[x(i)]$ f(x) = 1 if x(i) $\geq$ threshold 0 otherwise	It measures the mean value of myopulse output.
24	Simple Square Integral (SSI) [6]	$SSI = \sum_{i}^{N} x(i)^2$	It is summation of square values of EEG signal amplitude.
25	Willison Amplitude (WA) [1]	$WA = \frac{1}{N}\sum_{i=1}^{N}[f\lvert x(i) - x(i+1)\rvert]$ $f(x) = 1\ if\ x(i) \geq threshold$ $= 0\ otherwise$	Number of times the difference between EEG signal amplitude among two adjacent segments exceeds a predefined threshold to reduce noise effects.
26	Maximum Fractal Length(MFL) [6]	$MFL = log_{10}(\sqrt{(\sum_{i=1}^{N-1}(x(i+1) - x(i-1))^2})$	It measures the strength of contraction.
27	Slope Sign Change (SSC) [1]	$SSC = \sum_{i=2}^{N-1}[f[(x(i)-x(i-1))*(x(i)-x(i+1))]]$ $f(x) = 1\ if\ x(i) \geq threshold$ $= 0\ otherwise$	The number of changes between positive and negative slope among three consecutive segments.
28	Zero Crossing (ZC) [1]	$ZC = \sum_{i=1}^{N-1} sgn(\lvert x(i)*x(i+1)\rvert \cap \lvert x(i)*x(i+1)\rvert)$ $sgn(x) = 1\ if\ x(i) \geq threshold$ $= 0\ otherwise$	Number of times that the amplitude value of EEG signal crosses the zero y-axis.
29	Entropy E(p)	$E(p) = -\sum_i p_i * log(p_i)$ p_i = probability of occurrence	It quantifies the randomness in the information being computed.
30	Shannon Entropy (SE) [7]	$SE = -\sum_{i=0}^{N} p_i\ (\hat{x})\ .log\ (p_i\ (\hat{x}))$	It is a measure of degree of uncertainty of occurrence of a specific event.
31	Log Energy Entropy [7]	$LEE = -\sum_{i=0}^{N-1} log_2\ (p_i\ (x))^2$	It measures the complexity of the EEG signals.

TABLE II
LIST OF STATISTICAL FEATURES ALONG WITH THEIR RANKING FOR SEED DATABASE.

Sr#	Feature Name	P
1	Minimum	<0.0001
2	Maximum	<0.0001
3	Standard Deviation	<0.0001
4	Kurtosis	<0.0001
5	Mean Absolute Deviation	<0.0001
6	Median Absolute Deviation	<0.0001
7	Inter-Quartile Range	<0.0001
8	Shannon Entropy	<0.0001
9	Entropy	<0.0001
10	Enhanced Mean Absolute Value	<0.0001
11	Enhanced Wavelength	<0.0001
12	Mean Absolute Value	<0.0001
13	Root Mean Square	<0.0001
14	Log Detector	<0.0001
15	Modified Mean Absolute Value 1	<0.0001
16	Modified Mean Absolute Value 2	<0.0001
17	Myopulse Percentage Rate	<0.0001
18	Simple Square Integral	<0.0001
19	Willison Amplitude	<0.0001
20	Maximum Fractal Length	<0.0001
21	Slope Sign Change	<0.0001
22	Zero Crossing	<0.0001
23	Variance	<0.0001
24	Wavelength	<0.0001
25	Log-Entropy	<0.0001
26	Covariance	0.16
27	Average Amplitude Change	0.54
28	Skewness	0.59
29	Median	0.64
30	Mean	0.83
31	Difference Absolute STD Value	0.85

TABLE III
OPTIMAL HYPERPARAMETERS USING THE RANDOMSEARCHCV ALGORITHM FOR AROUSAL ON THE DREAMER DATABASE.

Algorithm	Hyperparameters	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
SVM	kernel	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf
	gamma	1	1	1	1	1	1	1	1	1	1
	degree	1	1	1	1	1	1	1	1	1	1
	C	2	2	2	2	2	2	2	2	2	2
LR	penalty	l2	l2	l2	l2	l2	l2	l2	l2	l2	l2
	C	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
LDA	solver	svd	svd	svd	svd	svd	svd	svd	svd	svd	svd
QDA	reg_param	0.1	0.1	0.5	0.4	0.1	0.5	0.1	0.5	0.1	0.1
KNN	p	2	2	2	2	2	2	2	2	2	2
	n_neighbors	6	6	1	6	6	1	1	1	1	1
	leaf_size	2	2	4	2	2	4	7	7	4	7
Bagging	n_estimators	1000	1000	200	500	1000	1000	500	500	1000	1000
RF	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_weight_fraction_leaf	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	min_samples_split	10	2	10	2	2	5	2	2	2	10
	min_samples_leaf	5	2	5	2	2	10	2	2	2	5
	max_leaf_nodes	40	40	40	40	40	40	40	40	40	40
	max_depth	5	5	5	5	5	4	5	5	5	5
AdaBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
RFE	n_estimators	100	100	100	100	300	100	100	100	100	100
	K	5	5	5	5	5	5	5	5	5	5
XGBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_child_weight	3	3	3	3	3	3	3	3	3	3
	max_depth	9	9	9	9	9	9	9	9	9	9
	learning_rate	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	gamma	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
	colsample_bytree	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6

TABLE IV
OPTIMAL HYPERPARAMETERS USING THE RANDOMSEARCHCV ALGORITHM FOR VALENCE ON THE DREAMER DATABASE.

Algorithm	Hyperparameters	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
SVM	kernel	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf
	gamma	1	1	1	1	1	1	1	1	1	1
	degree	1	1	1	1	1	1	1	1	1	1
	C	2	2	2	2	2	2	2	2	2	2
LR	penalty	l2	l2	l2	l2	l2	l2	l2	l2	l2	l2
	C	10000.0	1291.5	1291.5	10000.0	1291.5	166.8	1291.5	10000.0	1291.5	1291.5
LDA	solver	svd	svd	svd	svd	svd	svd	svd	svd	svd	svd
QDA	reg_param	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
KNN	p	2	2	2	2	2	2	2	2	2	2
	n_neighbors	1	1	1	1	1	1	1	1	1	1
	leaf_size	7	7	7	7	7	4	7	7	4	4
Bagging	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
RF	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_weight_fraction_leaf	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	min_samples_split	10	2	2	2	10	10	2	2	10	10
	min_samples_leaf	5	2	2	2	5	5	2	2	5	5
	max_leaf_nodes	40	40	40	20	40	40	40	20	40	40
	max_depth	5	4	4	4	5	5	5	4	5	5
AdaBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
RFE	n_estimators	100	100	100	100	300	100	100	100	100	100
	K	5	5	5	5	5	5	5	5	5	5
XGBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_child_weight	3	3	3	3	3	3	3	3	3	3
	max_depth	9	9	9	9	9	9	9	9	9	9
	learning_rate	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	gamma	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
	colsample_bytree	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6

TABLE V
OPTIMAL HYPERPARAMETERS USING THE RANDOMSEARCHCV ALGORITHM FOR AROUSAL ON THE SEED DATABASE.

Algorithm	Hyperparameters	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
SVM	kernel	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf
	gamma	1	1	1	1	1	1	1	1	1	1
	degree	1	1	1	1	1	1	1	1	1	1
	C	2	2	2	2	2	2	2	2	2	2
LR	penalty	l2	l2	l2	l2	l2	l2	l2	l2	l2	l2
	C	1291.55	1291.55	10000	10000	10000	1291.55	10000	10000	166.8101	1291.55
LDA	solver	svd	svd	svd	svd	svd	svd	svd	svd	svd	svd
QDA	reg_param	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
KNN	p	2	2	2	2	2	2	2	2	1	2
	n_neighbors	6	6	6	6	6	6	6	6	8	8
	leaf_size	2	2	2	2	2	2	2	2	9	3
Bagging	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
RF	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_weight_fraction_leaf	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	min_samples_split	10	10	10	10	10	10	10	10	10	10
	min_samples_leaf	5	5	5	5	5	5	5	5	5	5
	max_leaf_nodes	40	40	40	40	40	40	40	40	40	40
	max_depth	5	5	5	5	5	5	5	5	5	5
AdaBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
RFE	n_estimators	100	100	100	100	300	100	100	100	100	100
	K	5	5	5	5	5	5	5	5	5	5
XGBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_child_weight	1	7	1	1	7	7	7	1	7	1
	max_depth	7	7	7	7	7	5	7	7	7	7
	learning_rate	0.01	0.1	0.01	0.01	0.1	0.01	0.1	0.01	0.1	0.01
	gamma	0.1	0.2	0.1	0.1	0.2	0.2	0.2	0.1	0.2	0.1
	colsample_bytree	0.8	0.7	0.8	0.8	0.7	0.8	0.7	0.8	0.7	0.8

powerful tool that measures the discrimination ability of the ML classifier. The idea AUC should be unity in which the ROC curve passes through the upper left corner of the plot indicating both sensitivity and specificity of 100%.

In this study, we presented the average performance evaluation metrics over nine emotions (N=9), which are computed using the following mathematical expression.

$$\overline{PerformanceMetric} = \frac{1}{N}\sum_{i=1}^{N} PE \qquad (1)$$

IV. RESULT VALIDATION ON INTERFACES AND MUSEC DATABASE

To validate the performance of our proposed ML-based system, the same experimental protocol used for the DREAMER database is followed on the INTERFACES and MUSEC database. For both database, the same 31 statistical extracted features are then used to train our proposed ML-based system that performs emotion recognition using the EML and CML algorithms. The comparison of CML and EML algorithms for emotion recognition on the basis of six performance evaluation

TABLE VI

OPTIMAL HYPERPARAMETERS USING THE RANDOMSEARCHCV ALGORITHM FOR VALENCE ON THE SEED DATABASE.

Algorithm	Hyperparameters	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
	kernel	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf
SVM	gamma	1	1	1	1	1	1	1	1	1	1
	degree	1	1	1	1	1	1	1	1	1	1
	C	2	2	2	2	2	2	2	2	2	2
LR	penalty	l2	l2	l2	l2	l2	l2	l2	l2	l2	l2
	C	1291.5	10000.0	166.8	1291.5	166.8	10000.0	2.8	1291.5	1291.5	10000.0
LDA	solver	svd	svd	svd	lsqr	svd	svd	svd	svd	svd	svd
QDA	reg_param	0.2	0.1	0.1	0.1	0.1	0.1	0.2	0.2	0.2	0.2
KNN	p	1	1	1	2	1	1	2	1	2	2
	n_neighbors	8	8	8	8	8	8	8	8	8	8
	leaf_size	9	9	9	3	9	9	3	9	3	3
Bagging	n_estimators	1000	1000	1000	500	500	200	200	1000	1000	1000
RF	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_weight_fraction_leaf	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	min_samples_split	2	2	2	2	2	2	2	2	2	2
	min_samples_leaf	5	5	5	5	5	5	5	5	5	5
	max_leaf_nodes	20	20	20	20	20	20	20	20	20	20
	max_depth	3	3	3	3	3	3	3	3	3	3
AdaBoost	n_estimators	1000	1000	1000	900	1000	1000	900	1000	1000	1000
RFE	n_estimators	100	100	100	100	300	100	100	100	100	100
	K	5	5	5	5	5	5	5	5	5	5
XGBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_child_weight	3	3	3	3	7	7	3	3	3	3
	max_depth	9	9	9	9	7	5	9	9	9	9
	learning_rate	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	gamma	0.2	0.2	0.2	0.2	0.2	0.3	0.2	0.2	0.2	0.2
	colsample_bytree	0.6	0.6	0.6	0.6	0.7	0.7	0.6	0.6	0.6	0.6

TABLE VII

OPTIMAL HYPERPARAMETERS USING THE RANDOMSEARCHCV ALGORITHM FOR AROUSAL ON THE INTERFACES DATABASE.

Algorithm	Hyperparamters	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
SVM	kernel	linear	poly	linear	poly	poly	rbf	linear	rbf	poly	poly
	gamma	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	degree	2	3	2	3	3	1	2	1	3	3
	C	0.001	0.01	0.001	0.01	0.01	0.1	0.001	0.1	0.01	0.01
LR	penalty	l2	l2	l2	l2	l2	l2	l2	l2	l2	l2
	C	0.0001	0.0001	0.0001	0.0001	0.0001	2.782559	0.0001	0.0001	0.0001	0.0001
LDA	solver	svd	svd	svd	svd	svd	svd	svd	svd	svd	svd
QDA	reg_param	0.3	0.5	0.4	0.4	0.5	0.5	0.5	0.5	0.5	0.5
KNN	p	1	2	2	2	2	2	2	2	2	2
	n_neighbors	8	8	8	8	8	8	8	6	8	8
	leaf_size	9	3	3	3	3	3	3	2	3	3
Bagging	n_estimators	1000	200	200	1000	1000	1000	500	100	100	1000
RF	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_weight_fraction_leaf	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	min_samples_split	10	10	10	10	10	10	10	10	10	10
	min_samples_leaf	5	5	5	5	5	5	5	5	5	5
	max_leaf_nodes	40	40	40	40	40	40	40	40	40	40
	max_depth	5	5	5	5	5	5	5	5	5	5
AdaBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
RFE	n_estimators	100	100	100	100	300	100	100	100	100	100
	K	5	5	5	5	5	5	5	5	5	5
XGBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_child_weight	5	5	5	5	5	5	5	5	5	5
	max_depth	3	3	3	3	3	3	3	3	3	3
	learning_rate	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	gamma	0.2	0.3	0.3	0	0.2	0.2	0.2	0.3	0.3	0.2
	colsample_bytree	0.6	0.7	0.7	0.8	0.6	0.6	0.6	0.7	0.7	0.6

metrics for INTERFACES and MUSEC database is given in Table XII and Table XIII.

REFERENCES

[1] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for emg signal classification," *Expert systems with applications*, vol. 39, no. 8, pp. 7420–7431, 2012.

[2] K. Biju, H. A. Hakkim, and M. Jibukumar, "Ictal eeg classification based on amplitude and frequency contours of imfs," *Biocybernetics and Biomedical Engineering*, vol. 37, no. 1, pp. 172–183, 2017.

[3] S. K. Khare and V. Bajaj, "An evolutionary optimized variational mode decomposition for emotion recognition," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 2035–2042, 2020.

[4] S. Taran and V. Bajaj, "Emotion recognition from single-channel eeg signals using a two-stage correlation and instantaneous frequency-based filtering method," *Computer methods and programs in biomedicine*, vol. 173, pp. 157–165, 2019.

[5] A. Seal, P. P. N. Reddy, P. Chaithanya, A. Meghana, K. Jahnavi, O. Krejcar, and R. Hudak, "An eeg database and its initial benchmark emotion classification performance," *Computational and mathematical methods in medicine*, vol. 2020, 2020.

[6] A. Moura, S. Lopez, I. Obeid, and J. Picone, "A comparison of feature extraction methods for eeg signals," in *2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, 2015, pp. 1–2.

[7] S. Aydın, H. M. Saraoğlu, and S. Kara, "Log energy entropy-based eeg classification with multilayer neural networks in seizure," *Annals of biomedical engineering*, vol. 37, no. 12, p. 2626, 2009.

TABLE VIII
OPTIMAL HYPERPARAMETERS USING THE RANDOMSEARCHCV ALGORITHM FOR VALENCE ON THE INTERFACES DATABASE.

Algorithm	Hyperparameters	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
SVM	kernel	linear	linear	linear	linear	linear	linear	linear	linear	linear	linear
	gamma	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	degree	2	2	2	2	2	2	2	2	2	2
	C	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
LR	penalty	l2	l2	l2	l2	l2	l2	l2	l2	l2	l2
	C	0.4	0.4	0.4	2.8	2.8	2.8	2.8	0.4	2.8	21.5
LDA	solver	lsqr	lsqr	lsqr	lsqr	lsqr	lsqr	svd	svd	svd	lsqr
QDA	reg_param	0.5	0.3	0.5	0.5	0.4	0.5	0.5	0.5	0.3	0.4
KNN	p	2	2	2	1	1	2	2	2	2	2
	n_neighbors	8	6	6	8	8	6	8	6	6	6
	leaf_size	3	2	2	9	9	2	3	2	2	2
Bagging	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
RF	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_weight_fraction_leaf	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	min_samples_split	10	10	10	5	10	10	10	10	10	10
	min_samples_leaf	5	5	5	10	5	5	5	5	5	5
	max_leaf_nodes	40	40	40	40	40	40	40	40	40	40
	max_depth	5	5	5	4	5	5	5	5	5	5
RFE	n_estimators	100	100	100	100	300	100	100	100	100	100
	K	5	5	5	5	5	5	5	5	5	5
AdaBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
XGBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_child_weight	5	5	5	5	5	5	5	5	5	5
	max_depth	3	3	3	3	3	3	3	3	3	3
	learning_rate	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	gamma	0.2	0.3	0.2	0.3	0.2	0.2	0.3	0.2	0.2	0.3
	colsample_bytree	0.6	0.7	0.6	0.7	0.6	0.6	0.7	0.6	0.6	0.7

TABLE IX
OPTIMAL HYPERPARAMETERS USING THE RANDOMSEARCHCV ALGORITHM FOR MELODY ON THE MUSEC DATABASE.

Algorithm	Hyperparameters	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
SVM	kernel	linear	linear	linear	linear	linear	linear	linear	linear	linear	linear
	gamma	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	degree	2	2	2	2	2	2	2	2	2	2
	C	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
LR	penalty	l2	l2	l2	l2	l2	l2	l2	l2	l2	l2
	C	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
LDA	solver	svd	lsqr	svd	svd	svd	svd	svd	svd	svd	svd
QDA	reg_param	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
KNN	p	2	1	1	1	1	1	1	1	1	1
	n_neighbors	8	8	8	8	8	8	8	8	8	8
	leaf_size	3	9	9	9	9	9	9	9	9	9
Bagging	n_estimators	1000	1000	1000	1000	500	1000	1000	1000	1000	1000
RF	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_weight_fraction_leaf	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	min_samples_split	10	10	10	10	10	10	10	10	10	10
	min_samples_leaf	5	5	5	5	5	5	5	5	5	5
	max_leaf_nodes	40	40	40	40	40	40	40	40	40	40
	max_depth	5	5	5	5	5	5	5	5	5	5
AdaBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
RFE	n_estimators	100	100	100	100	300	100	100	100	100	100
	K	5	5	5	5	5	5	5	5	5	5
XGBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_child_weight	5	5	5	5	5	5	5	5	5	5
	max_depth	3	3	3	3	3	3	3	3	3	3
	learning_rate	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	gamma	0.2	0.2	0.2	0.3	0.2	0.2	0.2	0.3	0.2	0.2
	colsample_bytree	0.6	0.6	0.6	0.7	0.6	0.6	0.6	0.7	0.6	0.6

TABLE X
OPTIMAL HYPERPARAMETERS USING THE RANDOMSEARCHCV ALGORITHM FOR THE SONG ON THE MUSEC DATABASE.

Algorithm	Hyperparameters	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
SVM	kernel	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf	rbf
	gamma	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	degree	1	1	1	1	1	1	1	1	1	1
	C	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
LR	penalty	l2	l2	l2	l2	l2	l2	l2	l2	l2	l2
	C	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
LDA	solver	svd	svd	svd	svd	lsqr	svd	svd	svd	svd	svd
QDA	reg_param	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
KNN	p	1	1	1	1	1	1	1	1	1	1
	n_neighbors	6	6	6	6	6	6	6	6	6	6
	leaf_size	6	6	6	6	6	6	6	6	6	6
Bagging	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
RF	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_weight_fraction_leaf	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
	min_samples_split	5	5	5	5	5	2	2	5	5	5
	min_samples_leaf	10	10	10	10	10	5	5	10	10	10
	max_leaf_nodes	40	40	40	40	40	20	20	40	40	40
	max_depth	4	4	4	4	4	3	3	4	4	4
AdaBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
RFE	n_estimators	100	100	100	100	300	100	100	100	100	100
	K	5	5	5	5	5	5	5	5	5	5
XGBoost	n_estimators	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
	min_child_weight	5	5	5	5	5	5	5	5	5	5
	max_depth	3	3	3	3	3	3	3	3	3	3
	learning_rate	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	gamma	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
	colsample_bytree	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6

TABLE XI
MATHEMATICAL EXPRESSIONS FOR THE PERFORMANCE EVALUATION METRICS.

SN	Name of Metric	Mathematical Expression
1	Sensitivity, Se (%)	$Se(\%) = \frac{TP_i}{TP_i+FP_i} * 100$
2	Specificity, Sp (%)	$Sp(\%) = \frac{TN_i}{TN_i+FP_i} * 100$
3	F1-score, F1 (%)	$F1 = \frac{TP}{TP+\frac{1}{2}(FP+FN)}$
4	Accuracy, Acc (%)	$Acc = \frac{TP+TN}{TP+FP+TN+FN}$
5	Kappa score, k	$k = \frac{Acc-P_e}{1-P_e}$ $P_e = P_{YES} + P_{NO}$ $P_{YES} = \frac{FP+TP}{TP+FP+TN+FN} * \frac{FN+TP}{TP+FP+TN+FN}$ $P_{NO} = \frac{FP+TN}{TP+FP+TN+FN} * \frac{FN+TN}{TP+FP+TN+FN}$

TABLE XII
COMPARING THE CML AND EML ALGORITHMS FOR EMOTION RECOGNITION ON THE INTERFACES DATABASE.

Sr#	Type of Algorithm	Algorithm	Sensitivity (%)		Specificity (%)		Kappa-Score		F1-Score		AUC		Accuracy (%)	
			Arousal	Valence	Arousal	Valence	Arousal	Valence	Arousal	Valence	Arousal	Valence	Arousal	Valence
1	CML	QDA	27.97	45.68	74.5	56.11	0.03	0.02	0.30	0.41	0.52	0.51	56.4	51.59
2		KNN	32.26	39.07	71.88	64.36	0.04	0.04	0.37	0.42	0.53	0.52	56.46	53.42
3		LR	2.8	0.38	98.06	99.42	0.01	0	0.05	0.01	0.56	0.54	60.99	56.58
4		SVM	0.42	3.44	99.75	97.68	0	0.01	0.01	0.06	0.55	0.54	61.1	56.91
5		LDA	5.01	5.42	96.89	96.38	0.02	0.02	0.09	0.10	0.56	0.54	61.13	57.04
		Average	**13.69**	**18.80**	**88.22**	**82.79**	**0.02**	**0.02**	**0.16**	**0.20**	**0.54**	**0.53**	**59.22**	**55.11**
6	EML	XGBoost	32.99	38.15	73.7	65.54	0.07	0.04	0.38	0.42	0.56	0.53	57.86	53.69
7		RotationForest	31.8	37.13	76.49	66.34	0.09	0.04	0.38	0.41	0.56	0.53	59.1	53.71
8		Bagging	24.53	21.64	82.71	80.7	0.08	0.03	0.32	0.29	0.58	0.53	60.07	55.15
9		RF	25.73	33.33	82.42	72.49	0.09	0.06	0.34	0.39	0.58	0.55	60.36	55.55
10		AdaBoost	14.08	32.38	90.82	73.72	0.06	0.06	0.22	0.39	0.56	0.55	60.96	55.84
		Average	**25.83**	**32.53**	**81.23**	**71.76**	**0.08**	**0.05**	**0.33**	**0.38**	**0.57**	**0.54**	**59.67**	**54.79**

TABLE XIII

COMPARING THE CML AND EML ALGORITHMS FOR EMOTION RECOGNITION ON THE MUSEC DATABASE

Sr#	Type of Algorithm	Algorithm	Sensitivity (%)		Specificity (%)		Kappa- Score		F1-Score (%)		AUC		ACC (%)	
			Melody	Song	Melody	Song	Melody	Song	Melody	Song	Melody	Song	Melody	Song
1	CML	QDA	74.24	42.07	37.89	96.18	0.08	0.43	0.37	0.57	0.6	0.74	45.93	75.94
2		LDA	1.77	42.83	99.5	98.21	0.02	0.46	0.03	0.59	0.58	0.75	77.87	77.49
3		LR	1.27	51.78	99.74	93.19	0.02	0.49	0.02	0.63	0.57	0.76	77.95	77.7
4		SVM	3.33	49.36	99.84	95.13	0.05	0.49	0.06	0.63	0.66	0.75	78.49	78.01
5		KNN	34.8	65.19	92.92	95.73	0.32	0.65	0.44	0.76	0.74	0.89	80.06	84.31
		Average	**23.08**	**50.25**	**85.98**	**95.69**	**0.10**	**0.50**	**0.19**	**0.63**	**0.63**	**0.78**	**72.06**	**78.69**
6	EML	AdaBoost	8.91	48.84	97.84	94.94	0.1	0.48	0.15	0.62	0.64	0.75	78.16	77.69
7		XGBoost	33.41	51.17	95.84	96.79	0.36	0.53	0.45	0.65	0.79	0.78	82.03	79.72
8		RF	24.82	49.74	99.01	97.94	0.32	0.53	0.39	0.65	0.83	0.76	82.6	79.91
9		Bagging	26.59	66.4	98.77	97.97	0.34	0.68	0.41	0.78	0.82	0.93	82.8	86.16
10		RotationForest	37.16	72.4	97.59	97.1	0.43	0.73	0.51	0.82	0.79	0.76	84.22	87.86
		Average	**26.18**	**57.71**	**97.81**	**96.95**	**0.31**	**0.59**	**0.38**	**0.70**	**0.77**	**0.80**	**81.96**	**82.27**
Improvement for EML over CML (%)			**3.10**	**7.46**	**11.83**	**1.26**	**216.33**	**17.06**	**104.68**	**11.07**	**22.86**	**2.31**	**9.90**	**3.58**