Journal of Engineering and Technology for Industrial Applications

ISSN ONLINE: 2447-0228

ITEGAM-JETIA

Manaus, v.12 n.57, p. 671-683. January/February, 2026.

DOI: https://doi.org/10.5935/jetia. v12i57.2967

RESEARCH ARTICLE OPEN ACCESS

PREDICTING STUDENTS' CONCENTRATION IN COGNITIVE ACTIVITIES USING EEG AND DEEP LEARNING TECHNIQUES

Yogesh Patil[1], Snehal Mali[2], Harsha Gaikwad*[3], Sanil Gandhi[4], Manjushree Laddha[5] and Arvind Kiwelekar[6]

[1, 2, 3, 4, 5, 6]Department of Computer Engineering, Dr. Babasaheb Ambedkar Technological University, Lonere – 402 103, Raigad (M.S.), India.

[1]https://orcid.org/0009-0003-6128-5740, [2]https://orcid.org/0000-0003-2760-0009, [3]https://orcid.org/0000-0003-2760-0009
[4]https://orcid.org/0000-0003-0407-8975, [5]http://orcid.org/0000-0002-6265-1728, [6]https://orcid.org/0000-0002-3407-0221

Email: *harsha.gaikwad@dbatu.ac.in, ynpatil@dbatu.ac.in, smali@dbatu.ac.in, ssgandhi@dbatu.ac.in, mdladdha@dbatu.ac.in, awk@dbatu.ac.in

ARTICLE INFO

Article History
Received: November 18, 2025
Revised: December 20, 2025
Accepted: January 15, 2026
Published: February 28, 2026

Keywords:
Electroencephalogram (EEG),
Concentration Level Prediction,
Deep Learning Algorithms,
Cognitive Engagement

ABSTRACT

In an era of social media and online learning platforms, there are several opportunities for learning different technologies and topics that students do not easily understand. However, it also presents challenges by diverting students’ attention, such as notifications, multitasking activities, advertisements, etc. Assessing students' level of focus during cognitive tasks is crucial and complex. This study evaluates students' cognitive engagement through various activities, including arithmetic calculations, reading technical articles, listening to technical podcasts, reading transcripts, browsing the internet, and engaging in relaxation exercises, utilizing EEG signals. Concentration levels are classified using deep learning algorithms, specifically Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Artificial Neural Networks (ANN). The performance of these algorithms is also evaluated based on metrics such as accuracy, F1 Score, precision, and loss.

I. INTRODUCTION

With the advent of technology, the utilization of mobile phones has witnessed a surge across all age groups. Research has revealed that social networking applications offer users shorter material and can significantly increase their dopamine levels [1]. The floods of dopamine induced by these spikes lead individuals to develop an addiction to rapidly consuming stuff, so negatively impacting their capacity to sustain and concentrate on specific jobs for prolonged durations and acquire new knowledge. It is vital to evaluate the cognitive involvement of an individual when performing various jobs. This will help us determine their degrees of attentiveness to different cognitive focus requirements.

Within the realm of education, the ability to discern pupils’ degrees of concentration during the acquisition of new knowledge might potentially enable us to intervene to sustain their attention [2]. The concentration level of students can be inspected by examining their eye contact, body language, and response to a question or communication feedback [3]. The concentration examination techniques are subjective and can be influenced by the student’s mood or the external environment. Additionally, different students may have different concentration indicators, so applying one metric to all students will be accountable for inconsistency. It isn’t easy to manually assess the concentration level of a student while performing any cognitive task.

To address these issues, a data-driven, objective process is needed to assess concentration levels. EEG is a promising solution for quantifying cognitive focus and engagement [4]. EEG captures electrical signals caused by neuronal communication in the brain, representing the fine-granular, real-time cognitive states. Initially, the EEG device is used for medical diagnosis, for example, for diagnosing epilepsy [5], Parkinson’s disease [6], Alzheimer’s disease [7], sleep disorders [8], and emotion detection [9], [10]. EEG is now being seriously examined in non-medical fields, such as in education [11–13], human-computer interfaces [10], and cognitive workload estimation [14], [15].

The success of previous studies in classifying cognitive states [14], human emotions [9], and comprehension levels [11], [16] in learning exercises using EEG signals is a proven track record. Cognitive engagement is directly linked to brain activity; EEG can provide a more timely and accurate reading of levels of focus than a test of behaviour [17]. When humans perform different cognitive efforts, corresponding EEG patterns have different properties [16]. The complexity of EEG signals makes them challenging to analyze, as they are highly non-linear, multi-dimensional, and susceptible to noise. Although traditional machine learning algorithms have been proven capable of doing well in cognitive state classification based on EEG, it is hampered by their hand-engineered feature dependency [18]. The feature engineering process in analyzing EEG is traditionally complex and domain-specific, making scaling difficult.

The traditional algorithms also struggle to capture deep hierarchical representations in EEG signals [19], leading to a non-optimal outcome in concentration prediction. Deep learning is a revolutionizing tool for processing EEG signals, overtaking traditional machine learning algorithms [19]. The combined usage of ANNs and LSTMs in a joint model is highly accurate in cognitive state decoding, making it highly suitable for predictive levels of concentrations [16]. The intersection of deep learning and the processing of EEG signals is opening new paths for real-time monitoring of concentrations in learning, workplaces, and clinics [20]. In online learning contexts, EEG-based prediction of concentration can support adaptive learning systems reacting in a timely manner to learners' cognitive states [11–13], [21].

By identifying periods of lower levels of attention, one can adopt a smart tutoring system for teaching and improve the motivation and retention of the students. In occupational contexts, EEG-based concentration monitoring can enhance productivity by making tailored suggestions for inter-burst periods, task reconfigurations, and workloads [22], [23]. Additionally, in clinical contexts, prediction of levels of concentration can assist in diagnosing, in an early stage, attention-related diseases such as Attention-Deficit/Hyperactivity Disorder (ADHD), allowing for directed treatments for enhanced cognitive well-being [24], [25]. While it is potentially valuable, several challenges exist for EEG-based concentration prediction. The signals in an EEG can vary significantly between different individuals, necessitating robust feature extraction algorithms and adaptation in a specific domain [26], [27]. Computative efficiency is critical for real-time usage, especially for a wearable EEG device [28].

Transfer learning, data augmentation, and attention mechanisms have been tried to enhance a model's adaptability and generality in different populations [12], [21]. This study examines the EEG data collected during five different tasks. Each of these activities necessitates different degrees of cognitive involvement from an individual, necessitating differing amounts of concentration [12]. We employ deep learning algorithms to categorize these signals according to their concentration categories. The rest of the article is organized in the following sections: The related work in detailed and background about the several terminologies used in the experiment is described in section 2, followed by a detailed description of the methodology in section 3. The performance measures used for the described methods are in section 4. The results of the proposed methodologies are discussed in section 5, and finally, the article concludes in section 6.

II. RELATED WORK AND BACKGROUND

This section reviews previous research work using EEG signals to predict concentration, attention level, and cognitive load using deep learning algorithms. The authors [19] presented a comprehensive review of recent advancements in deep learning methodologies for classifying EEG signals. The paper also elucidated the efficacy of various deep learning architectures, including ANN, CNN, LSTM, and MLP, in the processing and analysis of EEG data and compares deep learning approaches with the traditional machine learning techniques, highlighting the superior accuracy and efficiency achieved through deep learning in EEG classification tasks. The authors [16] preferred EEG recordings to classify students' cognitive states during learning activities automatically. The data preprocessing techniques included noise reduction and artefact removal from raw EEG signals, followed by feature extraction using summary statistics from various frequency bands.

Several classification algorithms were used in the study, with logistic regression serving as the primary technique for model training. These algorithms include local models, global models, and multi-task learning approaches. According to the data, global models—which take into account more contextual information—predict recall levels more accurately than local models, which predict concentration levels with an accuracy of 91.88%. To address the impact of distractions on cognitive performance, the authors used EEG signals to classify human concentration levels [21], [29]. The dataset comprises brain signals collected using an EEG device from individuals in both relaxation and concentration states. This data was preprocessed using high-pass filtering to eliminate low-frequency noise and segmentation of signals into the delta, theta, alpha, and beta frequency bands.

Statistical features were calculated using the PyEEG library to extract features. The classification of concentration levels is executed using the Support Vector Machine (SVM) algorithm with a radial basis function (RBF) kernel, achieving an overall accuracy of 84% to classify EEG signals into low and high concentration levels. The research findings of [30] suggested that effective brain-machine interfaces can be developed for distinguishing mental states, even with inexpensive EEG devices. The authors used a low-cost, toy-grade EEG device to collect EEG signals to classify relaxed and concentrated mental states. Data was preprocessed by converting raw EEG signals into time-frequency representations and aggregating spectral components into EEG bands such as alpha, beta, and gamma.

The SVM and multilayer feedforward neural networks (BPNN) classification algorithms were preferred in the study. Experimental results demonstrated that using features from the combined alpha, beta, and gamma bands with a bandwidth of 4Hz yields an average accuracy exceeding 80% across subjects, with some individuals achieving accuracies above 90% using the SVM classifier. The authors put forward a strong correlation between students' concentration levels and learning outcomes by analysing brain signals captured by an EEG device [31], [32]. The brain wave data was collected from junior high school students performing cognitive tasks, specifically the Culture Fair Intelligence Test (CFIT) and the Indonesian Competency Test (CT). EEG signals were extracted using Fast Fourier Transform (FFT) followed by normalization to standardize the data range.

The study used the K-Nearest Neighbor (KNN) algorithm, with a K-value of 5 for identifying patterns in the extracted data and reports an impressive classification accuracy of 94.59% to classify concentration levels. The classification accuracy was evaluated using K-Fold Cross Validation with a K-value of 11. The study by [22] investigated young adults' (ages 18-30) cognitive activity and concentration levels using EEG data. At the same time, participants performed IQ tests based on the Standard Progressive Matrices

(SPM). The data preprocessing involved signal reduction and filtering beta waves in the frequency range of 13-30Hz to enhance the data quality. The classification of concentration levels was performed using the k-nearest Neighbor (KNN) algorithm into three categories: high, medium, and low. The model achieved an accuracy of 70%, providing insights into the correlation between brain activity and cognitive performance. The paper presented by [33] gave a method for classifying EEG signals into three mental states: relaxation, moderate, and high concentration.

The brain signals were captured using an EEG device with four electrodes and preprocessed by applying a large Laplacian filter, a notch passive filter, and a wavelet filter. Mental states in the time domain were examined by segmenting the measured signal into windows, from which various features are extracted. These extracted signals were utilized in machine learning algorithms to classify the different mental states. The two-fold cross-validation method was used to evaluate the classification. A novel model tailored for embedded systems to classify attention levels while performing specific tasks. This system used a single-channel EEG device to capture brain signals that were subsequently increased through spectrum analysis to increase the dimensionality of the features. The authors used simulated annealing and geometric particle swarm optimization algorithms to optimize feature selection. Classification of mental states (concentrated or relaxed) was achieved using an SVM. Furthermore, Echo State Networks, a class of neural networks well-suited for embedded systems, demonstrated satisfactory accuracy in meeting the model’s requirements [34].

The research paper [13] tackled the critical issue of assessing student attention in online learning, a challenge intensified by the COVID-19 epidemic. The authors proposed a brain-computer interface (BCI) system that uses EEG data to evaluate student involvement and attention levels. A publicly available EEG dataset was preferred, and the quick Fourier transform extracted power spectral density (PSD) features and various attention metrics were calculated. The study encompassed three classification algorithms: KNN, SVM, and random forest (RF), with the RF model achieving the highest accuracy of 96% in holdout validation. The method encompassed data collection, preprocessing, feature extraction, and model training, culminating in classification performance assessment. The results indicated that the suggested RF model successfully differentiates attention states, presenting a viable alternative for improving online learning experiences.

The classification of normal and pathological EEG patterns using deep learning models enhanced with attention mechanisms is investigated in the paper [26]. A systematic approach was used to preprocess EEG data to filter and normalize signals, extract time-domain and frequency-domain features, and segment the data into frames for study. The authors tested InstaGATs, LSTM with Attention, and CNNs with Attention against baseline models without attention mechanisms. The CNNs with Attention model has the highest accuracy of 96.98% on the TUH EEG Seizure dataset. The research faced intrinsic noise in EEG recordings, the difficulty of distinguishing anomalies, and the need for efficient feature extraction from varied datasets. Attention processes can improve classification performance by stressing relevant data items and enhancing clinical EEG analysis interpretability and efficacy. Attention-enhanced deep learning models can significantly improve EEG classification tasks, enabling more robust implementations in real applications.

II.1 BACKGROUND

II.1.1 EEG Device

The neuron is the smallest unit of the human brain. Electric signals are generated when neurons communicate with each other. An electroencephalogram is a device used to capture the electrical signals generated by neurons in the brain. These electrical signals are recorded with the electrodes of an EEG device. With the help of electric signals captured by EEG devices, abnormal changes in the brain can be detected, which is why EEG devices are widely used in the medical sector to diagnose brain-related diseases such as epilepsy, sleep disorders, and so on. In recent years, EEG devices have also been utilized in the education sector to solve various research problems, such as determining attention levels, mental engagement, and cognitive load. This study uses a MUSE EEG device to record brain wave data while performing various cognitive tasks. This device has four electrodes named FP1 and FP2, located in the frontal lobe of the brain, and TP9 and TP10, located in the parietal lobe of the brain, as shown in Figure 1.

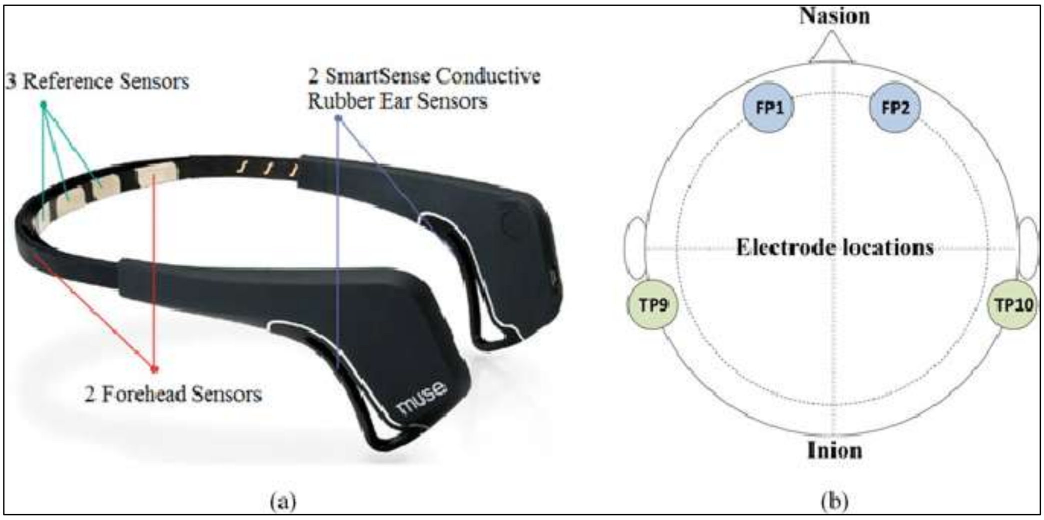


Figure 1: MUSE EEG Device.
Source: Authors, (2026).

Table 1: Types of Brain Waves.

Types of Wave	Frequency (Hz)	Task
Delta	0.5 – 4	Deep sleep
Theta	4 – 8	Relaxing or meditating
Alpha	8 – 12	Just after woke up and before going to sleep
Beta	12 – 35	Routine work
Gamma	> 35	Deep concentration on some task or doing cognitive activity

Source: Authors, (2026).

II.1.2 EEG Signals

Five different types of brain waves are emitted by the brain and recorded using an EEG device, as shown in Table 1. Delta and theta waves are small amplitude waves emitted, implying an individual's deep sleep or relaxation. Their frequency ranges from 0.5Hz to 8Hz. These waves are produced in large quantities as compared to other types of waves while doing tasks such as relaxing with closed or open eyes, browsing the internet, and so on. When humans are in partial sleep, waking up, or transitioning into a deep sleep state, the brain emits alpha waves with a frequency range of 8Hz to 12Hz. The beta and gamma waves indicate high cognitive activity. Their frequency range is 12Hz to 35Hz and >35Hz, respectively. Among all the waves, gamma and beta waves are produced in large quantities with more intensity while doing high-concentration tasks such as doing arithmetic calculations or reading technical articles.

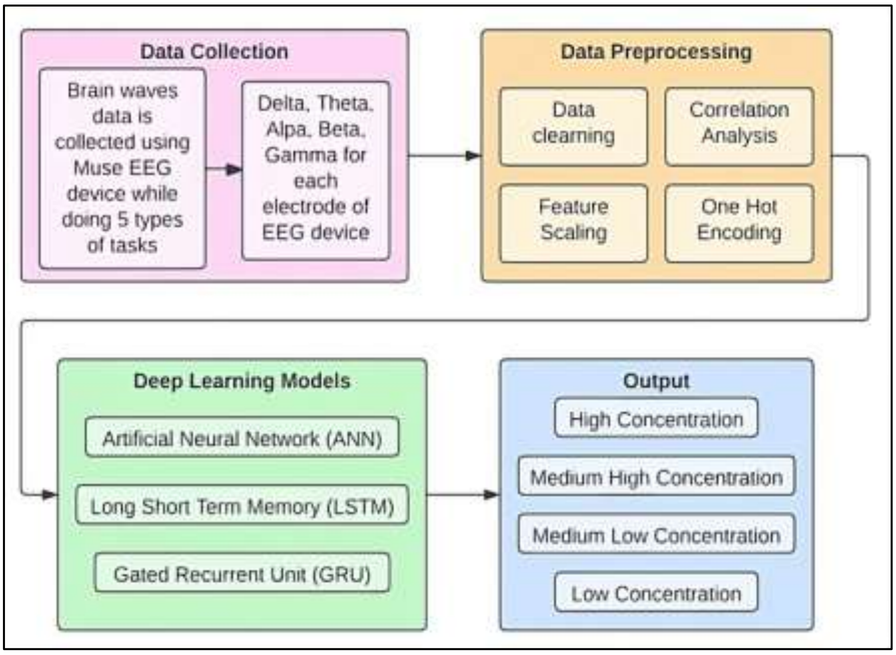


Figure 2: Flowchart Diagram of the Proposed System.

Source: Authors, (2026).

III. METHODOLOGY

The proposed methodology for predicting concentration levels utilizing EEG signal inputs and deep learning algorithms is presented in Figure 2. The procedure has four major phases: Data Collection, Data Preprocessing, Deep Learning Models, and Output Classification. Data collection refers to the methodology employed in acquiring the necessary brain wave data for forecasting concentration levels using an EEG device, the parameters included in the dataset, the number of features present, and the associated concentration levels for each recorded instance. The subsequent phase is data preprocessing, which involves preparing the data as input for deep learning algorithms. The process entails cleaning the data to eliminate redundant data, correlation analysis to identify pertinent features for predicting their target variable, feature scaling to prevent bias in weight assignment during deep learning model training, and one-hot encoding for the transformation of the target variable. The subsequent phase involves training various deep learning models—ANN, LSTM, and GRU—and evaluating their accuracies on test data to assess their performance on unknown data. The data input for these algorithms is categorised into four distinct classes: high concentration, medium-high concentration, medium-low concentration, and low concentration.

Table 2: Concentration Level for Corresponding Tasks.

Task	Concentration Level
Performing Arithmetic Calculations	High
Reading Technical Articles, listening to technical podcasts	Medium-High
Reading the transcripts	Medium-High
Browsing the Internet	Medium-Low
Just sitting there, eyes open or closed	Low

Source: Authors (2026).

III.1 DATA COLLECTION

The dataset used for predicting an individual's concentration level during activities of varying levels of attention is shown in Figure 3. An individual wears the Muse EEG device on their scalp while participating in five activities. Table 2 gives the activities' details and associated concentration levels. The EEG device has four electrodes, to which we assign the numbers 0, 1, 2, and 3. Five different types of brain waves—delta, theta, alpha, beta, and gamma—are captured from each electrode while performing each task.

The captured brain waves are labelled with their names and subscripted with the electrode number to neatly identify the features in the dataset. For instance, if electrode 0 captures waves in the frequency range of 0.5 to 4Hz (delta wave frequency), we will refer to that frequency as "delta0". In this way, all 20 features (five types of brain waves from each of the four electrodes) in the dataset are captured by an EEG device and labelled according to the given method. For each record, the concentration level (low, medium low, medium high, and high) has been assigned as shown in Table 2. This dataset serves as the basis for identifying an individual's cognitive engagement.

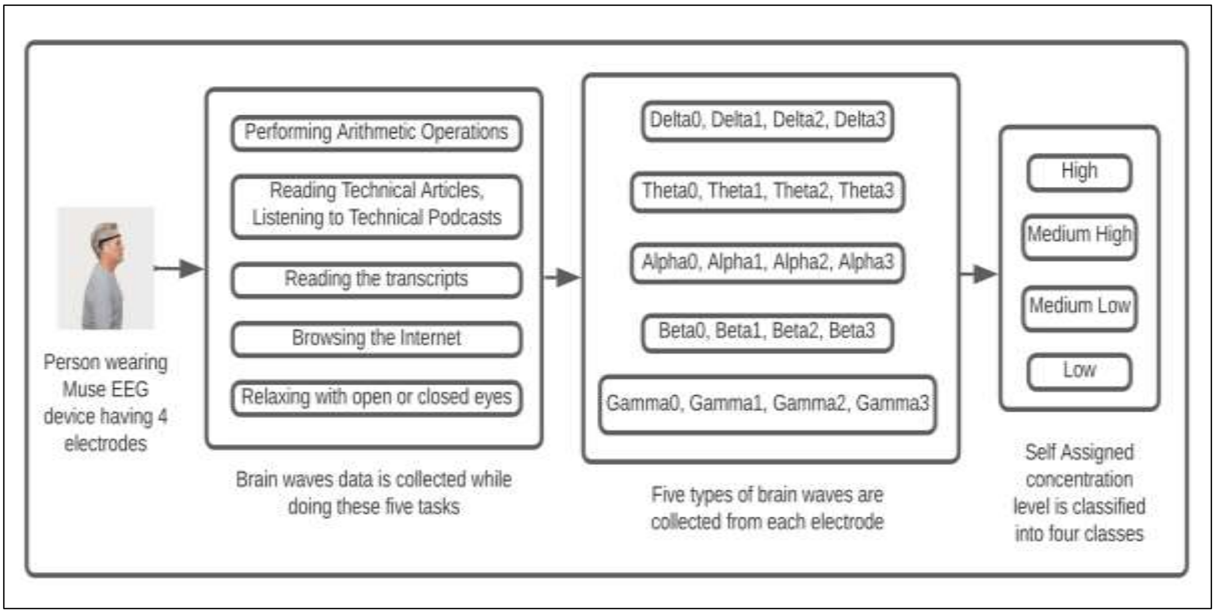


Figure 3: Data Collection Process.
Source: Authors, (2026).

III.2 DATA PREPROCESSING

Data preprocessing is the most crucial step after data collection and before applying any deep learning algorithm. The following four steps are performed to prepare the collected brain wave data in the required format to provide as input to the algorithms.

a) **Data Cleaning:** When an EEG device's electrodes capture brain wave data, there is a strong possibility that eye movements or even slight head movements could introduce noise into the data. Eliminating that noise is important before going ahead to ensure the correct predictions. The integrated software on the device has already removed noise from the collected dataset.

b) **Correlation Analysis:** Correlation analysis involves finding the direct relationship between the target variable and the input features of the dataset. There are two types of correlation: positive correlation, which means if input features are increasing, then the value of the target variable will also increase, and vice-versa. Another type is negative correlation, which means if input features are increasing, then the value of the target variable will decrease, and vice versa. In the current study, the target variable is an individual's concentration level, while the input features are five types of brain waves captured from four electrodes. Correlation analysis proves no positive or negative correlation between the input features and target variables.

c) **Feature Scaling:** Due to the varied frequency ranges in brain wave data, deep learning models might give greater importance to higher-magnitude values during training and less importance to lower-magnitude ones. To address this possible bias, a scaling approach is used to normalize the data by adjusting each feature's minimum and maximum values to 0 and 1, respectively. The following formula will calculate the new scaled value:

$$X_{scaled} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

This formula normalizes the values of all features. This process ensures that all features are treated equally, preventing magnitude-based bias in the training process.

d) **One Hot Encoding:** The concentration level feature is converted using one-hot encoding, a technique that transforms categorical data into numerical arrays with a length equal to the number of unique categories as stated by [12]. For this case, we have four distinct categories of concentration levels, i.e., High concentration, Medium high concentration, Medium low concentration, and Low concentration, resulting in an array length of four.

This procedure produces binary vectors, in which the element matching the appropriate class is assigned a value of one while all other elements are set to zero. For example, if we had a brain wave record that indicates high concentration during a task, the corresponding one-hot encoded array would be [1 0 0 0], where 1 will denote the current record is for the High concentration level task, and other classes will be set to 0. In this way, brain wave records are encoded using one-hot encoded arrays and then used as input for future processing and analysis.

III.3 DEEP LEARNING ALGORITHMS

Preprocessed data is provided as input to three deep learning algorithms: Artificial Neural Network (ANN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU). All the models use one input layer, three hidden layers, and one output layer. The output layer contains four neurons and uses the softmax activation function, which is best for multiclass classification problems. All three models are trained with 70% and tested on 30% of the data. A categorical cross-entropy loss function is used since predicting the concentration level is a multi-class classification problem. To update the weights while training these models, the Adam optimizer is used as an optimization algorithm instead of the classical stochastic gradient descent procedure.

a) **Artificial Neural Network (ANN):** The neuron is the fundamental component of an Artificial Neural Network (ANN). ANN consists of several linked neurons that process information by applying weights and activation functions. Each time the neuron receives an input, a random weight is assigned to that value, which shows the significance of that value in the prediction process. Subsequently, all the data is combined and given to an activation function. This activation function introduces non-linearity in the data. Lastly, the output is computed. The error is backpropagated, and the weights are modified based on the difference between the calculated and actual output to minimize the difference. The accuracy of an ANN increases as the difference decreases. The preprocessed data of the brain waves is given as input to the ANN to predict the concentration level while doing different cognitive tasks. The architecture of the ANN model is illustrated in Figure 4. The input layer of the ANN consists of 12 neurons and employs the ReLU activation function. It features three hidden layers with 128, 256, and 128 neurons, each utilizing the ReLU activation function. The output layer comprises four neurons, which classify the data into four different concentration classes and use the softmax activation function, which is optimal for multiclass classification tasks.

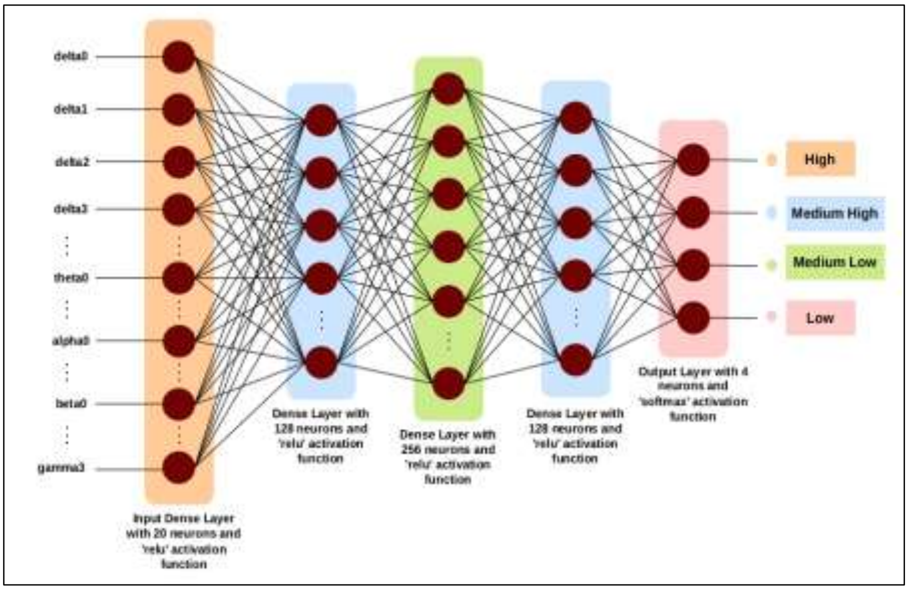


Figure 4: Architecture of Proposed Artificial Neural Network.
Source: Authors, (2026).

b) **Long Short Term Memory (LSTM):** Brain wave data refers to the frequency of electrical activity recorded over time, similar to time series data. Recurrent neural networks do well in processing time series data, but cannot learn long-term dependencies. LSTM networks are a specific form of recurrent neural networks that successfully overcome this drawback. The fundamental structure of LSTM includes the cell state, the input gate, the output gate, and the forget gate. The cell state is the central part of the LSTM network, which stores long-term dependencies and decides which information to let go of and which information to send for further analysis. The other three gates regulate the flow of information: the input gate admits new data, the output gate outputs the information, and the forget gate erases redundant data. The LSTM network uses the truncated backpropagation of error to enhance its accuracy and train it to understand long-term data dependencies. The architecture of the LSTM has been illustrated in Figure 5.

It consists of three LSTM layers, each containing 128, 256, and 128 neurons. The activation function used in these layers is the Rectified Linear Unit (ReLU). The input comprises preprocessed and reshaped data that is suitable for LSTM cells. After undergoing additional processing through three LSTM layers, the data ultimately reaches the output layer, which consists of four neurons for each concentration level class. The softmax activation function is employed as it is the most suitable choice for multiclass classification tasks.

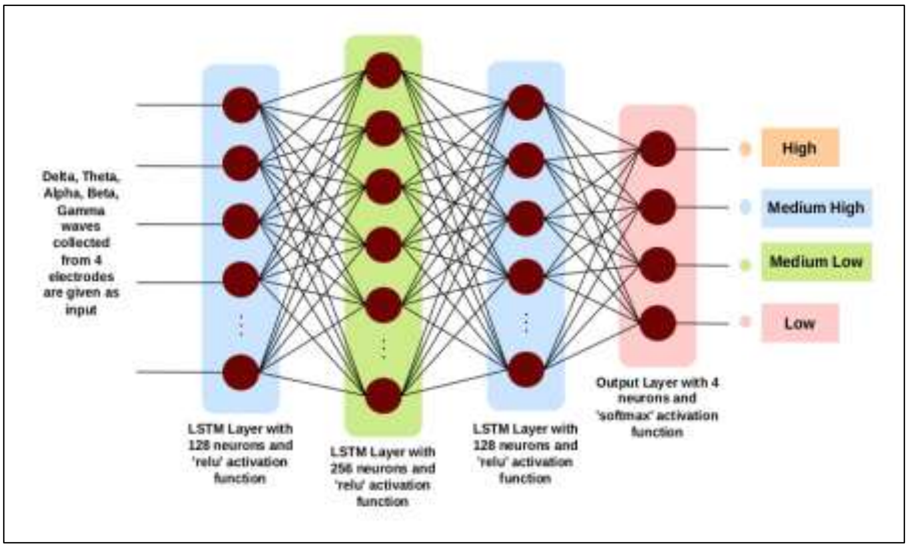


Figure 5: Architecture of Proposed Long Short-Term Memory.
Source: Authors, (2026).

c) **Gated Recurrent Unit (GRU):** GRUs is a kind of RNN that addresses the limitations of LSTM networks and traditional RNNs. Unlike LSTM, which is a cell state, the basic unit of GRU is the hidden state. Compared to LSTM, the GRU has fewer gates. The gates include the reset gate, which forgets redundant information, and the update gate, which balances past information and new important data for longer. Due to its reduced number of gates and parameters, the GRU model can be trained more quickly than the LSTM. The architecture of the GRU model has been illustrated in Figure 6. It consists of one input layer, three dense GRU layers, each containing 128, 256, and 128 neurons, and the output layer, which consists of four neurons for each concentration level class. The ReLU is the activation function in these layers. The input comprises preprocessed and reshaped data that is suitable for GRU cells. After undergoing additional processing through three GRU layers, the data ultimately reaches the output layer, where the softmax activation function is employed because it is the most suitable choice for multiclass classification tasks.

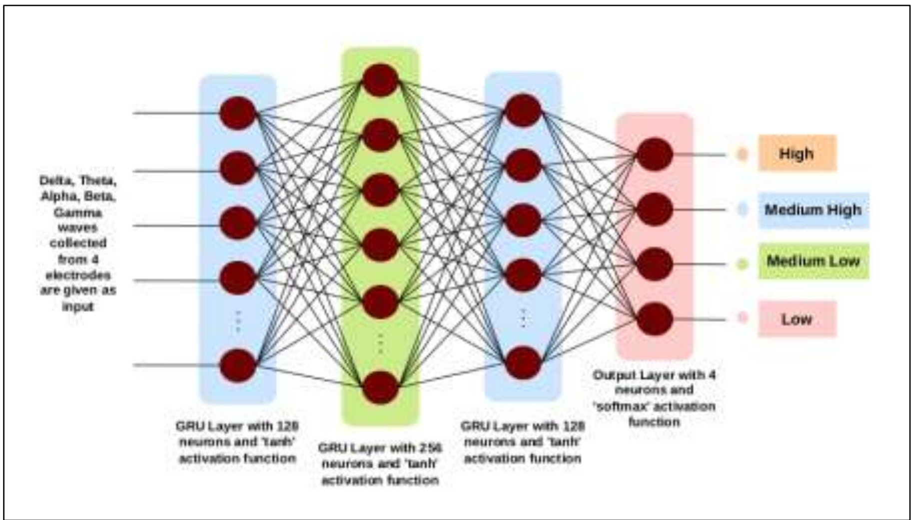


Figure 6: Architecture of Proposed Gated Recurrent Unit.
Source: Authors, (2026).

IV. PERFORMANCE MEASURES

Performance evaluation ensures the proposed methodology's efficiency, validity, and applicability. The proposed models are evaluated using the following metrics.

a) **Confusion Matrix:** A confusion matrix is a crucial performance evaluation tool that encapsulates the accuracy of a classification model. It helps to determine where our model underperforms when differentiating between classes. The matrix enumerates the counts of true positives (TP), which reveals how frequently a model accurately identifies a positive sample as positive; true negatives (TN), which means the frequency with which a model labels a positive sample as negative; false positives (FP), which indicate how frequently a model wrongly labels a negative sample as positive; and false negatives (FN), which reflects how often a model accurately labels a positive sample as Negative. According to data analysis, these four categories can be specified for the high concentration class as follows: TP refers to the number of instances in which the recorded brain waves indicative of high concentration were accurately classified by proposed algorithms.

TN relates to cases where the proposed algorithms accurately predict brain wave recordings classified as non-high concentrations belonging to other categories, such as medium-high, medium-low, or low concentrations. FP denotes false positives, indicating the frequency of brain wave records that belong to different classes but are incorrectly classified by the proposed system as belonging to the high-concentration class. FN denotes false negatives, indicating the frequency of brain wave recordings associated with high-concentration tasks misclassified as belonging to concentration classes other than high-concentration activities when processed by the proposed system. Specifically, a confusion matrix is an N × N matrix employed to assess the performance of a classification model, where N denotes the total number of target classes. It compares the actual values of the target class in the dataset with the predicted target class by the proposed model; if both match, then the prediction is correct; otherwise, it is not. It helps to identify how well our classification model performs and what kinds of errors it makes.

b) **Accuracy**: Accuracy metrics provide insight into how well the proposed model classifies a specific data set into the appropriate class. Accuracy is calculated as the percentage of accurate predictions over all other predictions.

$$Accuracy = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{2}$$

Equation 2 provides the formula for calculating accuracy and finds the per centage of correctly classified brain wave records to each class among all the predictions.

c) **Precision:** Precision measures the percentage of true predictions within all positive predictions. The proportion of accurately categorized positive samples (TP) to the total number of positively classified samples (either correctly or incorrectly) is known as precision, as given in Equation 3.

$$Precision = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{3}$$

Regarding the classification problem, precision is the rate at which the proposed model accurately classifies a given brain wave record to the actual concentration class. The precision value increases as the number of FP decreases. FP refers to the number of records that do not belong to a specific class but are falsely classified as belonging to that class. FP increases when the model has not accurately learned the parameters to detect records that do not belong to the current class. That is why the model erroneously classifies the brain record as positive. Increased precision indicates that the model has effectively learned the parameters that allow it to classify records accurately, distinguishing which belong to a specific class.

d) **Recall:** The recall is determined as the proportion of correctly identified positive samples, i.e., TP to the total number of actual positive samples, which encompasses both TP and those that are positive but incorrectly identified as negative (FN), as given in Equation 4 as shown below:

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{4}$$

The recall value increases as the FN decreases. FN represents the count of records that belong to a specific class but are misclassified as belonging to another class. FN will increase when the model has not accurately learned the parameters necessary to identify data belonging to the current class. That is why the model erroneously categorises the brain record as negative. A higher recall indicates the model has not effectively learned the parameters to classify this data into the right class.

e) **F1 Score:** The F1 Score, defined as the harmonic mean of precision and recall, offers a consolidated metric that integrates precision and recall, thus equilibrating the trade-off between these two measures. The formula to calculate the F1 Score using precision and recall is given in Equation 5.

$$F1\,Score = 2 * \frac{\mathrm{Precision} * \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{5}$$

Since FP influences precision, and FN influences recall. Evaluating both precision and recall when computing the F1 Score enables one to assess the model's capacity to identify FP and FN. As the number of data points for each class differs in an unbalanced dataset, the proposed model must discern the parameters that accurately classify each record as belonging or not belonging to a specific class. If precision is the sole performance indicator, the model's capacity to accurately identify positive records may be overlooked. On the other hand, focusing solely on recall could hinder the model's capacity to identify bad records. Both features are essential for an effective model. The F1 Score is an excellent indicator for assessing model performance, as it mitigates the substantial bias inherent in precision and recall.

V. RESULT ANALYSIS AND DISCUSSION

This section presents the results of implementing deep learning methods for predicting concentration levels, demonstrating how an individual engages cognitively in different attention activities. The dataset has 20 features, including five unique brain wave frequencies, each recorded from four electrodes. The data set undergoes data preprocessing, which includes cleaning, correlation analysis, feature scaling, and one-hot encoding of the target variable. After this, the dataset is fed into three deep-learning algorithms to predict concentration levels in four categories. The preprocessed dataset is partitioned into two subsets, with 70% of the data utilized for training the model and the remaining 30% used for evaluating its performance.

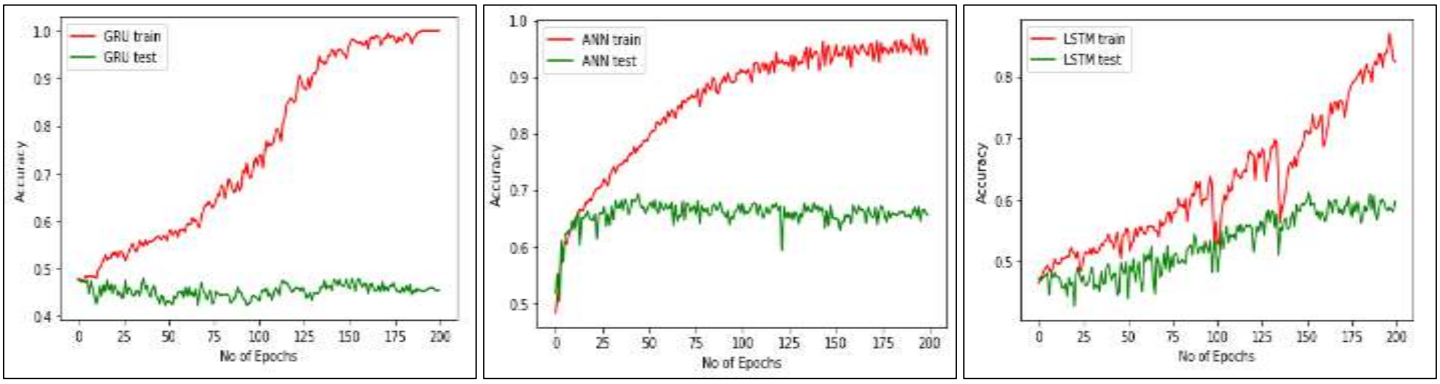
The models, including ANN, LSTM, and GRU, are trained for 200 epochs using different batch sizes. The ANN model uses a batch size of 5, the LSTM model uses a batch size of 64, and the GRU model uses a batch size of 32. The additional parameters for training these models are specified in Table 3. The performance of models is evaluated using various measures, including accuracy, precision, recall, and F1 Score.

Table 3: Parameter values used to train deep learning models.

Parameter	Values
Optimizer	Adam
Loss Function	Categorical Cross-entropy
Metric	Accuracy
Epoch	200

Source: Authors (2026).

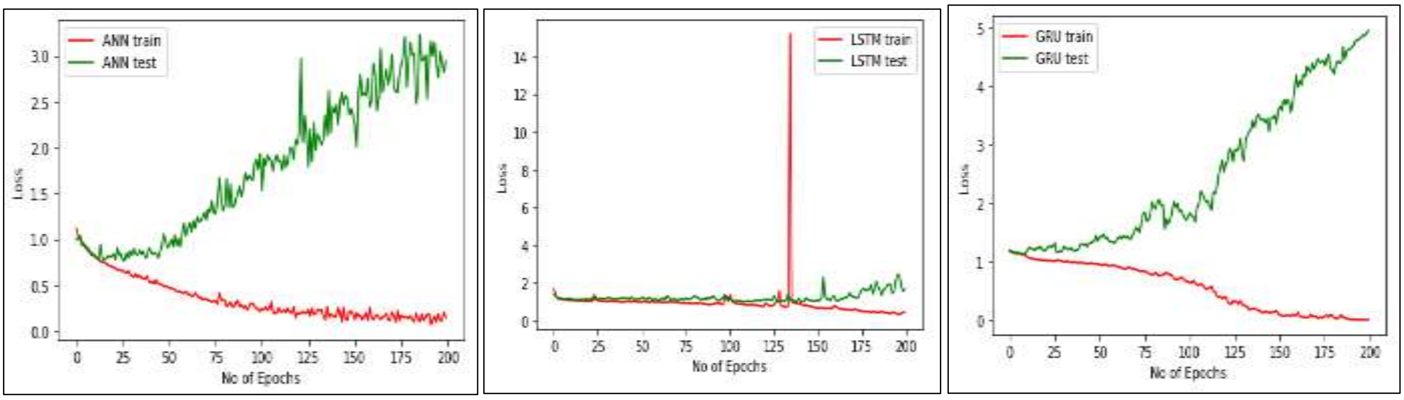
Figure 7 shows that ANN's training and testing accuracies steadily increase up to 50 epochs. After that, although there is an increase in training accuracy, the testing accuracy shows only a slight improvement. In the case of LSTM, both the training and testing accuracies show fluctuations until the 125th epoch. However, after this point, the testing accuracy remains constant, regardless of the continued increase in training accuracy. The GRU model consistently improves training accuracy up to 100 epochs, whereas testing accuracy remains relatively stable. The ANN, LSTM, and GRU models demonstrate overfitting of the data at epochs 50, 125, and 50, respectively.



(a) GRU (b) ANN (c) LSTM

Figure 7: Training and Testing Accuracy of Deep Learning Algorithms.
Source: Authors (2026).

The same phenomena are observed from the loss graphs in Figure 8. The training and testing loss for ANN reduces up to 50 epochs, LSTM up to 125 epochs, and GRU up to 100 epochs. However, after that point, only the training loss decreases while the testing loss increases, suggesting that the models are overfitting the data. Figure 9 compares the classification accuracy of algorithms ANN, LSTM, and GRU. The ANN achieves maximum accuracy, which is 65.63%, followed by LSTM, which is 59.67%, and GRU networks, which is 45.39%.



(a) ANN (b) LSTM (c) GRU

Figure 8: Training and Testing Loss of Deep Learning Algorithms.
Source: Authors (2026).

The imbalance of the data set, as indicated in Table 4, compromises the reliability of accuracy as a sole metric for assessing the performance of deep learning models. There is a need to determine the model's ability to correctly classify the major class (containing a more significant number of records) and the minor class (containing a smaller number of records). The accuracy of the models for each class is provided in the confusion matrix for each deep learning technique mentioned in Table 5. The confusion matrix indicates that the ANN model exhibits a higher count of TP and TN than the LSTM and GRU models. Precision verifies the model's quality, determining the accurate classification of positive samples from the total number of positive samples.

As depicted in Figure 10a, the ANN model exhibits marginally superior precision to the LSTM network for the low and medium high concentration classes. Conversely, the LSTM network demonstrates higher precision than the ANN for the high and medium low concentration classes. Recall is a measure of quantity that identifies the number of correctly classified positive samples from the total number of positive samples. A model with a high recall value will yield more relevant results. Based on the data presented in Figure 10b, it is evident that ANN exhibits superior performance in tasks involving high and medium-high concentrations. On the other hand, LSTM networks demonstrate better performance in tasks involving low concentrations. ANN and LSTM networks exhibit similar recall values for tasks involving medium-low concentrations.

Table 4: Number of records for each concentration level class.

Class	No. of Records
High Concentration	3410
Medium_High Concentration	501
Medium_Low Concentration	909
Low Concentration	2423

Source: Authors, (2026)

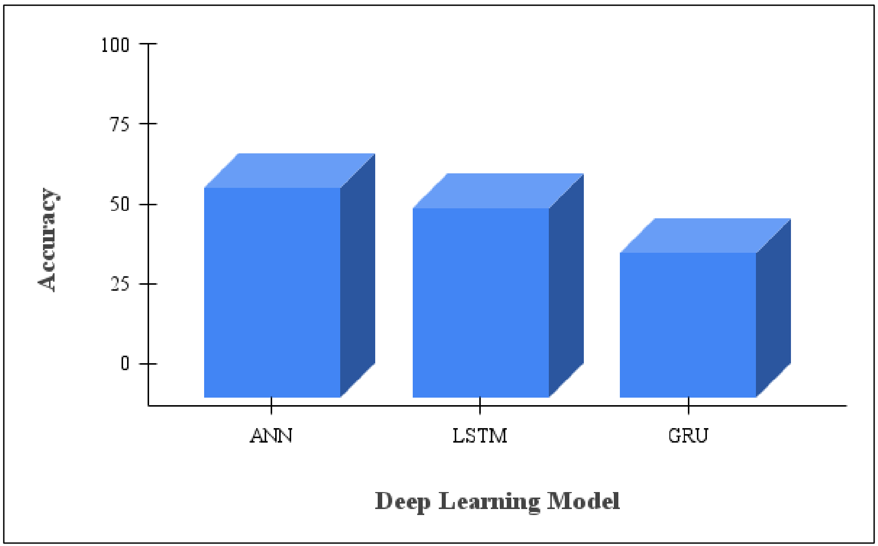


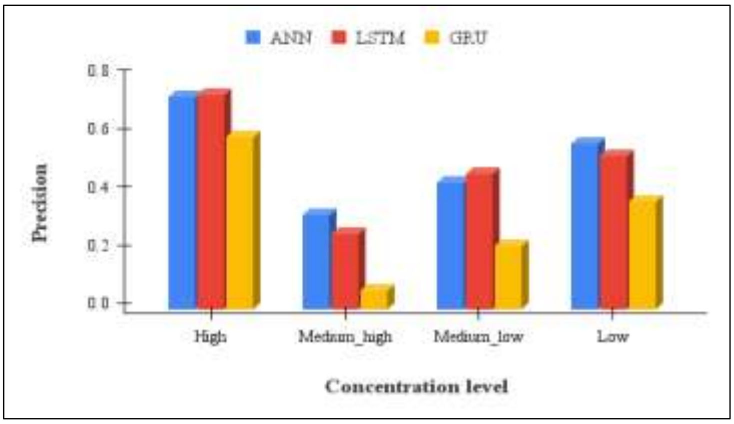
Figure 9: Comparison of Accuracies of Deep Learning Algorithms.
Source: Authors, (2026).

Support is a measure of the class frequency that appears in the dataset. Due to the imbalance in the dataset, which can be seen in Figure 10d, the precision and recall values are likely to be biased. To address this issue, the F1 Score metric, the harmonic mean of precision and recall, is computed to measure the model's overall performance. Figure 10c shows that the ANN model has a higher F1 Score for high and medium high concentrations classes than other models. For low and medium low concentration classes, LSTM networks have a slightly higher F1 Score than ANN.

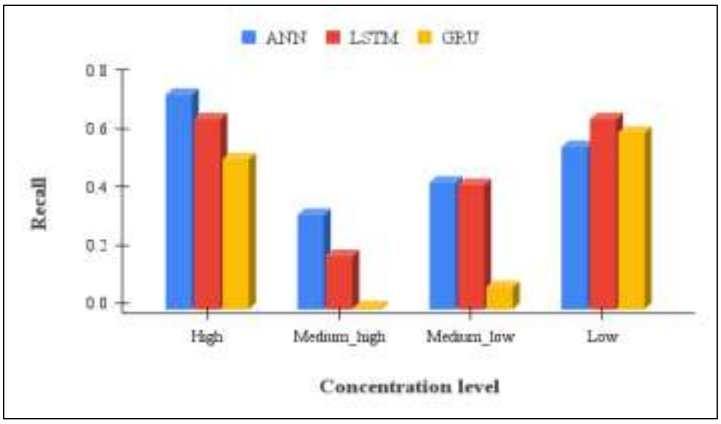
Table 5: Confusion Matrix of Deep Learning Algorithms.

Algorithm	True Value	Predicted Value			
		High	Medium_High	Medium_Low	Low
ANN	High	575	38	27	74
	Medium_High	36	32	8	19
	Medium_Low	32	13	56	77
	Low	117	11	46	288
LSTM	High	547	34	44	209
	Medium_High	53	22	10	34
	Medium_Low	27	10	100	96
	Low	114	18	60	380
GRU	High	430	6	45	353
	Medium_High	44	1	9	65
	Medium_Low	42	2	18	171
	Low	207	6	10	349

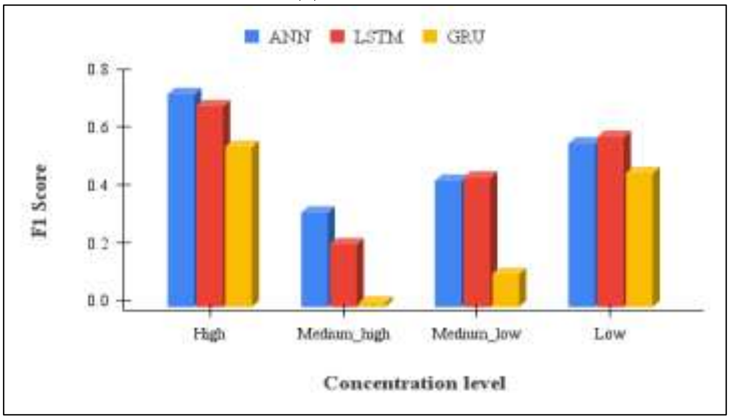
Source: Authors, (2026).



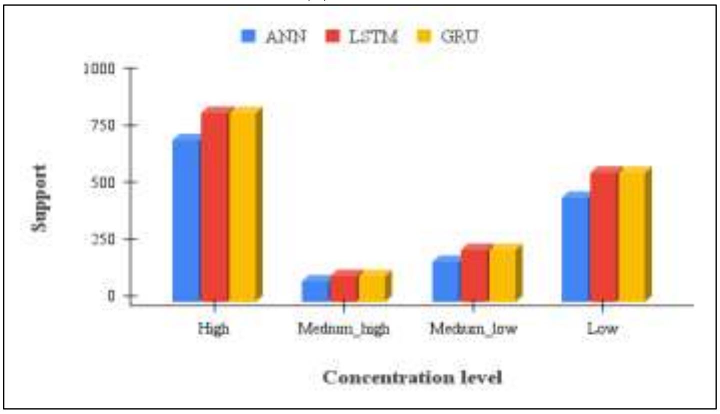
(a) Precision



(b) Recall



(c) F1 Score



(d) Support

Figure 10: Performance measures of Deep Learning Algorithms.
Source: Authors, (2026).

VI. CONCLUSION

This study used recorded brain signals from the Muse EEG device to compare the effectiveness of three deep learning algorithms in predicting concentration levels while performing various cognitive tasks such as arithmetic calculations, reading technical articles, listening to technical podcasts, reading transcripts, browsing the Internet, and simply relaxing with closed or open eyes. Deep learning algorithms, such as ANN, LSTM, and GRU, are used to analyse the brain signals; of these, ANN provides the best accuracy (65.63%), followed by LSTM (59.67%) and GRU (45.39%). ANN outperformed LSTM and GRU networks in accurately predicting concentration levels by analysing EEG data.

When evaluating the F1 Score, ANN demonstrated superior performance in classes with high and medium-high concentrations. Still, LSTM networks perform better for classes with low and medium-low concentrations than the other two networks. This study utilizes the brainwave data obtained from the Muse EEG equipment, equipped with four electrodes, to predict the degree of concentration. To further enhance this research in the future, one can augment the data collection process by collecting brainwave data from several people using an EEG device equipped with many channels. Subsequently, the data collected can be examined to identify the underlying trends in the brainwave data and to build a reliable deep-learning model to evaluate the students' cognitive involvement.

VII. AUTHOR'S CONTRIBUTION

Conceptualization: Yogesh Patil, Snehal Mali, and Harsha Gaikwad.
Methodology: Harsha Gaikwad, Manjushree Laddha, and Arvind Kiwelekar.
Investigation: Yogesh Patil, Snehal Mali.
Discussion of results: Yogesh Patil, Snehal Mali, and Harsha Gaikwad.
Writing – Original Draft: Yogesh Patil, Snehal Mali, and Harsha Gaikwad.
Writing – Review and Editing: Sanil Gandhi and Harsha Gaikwad.
Resources: Harsha Gaikwad, Manjushree Laddha, Yogesh Patil, and Snehal Mali.
Supervision: Manjushree Laddha, and Arvind Kiwelekar.
Approval of the final text: Yogesh Patil, Snehal Mali, Harsha Gaikwad, Sanil Gandhi, Manjushree Laddha, and Arvind Kiwelekar.

VIII. REFERENCES

[1] Q. He, O. Turel, and A. Bechara, “Brain anatomy alterations associated with social networking site (sns) addiction,” Scientific Reports, vol. 7, no. 1, p. 45064, Mar 2017.

[2] H. Gardner, The Disciplined Mind: What All Students Should Understand. New York: Simon & Schuster, 1999.

[3] M. A. A. Dewan, M. Murshed, and F. Lin, “Engagement detection in online learning: a review,” Smart Learning Environments, vol. 6, no. 1, p. 1, Jan 2019.

[4] M. Piwowarski, U. S. Singh, and K. Nermend, “Application of eeg metrics in the decision-making process,” in Experimental and Quantitative Methods in Contemporary Economics, K. Nermend and M. Latuszy´nska, Eds. Cham: Springer International Publishing, 2020, pp. 187–199.

[5] Aayesha, M. B. Qureshi, M. Afzaal, M. S. Qureshi, and M. Fayaz, “Machine learning-based eeg signals classification model for epileptic seizure detection,” Multimedia Tools and Applications, vol. 80, no. 12, pp. 17 849–17 877, 2021..

[6] A. M. Maitin, J. P. Romero Mu˜noz, and ´A. J. Garc´ıa-Tejedor, “Survey of machine learning techniques in the analysis of eeg signals for parkinson’s disease: A systematic review,” Applied Sciences, vol. 12, no. 14, p. 6967, 2022.

[7] D. Pirrone, E. Weitschek, P. Di Paolo, S. De Salvo, and M. C. De Cola, “Eeg signal processing and supervised machine learning to early diagnose alzheimer’s disease,” Applied sciences, vol. 12, no. 11, p. 5413, 2022.

[8] T. Sudhakar, G. H. Krishnan, N. Krishnamoorthy, B. Janney, M. Pradeepa, and J. Raghavi, “Sleep disorder diagnosis using eeg based deep learning techniques,” in 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII). IEEE, 2021, pp. 1–4.

[9] R. Alhalaseh and S. Alasasfeh, “Machine-learning-based emotion recognition system using eeg signals,” Computers, vol. 9, no. 4, p. 95, 2020.

[10] J. J. Bird, A. Ekart, C. D. Buckingham, and D. R. Faria, “Mental emotional sentiment classification with an eeg-based brain-machine interface,” in Proceedings of theInternational Conference on Digital Image and Signal Processing (DISP’19), 2019.

[11] H. Gaikwad, S. Gandhi, A. Kiwelekar, and M. Laddha, “Analyzing brain signals for predicting students’ understanding of online learning: A machine learning approach,” International Journal of Performability Engineering, vol. 19, no. 7, p. 462, 2023.

[12] V. T. Lokare and L. D. Netak, “Concentration level prediction system for the students based on physiological measures using the eeg device,” in Intelligent Human Computer Interaction: 12th International Conference, IHCI 2020, Daegu, South Korea, November 24–26, 2020, Proceedings, Part I 12. Springer, 2021, pp. 24–33.

[13] A. Al-Nafjan and M. Aldayel, “Predict students’ attention in online learning using eeg data,” Sustainability, vol. 14, no. 11, p. 6553, 2022.

[14] N. Friedman, T. Fekete, K. Gal, and O. Shriki, “Eeg-based prediction of cognitive load in intelligence tests,” Frontiers in human neuroscience, vol. 13, p. 191, 2019.

[15] A. Farkish, A. Bosaghzadeh, S. H. Amiri, and R. Ebrahimpour, “Evaluating the effects of educational multimedia design principles on cognitive load using eeg signal analysis,” Education and Information Technologies, vol. 28, no. 3, pp. 2827–2843, 2023.

[16] X. Liu, P.-N. Tan, L. Liu, and S. J. Simske, “Automated classification of eeg signals for predicting students’ cognitive state during learning,” in Proceedings of the International Conference on Web Intelligence, 2017, pp. 442–450.

[17] N. Jamil and A. N. Belkacem, “Advancing real-time remote learning: A novel paradigm for cognitive enhancement using eeg and eye-tracking analytics,” IEEE Access, vol. 12, pp. 93 116–93 132, 2024.

[18] F. J. Ram´ırez-Arias, E. E. Garc´ıa-Guerrero, E. Tlelo-Cuautle, J. M. Colores-Vargas, E. Garc´ıa-Canseco, O. R. L´opez-Bonilla, G. M. Galindo-Aldana, and E. Inzunza Gonz´alez, “Evaluation of machine learning algorithms for classification of eeg signals,” Technologies, vol. 10, no. 4, p. 79, 2022.

[19] S. K. Gatfan, “A review on deep learning for electroencephalogram signal classification,” Journal of Al-Qadisiyah for Computer Science and Mathematics, vol. 16, no. 1, pp. 137–151, 2024.

[20] E.-A. Paraschiv, L. B˘ajenaru, C. Petrache, O. Bica, and D.-N. Nicolau, “Ai-driven neuro-monitoring: Advancing schizophrenia detection and management through deep learning and eeg analysis,” Future Internet, vol. 16, no. 11, 2024.

[21] B. Siregar, G. Florence, S. Seniman, F. Fahmi, and N. Mubarakah, “Classification of human concentration levels based on electroencephalography signals,” JOIV: International Journal on Informatics Visualization, vol. 8, no. 2, pp. 923–930, 2024.

[22] A. Azhari and F. I. Ammatulloh, “Classification of concentration levels in adult-early phase using brainwave signals by applying k-nearest neighbor,” Signal and Image Processing Letters, vol. 1, no. 1, pp. 14–24, 2019.

[23] H. Zeng, C. Yang, G. Dai, F. Qin, J. Zhang, and W. Kong, “Eeg classification of driver mental states by deep learning,” Cognitive neurodynamics, vol. 12, pp. 597–606, 2018.

[24] J. Slater, R. Joober, B. L. Koborsy, S. Mitchell, E. Sahlas, and C. Palmer, “Can electroencephalography (eeg) identify adhd subtypes? a systematic review,” Neuroscience Biobehavioral Reviews, vol. 139, p. 104752, 2022.

[25] A. Rehman, I. Heldal, and J. C.-W. Lin, “Ssrepl-adhd: Adaptive complex representation learning framework for adhd detection from visual attention tasks,” in 2024 IEEE International Conference on Big Data (BigData). IEEE, 2024, pp. 8164–8172.

[26] G. Cisotto, A. Zanga, J. Chlebus, I. Zoppis, S. Manzoni, and U. Markowska-Kaczmar, “Comparison of attention-based deep learning models for eeg classification,” arXiv preprint arXiv:2012.01074, 2020.

[27] C¸ . ˙I. Acı, M. Kaya, and Y. Mishchenko, “Distinguishing mental attention states of humans via an eeg-based passive bci using machine learning methods,” Expert Systems with Applications, vol. 134, pp. 153–166, 2019.

[28] N. S. Williams, G. M. McArthur, and N. A. Badcock, “10 years of epoc: A scoping review of emotiv’s portable eeg device,” BioRxiv, 2020.

[29] L. E. Ismail and W. Karwowski, “Applications of eeg indices for the quantification of human cognitive performance: A systematic review and bibliometric analysis,” PLOS ONE, vol. 15, no. 12, pp. 1–33, 12 2020.

[30] S. D. You, “Classification of relaxation and concentration mental states with eeg,” Information, vol. 12, no. 5, p. 187, 2021.

[31] D. C. E. Saputra, A. Azhari, and A. Ma’arif, “K-nearest neighbor of beta signal brainwave to accelerate detection of concentration on student learning outcomes.” Engineering Letters, vol. 30, no. 1, 2022.

[32] A. U. Rehman, X. Shi, F. Ullah, Z. Wang, and C. Ma, “Measuring student attention based on eeg brain signals using deep reinforcement learning,” Expert Systems with Applications, vol. 269, p. 126426, 2025.

[33] M. S. Dehnavi, V. S. Dehnavi, and M. Shafiee, “Classification of mental states of human concentration based on eeg signal,” in 2021 12th International Conference on Information and Knowledge Technology (IKT). IEEE, 2021, pp. 78–82.

[34] S. Basterrech and P. Kr¨omer, “A nature-inspired biomarker for mental concentration using a single-channel eeg,” Neural Computing and Applications, vol. 32, no. 12, pp. 7941–7956, 2020.