Journal of Scientific & Industrial Research
Vol. 84, January 2025, pp. 36-47
DOI: 10.56042/jsir.v84i1.10829

Ensemble Learning based EEG Classification – Investigating the Effects of Combined Yoga and Rajyog Meditation

Shobhika Madhu[1,2]*, Prashant Kumar[1,2] & Sushil Chandra[3]

[1]CSIR-Central Scientific Instruments Organisation, Sector 30 C, Chandigarh 160 030, India

[2]Academy of Scientific & Innovative Research (AcSIR), Ghaziabad 201 002, India

[3]Rishihood University, Sonipat 131 021, Haryana, India

Received 01 May 2024; revised 11 September 2024; accepted 19 October 2024

The ability to detect and prevent mental health deterioration has been one of the major achievements of digital psychiatry using artificial intelligence and machine learning. The aim of this paper is to address the issue of preventing the mental health disorders of young generation by developing a system to predict the changes in an individual's states of psychological health. Pre-and post-yoga and Rajyoga meditation states were analyzed for classification of data. Also, the paper investigates if bidirectional long-short-term memory BiLSTM-based ensemble models outperform the CNN-based models in prediction modeling. The EEG data was collected from 69 students for pre- and post-intervention. To determine an objective marker for yoga and meditation, collected data were analyzed using spectrum analysis, and classification. The post meditation group exhibited highest band powers and wavelet coefficients, indicating the differences in meditation and control conditions. Additionally, in this study, an ensemble model classifier has been developed utilizing EEG data that was more accurate (82%) than other models at differentiating between meditation and control situations. To the best of the knowledge of the authors, this is the first research to apply ensemble model-based classifiers to distinguish between states of meditation and non-meditation. The performance of BiLSTM-DT was the highest among all other models in terms of precision, recall, f-measure, and accuracy. Therefore, the BiLSTM-DT ensemble model is a viable objective marker for psychological health states.

Keywords: BiLSTM, CNN, Ensemble models, Mental health, Rajyoga meditation

Introduction

It is crucial to recognize and identify vulnerable people in the early phases of acute stress in order to stop the emergence of more severe long-term mental health conditions including depression, suicidal thoughts and behaviors, and PTSD. Mental illnesses are hard to diagnose, and much harder to forecast because of lack of bio markers and subjectivity in humans. One of the most significant effects of digital psychiatry, equipped with artificial intelligence and machine learning algorithms, in these situations is its capacity to detect and anticipate the decline in individual’s mental health, which can result in chronic mental health illnesses. Psychiatry powered by AI might assist mental health professionals in objectively redefining mental diseases. Healthcare system applications of machine and deep learning techniques have shown tremendous promise. These cutting-edge prediction algorithms have been effectively used for various tasks, including managing medical data, mental health analysis, and illness prediction. The effectiveness and robustness of predictions have been significantly increased by using ensemble learning, a potent approach that mixes numerous models. In recent years, addressing complicated prediction tasks has shown tremendous promise by combining deep learning and machine learning models inside ensemble frameworks.[1] Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), in particular, have revolutionized several areas by successfully extracting complex patterns and representations from high-dimensional data. These models are exceptional in detecting, recognizing, and classifying data. In contrast, machine learning algorithms like K-Nearest Neighbors (KNN), decision trees, and Random Forests (RF) have a long history of producing reliable and understandable predictions for various applications. The complementing advantages of both methods may be taken advantage of by mixing deep learning and machine learning models in an ensemble. In contrast to machine learning models, which offer interpretability,

*Author for Correspondence
E-mail: madhushobhika@gmail.com, shobhika.csio19a@acsir.res.in

generalization abilities, and resistance to noise and outliers, deep learning models excel in extracting complicated connections and feature representations from raw data. By utilizing the diversity and complementary character of the models, the ensemble framework produces a synergistic effect that enhances prediction accuracy.[2]

Complementary health interventions such as yoga and meditation have gained popularity as a productive way to improve cognitive function, lower stress, and promote mental health.[3] To offer unbiased insights into yoga's advantages, it is necessary to analyze the changes in mental health before, objectively and after practice. Tests are conducted on a range of physiological and psychological parameter in order to comprehend the many sensations and consequences linked to meditation. EEG readings, which record the brain's electrical activity, provide helpful information on neurodynamic underlying mental health disorders. Researchers are now a days able to assess the ability of meditation using continuous EEG data. In order to comprehend the neurological correlations involved in the meditation process, it is crucial to investigate the viability of employing machine learning algorithms to compare the states of meditation with non-meditative states.

The major contributions of this study are: 1) To evaluate EEG during control and meditation condition as potential markers of meditation ability and mental health of college students, 2) to evaluate the effectiveness of hybrid machine learning and deep learning-based classification algorithms using the provided data, and 3) to propose the most efficient algorithm for prediction of varying psychological health states depending on the band powers yielded by EEG data. While a large number of research have documented the categorization of mental health states using different models, classification studies using ensemble model classifiers are rare.

Related Work

Various machine learning and deep learning algorithms have been deployed for prediction of mental health states in research studies. Recent research on existing artificial intelligence based mental health prediction systems has been summarised in Table 1.

Khare *et al.*[4] investigated EEG data to identify Attention Deficit Hyperactivity Disorder (ADHD). The study suggests using a technique known as VMD-HT—variational mode decomposition combined with the Hilbert transform—to extract hidden information from EEG data. The Explainable Boosted Machine (EBM) model, which enables interpretability and explainability of the model's predictions, is used to classify the extracted features. In automatically identifying ADHD, the model achieves excellent accuracy, sensitivity, specificity, and precision. The interpretability analysis emphasizes the significance of several brain areas in identifying ADHD. The results indicate that the created model is accurate, robust, interpretable, and explicable, offering medical professionals helpful information in identifying ADHD in youngsters.

Balasubramanian *et al.*[5] presented Adaptive Neuro-Fuzzy Inference System (ANFIS) to accurately forecast schizophrenia from multi-channel EEG data. Pre-processing the EEG data, extracting features, choosing distinguishing features, and using the ANFIS for classification are all steps in the procedure. The Hybrid Grey Wolf-Bat Algorithm increases accuracy and efficiency by optimizing the ANFIS settings. Results from experiments reveal how the suggested method is superior to previous approaches, highlighting its potential for computer-assisted diagnosis and better patient outcomes in schizophrenia.

Emre *et al.*[6] worked upon the machine learning techniques to distinguish and categorize the illnesses using a dataset of EEG recordings of people with various mental diagnoses. Random Forest (RF), C5.0, Support Vector Machine (SVM), and Artificial Neural Networks (ANN) are just a few of the machine learning models that were used. The hyperparameters in the models were additionally optimised by the use of the 5-fold Cross Validation (CV) approach, which iterates three times across the training dataset. Various models have been created using the chosen hyperparameters, and the test dataset has been used to evaluate each model's performance. The classification of the illness categories showed great accuracy, with schizophrenia, ADHD, and depression performing exceptionally well in distinction. The results point to the potential of EEG data as a valid biomarker for psychiatric disorder detection and diagnosis, advancing the field of psychiatry and facilitating further machine-learning research.

Ksibi *et al.*[7] identified depressive tendencies using EEG data available from open dataset. To enhance the generalization of the detection method, the study includes demographic variables that affect the occurrence of depression, such as gender and age.

Table 1 — Mental health classifiers using machine and deep learning algorithms

Data Source	Features	Prediction	Classifier[Ref]	Results
Publicly available EEG dataset of healthy and ADHD children from IEEE data port	Entropy, statistical, and nonlinear measures from VMD-HT of EEG dataset	Attention deficit hyperactivity disorder (ADHD)	Explainable boosted machine (EBM) mode[4]	The model yielded 9.81% accuracy, 99.78% sensitivity, 99.84% specificity, 99.83% F-1 measure, 99.87% precision, 0.13% false detection rate
Public datasets of depressive individuals	Statistical, time domain, frequency domain, and spectral features from EEG dataset	Psychiatric disorder Schizophrenia	Adaptive neuro-fuzzy inference system (ANFIS) based on the Hybrid Grey Wolf- Bat Algorithm[5]	The results show an accuracy of 99.54% and 99.35%, for datasets 1 and 2 respectively, with a respectable F1-score and MCC.
Dataset of people diagnosed with different diseases and healthy people	Absolute power values from EEG signals	Psychiatric diseases	Random forest (RF), support vector machine (SVM) and C5.0[6]	Patients were classified with accuracy of 84.1% for C5.0 for SVM, and 76.2% for RF algorithms.
Multi-modal open dataset MODMA	Alpha, beta, delta, and theta powers, mean, median, max and min amplitude, spectral and singular value deposition entropy	Depression detection	XGBOOST; Random Forest; 1D CNN model[7]	The CNN model produced the greatest accuracy across 25 training epochs, at 97%.
Data set of children taken from Clinical psychologist	Attributes from the dataset e.g. academic performance, relationship formation etc.	Five basic mental health problems,	AODE, Multi-Layer Perceptron, RBF Network, IB1, K Star, Multi-Class Classifier, FT, LAD Tree[8]	MLP, MCC, and LAD Tree provided more accurate results
Open Sourcing Mental Illness Survey of working individuals	Labels	Onset of mental illness	Decision Tree, Random Forest, Naïve Bayes[9]	Better Performance by Decision Tree (82.2%)
Depression, Anxiety and Stress Scale Questionnaire of participants aged between 20 and 60 years	Scales from DASS questionnaire	Anxiety, stress, depression	Decision Tree, Random Forest Tree, Naïve Bayes, Support Vector Machine and K-Nearest Neighbour[10]	Although Random Forest was determined to be the best model, the accuracy of naive Bayes was found to be the greatest.
Strengths and Difficulties Questionnaire of Twins from the Child and Adolescent Twin Study	Binary variable created from parent reported subscales.	Mental health problems in mid adolescence	Random Forest, Support Vector Machines, Neural Network, and XG Boost[11]	All the models had a respectable AUC, but none performed statistically significantly better than the others.
EEG signal Analysis of stressed participants	Absolute power, amplitude asymmetry, coherence, and phase lag from EEG signal	Levels of stress	Logistic regression, Naïve Bayes, SVM[12]	The proposed framework produced 94.6% accuracy for stress identification.
Twitter API of twitter users	Structural features (words in tweet) and behavioral features (number of connections and hashtags etc.)	Normal or alarming tweet	Logistic Regression MLP Classifier, RF, KNN, Ada and Gradient Boosting[13]	This model has an accuracy as high as 89%.

Using EEG data, machine learning algorithms have been utilized to identify depressed patients. The study uses a convolutional neural network (CNN) trained with 25 epoch repetitions, obtaining a high accuracy rate of 97%. The analysis considers several mental conditions, such as Major Depressive Disorder (MDD) and other disorders. According to the research, detecting depression may be possible when demographic information is combined with EEG signals.

Sumathi *et al.*[8] used a variety of machine learning methods to predict children's mental health concerns. This study examined the accuracy of eight machine learning approaches in identifying five common mental health issues using different metrics of accuracy. For training and assessing the performance of the methods, a data set of sixty instances was gathered. Twenty-five characteristics were found to be critical in diagnosing the problem from the papers.

Feature Selection methods were used to the whole attribute data set to minimise the attributes. The results show that Multilayer Perceptron, LAD Tree, and Multiclass Classifier yielded more accuracy, with just a small variation in performance between the entire attribute set and the chosen attribute set.

Kavitha *et al.*[9] presented a mental health prediction system for working individuals based on three machine learning algorithms and discovered that the decision tree algorithm outperforms the Random Forest and Nave Bayes techniques. The authors have created a web-based environment in which users must complete a questionnaire provided by the researchers, after which the researchers will suggest certain activities to the user. For prediction, the system had an accuracy of about 82%.

Priya *et al.*[10] proposed a prediction method for anxiety, depression, and stress based on machine learning. Working and jobless persons from various cultures and groups filled out the DASS-21 questionnaire. Due to their excellent accuracy, five different machine learning algorithms were able to predict anxiety, sadness, and stress on five different severity levels. These algorithms are especially well-suited to predict psychiatric diseases. The classes in the confusion matrix were not in balance. Consequently, the f1 score metric was introduced, which helped the Random Forest classifier choose the best model.

Another method developed by Tate *et al.*[11] utilises machine learning techniques to anticipate mental health issues. They used Random Forest, SVM, Neural Network, and XG Boost algorithms in this study, and it was particularly evaluated for teens. Although all of the models had a respectable AUC, none of them performed statistically significantly better than the others. Finally, they concluded that while machine learning approaches are promise for integrating risks across domains to predict mental health issues in adolescents, clinical adoption appears premature.

Subhani *et al.*[12] proposed a machine learning framework for stress level identification that includes feature extraction from EEG data of stressed individuals, selection of features using Bhattacharya distance, ROC curve, and t test. Logistic regression, Naïve Bayes, and SVM are used as classifiers. Montreal Imaging Stress Task (MIST) was used to create stress among the subjects and the task performance and subjective feedback confirmed the introduction of stress. The proposed framework was able to detect stress with a maximum accuracy of 94.6% between two stress levels and the control.

Joshi *et.al.*[13] have blended deep learning feature extraction methods, such as phrase embedding, with classic machine learning algorithms to analyse people's mental health from their social media posts and behavioural variables. Apart from categorising tweets as either normal or concerning, the authors have gathered tagged data from several Twitter users and made an effort to divide them into two groups: individuals who are typically healthy and those who are vulnerable to mental health issues. The ensemble model has classified the Twitters users with an accuracy of 89%.

From the literature review, it can be seen that most of the studies have relied upon single classifiers in order to produce mental health predictions. Additionally, mostly studies involving EEG datasets utilized online available datasets. The fact that the current study compared the psychological states among four classes, i.e. pre and post states of intervention and control groups, using ensemble of both machine and deep learning classifiers, makes this study unique and versatile.

Experimental Details

Participants

Twenty-five PG Diploma in Yoga Therapy students and twenty B.Ed. students in Yoga Education from GCYEH, Chandigarh, made up the intervention group. Twenty-four students from Dev Samaj College, Chandigarh, who had never done yoga, made up the control group. Students who volunteered for the research were between the ages of 18 and 30 and had no prior history of mental illness. Those having a history of respiratory disease, known psychiatric disorder, current or chronic medical conditions, or experience with yoga or Rajyoga meditation were excluded. The intervention group learned yoga and meditation for eight weeks, six days per week and 120 hours per day, whereas control group was not provided with any yoga or meditation practices.

Data Collection

Pankhtech's Neuphony setup was used to record the EEG data. This device, as seen in Fig. 1, has 8 sensors placed at the parietal (Pz), medium frontal (Fz), anterior frontal (Fp1, Fp2), lateral frontal (F3, F4), and temporal (T3, T4) lobes. Each of the dry sensors has an Ag/AgCl coating on conductive polycarbonate passive electrodes. EEG signals were recorded using a sample frequency of 250 Hz, cutoff frequency of 125 Hz, and a time constant of four ms. Every electrode has its

Fig. 1 — Neuphony EEG headband demonstration

impedance adjusted to be about less than 10 KHz. The device records and stores EEG signals by connecting to the laptop via a USB port. The pre and post EEG data was recorded for intervention and control groups.

Pre-processing

Pre-processing of EEG signals is necessary to improve signal quality, eliminate noise and artifacts in the raw data, and prepare the data for analysis.[14] Filtering to eliminate undesirable signals, segmenting the data using epochs, artifact rejection to find and eliminate artifacts, reference to take electrode location into account, and interpolation to approximate missing values are a few standard pre-processing techniques. This research involved the development of a minimum-order lowpass FIR filter with a normalized passband frequency of 0.25 rad/s, stopband frequency of 0.35 rad/s, passband ripple of 0.5 dB, and stopband attenuation of 65 dB.[15] In addition to that, a Kaiser window with adjustable frequency response properties was used to improve the filter.[16]

Statistical Analysis

A Windows version of SPSS software was utilised to do statistical analysis. EEG data were compared using the Wilcoxon Signed Rank Test for three bands because of non-normality of the results. The nonparametric equivalent of the parametric paired t-test is the Wilcoxon Signed Rank Test. P 0.05 was used as the lowest significant criteria. The significant changes between the baseline, intervention, -meditation, and control group were compared.

Feature Extraction

Analyzing EEG signals entails finding pertinent patterns or qualities in the data, known as feature extraction. Standard features are power spectrum density, time-domain characteristics, coherence, entropy, and event-related potentials.[17] This study extracts the signal's band power and wavelet approximation coefficients as a feature for EEG signal data.

Band Power: Band power is a frequently utilized characteristic of EEG data. Within a particular frequency band, such as the delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–100 Hz) bands, it assesses the power or energy of the EEG signal.[18] The band power feature is frequently used to examine the power distribution across various frequency bands and can offer insightful information about the underlying neuronal activity.

Approximation Coefficients using Wavelet: It is common practice to extract features from EEG signals using wavelet analysis, which divides EEG data into several frequency sub-bands to find frequency-specific patterns.[19] Wavelet coefficients, energy, entropy, and coherence are typical wavelet-based characteristics. These characteristics examine how complexity, power, and functional connectedness are distributed across various frequency bands and scales. Wavelet analysis of EEG signals frequently uses the Daubechies wavelet family, the db8 wavelet, used in this work.[20] The orthogonal wavelet known as the db8 has eight vanishing moments, making it capable of efficiently capturing fast transitions in the EEG signal while minimizing noise and other artifacts. The discrete wavelet transforms (DWT) using the db8 wavelet breaks down the signal into several frequency sub-bands before applying the db8 wavelet to EEG recordings. Each level of decomposition's detail and approximation coefficients can be utilized as features in later analyses like classification or clustering.[21]

Feature Cleaning

To ensure the accuracy of the analysis, characteristics that may be noisy, redundant, or unnecessary are removed from the EEG data during feature cleaning.[22] Techniques include outlier identification, feature scaling, expert knowledge, PCA, correlation-based feature selection, and recursive feature removal.[23] To appropriately represent the EEG data and increase the accuracy and reliability of the later analysis, it is crucial to clean the characteristics that will be used properly. Here in this work, the feature scaling method is used for feature cleaning, where any missing, negative, and zero feature value is scaled to the average value of that particular feature.[24] It will help to maintain the balance between the data and minimizes false positives.

Labeling

In a machine learning environment, data labeling puts one or more preset categories or labels on a dataset. A model is trained on a labeled dataset as part of supervised machine learning to discover the link between input characteristics and output labels.[25] Human annotators can execute it manually, or it can be done automatically with pre-established rules or algorithms. In contrast to automatic labeling, which can be quicker and less expensive but potentially less accurate, especially for complex or ambiguous datasets, manual data labeling is frequently more accurate but can also be time-consuming and expensive.[26] In this work, the dataset is already segregated by pre and post data so the binary labels are assigned with the help of an automatic script that checks the location label in the data and sets label '0' for pre and label '1' for post-data.

Classification

The purpose of this study is to classify meditation abilities using Machine Learning (ML) and Deep Learning (DL) methods. It aims to create a model that can correctly forecast the label of new, unforeseen data items. Decision trees, logistic regression, k-nearest neighbors, and support vector machines are just a few of the ML methods that may be utilized for classification.[27] These algorithms determine a border or rules dividing various classes in the feature space. Artificial Neural Networks (ANNs), a collection of linked nodes or neurons capable of learning sophisticated representations of the input data, are frequently used in DL to categorize. The Convolutional Neural Network (CNN), intended to extract features automatically and build hierarchical representations of the input, is one of the most well-liked designs for classification in DL.[28] This work proposes hybrid machine learning and deep learning-based classification algorithms and compares their performance for the given data.

Hybrid Classification Approach

In machine learning, a hybrid method integrates many algorithms or approaches from other disciplines to address a specific issue. It is significant because it capitalizes on the advantages of many approaches and eliminates their disadvantages, resulting in better performance and more durable solutions. In addition, it may deal with complex issues, mix various data kinds, overcome technique restrictions, and enhance generalizability.[29] On the other hand, a hybrid strategy that blends DL and ML utilizes the advantages of both fields to address complicated issues. It uses ML's interpretability for generating predictions and offering insights, as well as DL's capacity for handling massive, complicated datasets. As a result, different data types may be handled, individual technique shortcomings can be addressed, and transfer learning is improved using a hybrid approach.[30] This work proposes different hybrid classification approaches to address the classification of psychologically healthy and unhealthy people using their pre-and post-intervention data.

Deep Learning Algorithms

Two different deep learning algorithms are proposed for EEG data classification in this paper. The first model is Convolution Neural Networks (CNN), and the second is Bidirectional Long Short-Term Memory (BiLSTM) Network. This section describes the architecture and layers of these models' architecture.

CNN: In this work, the proposed CNN takes features as input and is accepted by the input layer for further analysis. As shown in Fig. 2, the input layer is of size [1 × 9 × 1], which means for each input data, the input layer uses 9 features as input. Further, these input features are passed to the first convolutional

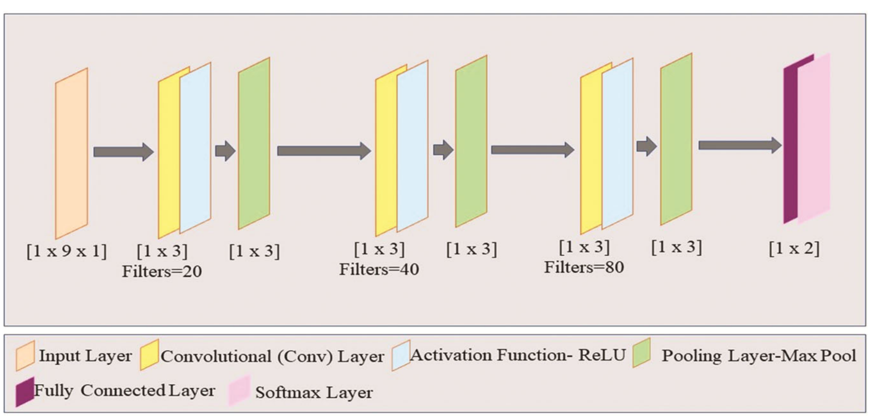

Fig. 2 — CNN model architecture

(Conv) layer, which utilizes the input features and extracts new features with different filters. This work uses three Conv layers of the same size [1 × 3] with 20, 40, and 80 filters, respectively. Further, a Rectified Linear Unit (ReLU) activation function introduces non-linearity to the network. It extracts only positive features from the Conv Layers by converting negative features to zero. In CNN, the pooling operation is generally used to down sample the feature maps by reducing their spatial dimensions. It is mainly of two types: Average Pooling, which computes the average value and Max Pool selects the maximum value from the given pooling window. This work uses a Max Pool layer of size [1 × 3] after every Conv Layer. Finally, for predictions, a fully connected layer with a softmax function is used, which computes the class probabilities to identify the class for given predictions.

BiLSTM: In a BiLSTM, the input features are fed into two independent LSTM layers, one of which does forward processing and the other backward processing. The concealed states and cell states are kept separately for each layer. Each time step involves concatenating the hidden states of both directions, resulting in a fused representation that incorporates knowledge from both the past and the future. Furthermore, the forward and backward representations are frequently passed via activation layers to generate the outcomes. The architecture of the BiLSTM network model used in this work is presented in Fig. 3.

Hybrid Algorithms

This study combines different machine learning algorithms with the above-defined deep learning architectures for better classification results. An AdaBoost (Adaptive Boosting) method integrates deep learning and machine learning methods, which iteratively trains the data with both algorithms and generates the outcomes using the voting method, as shown in Fig. 4. The hybrid models are trained with 100 iterations and based on it, final predictions are made. AdaBoost here is used to build models that combine deep learning with machine learning, allowing for better performance by utilizing the advantages of both techniques.

The above hybrid architecture combines two different deep-learning architectures with the three most popular machine-learning algorithms, decision tree, KNN, and random forest. In total, six hybrid models are proposed to classify pre- and post-intervention data.

Integration with KNN: Two deep learning algorithms discussed earlier are integrated with the K-nearest Neighbor (KNN) algorithm. It is named CK-NN and BiLSTM-KNN where KNN is a straight forward and efficient classifier that places a data point in the feature space among its k closest neighbors according to the majority class. The combination of CNN and KNN (CK-NN) has advantages due to CNN's capacity to learn intricate and discriminative features and KNN's strong generalization capabilities based on these features. Similarly, BiLSTM with KNN also integrates the benefits of both models and achieves the required performance. The ensemble model may enhance accuracy and generalization by integrating their predictions for both deep learning architectures.

Integration with DT: This architecture integrates CNN and BiLSTM with the decision tree algorithm in different architectures. A decision tree is a tree-structured model used to classify data based on feature values and make decisions. It employs a hierarchical structure of if-else conditions. To generate a tree-like structure for classification, it divides the data into subgroups according to the selected attributes and their thresholds. The integration of CNN and the decision tree improves the decisions of CNN and provides better results; similarly, the case with BiLSTM.

Integration with RF: Combining Random Forest with CNN and BiLSTM enables the learning of complicated and discriminative features by deep learning algorithms with the handling of a variety of non-linear decision

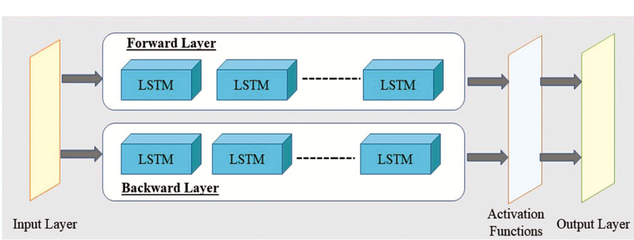

Fig. 3 — BiLSTM architecture

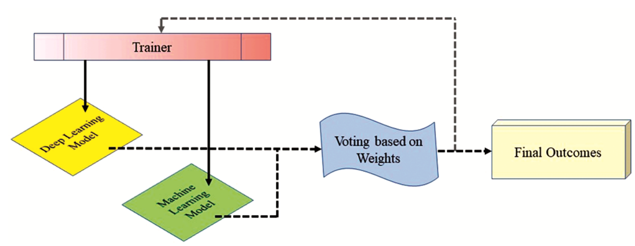

Fig. 4 — Proposed hybrid architecture

boundaries by Random Forest. Here, Multiple decision trees are used to generate a Random Forest. By randomly picking subsets of traits and samples, it generates a varied collection of trees. The accuracy and robustness of the prediction are increased by combining the predictions of several trees.

Result Analysis

The proposed system is simulated using a MATLAB simulator with a deep learning toolbox to utilize the benefits of the existing libraries. Moreover, the simulations are run on a CPU system with an integrated graphics card, 16 GB RAM, and an Intel i7 processor.

Training Parameters

The deep learning and proposed integrated hybrid models are trained with 70% of the data features. The training parameters are demonstrated in Table 2 for different models as per their requirements.

Spectral Analysis

The wavelet characteristics were plotted for eight electrodes from three bands (alpha, beta, and theta) for the baseline, intervention, meditation, and control groups as shown in Fig. 5.

The post-meditators had higher band powers across all channels than in pre-meditators, as the figures demonstrate. With the exception of temporal (T3 and T4) in the theta band and lateral frontal (F3) in the alpha and beta bands, the post-meditators group's feature values were higher than those of the control group.

Performance Evaluation of Models

The efficacy of classification or prediction models can be evaluated using performance assessment measures, including precision, recall, F-measure, accuracy, and confusion matrix. By comparing the model's predictions to the actual results, these metrics offer insightful information about the model's performance. While recall assesses the model's capacity to record all pertinent positive cases accurately, precision examines the model's accuracy in forecasting positive instances. The F-measure provides a balanced overall performance indicator by combining recall and accuracy into a single score. Contrarily, accuracy evaluates the model's correctness by determining the percentage of accurate predictions among all occurrences. The model's predictions are compared to the actual results in a table format called the confusion matrix. It thoroughly studies the model's performance across several classes by dividing the predictions into true positives,

The performance of the proposed ensemble models is compared with their base models CNN and BiLSTM. First, the confusion matrix infers the results in the matrix form, as shown below in Fig. 6. True positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) are the four primary subsets of the above-given matrix. The instances when the model accurately predicted the positive class are true positives. The situations where the model correctly predicts the negative class are true negatives. False positives happen when the model predicts the wrong class, usually a positive one when the true class is negative. False negatives refer to situations where the model erroneously predicts a negative class when the true class is positive. The confusion matrix's values are examined, and several

Table 2 — Training Parameters

Models	Training parameters	Values
Deep learning and Hybrid models	Learning rate	0.01
	Epochs	100
	Optimizer	SGDM
BiLSTM	Hidden units	100
KNN	Number of neighbours	05
	Distance method	Euclidean
Hybrid models	Ensemble approach	Adaboost
	Number of cycles	100

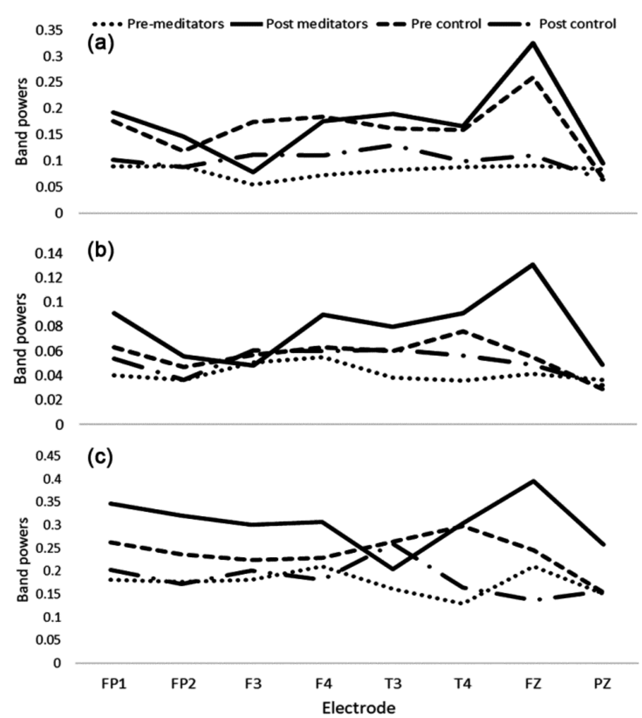

Fig. 5 — Distribution of wavelet features from EEG analysis for (a) alpha band (b) beta band, and (c) theta band

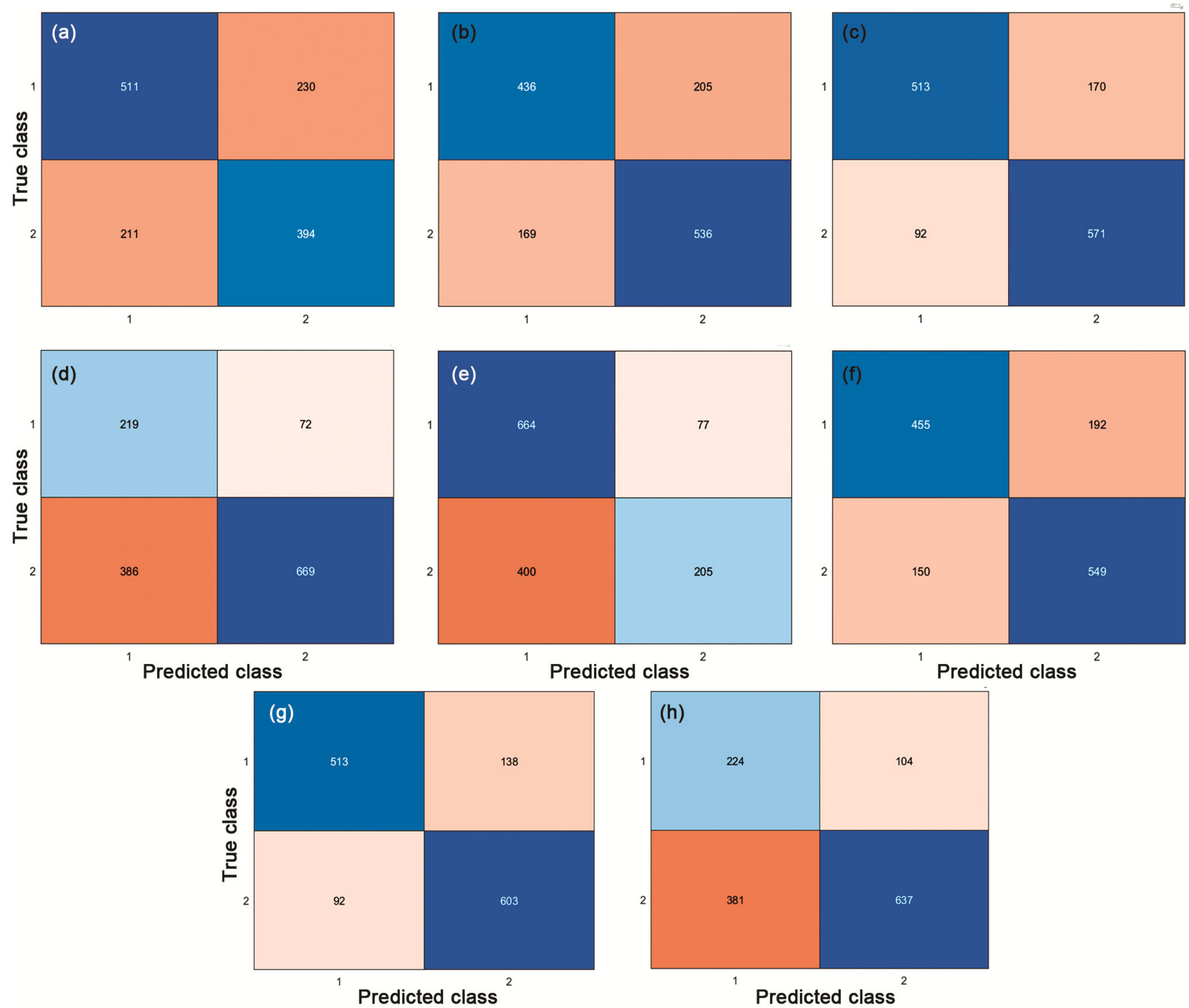

Fig. 6 — Confusion Matrix for (a) CNN model, (b) CK-NN model, (c) CNN-DT model, (d) CNN-RF mode, (e) BiLSTM model, (f) BiLSTM-KNN, (g) BiLSTM-DT model, and (h) BiLSTM-RF model

assessment metrics are generated. Precision, recall, and F-measure are some of these measurements. These measures reveal the preciseness of the model's capacity to differentiate between positive and negative occurrences, the general accuracy of its predictions, and the trade-off between precision and recall. The computed Precision, Recall, and F-measure comparison is shown in Table 3.

Table 3 — Performance Evaluation based on Precision, Recall, and F-Measure

	Precision	Recall	F-Measure
CNN	0.67	0.67	0.67
BiLSTM	0.67	0.62	0.59
CK-NN	0.72	0.72	0.72
BiLSTM-KNN	0.75	0.74	0.74
CNN-DT	0.81	0.80	0.80
BiLSTM-DT	0.83	0.83	0.83
CNN-RF	0.63	0.69	0.61
BiLSTM-RF	0.61	0.65	0.60

The models' precision, recall, and F-measure can all be shown to vary from one another in their results. The CNN model's precision, recall, and F-measure are all 0.6696, 0.6704, and 0.6699, respectively. The BiLSTM model has a lower recall of 0.61745 but a slightly higher precision of 0.67555, resulting in a lower F-measure of 0.59895. The CK-NN model performs better with a precision of 0.722, recall of 0.72025, and F-measure of 0.7206. The BiLSTM-KNN combination exhibits additional improvement with precision, recall, and F-measure of 0.7465, 0.7443, and 0.7465, respectively. Compared to the other models, the models integrating Decision Tree (DT) classifiers exhibit greater accuracy, recall, and F-measure. The CNN-DT model's accuracy, recall, and F-measure are all 0.80925, 0.80615, and 0.805, respectively. With an accuracy of 0.83085, recall of 0.8278, and F-measure of 0.82835, the BiLSTM-DT model performs even better. However, the CNN-RF and BiLSTM-RF models perform less well than the other models in terms of precision, recall, and F-measure.

Furthermore, the accuracy of the models is also computed, and the comparison is made between CNN and its ensemble models and BiLSTM and its ensemble models, as shown in Fig. 7. The accuracy of the CNN model is 67.2363%, according to the above findings. With an accuracy of 72.214%, the CK-NN ensemble model performs marginally better. The accuracy of the CNN DT ensemble model has significantly increased, reaching 80.5349%. In contrast to the other models, the CNN-RF (CNN-Random Forest) ensemble model has a lower accuracy of 65.9733%. Given these findings, it is clear that the CNN-DT ensemble model outperforms the other models in terms of accuracy. It outperforms the CK-NN and CNN-RF ensemble models and the solo CNN model. On the other hand, the BiLSTM model achieves an accuracy of 64.5617%. The accuracy of the BiLSTM-KNN ensemble model is superior, at 74.5914%. With an accuracy of 82.9123%, the BiLSTM-DT ensemble model stands out as having significantly improved. Unlike the other models, the BiLSTM-RF ensemble model has a lower accuracy of 63.9673%. The BiLSTM-DT ensemble model is the most accurate of the studied models in light of these findings. It performs better than the other ensemble models, such as the BiLSTM-KNN and BiLSTM-RF models and the individual BiLSTM model.

The above results examine the effectiveness of ensemble models when paired with deep learning models utilizing several machine learning methods, emphasizing the KNN, DT, and RF algorithms. These findings show that for both the CNN and BiLSTM models, the ensemble models using DT consistently beat the other algorithms in terms of accuracy. It shows that a possible method for enhancing ensemble learning performance is the combination of deep learning models with decision trees.

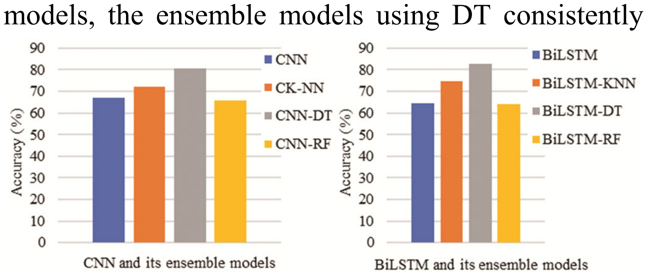

Fig. 7 — Accuracy of different ensemble models

Discussion

The objective of this work is to provide the subjective measures of potential mental health issues. The inclusion of neuro physiological traits enhances reliability in detection of mental health disorders among vulnerable population.

The results of the current study synchronize with the previous studies where meditation traits have been classified with the help of machine learning algorithms. An increase in alpha, beta and theta powers has been observed as a result of yoga and Rajyoga meditation. A joint classifier using both EEG and respiration data can also be used as a marker for different states of mind.[31] After completing a six-week mindfulness meditation intervention, the EEG and respiratory signals of beginner meditators were monitored during meditation and in a control condition. The spectral analysis of EEG data showed increase in alpha and theta powers and decrease in respiration rate during meditation. Support vector machine was used to classify the meditation and control conditions and it proved to be a feasible objective marker for meditation ability. The unsupervised Self-Organizing Map (SOM) algorithms can also be used to cluster the brain mappings of centroid frequency of brain function.[32] The spectral properties of resting Zen meditation EEG were classified using SOM. The Zen meditators exhibited global high beta frequency, thus implying the state of mindfulness concentration. Thus, the cluster features of brain mappings of centroid frequency describe the state of brain function related to the multi-channel EEG through SOM categorization, which aids in interpretation of EEG and mental health diagnosis. The deep learning algorithms can act as good classifiers for recognizing state and trait effects in meditators.[33] The effects of short-term mindfulness-based stress reduction training were examined using convolutional neural networks based deep learning methods and traditional machine learning methods. The classification performance was improved using deep learning methods. Machine learning classifiers were also used to classify the various levels of Heartfulness meditation.[34]. According to the classifier results, functional connectivity measures were more effective at detecting alterations brought on by meditation. The machine learning models exhibited high accuracy of classification, ranging between 80 to 100 percent.

Limitations

This study comes with some limitations. The research included novice meditators only. The students of Yoga college in Chandigarh were taken as subjects, because these subjects were easily available as a result of signed agreement for research studies between the research wing of Brahma Kumaris Organization (teaching Rajyoga meditation) and the institute. Expert meditators should be included in a future study as well. Data collection involves multiple collaborations, especially if expert meditators are to be assessed, collaborations of the research institute with the spiritual organization has to be considered. Multi center collaborations are therefore suggested for enhancing the subject availability. In this study, EEG data has been acquired before and after the eight weeks of intervention. Future studies perform longitudinal studies with the same subjects to examine the students for changes that may occur over a period of time if meditation is not practiced regularly. Only EEG measures were utilized in this investigation. Future research will incorporate more physiological cues into the classifier to determine whether accuracy can be increased.

Future Research

To handle small datasets, future research may look at more sophisticated deeper learning systems, specifically designed for fewer parameters. Data augmentation and self-supervised learning techniques can also help in coping up with the issues related to small datasets. The performance of the classifiers may also be enhanced by increasing the dataset. More physiological variables such as respiration data, speech features, smartphone data, eye gaze variables etc. can be included to build multi modal systems. Feature engineering helps increase the accuracy of the model by offering more pertinent and useful data. The generalization ability of the models can be enhanced by selecting the optimal features through hyperparameter tuning. Domain adaptation methods may also be adopted to make the model generalize across different subjects and recording setups.

Conclusions

An ensemble of machine and deep learning algorithms for distinguishing meditation state from baseline and control states has been presented in this study. The spectral analysis of EEG data developed a relationship between meditation and brain dynamics. With the given data, the maximum accuracy (82.91%) has been achieved by the proposed BiLSTM-DT ensemble model. In future, the performance of these algorithms in classifying multi modal data will be examined. The investigation can be assessed more thoroughly by augmenting the quantity of participants in both the control and research groups.

Acknowledgement

The authors are thankful to all the students who participated in the study. The authors also extend their gratitude to the college staff a SpARC wing of Brahma Kumaris for providing their support to conduct this study.

References

1 Cerchia C & Lavecchia A, New avenues in artificial-intelligence-assisted drug discovery, *Drug Discov Today*, **28** (2023) 103516, https://doi.org/10.1016/j.drudis.2023.103516.

2 Liu X, Song C, Liu S, Li M, Zhou X & Zhang W, Multi-way relation-enhanced hypergraph representation learning for anti-cancer drug synergy prediction, *Bioinformatics*, **38** (2022) 4782–4789, https://doi.org/10.1093/bioinformatics/btac579.

3 Yadav A, Verma S, Panwar M & Yadav N K, Role of Yoga practices on cognitive functions, *Int J Health Sci*, **6** (2022) 3288–3304, https://doi.org/10.53730/ijhs.v6nS3.6341.

4 Khare S K & Acharya U R, An explainable and interpretable model for attention deficit hyperactivity disorder in children using EEG signals, *Comput Biol Med*, **155** (2023) 106676, https://doi.org/10.1016/j.compbiomed.2023.106676.

5 Balasubramanian K, Ramya K & Gayathri Devi K, Optimized adaptive neuro-fuzzy inference system based on hybrid grey wolf-bat algorithm for schizophrenia recognition from EEG signals, *Cogn Neurodyn*, **17** (2023) 133—151, https://doi.org/10.1007/s11571-022-09833-y.

6 Emre İE, Erol Ç, Taş C & Tarhan N, Multi-class classification model for psychiatric disorder discrimination, *Int J Med Inform*, **170** (2023), 104926, https://doi.org/10.1016/j.ijmedinf.2022.104926.

7 Ksibi A, Zakariah M, Menzli L J, Saidani O, Almuqren L & Hanafieh R A M, Electroencephalography-based depression detection using multiple machine learning techniques, *Diagnostics*, **13** (2023) 1779, https://doi.org/10.3390/diagnostics13101779.

8 Sumathi M R & Poorna B, Prediction of mental health problems among children using machine learning techniques, *Int J Adv Comput Sci Appl*, **7** (2016) 552–557, http://dx.doi.org/10.14569/IJACSA.2016.070176.

9 Kavitha M, Pingili M, Spurthi M, Kirthan K & Santhosh S, Classification algorithm based mental health, *Turk J Comput Math Educ*, **13** (2022) 1168–1175, https://doi.org/10.17762/turcomat.v13i2.13708.

10 Priya A, Garg S & Tigga N P, Predicting anxiety, depression and stress in modern life using machine learning algorithms, *Procedia Comput Sci*, **167** (2020) 1258–1267, https://doi.org/10.1016/j.procs.2020.03.442.

11 Tate A E, McCabe R C, Larsson H, Lundström S, Lichtenstein P & Kuja-Halkola R, Predicting mental health problems in adolescence using machine learning techniques, *PLoS One*, **15** (2020) 1–13, https://doi.org/10.1371/journal.pone.0230389.

12 Subhani A R, Mumtaz W, Saad M N B M, Kamel N & Malik A S, Machine learning framework for the detection of mental stress at multiple levels, *IEEE Access*, **5** (2017) 13545–13556, https://doi.org/10.1109/ACCESS.2017.2723622.

13 Joshi D J, Makhija M, Nabar Y, Nehete N & Patwardhan M S, Mental health analysis using deep learning for feature extraction, in *Proc ACM India Joint Int Conf Series*, (2018) 356–359, https://doi.org/10.1145/3152494.3167990.

14 Desjardins J A, van Noordt S, Huberty S, Segalowitz S J & Elsabbagh M, EEG ntegrated Platform Lossless (EEG-IP-L) pre-processing pipeline for objective signal quality assessment incorporating data annotation and blind source separation, *J Neurosci Methods*, **347** (2021) 108961, https://doi.org/10.1016/j.jneumeth.2020.108961.

15 Mohammadi A, Fakharzadeh M & Baraeinejad B, An integrated human stress detection sensor using supervised algorithms, *IEEE Sens J*, **22** (2022) 8216–8223, https://doi.org/10.1109/JSEN.2022.3157795.

16 Leach S, Sousouri G & Huber R, High-Density-Sleep Cleaner: An open-source, semi-automatic artifact removal routine tailored to high-density sleep EEG, *J Neurosci Methods*, **391** (2023) 109849, https://doi.org/10.1016/j.jneumeth.2023.109849.

17 Arpaia P, Covino A, Cristaldi L, Frosolone M, Gargiulo L, Mancino F, Mantile F & Moccaldi N A, A systematic review on feature extraction in electroencephalography-based diagnostics and therapy in attention deficit hyperactivity disorder, *Sensors*, **22** (2022) 4934, https://doi.org/10.3390/s22134934.

18 Hussain I, Hossain M A, Jany R, Bari M A, Uddin M, Kamal A R M, Ku Y & Kim J S, Quantitative evaluation of EEG-biomarkers for prediction of sleep stages, *Sensors*, **22(8)** (2022) 3079, https://doi.org/10.3390/s22083079.

19 Sunaryono D, Sarno R & Siswantoro J, Gradient boosting machines fusion for automatic epilepsy detection from EEG signals based on wavelet features, *J King Saud Univ – Comput Inform Sci*, **34** (2022) 9591–9607, https://doi.org/10.1016/j.jksuci.2021.11.015.

20 Abukhettala K & Ata O, Analyzing of EEG signals with deep learning and discrete wavelet transform, *Eur J Sci Technol*, **35** (2022) 514–521, https://doi.org/10.31590/ejosat.953576.

21 Malviya L & Mal S, A novel technique for stress detection from EEG signal using hybrid deep learning model, *Neural Comput Appl*, **34** (2022) 19819–19830, https://doi.org/10.1007/s00521-022-07540-7.

22 Zhu X, Rong W, Zhao L, He Z, Yang Q, Sun J & Liu G, EEG emotion classification network based on attention fusion of multi-channel band features, *Sensors*, **22(14)** (2022) 5252, https://doi.org/10.3390/s22145252.

23 Bi J, Wang F, Yan X, Ping J & Wen Y, Multi-domain fusion deep graph convolution neural network for EEG emotion recognition, *Neural Comput Appl*, **34** (2022) 22241–22255, https://doi.org/10.1007/s00521-022-07643-1.

24 Tarafder S, Badruddin N, Yahya N & Nasution A H, Drowsiness detection using ocular indices from EEG signal, *Sensors*, **22(13)** (2022) 4764, https://doi.org/10.3390/s22134764.

25 Wang X, Ma Y, Cammon J, Fang F, Gao Y & Zhang Y, Self-supervised EEG emotion recognition models based on CNN, *IEEE Trans Neural Syst Rehabil Eng*, **31** (2023) 1952–1962, https://doi.org/10.1109/TNSRE.2023.3263570.

26 Bailey N W, Biabani M, Hill A T, Miljevic A, Rogasch N C, McQueen B, Murphy O W & Fitzgerald P B, Introducing RELAX: An automated pre-processing pipeline for cleaning EEG data - Part 1: Algorithm and application to oscillations, *Clin Neurophysiol*, **149** (2023) 178–201, https://doi.org/10.1016/j.clinph.2023.01.017.

27 Suthanan A M, Rathee S & Kumar A, Comparative analysis of machine learning algorithms for classification of Alzheimer's disease, *2022 International Conference on Emerging Trends in Engineering and Medical Sciences* (IEEE) 2022, 454–460, https://doi.org/10.1109/ICETEMS56252.2022.10093394.

28 Malviya L & Mal S, CIS feature selection based dynamic ensemble selection model for human stress detection from EEG signals, *Cluster Comput*, **26** (2023) 2367–2381, https://doi.org/10.1007/s10586-023-04008-8.

29 Chen X, Li C, Liu A, McKeown M J, Qian R & Wang Z J, Toward open-world electroencephalogram decoding via deep learning: A comprehensive survey, *IEEE Signal Process Mag*, **39(2)** (2022) 117–134, https://doi.org/10.1109/MSP.2021.3134629.

30 Peirelinck T, Kazmi H, Mbuwir B V, Hermans C, Spiessens F, Suykens J & Deconinck G, Transfer learning in demand response: A review of algorithms for data-efficient modelling and control, *Energy and AI*, **7** (2022) 100126, https://doi.org/10.1016/j.egyai.2021.100126.

31 Ahani A, Wahbeh H, Nezamfar H, Miller M, Erdogmus D & Oken B, Quantitative change of EEG and respiration signals during mindfulness meditation, *J Neuro Eng Rehabil*, **11** (2014) 1–11, https://doi.org/10.1186/1743-0003-11-87.

32 Lo P C & Hussain N, Comparison of spatio-spectral properties of zen-meditation and resting EEG based on unsupervised learning, *J Behav Brain Sci*, **11** (2021) 58–72, https://doi.org/10.4236/jbbs.2021.112005.

33 Laan Tom van der, A deep learning approach to classifying the level of meditation expertise using EEG data, PhD Thesis, Tilburg University, The Netherlands, 2022.

34 Shrivastava A, Singh B K, Krishna D, Krishna P & Singh D, Effect of heartfulness meditation among long-term, short-term and non-meditators on prefrontal cortex activity of brain using machine learning classification: A cross-sectional study, *Cureus*, **15(2)** (2023) e34977, https://doi.org/10.7759/cureus.34977.