EEG-based deep learning framework for the plausible solution of current E-learning challenges

[1]Dhamendra Pathak
Amity University Chhattisgarh
Dept. of Computer Science and Engg.

[2]Ramgopal Kashyap
Amity University Chhattisgarh
Dept. of Computer Science and Engg.

Abstract

There is a high surge in usage of online e-learning platforms due to the current ongoing Covid-19 scenario. There are various online learning platforms available, i.e., MOOC Courses, NPTEL, Swayam Portal, Coursera, etc. There are specific problems that persist in the current E-learning online models, i.e., validations and tracking of students learning curve, validation of presented course material, content-based personalization as per the requirements of the students, identification of learning disabilities among students, etc. Most of the reviewed solutions have suggested various machine learning models to address the ongoing E-learning problems. The main issue associated with these types of models is that they are not based on real-time E-learning data. Our paper proposes the deep learning model to solve the issues related to existing machine learning models of manual feature extraction and training on limited data. Also, real-time E-learning data will be collected from students wearing EEG-headband while taking online classes. It solves the issues associated with conventional machine learning models and historical data. Later, preprocessed EEG data will be fed to 7-layers Convolution Neural Network (CNN) deep learning model for better training and testing purposes. The proposed CNN model will classify the students on different grades and help in the development of an automated framework for the tracking of a student learning curve, providing recommendations for the betterment of E-learning course materials and personalization of course material.

Keywords:-Automated Framework, CNN, Deep Learning, EEG Data, E-learning, Feature extraction, Machine Learning.

I.INTRODUCTION

In the current scenario, around 1.2 billion children in 186 countries worldwide are being affected by school and college closures due to the ongoing Covid pandemic. Even before the pandemic, high growth and latest technology adoption in the education domain were noted with global edtech investments of US$18.66 billion in 2019 and a tentatively $350 Billion investment projected by 2025 in the online education market. There is a significant growth in all the verticals of online education, including virtual e-tutoring, language apps, conferencing tech, various online learning platforms, etc., since the pandemic.

Online learning can be very effective in providing the user the right tech to embed with it. Some research study has already shown that students retention rate exceeds 25-60% more than traditional classroom teaching. This happens because students learning ability excels more and faster in online education. E-learning requires relatively less time to learn, around 40-60% less than classroom teaching [1]. In E-learning, students choose to learn at their own pace supported by re-reading, looping, accelerating, and skipping through various concepts taught as per their convenience. Though, the effectiveness of e-learning varies among age groups due to their ability to focus and concentrate at different levels. A more structured and controlled environment is required for younger students due to their inability to focus on a particular concept for a longer time [2]. Presently, there is no tool or method by which validation of the leanings of students can be done. The same course/e-learning materials are taught to every student in the same way. There are no student-based customizations based on their learning capabilities [3].

Electroencephalographic signals (EEG signals) are the imprints of electrical activities of the brain. Those signal intensities are very small but measurable in microvolt (μV). Followings are some of the main frequencies associated with EEG signals:

i. Delta Frequency (range varies from 0 Hz to 4 Hz): It denotes Deep sleep, Natural Healing, and Immune System.
ii. Theta Frequency (range varies from 4 Hz to 8 Hz): It is responsible for creativity, emotional connection, and relaxation.
iii. Alpha Frequency (range varies from 8 Hz to 12 Hz): It deals with the ability to focus and relax.
iv. Beta Frequency (range varies from 12 Hz to 40 Hz): It represents conscious focus and problem-solving ability.
v. Gamma Frequency (range varies from 40 Hz to 100 Hz): It associates with acute senses, cognition, and learning.

There are various approaches that can be used for the recording and collection of EEG signals in which most common methods are collecting EEG signals with the help of placing electrodes all over the brain scalp following 10/20 system or using various EEG signals recording devices, i.e., Neuro Mind Headsets, Muse Head Sets, etc. or any other consumer-grade BCI device [4]. After a successful collection of EEG signals, the next important step is feeding off these raw signals to deep learning networks for further processing and classifications of EEG signals.

Deep learning is a subclass of machine learning, in which learning takes place with the help of deep neural networks. All the deep neural networks aim is to simulate the working of the human brain, which is quite efficient for classification tasks [5].

Each layer that is shown in figure 1, i.e., Input Layer, Hidden Layers, and Output Layer, consists of neurons. Those neurons are the processing units, mainly responsible for feature extraction, classification, etc. The strength of the signal forwarded to the subsequent layer neurons depends on the parameters, i.e., weight, bias, and activation function.

In deep neural networks, input to the current layer is based on the previous layer's output and trained accordingly. As we move further to the layers, the complexity of feature extractions and recognitions are increased due to the consolidations and reusability of the features based on the previous layer.

Figure 2 shows the concept of feature hierarchy, which enables deep neural networks to work on very massive, complex datasets with highly dense parameters through non-linear functions. It makes deep learning very popular for the processing and clustering of datasets, identifying unknown patterns, and tracking inconsistencies that are very difficult to recognize.

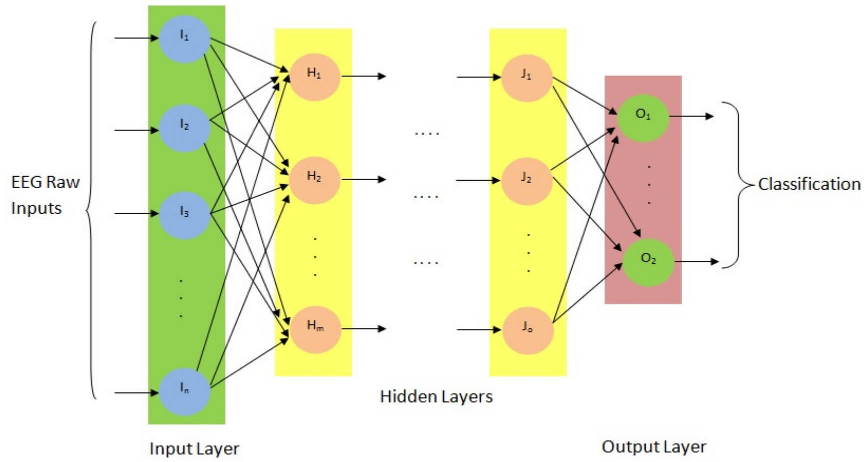

Fig. 1. The architecture of Deep Neural Network.

Fig. 2. Deep Neural Processing Pipelines.

This paper aims to propose a framework to identify the attention level and learning patterns of the students while attending and listening to online classes by collecting electroencephalogram (EEG) brain signals in real-time and later on feed them to the deep neural networks for efficient classification. The study will be further supported by collecting the EEG data while attempting relevant quizzes based on the online classes and match them with processed EEG data collected during online classes to validate their learning. On the next level, provide recommendations for better e-learning materials and better feedback to the individual and collaborative learning. Also, propose framework may also be helpful for the identification of learning disabilities in students. Apart from this, it will also support real-time monitoring of user learning curves and customized user-based recommendations.

II.Challenges in the current E-learning scenario

Due to the ongoing pandemic, millions of students have shifted to online education primarily for their learning. The flexibilities associated with Online Education concerning time, place, and on-demand learning influenced many scholars to opt for the solution. Though, it is not easy for everyone to adopt this technological shift. It requires immense time and effort at the teacher's end to design and deliver quality materials. This changing paradigm has raised its challenges and issues that require immediate attention [6]. Here is the list of most common E-learning challenges faced by the students:

The following problem arises in our present teaching/learning E-Learning model:

i. *Lack of student interest in studies*

Many factors can contribute to it, i.e., comfortable with the teaching platform, mode of communication, communication media, internet speed, delivery experts, timings, student interest/expertise in the subjects, practical/case studies involved, etc. The issue is how to overcome those challenges and motivate students towards the E-learning platforms [7].

ii. *Unfruitful learning outcomes despite great teaching efforts*

It has been concluded through many studies that despite providing the best teaching materials and great content deliveries, the actual student learning outcomes are not up to the mark. Here student interest in the subject and teaching methodology plays a very crucial role.

iii. *Unable to track the student learning curve*

There is a need for an automated framework where the actual student learning curve can be tracked and convey to students and their parents. With the help of this, the real learning problems can be identified in the early stages of online courses. The same can be achieved by conducting quizzes and tests after each session to validate the students learning [8].

iv. *Justification of teaching methodologies*

It is not so easy to justify the teaching methodologies with respect to the listening audience. If the speaker delivers the lecture in a monotonous tone, then the sessions will not be that effective. It has been observed that two-way communication puts more impressions on the students as compared to one-way communication.

v. *Unable to identify the root cause of actual problems in teaching-learning concerning individual student*

Since online course materials are designed by keeping mass students in mind, it is challenging to identify the root cause of poor learning concerning the individual students. To address this issue, those students should be evaluated on multiple parameters, i.e., their last performance, feedbacks, focus level, attainment level, learning curve, etc.

vi. *Incorrect feedbacks/reports regarding student learning growth to respective parents/guardians*

This challenge is very crucial concerning online education. There is a need for an automated platform that the students and parents can access, which is regularly monitored and updated by the respective course coordinator. On that platform, session-wise evaluation reports, quizzes, MCQs, mock tests, etc., should be conducted and timely maintained. So that actual student feedback and performance reports should be conveyed to both parties [9].

vii. *Capabilities to form efficient collaborative/group learning exercises*

Role-playing kind of exercises is complicated to perform in online education. Though group-based assignments and exercises can be provided to students to validate collaborative learning, they can also be used to validate online materials for a group of students [10].

viii. *Formations of good e-learning materials/courses as well as delivery methods as per the requirement of students*

Each student is having a different pace of learning. One common E-material cannot be circulated among all the students. So, there is a need for customized E-learning materials as per the learning curve of students. Also, for students facing difficulties while attending online classes, a separate or extra time slot should be created to set every student on an equivalent learning level [11].

ix. *Unable to track the level of concentration/focus of individual student*

There is a need for devices and a platform that can detect students' focus and attention level while attending online classes. So that the same can be further conveyed to students and used for learning evaluation. Some of the latest published reviewed papers have discussed the methods based on machine learning algorithms for the detection of focus, attention, and medication levels of subjects. But they are subjected to limited EEG data sets and lesser accuracy due to the non-implementation of deep learning models [12][13][14].

x. *Unable to provide better suggestions/area of improvements as per the taught study materials to individual student*

Again, there is a high demand and need for a platform that can suggest to the students the plausible areas of improvements based on attended classes and attempted online tests. Those suggestions may be related to online learning materials, online course instructors, students learning behavior, etc.

In the next section, we will discuss the plausible solutions to the problems mentioned above and the methodology that overcome most of the challenges of E-learning and provide recommendations for the betterment of online education.

III. Solution approach to address online education issues

In our previous section, we have discussed the critical challenges faced while adapting to online education. Next, fundamental challenges are further narrowed down in following five issues associated with e-learning:

i. Lack of automated framework for the validation of students learning.
ii. E-learning materials formation as per the feedbacks of students.
iii. Student-based course material customizations based on individual learner capabilities.
iv. Tracking and monitoring the students learning curve in real-time and proper feedback generation.
v. Better solution for the identification of student learning disabilities

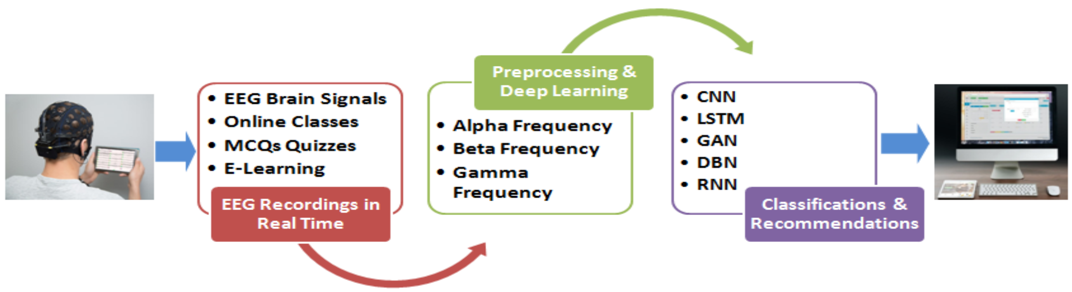

Fig. 3. EEG BCI Model functional diagram.

Now based on the above-concluded challenges following solution approach have been proposed as shown in figure 3:

i. The objective is to identify the students' attention level and learning patterns while attending and listening to online classes by collecting electroencephalogram (EEG) brain signals in real-time and later on feeding them to the deep neural networks for efficient classification.
ii. The study will be further supported by collecting the EEG data while attempting relevant quizzes based on the online classes and match them with processed EEG data collected during online classes to validate their learning.
iii. On the next level, provide recommendations for better e-learning materials and better feedback to the individual as well as collaborative learning
iv. Design a solution for the identification of learning disabilities in students
v. Provide real-time monitoring of user learning curve and customized user-based recommendations

Technological stack plays a crucial role in the successful completion of identified objectives. Following are the key technological stack and aspects associated with the proposed solutions:

i. EEG signals are very complex and require so much training to understand, but recently Deep Learning has shown significant progress in understanding and extracting the features out of it.
ii. Deep neural networks have demonstrated performance gains in many challenging tasks such as emotion detection, mental state recognition, cognitive workload detection, etc.
iii. Different classifiers have been applied to extracted features to recognize human brain states from raw EEG signals.
iv. The combination of advanced processing techniques with deep learning, through clustering and CNNs, respectively, has been proved to be quite interesting, providing efficient classification performances.
v. Deep learning approach has already shown better results in most domain applications, i.e., medical, agriculture, industrial, educational, cognition, etc. [15].

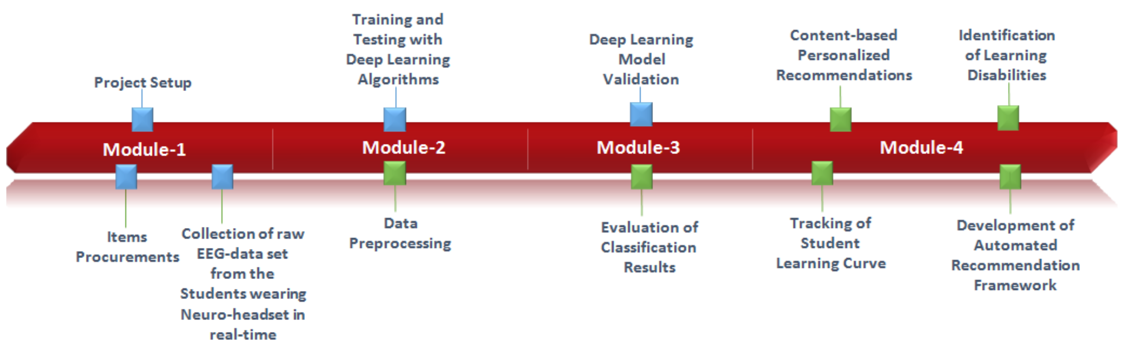

Fig. 4. Essential Processes and Modules.

The above figure 4 describes the essential processes and modules to be carried out as follows:

i. *Module 1:*

It deals with the initial setup of equipment and the preparation of selected online education material with the respective tests for the assessment process. Later, EEG data would be collected from the students wearing Neuro-headset in real-time while attending online classes and attempting respective quizzes.

ii. *Module 2:*

Preprocessing of the collected EEG data will be performed to train and test using deep learning algorithms. Preprocessed EEG data may be trained and tested on various deep learning algorithms, i.e., CNN-LSTM, for better learning classifications and comparative analysis. Deep learning models have already shown significant improvements in emotion detection, medication, focus, and attention level detection [9]. The training and testing output will classify students based on their learning capabilities, i.e., Grade A (Excellent), Grade B (Very Good), Grade C (Good), and Grade D (Needs improvements).

iii. *Module 3:*

In this module, the trained models will be validated against model validation techniques, and calculations of various performance parameters, i.e., accuracy, precision, recall, F-measure, etc., will be carried out to find the efficiency of classification.

iv. *Module 4:*

The results and findings of the deep learning algorithm will help for the further development of automated platform suggesting the following improvements:

- Content-based personalized recommendations

- Customization of E-learning material as per the student and group requirements
- Better Feedbacks and reporting to both the parties, students as well as parents
- Tracking of individual student learning curve/track/pattern
- Identification of learning disabilities

IV. Proposed Methodology to overcome the E-learning challenges

The following describes the critical components and processes used in the proposed methodology:

i. *E-Learning Raw EEG Data Collection:*

In the first process, EEG data would be collected from the students wearing EEG Headband in real-time while performing attending the online classes and attempting quizzes. For efficient training and testing, the recording limit can be set to 10-15 minutes. The subjects for the EEG Data Collections can also be specified based on course type, stream, groups, year of study, etc. The next process is to exact the three frequencies, i.e., Alpha, Beta, and Gamma, responsible for focus, attention, problem-solving ability, cognition, and learning.Signal Discretization (i.e., conversion of brain signal frequencies from analog to digital form) will be performed, and digital values of EEG signals can be used for further model training.

ii. *Data Preprocessing:*

Preprocessing on the collected data will be performed on the collected data to train and test using deep learning algorithms. Before feeding the raw data to deep networks, it must be preprocessed with techniques like normalization and cropping. Artifact handling is the process of removing the specific type of noise from the EEG signals like ocular noise. This process is dependent on the domain for which signals will be classified. This will be implemented by applying frequency amplitude thresholding, blinking-related noise, etc. [16].

iii. *Convolution Neural Network (CNN) Model for classification:*

CNN architecture is a class of models that is both spatially and temporally deep, and has the flexibility to be applied to various tasks involving sequential inputs and outputs. Key is using a CNN that is pre-trained on a challenging classification task that is re-purposed as a feature extractor for the EEG classification problem.The CNN architecture involves convolution layers for feature extraction on input data combined. Automatic feature extraction is one of the main reasons for the deep learning technique to become popular nowadays since it entirely or partially removed the manual feature selection process as required in conventional machine learning algorithms. In some of the papers studied, researchers had taken short time Fourier transformations, in the case of epilepsy spatial or temporal frequency domain and in case of EEG raw inputs alpha, beta, gamma, for brain-computer mapping (BCI) [17].

This architecture is appropriate because of the temporal and spatial nature of EEG signals. In the proposed work, CNN will be used for feature extraction and further classification of EEG signals. Preprocessed EEG data will be trained and tested on seven layers of CNN deep learning algorithms for better classifications and calculations of various performance parameters, i.e., accuracy, precision, recall, F-measure, etc. [18]. The activation function used in each of the CNN layers is 'ReLU' to overcome the exponential growth and prevent the vanishing gradient problem. Next, for downsampling Max Pooling is performed. To overcome the problem of overfitting, dropout operations are performed, which deactivates some neurons towards the next successive layer. The output of one CNN layer will be provided as input to the next subsequent layer. More than one number of CNN layers will be used for better feature extraction and efficient classification. Finally, the output of CNN layers will be fed to the fully connected layer containing a soft-max activation function for further classifications.

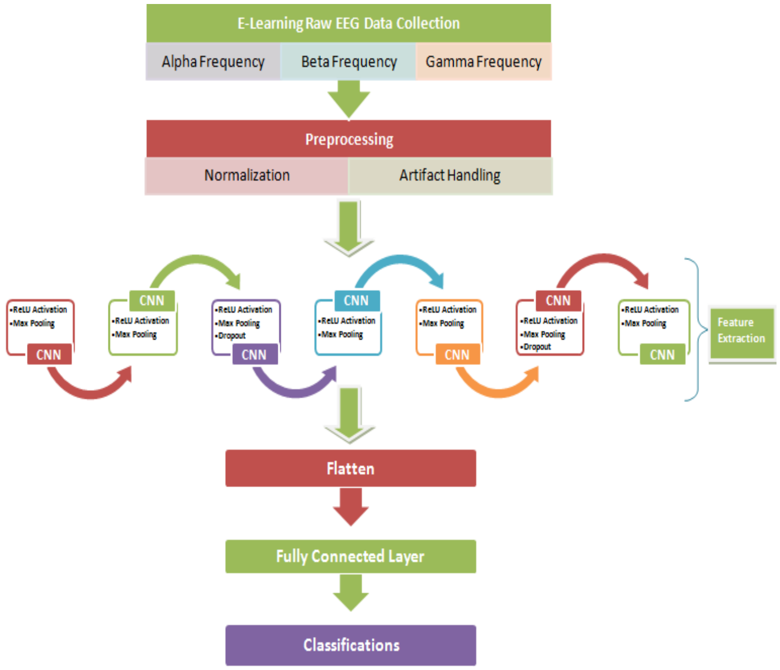

Fig. 5. Proposed deep learning-based methodology.

iv. Evaluation:

The ultimate goal of any classification research study is to evaluate and interpret the result of the research findings successfully. In our case, that is to classify the students based on learning capabilities, i.e., Grade A (Excellent), Grade B (Very Good), Grade C (Good), and Grade D (Needs improvements). We recommend calculating various performance evaluation criteria, i.e., accuracy, sensitivity, specificity, etc., to validate the efficiency of the trained model.

v. Classification Validation:

For classification and training validation, the following parameters will be considered:

a. MCQs based Test:

This will be conducted after the commencement of the online class to validate the student's learning concerning the attended lecture. Later on, test results will be used in the classification of EEG data while training the model.

b. Student Feedback after the commencement of Online Class, i.e., Excellent, Good, Average, Poor:

This particular attribute will be helpful in online content optimizations concerning learning outcomes and will be used in developing an automated learning framework.

Apart from the EEG signals, i.e., Alpha, Beta, Gamma features, the following factors may also be considered while training the model for better classification [19]:

i. Content-Type (PPT, Audio, Video, Animation)
ii. Content-Length/Size (Number of Slides)
iii. Topic Knowledge (On the Scale of 1 to 10)
iv. Exercise/Problems Based on the topic (Yes/No)
v. Age Group (5-10, 11-15, 16-21, etc.)

V. Discussions

We have also reviewed various research papers on implementing a deep learning framework on EEG-based data for the classifications in different domains, i.e., epilepsy, cognition, mental state, sleep stages, etc. [20]. Based on that, we propose the following figure represents the generalized framework that can be recommended to overcome the various challenges faced while the classification of EEG signals in multiple domains:

i. Dataset:

In most of the reviewed papers, a limited dataset was used for EEG classification that further caused problems in model validation, pre-trained models, and resulting classification. Therefore it is recommended to opt for massive open data set for training and testing of EEG signals for various problem domains.

ii. Methodology:

In the case of machine learning implementation, it is not possible to manually extract the end-to-end features from the EEG Signals and further classify them efficiently due to their very complex nature. Therefore it is highly recommended to go for Deep learning methodology and implement different algorithms like CNN, R-CNN, DBN, GAN, etc. and perform the comparative analysis for better performance.

iii. Model validation:

With the help of model validation techniques, we can validate the built model and reduce the problems of overfitting and underfitting. Since most of the reviewed papers have not implemented any model validation techniques, it is also recommended to validate the proposed model with either model validation technique, i.e., k-fold cross-validation leave-one-out, bootstrapping, etc.

iv. Evaluation:

The ultimate goal of any classification research study is to evaluate the result of the research findings successfully. We recommend calculating various performance evaluation criteria, i.e., accuracy, sensitivity, specificity, F-measure, etc.

Apart from the above recommendations, using a pre-trained model is also recommended since it has increased the accuracy of feature extraction and classification significantly in some of the studies.

VI. Conclusion

Online education and E-learning have now become inseparable parts of current teaching-learning methodologies. Many schools and institutions have adopted the same all over the world. Proposed research work is the plausible application of EEG-BCI model in the domain of E-learning would be the milestone for overall improvements and advancements in teaching-learning experience followed by the preparation of better online study materials, better content delivery, better understanding of students learning patterns, and the recommended customization as per the requirements of an individual student, conveyed through a mobile app. The future work in the classification of EEG signals may be the analysis of user brainwave data to predict learning disabilities based on data collected.

References

1. S. Keskin and H. Yurdugül, "Factors Affecting Students' Preferences for Online and Blended Learning: Motivational Vs. Cognitive", European Journal of Open, Distance and E-Learning, vol. 22, no. 2, pp. 72-86, 2020.
2. A. Shahzad, R. Hassan, A. Aremu, A. Hussain and R. Lodhi, "Effects of COVID-19 in E-learning on higher education institution students: the group comparison between male and female", Quality & Quantity, 2020.

3. S. Leahy, C. Holland and F. Ward, "The digital frontier: Envisioning future technologies impact on the classroom", Futures, vol. 113, p. 102422, 2019.
4. V. Peterson, C. Galván, H. Hernández and R. Spies, "A feasibility study of a complete low-cost consumer-grade brain-computer interface system", Heliyon, vol. 6, no. 3, p. e03425, 2020.
5. Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. Falk and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review", Journal of Neural Engineering, vol. 16, no. 5, p. 051001, 2019.
6. J. Valverde-Berrocoso, M. Garrido-Arroyo, C. Burgos-Videla and M. Morales-Cevallos, "Trends in Educational Research about e-Learning: A Systematic Literature Review (2009–2018)", Sustainability, vol. 12, no. 12, p. 5153, 2020.
7. T. Alrawashdeh, "Factors Affecting Acceptance of Web-Based Training System: Using Extended Utaut and Structural Equation Modeling", International Journal of Computer Science, Engineering and Information Technology, vol. 2, no. 2, pp. 45-54, 2012.
8. A. Mousavi, A. Mohammadi, R. Mojtahedzadeh, M. Shirazi and H. Rashidi, "E-learning educational atmosphere measure (EEAM): a new instrument for assessing e-students' perception of educational environment", Research in Learning Technology, vol. 28, no. 0, 2020.
9. S. Ghosh, "An Approach to Building a Learning Management System that Emphasizes on Incorporating Individualized Dissemination with Intelligent Tutoring", Journal of The Institution of Engineers (India): Series B, vol. 98, no. 1, pp. 1-8, 2016.
10. D. Onah and J. Sinclair, "Assessing Self-Regulation of Learning Dimensions in a Stand-alone MOOC Platform", International Journal of Engineering Pedagogy (iJEP), vol. 7, no. 2, p. 4, 2017.
11. S. Salloum, A. Qasim Mohammad Alhamad, M. Al-Emran, A. Abdel Monem and K. Shaalan, "Exploring Students' Acceptance of E-Learning Through the Development of a Comprehensive Technology Acceptance Model", IEEE Access, vol. 7, pp. 128445-128462, 2019.
12. Z. Mohamed, M. El Halaby, T. Said, D. Shawky and A. Badawi, "Characterizing Focused Attention and Working Memory Using EEG", Sensors, vol. 18, no. 11, p. 3743, 2018.
13. C. Peng, Y. Chen, C. Chen, S. Chen, B. Cagneau and L. Chassagne, "An EEG-Based Attentiveness Recognition System Using Hilbert–Huang Transform and Support Vector Machine", Journal of Medical and Biological Engineering, vol. 40, no. 2, pp. 230-238, 2019.
14. K. Yoshida et al., "Focused attention meditation training modifies neural activity and attention: longitudinal EEG data in non-meditators", Social Cognitive and Affective Neuroscience, vol. 15, no. 2, pp. 215-224, 2020.
15. M. Alom et al., "A State-of-the-Art Survey on Deep Learning Theory and Architectures", Electronics, vol. 8, no. 3, p. 292, 2019.
16. H. Zeng et al., "EEG Emotion Classification Using an Improved SincNet-Based Deep Learning Model", Brain Sciences, vol. 9, no. 11, p. 326, 2019.
17. X. Zhang, L. Yao, X. Wang, J. Monaghan, D. McAlpine and Y. Zhang, "A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers", Journal of Neural Engineering, vol. 18, no. 3, p. 031002, 2021.
18. A. Mansoor, M. Usman, N. Jamil and M. Naeem, "Deep Learning Algorithm for Brain-Computer Interface", Scientific Programming, vol. 2020, pp. 1-12, 2020.

19. D. Lu, H. Le and T. Vu, "The Factors Affecting Acceptance of E-Learning: A Machine Learning Algorithm Approach", Education Sciences, vol. 10, no. 10, p. 270, 2020.
20. M. Rashid et al., "Current Status, Challenges, and Possible Solutions of EEG-Based Brain-Computer Interface: A Comprehensive Review", Frontiers in Neurorobotics, vol. 14, 2020.