EEG Dataset Based On Meditation

D Nitish Sunit Reddy
Department of Computer Science and Engineering
Amrita School of Computing
Amrita Vishwa Vidyapeetham
Bengaluru, India
bl.en.u4aie23139@bl.students.amrita.ed
u

***Abstract*—This study utilizes five diverse EEG datasets collected from various sources to explore brain activity under different cognitive and emotional states. The datasets include recordings from meditation, relaxation, concentration, and brain-computer interface (BCI) tasks. Data were collected using devices like Neurosky Mindwave, Emotiv EEG, and high-density EEG systems. Formats such as CSV, EDF, and BIDS are used, supporting flexible preprocessing and model training. The OpenNeuro and Figshare datasets provide high-resolution, multi-session EEG data suitable for deep learning applications. Kaggle and Mendeley datasets offer practical data for emotion classification and mental state analysis. These datasets enable robust comparison between traditional ML and deep learning models. Overall, the combined resource supports EEG-based emotion recognition, BCI development, and neurofeedback systems.**

***Terms*—Neurofeedback System, Electroencephalography (EEG), Machine Learning (ML), Feature Extraction, EEG Signal Processing, Frequency Bands (Alpha, Beta, Theta, Delta).**

I. Introduction

Electroencephalography (EEG) datasets are critical for advancing research in brain-computer interfaces, emotion recognition, and mental state monitoring. In this work, five publicly available EEG datasets are utilized, each offering unique insights into brain activity across various cognitive and emotional states. The Neurosky Mindwave dataset captures EEG signals during general tasks using a single-channel consumer-grade device. The Kaggle meditation EEG dataset focuses on brainwave patterns during mindfulness and relaxation exercises. Mendeley dataset 1 presents data collected under relaxation and concentration moods using Emotiv EEG headsets, ideal for binary classification tasks. Mendeley dataset 2 provides detailed multi-channel EEG recordings in BCI contexts, offering rich spatiotemporal data. The OpenNeuro ds001787 dataset follows the BIDS format and includes EEG data collected during meditation studies using high-resolution systems. These datasets vary in sampling frequency, channel count, and application domain, allowing for diverse analytical approaches. They support real-world use cases such as mental health monitoring, cognitive load assessment, and neurofeedback systems. By combining low-cost and clinical-grade data sources, the study ensures a broad understanding of EEG signals. All datasets are open-access, promoting reproducibility and collaboration. Together, they serve as a strong foundation for building and evaluating machine learning and deep learning models.

II. System Overview And Methodology

A. EEG Signal Acquisition and Preprocessing

EEG signal acquisition is the foundational step in brainwave analysis, where electrical activity generated by the brain is captured using sensors placed on the scalp. In this study, EEG data were sourced from five publicly available datasets collected using various devices including Neurosky Mindwave, Emotiv EEG, and high-density EEG systems with up to 62 electrodes. Each device offers different sampling frequencies and channel configurations, enabling analysis across both low and high-resolution signals.

Preprocessing was carried out to clean and standardize the raw EEG signals. Common steps included removing artifacts such as eye blinks, muscle noise, and power line interference using filters like bandpass (1–50 Hz) and notch (50/60 Hz). Signals were then normalized or standardized, and irrelevant or noisy channels were excluded. For datasets in EDF or BIDS format, the MNE-Python library was used to parse, visualize, and extract features. For CSV-based datasets, Pandas and SciPy were utilized for signal filtering and transformation..

B. Feature Extraction

The initial phase of EEG signal processing starts with feature extraction since it accomplizes both data simplification and preservation of essential cognitive state markers. Different extraction methods used for analyzing EEG signals demonstrate their effectiveness in evaluation processes. The extraction methods reveal relevant patterns from raw EEG signals which lead to better results in interpretation and classification processes.

(i) Time Domain Features: Time-domain features capture the overall statistical properties of the EEG signal over time. Mean represents the average signal value across a time window, providing insight into the general signal level. Standard deviation measures the variability of the signal, indicating how much the signal fluctuates. Skewness describes the asymmetry of the signal's distribution, showing whether the waveform is tilted to one side. Kurtosis measures the "peakedness" or flatness of the waveform, reflecting the extremity of signal fluctuations. Zero-crossing rate counts how often the signal changes its sign, which indicates the frequency of signal oscillations. Root mean square (RMS) represents the overall signal energy and reflects the intensity of brain activity over time.

(ii) Frequency Domain Features: Frequency-domain analysis involves breaking down EEG signals into different frequency components. Delta waves (0.5–4 Hz) are linked to deep sleep and unconsciousness. Theta waves (4–8 Hz) reflect relaxed or meditative states, often appearing in light sleep or daydreaming. Alpha waves (8–13 Hz) are present when a person is relaxed but awake, especially with closed eyes. Beta waves (13–30 Hz) represent active thinking, focus, or stress, and are common in concentration tasks. Gamma waves (30–100 Hz) are related to higher cognitive functions such as memory, learning, and problem-solving. Band power measures the energy concentration in each frequency band, giving an indication of dominant brain activity. Spectral entropy quantifies the unpredictability or disorder within the signal's frequency content, with higher values representing more complex mental states.

(iii) Statistical Features: Statistical features are essential for summarizing the general properties of EEG signals. Mean provides the average value of the EEG signal, indicating baseline brain activity. Variance measures the spread of the EEG signal around the mean, highlighting signal fluctuations. Skewness assesses the asymmetry of the signal distribution, which can reveal information about signal distortion or the presence of certain brain states. Kurtosis measures the "peakedness" of the signal's distribution, indicating whether the signal has sharp peaks or is more uniform. Entropy, such as Shannon entropy or sample entropy, quantifies the randomness or unpredictability of the EEG signal, offering insights into the signal's complexity. Maximum and minimum values provide information about the extreme levels the signal reaches. Root Mean Square (RMS) is used to measure the signal's energy, and Standard Deviation (SD) indicates the variability in the signal, both of which are critical for assessing the strength and consistency of brain activity.

(iv) Deep Feature Extraction: Deep feature extraction leverages advanced neural network models to capture high-level, complex representations of EEG data. Convolutional Neural Networks (CNNs) are frequently used to automatically extract spatial and temporal patterns from raw EEG signals. Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), are effective in capturing temporal dependencies and sequential patterns in EEG signals.

(v) Hjorth Parameters: Hjorth parameters offer compact measures of EEG signal characteristics. Activity reflects the overall signal power, indicating the level of brain activity. Mobility measures the frequency variation in the signal, capturing how fast the brainwaves change. Complexity quantifies the changes in frequency over time, reflecting how dynamically the brain is functioning. These parameters are particularly useful for real-time analysis of brain activity.

C. *Machine Learning-Based Analysis*

Machine learning (ML) models are widely used for analyzing EEG signals, enabling the classification of different mental states, emotions, or cognitive tasks. These models learn patterns in the EEG data by training on labeled datasets and can then predict or classify unseen data. Some of the ML techniques that are a par with EEG systems are:

(i) Support Vector Machine (SVM):

Support Vector Machines are highly effective for classifying EEG signals, especially when the classes (e.g., relaxed vs. focused) are well-separated in high-dimensional space. SVMs can work with linear or non-linear kernels to effectively separate meditation and non-meditation states based on the frequency characteristics of the EEG signal. SVMs are ideal for binary classification tasks, such as distinguishing between relaxed states (during meditation) and active or stressed states.

(ii) K-Nearest Neighbors (KNN):

KNN is a simple and intuitive model that classifies EEG signals based on the closest neighbors. In meditation-related EEG analysis, KNN can be used to classify states like relaxed, focused, or stressed by comparing the current signal's features with similar signals in the training dataset. It is effective when the dataset is not very large and is often used with features like the power of different EEG frequency bands (e.g., alpha and theta waves during meditation).

(iii) Random Forest (RF):

Random Forests are ensemble learning methods that use multiple decision trees to classify EEG signals. By combining the results from many trees, RF can reduce overfitting and improve the model's generalization on unseen data. For meditation analysis, Random Forests can handle the complexity of EEG signals, especially when dealing with various features (e.g., power in alpha, beta, and theta bands) that distinguish between different mental states such as meditation, relaxation, or stress.

(iv) Convolutional Neural Networks (CNN):

CNNs are highly effective for feature extraction from raw EEG data and are increasingly used in EEG signal classification, including meditation. These networks automatically learn spatial features from the raw EEG signal, capturing patterns that may indicate relaxation or focus during meditation. CNNs can be used for both classification and feature extraction, providing a hierarchical understanding of EEG signals associated with different mental states.

(v) k-Means Clustering:

k-Means is a clustering technique that is useful for unsupervised learning tasks. In the context of EEG signals for meditation, k-Means can be used to segment EEG data into clusters based on similarities in signal characteristics, such as specific frequency band activities during different meditation stages. This method is often used to identify distinct mental states during the meditation process, like focusing, relaxation, and deeper meditative states.

D. *Threshold Function and Loop*

The threshold function is a crucial component in many EEG signal processing systems, particularly when classifying signals or detecting specific mental states such as relaxation, focus, or meditation. It is used to distinguish between different states or levels of activity in the EEG signal. The function works by setting a predefined threshold value that the EEG signal must exceed or fall below in order to trigger a specific action or classification. The loop structure in EEG signal processing allows for continuous analysis of time-series data and iterative refinement of results, especially in real-time monitoring and long-duration meditation sessions.

(i) Fixing Thresholds: Fixing thresholds in EEG signal processing is a critical step to ensure the system can reliably classify and detect specific mental states (e.g., relaxation, focus, or meditation). The threshold values are used to determine when the EEG signal exceeds certain predefined criteria, such as when a person enters a meditative state or reaches a certain level of relaxation. However, choosing appropriate threshold values is essential to avoid misclassifications or errors due to noise or variations in the EEG data.

(ii) *Adaptive thresholds:*Adaptive thresholds in EEG signal processing dynamically adjust based on the characteristics of the incoming EEG signal. Unlike fixed thresholds, which may not account for individual variability or changes over time, adaptive thresholds continuously update to improve the detection of mental states. Techniques like moving averages, standard deviation, and machine learning models are commonly used to adjust thresholds based on recent signal data. A sliding window approach, where a fixed segment of EEG data is used to calculate thresholds, is often employed for real-time adaptation. However, they face challenges such as noise sensitivity and increased computational complexity. Despite these issues, adaptive thresholds are crucial for real-time EEG systems, including meditation, focus, and brain-computer interface applications. Calibration is needed initially to set the baseline, but the system remains responsive as it adjusts to signal variations.

(iii) Z-score Normalization: Z-score normalization plays a crucial role in EEG meditation analysis by standardizing EEG features. It removes the mean and scales the data to have unit variance, making it easier to compare EEG signals across individuals and sessions. This preprocessing step ensures that all features contribute equally, which is essential for machine learning models like SVM or KNN. Z-score normalization also helps manage the natural variability in EEG signals due to individual differences and external factors, providing a consistent input for model training. Additionally, it makes outlier detection more effective, as extreme values in the normalized data are easily identified. For real-time EEG systems, such as brain-computer interfaces, this normalization ensures that incoming data is adjusted quickly and accurately, without manual recalibration. By applying Z-score normalization, the analysis of meditation states becomes more reliable, enabling precise identification and classification of different mental states.

(iv) Non-Linear Thresholds and Sigmoidal Functions:

In EEG analysis, non-linear thresholds and sigmoidal functions play a key role in improving the detection and classification of mental states, such as meditation. Non-linear thresholds adapt dynamically to changes in the EEG signal, making them more suitable for real-time applications, where brain activity fluctuates over time. These thresholds help capture subtle shifts in mental states, offering better sensitivity compared to linear methods. Sigmoidal functions, such as the logistic sigmoid, are commonly used in machine learning models for classification tasks. They map input signal values to a continuous range between 0 and 1, making them ideal for binary or multi-class state detection. In EEG datasets, sigmoidal functions help classify signals into categories like relaxed or meditative states with probability outputs. These methods enhance the accuracy of mental state classification by providing smooth transitions between different states. Their flexibility allows models to handle the complexity of EEG signals, which are non-linear by nature. Non-linear thresholds and sigmoidal functions are thus crucial for robust EEG signal processing in meditation studies. Overall, they improve the real-time analysis and reliability of EEG-based systems.

E. *Adaptive Feedback Generation*

Adaptive feedback generation in EEG-based meditation systems involves providing real-time feedback to users based on their mental state, detected through EEG signals. In the context of EEG datasets used for meditation, this process helps users optimize their mental states, such as relaxation or focus, by dynamically adjusting the feedback based on their real-time brain activity.

(i) Concept of Adaptive Feedback:

Adaptive feedback systems adjust their responses based on the user's EEG signal patterns. These systems continuously monitor the EEG data to assess mental states, such as whether the user is in a meditative, relaxed, or focused state. When the system detects a shift in mental state, it generates feedback that helps guide the user towards a desired state.

(ii) Mechanisms for Adaptive Feedback:

Signal Processing: Continuous monitoring of EEG data allows the system to detect patterns indicative of specific mental states. Features like brain wave frequency bands (alpha, beta, theta) are commonly used to track relaxation and focus levels.Real-Time Adjustment: Feedback is dynamically adjusted based on the detected state. For instance, if a user is not in a relaxed state, the system may provide calming sounds, visual cues, or haptic feedback to encourage relaxation.Thresholding and Adaptation: Adaptive thresholds ensure that the feedback remains relevant to each user's specific EEG patterns, accounting for individual differences. As the user progresses through different mental states, the feedback intensity or type may change in response.

(iii) Applications in Meditation:

Relaxation and Focus Enhancement: In meditation applications, adaptive feedback can guide users by providing prompts when their EEG signals indicate they are losing focus or drifting from a meditative state. This feedback could include auditory signals (like calming music) or visual cues (such as a calming light intensity).

Continuous Learning: Over time, the system learns the user's unique EEG signal patterns, allowing for increasingly personalized feedback. This feedback mechanism adapts to the user's progress, improving their meditation experience by making the feedback more relevant as they develop.

(iv) Benefits:

User Engagement: By providing personalized, real-time feedback, users are more likely to stay engaged and achieve the desired mental state.Real-Time Adaptation: Adaptive feedback ensures that the system responds to changing mental states without needing manual intervention, promoting a smooth and continuous experience.Enhanced Meditation Outcomes: As the system fine-tunes feedback based on the user's specific needs, it helps improve relaxation, focus, and meditation practice.

III. Experimental Setup And Results

The experimental setup for EEG meditation analysis involves signal acquisition using EEG headsets, followed by preprocessing to remove noise and artifacts through filtering and normalization. Key features, including time-domain, frequency-domain, and statistical measures, are extracted from the signals. Machine learning models like SVM, Random Forest, and Deep Learning techniques are then used to classify mental states such as relaxation and meditation. The models are evaluated for classification accuracy, precision, and recall, with results showing high accuracy (80-90%) in identifying meditation states. Real-time adaptive feedback, based on classification results, is provided to users to enhance their meditation experience, and its effectiveness is measured by analyzing EEG signal transitions. Model comparison shows that deep learning models, such as CNN and LSTM, may outperform traditional models in some cases. User feedback indicates improved meditation outcomes when receiving adaptive feedback. Challenges include handling noise and artifacts, which can affect signal quality. The system demonstrates the potential for real-time mental state monitoring and applications in meditation training and mental health.

IV. Conclusion

EEG-based meditation analysis using machine learning models and adaptive feedback systems shows promising potential for real-time mental state monitoring. The system effectively classifies mental states like meditation and relaxation with high accuracy. Adaptive feedback enhances user engagement and helps guide individuals towards their desired mental state. While challenges such as signal noise remain, the results demonstrate that advanced preprocessing and model techniques can significantly improve classification performance. The integration of deep learning models further strengthens the system's reliability. This research opens avenues for applications in mental health, meditation training, and brain-computer interfaces. Overall, the combination of EEG signals and adaptive feedback holds great promise for improving user experiences in various domains.