ISSN: 2765-7213 © 2023 KODISA & KFIA.
https://acoms.kisti.re.kr/fir & http://www.fir.or.kr
doi: http://doi.org/10.20498/fir.2023.3.1.13

A Study on the Development of Adaptive Learning System through EEG-based Learning Achievement Prediction

Jinwoo KIM[1], Hosung WOO[2]

[1 First Author] Student, Dept. of E-learning, Graduate School, Korea National Open University, South Korea.
E-mail: tggksnu@knou.ac.kr

[2 Corresponding Author] Assistant Professor, Dept. of E-learning, Graduate School, Korea National Open University, South Korea.
E-mail: hosungwoo@knou.ac.kr

Received: February 14, 2023. Revised: March 17, 2023. Accepted: March 23, 2023.

Abstract

Purpose – By designing a PEF(Personalized Education Feedback) system for real-time prediction of learning achievement and motivation through real-time EEG analysis of learners, this system provides some modules of a personalized adaptive learning system. By applying these modules to e-learning and offline learning, they motivate learners and improve the quality of learning progress and effective learning outcomes can be achieved for immersive self-directed learning

Research design, data, and methodology – EEG data were collected simultaneously as the English test was given to the experimenters, and the correlation between the correct answer result and the EEG data was learned with a machine learning algorithm and the predictive model was evaluated..

Result – In model performance evaluation, both artificial neural networks(ANNs) and support vector machines(SVMs) showed high accuracy of more than 91%.

Conclusion – This research provides some modules of personalized adaptive learning systems that can more efficiently complete by designing a PEF system for real-time learning achievement prediction and learning motivation through an adaptive learning system based on real-time EEG analysis of learners. The implication of this initial research is to verify hypothetical situations for the development of an adaptive learning system through EEG analysis-based learning achievement prediction.

Keywords: EEG-based Edtech, Adaptive Learning System, PEF(Personalized Education Feedback) System, Machine Learning, Neuroengineering

JEL Classification Code: I20, I23, C80

1. Introduction

The adaptive learning system effectively provides an individualized learning system suitable for the learner's individual education level and environment through data provided via learning analysis, thereby improving the efficiency of individual learning, and enabling self-directed learning, It is getting attention. Various components and modules of these adaptive learning systems have recently been established as an applicable environment due to technological innovation in edutech, and interest in adaptive learning in the field of education is greatly increasing, and things that were difficult to implement in the past have recently been shown to be actually implemented through edutech technology innovation. (Vesin & Giannakos, 2018)

In particular, online learning has become important due to COVID-19 and the expansion of e-learning and MOOC(Massive Open Online Courses), and online education has emerged as an alternative to a new education method. At the same time, although fMRI was the only device that could objectively grasp the mental states of learners, which were difficult to grasp objectively, portable EEG devices recently have been gradually commercialized, the price has been lowered, and the distribution is also expanding, so studies using it are actively being conducted. (Lin & Kao, 2018).

The current theme in EEG-based edtech is towards the development of non-invasive and wearable EEG devices, which can be used to measure brain activity in real-time during educational activities. This allows for the creation of more immersive and interactive learning experiences that can be tailored to the individual needs of each student. (Koelstra, 2019; Lim, 2018)

Another current theme in EEG-based edtech is the use of machine learning(ML) and artificial intelligence (AI) to analyze EEG data and provide personalized feedback to students. This is being done to enhance the effectiveness and efficiency of educational experiences, by using AI to analyze and interpret brain activity data in real-time and provide real-time feedback to students. Overall, the current theme in EEG-based edtech is the use of brain activity data to create more personalized, effective, and engaging educational experiences. (Bashir, 2021)

EEG-based edtech that utilizes EEG data has the potential to revolutionize the way we learn and teach. EEG technology allows us to measure brain activity in real-time, which can provide valuable insights into how our brains process information and learn. Especially by understanding learning processes, EEG-based edutech can lead to the development of more effective teaching methods that are tailored to the individual needs of each student. Also by utilizing EEG-based edtech, students can learn more efficiently and effectively, resulting in enhanced learning outcomes.

Therefore, this study provides some modules of a personalized adaptive learning system that can more efficiently complete immersive self-directed learning by designing a PEF system for real-time learning prediction and learning motivation through an adaptive learning system through real-time EEG analysis of individual learners. By applying these modules to e-learning and offline learning, they motivate learners and improve the quality of learning progress and effective learning outcomes can be achieved for immersive self-directed learning.

2. Literature Review

2.1. EEG (Electroencephalogram) and BCI(Brain-Computer Interface) and Analysis Algorithm

EEG(Electroencephalograph) measurement is an electrophysiological monitoring method that records electrical activity in the brain. EEG data measured using a potential difference measured non-invasively in the scalp electrode is used to measure brain waves. EEG is classified into Delta, Theta, Alpha, Beta, and Gamma waves according to frequency characteristics, and the types and general characteristics of EEG are shown in Table 1.

Currently, commercial portable EEG equipment mainly measures brain waves of 1-4 channels, as shown in Table. 2, and most of them are non-invasive dry sensor types and are in the form of headbands for portability. A typical EEG-based BCI configuration is a typical method in which brainwave data measured through headband-type equipment is transferred to a computer in a raw data manner through an API or SDK provided by the manufacturer, and this data is analyzed and used for research.

Previous studies have conducted the feature extraction of brain waves in the Power Spectral Density format through Welch's method, which is most commonly used in EEG analysis studies. (Al-Nafjan & Aldayel, 2022). In addition, several previous studies have shown that the accuracy of predicted value results based on learners' brain wave state discrimination or brain wave analysis was improved when the Artifact was removed through Independent Component Analysis(ICA). (Lee, 2019)

Table 1: Type of EEG

Wave	Frequency range	Major function
Delta	~3Hz	Sleep, found in attention tasks
Theta	4~7Hz	Idling, inefficiency, related to ADHD
Alpha	8~15Hz	Relax, eyes closing
Beta	16~30Hz	Focus, anxious thinking
Gamma	31Hz~	Cognition

Table 2: Commercial portable EEG device example

Wave	Neurosky Mindwave	Interaxon Muse 2	OpenBCI Biosensing Kit
Price	$100	$250	$800
# of channel	1	4	4
Sampling rate	512Hz	256Hz	16kHz
Bandwidth	3~100Hz	2~50Hz	0.01~70Hz

EEG research related to online learning is mainly a basic level of progress for identifying the status of whether students focus or not during learning. As shown in Table.3, for the machine learning(ML) and artificial intelligence(AI) algorithm used for EEG analysis, artificial neural network(ANN), k-nearest neighbors(kNN), support vector machines (SVM), and Random Forrest(RF) were used, and some studies used Linear Discriminant Analysis(LDA) and Logistic Regression. (Xu & Zhong, 2018)

2.2. EEG-based PEF(Personalized education feedback) System

A PEF(personalized education feedback) system is a technology-based system designed to provide customized feedback and recommendations to students based on their individual learning styles, strengths, weaknesses, and progress. This type of system collects data on students' performance, preferences, and progress, and then uses machine learning algorithms to generate personalized feedback and guidance to help them improve their skills and achieve their learning goals. This type of system has the potential to enhance the educational experience by providing students with tailored and engaging feedback, which can help to motivate and empower them to take control of their own learning.

The PEF system method related to EEG in previous studies was mainly conducted by analyzing learners' brain wave status and sending simple instant messages immediately as soon as it was determined that feedback was needed. For example, monitoring and alarm messages use short instant messages such as "Your concentration is better now! Pay more attention in class! cheer up!". Through this PEF system, it was analyzed that the learning effect was significantly higher in the group applied than in the group not applied. (Chen & Wang, 2018)

Another study suggested the 1:1 adaptive learning content matching module that detects students who take open online lectures(MOOCs) in real-time and delivers them to instructors when they are determined to be unable to concentrate or understand during online learning hours as shown in Figure 1. (Lin & Kao, 2018)

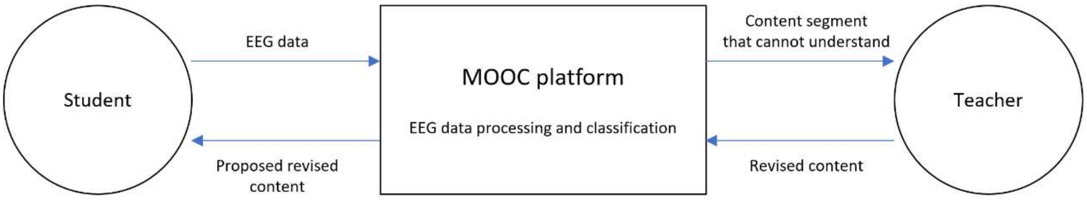

Figure 1: An application scenario of adaptive content matching (Lin & Kao, 2018)

3. Methodology

3.1. Research Model Design

The research procedure in this study is as follows as shown in Figure.2.

First, the method of the experiment was explained to the subjects and test papers were distributed.

Second, after checking Muse2 headset was used to test whether EEG data was transmitted and received well, the duration of the experiment was prepared by checking the status of EO(Eye open) and EC(Eye closed) for about 1 minute. Third, the test subjects were asked to solve the TOEIC short sentence grammar test of around 10 questions, and the experiment facilitator recorded the time required to solve each question (by 1 second) and the correct/incorrect answer results. Finally, the time required to solve each question (by 1 second), correct/incorrect answer results, and EEG data obtained through the Muse2 headset were analyzed.

In this study, as a study to confirm the possibility of a full-fledged in-depth study in the future, a total of three experiments were conducted with two subjects. One person conducted the experiment twice in total, and one person conducted the experiment once. Through this, a total of 2,731(rows) data was obtained.

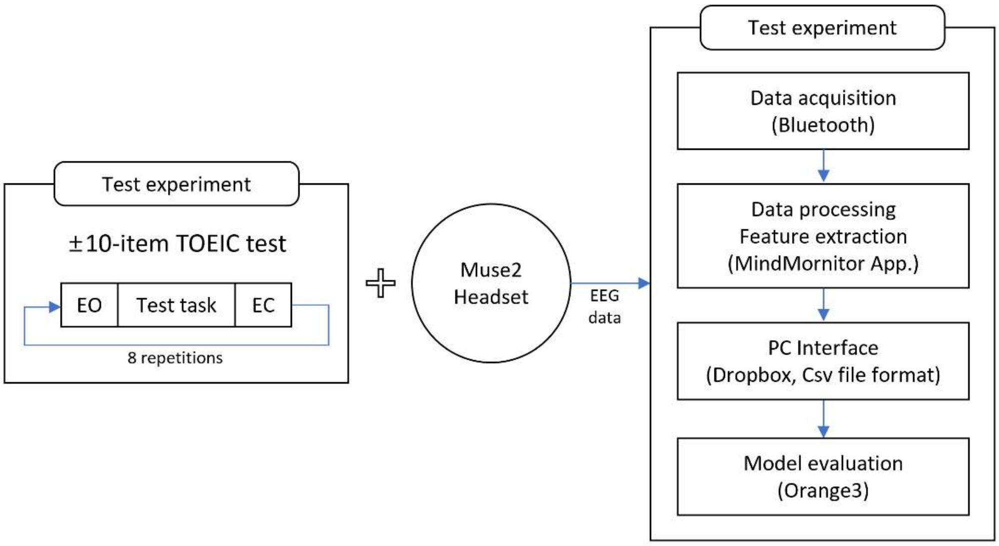

Figure 2: Research procedure

3.2. Data Acquisition

For this study, it was decided to use Interaxon's Muse 2(Canada company), which is the most widely used in research among the 4 electrode channel portable equipment among representative commercial products with good portability and improved performance. As shown in Figure 3, Muse supports four electrode channel data at AF7, AF8, TP9, and

TP10 locations in the Modified Combinatorial Nomenclature(MCN) system, an international system that measures higher resolution than the International 10-20 system. Since the potential difference of each delta, theta, alpha, beta, and gamma wave of the corresponding electrode is measured, a total of 20 individual data measurements (4 channels X 5 EEG waveforms) related to EEG are possible.

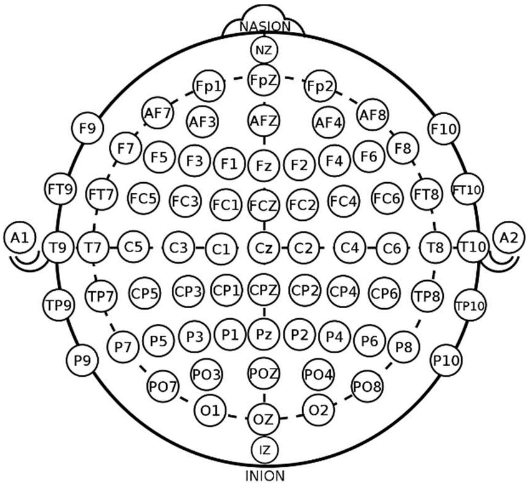

Figure 2: International MCN system for EEG

Data transmission and reception from Muse 2 to PC is performed by measuring brain waves, sending and receiving data using the MindMonitor app, and then transmitting to PC in the form of a CSV file. When measuring, there is a transmission/reception delay (±20ms) due to the characteristics of Bluetooth, but the accuracy of the standard time based on the mobile phone may differ by several seconds, and most problem-solving experiments show that it takes at least 20 seconds per problem. Therefore, it was decided to ignore the delay in the data measurement time. Muse equipment measures data of 12 bits size at a rate of 256Hz. In this study, Fast Fourier Transform(FFT) was applied through the MindMonitor app, and in the settings, the result information acquisition according to the timestamp of the data was applied twice per second, and the value was set as the average value. In the case of the window size and overlap, assuming that there is no appropriate size applied to all brain waves, Fast Fourier Transform(FFT) was applied with a window size of 3 seconds and an overlap of 2 seconds set in related studies. To remove noise applied to signal processing, a notch filter was applied at 60Hz, and a bandpass filter was set at 0.5~50Hz to measure 5 EEG bands.

3.3. Test method and Data Processing, Model Evaluation

In this study, in order to continuously collect the experimenter's learning data in a relatively short period of time, TOEIC's short grammar test items were randomly selected from the Internet and used as an experimental tool.

There are about 10 questions in one trial (1 page), and the experiment facilitator wrote down each start time on the question paper before the start and had the experimenter write down the time (in 1-second units) when the problem was solved. Until the next problem was solved, the experimenter closed the eyes (EC state, Eye-Closed) and took a rest until the start time of the next problem. In this way, the time (seconds) it took the experimenter to solve each problem, the time (seconds) that the experimenter had to close their eyes(EC), and the data of the question answer sheet were obtained from the experiment supervisor through the test result sheet.

Among the data obtained through portable EEG device, the EEG values of the four electrode channels were used, and the additional three-dimensional acceleration sensor values and gyro sensor values provided were not used. Thereafter, when the correct answer was matched to the SCR cell(=Score) according to the time stamp according to the problem-solving time, 1 insertion treatment was performed, and if the wrong answer was inserted as 0. As a result of verifying the data measured from the actual test, some electrodes did not measure for several seconds or found abnormalities in which the measured values were fixed, but they were used as they were without being especially deleted or replaced with other values. However, data with error messages such as 'Blink' and 'Jaw Clench' in the index cell were filtered and all corresponding data were deleted.

Recent education-related studies have achieved over 90% accuracy using Support vector machines(SVM), Artificial neural networks(ANN), and Random forests(RF) as machine learning(ML) and artificial intelligence(AI) algorithms, so it decided to apply these three algorithms. (Ramírez-Moreno et al., 2021)

In the achievement prediction discrimination, in this study, the learning achievement was defined as the probability of a correct answer through EEG data analysis, and the predicted probability of obtaining the correct answer was regarded as the learner's understanding. Therefore, the EEG data of each individual learner becomes an independent variable(X), and the correct answer is set as a dependent variable(Y), and through the training of X and Y, Y's prediction is determined to determine the learner's understanding of the question. do. For training, 70% of the collected experimenters' data was verified for training and 30% for testing, and the number of training tests was set to 20 times. The application and performance evaluation of the algorithm used the Orange3 program.

4. Result

The information of the subjects related to the experimental test is shown in Table. 3, and a total of 2,731(rows) data were analyzed.

Table 3: Subjects list

	A-1st test	**B-1st test**	**A-2nd test**
Personal data	Male in his 40s	Woman in her 30s	Male in his 40s
# of questions	11	6	6
Total data (rows)	1,284	711	736
Total time (min)	11 min	6 min	6 min

In the results of the 1st test through subject A, the performance evaluation of the three algorithms for the learning achievement prediction system only through EEG data analysis was shown in Table. 4. All three algorithms showed more than 91% accuracy, and in terms of accuracy, Artificial neural network(ANN) was the best at 93.7%, followed by Support vector machine(SVM) and Random forests(RF). In addition, except for Artificial neural network(ANN), which require a considerable amount of time for data processing speed and are somewhat difficult to apply to real-world systems, the achievement prediction showed that the Support vector machine(SVM) algorithm has good accuracy.

Table 4: EEG-based prediction performance (A-1st test)

	SVM	**RF**	**ANN**
Accuracy	93.3%	91.8%	**93.7%**
Recall	93.3%	91.8%	93.7%
Precision	93.2%	91.7%	93.6%
F-measure	93.2%	91.8%	93.7%

In the results of the 1st test through subject B, the performance evaluation of the three algorithms for the achievement prediction system was shown in Table. 5. In terms of accuracy, unlike the 1st test results of subject A, Support vector machine(SVM) was the best with 98.3%, followed by Artificial neural network(ANN) and Random forest(RF).

All three algorithms showed more than 95% accuracy, which was higher than the predictive performance of subject A. The fact that the predictive performance algorithms according to personal brainwave analysis are different for each individual is important for developing an adaptive learning system through prediction of achievement in the future, especially the fact that the amount of training data is 1/2 of that of the first subject.

Table 5: EEG-based prediction performance (B-1st test)

	SVM	RF	ANN
Accuracy	**98.3%**	97.7%	98.1%
Recall	98.3%	97.7%	98.1%
Precision	98.2%	97.7%	98.1%
F-measure	98.2%	97.7%	98.1%

In the combined test results of the 1st and 2nd rounds of data through subject A, the performance evaluation of the three machine learning algorithms for the achievement prediction system was shown in Table. 6. In terms of accuracy, Artificial neural network(ANN) was the best with 95.2%, followed by Support vector machine(SVM) and random forests. In terms of accuracy, all three algorithms show improved accuracy as the dataset increases, indicating that the accuracy of the system's achievement prediction improves as the cumulative brainwave data for each individual increases as expected. Therefore, we show that it may be more useful to increase accuracy while learning by increasing the amount of individual brainwave data rather than improving artificial neural network algorithms with slow processing speed in actual commercialization system modules.

Table 6: EEG-based prediction performance (A-1st and 2nd test)

	SVM	RF	ANN
Accuracy	94.4%	93.8%	**95.2%**
Recall	94.4%	93.8%	95.2%
Precision	94.3%	93.5%	95.1%
F-measure	94.2%	93.6%	95.1%

5. Discussion

In this study, outcomes appear that learning achievement can be predicted in real time with an accuracy of over 91% only by analyzing EEG data even with existing algorithms.

This research provides some modules of personalized adaptive learning systems that can more efficiently complete by designing a PEF system for real-time learning achievement prediction and learning motivation through an adaptive learning system based on real-time EEG analysis of learners. By applying these modules to e-learning and offline learning, they motivate learners and improve the quality of learning progress and effective learning outcomes can be achieved for immersive self-directed learning

Through this research, the suggestions for the PEF system for the development of an adaptive learning system are as follows.

First, it is the development of scenarios for personalized adaptive learning feedback.

To date, the percentage of the question that the learner has understood and solved through the test has no choice but to rely on qualitative data such as a survey through the learner's response. In addition, the value of use is reduced due to the large personal deviation between learners with improved metacognitive skills and learners who do not, and since learners can use figures only when they evaluate and fill in their understanding every time, there is no possibility of reality.

However, if the predictions of the system used in this research are provided as reasonable achievement predictions through sufficient EEG dataset, learners' understanding of subjects, units, and question types can be continuously tracked and supplemented based on this. In the case of teachers, by easily understanding learner’s status, it can be expanded to a more sophisticated 1:1 adaptive learning system. (Bevilacqua et al., 2019)

Second, it is the development of a self-motivation dashboard that enables immersive self-directed learning.

In the current education industry, the biggest aspects of relying on learners' motivation to learn are only test scores and study time. However, since test scores are the result of a combination of various comprehensive factors, it is difficult to motivate individual and detailed learning. In the case of study time, only the quantitative aspect of learning is considered, and the efficiency of learning motivation for "Why should I study more and focus on learning?" is low.

In order to compensate for these weaknesses, the following dashboard can be configured. If learning time (A) is placed on the horizontal axis and learning achievement (B) is placed on the vertical axis, the area (A×B) eventually represents the quantitative factor (Quantity) of learning time and learning achievement, that is, the quality of learning (Quality). It is the total amount of learning input and output reflected in all. In particular, by providing the total amount of these learning inputs and results in a visually intuitive form to learners like a heat map, immersive motivation for self-directed learning will be possible. (Acı et al., 2019)

The implication of this initial research is to verify hypothetical situations for the development of an adaptive learning system through EEG analysis-based learning achievement prediction.

The limitations of this study are the number of experimental subjects and dataset. Additional experiments with a larger number of experimenters are needed in terms of both the number of experimenters and the amount of data. In particular, as a follow-up study, to what extent the accuracy of the learning achievement prediction model according to the number of items in the training dataset reaches a critical value at the time of the number of items in the training dataset, to what extent the accuracy improves thereafter, and what differences exist between individuals and what factors It is necessary to establish a hypothesis about cognition, and to proceed and verify it in future research. In addition, algorithms with high accuracy are different for each individual, and the initial accuracy is also different. It is necessary to verify through follow-up research whether this is due to differences in age, gender characteristics, or other characteristics according to brain waves.

References

Acı, Ç. İ., Kaya, M., & Mishchenko, Y. (2019). Distinguishing mental attention states of humans via an EEG-based passive BCI using machine learning methods. *Expert Systems with Applications, 134*, 153-166.

Al-Nafjan, A., & Aldayel, M. (2022). Predict Students' Attention in Online Learning Using EEG Data. *Sustainability, 14*(11), 6553.

Bashir, F., Ali, A., Soomro, T. A., Marouf, M., Bilal, M., & Chowdhry, B. S. (2021). Electroencephalogram (EEG) Signals for Modern Educational Research. *In Innovative Education Technologies for 21st Century Teaching and Learning* (pp. 149-171). CRC Press.

Bevilacqua, D., Davidesco, I., Wan, L., Chaloner, K., Rowland, J., Ding, M., ... & Dikker, S. (2019). Brain-to-brain synchrony and learning outcomes vary by student–teacher dynamics: Evidence from a real-world classroom electroencephalography study. *Journal of cognitive neuroscience, 31*(3), 401-411.

Chen, C. M., & Wang, J. Y. (2018). Effects of online synchronous instruction with an attention monitoring and alarm mechanism on sustained attention and learning performance. *Interactive Learning Environments, 26*(4), 427-443.

Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., ... & Patras, I. (2011). Deap: A database for emotion analysis; using physiological signals. *IEEE transactions on affective computing, 3*(1), 18-31.

Lee, H., Shin, D., & Shin, D. (2019). A research on the emotion classification and precision improvement of EEG (Electroencephalogram) data using machine learning algorithm. *Journal of Internet Computing and Services, 20*(5), 27-36.

Lim, W. L., Sourina, O., & Wang, L. P. (2018). STEW: Simultaneous task EEG workload data set. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26*(11), 2106-2114.

Lin, F. R., & Kao, C. M. (2018). Mental effort detection using EEG data in E-learning contexts. *Computers & Education, 122*, 63-79.

Ramírez-Moreno, M. A., Díaz-Padilla, M., Valenzuela-Gómez, K. D., Vargas-Martínez, A., Tudón-Martínez, J. C., Morales-Menendez, R., ... & Lozoya-Santos, J. D. J. (2021). Eeg-based tool for prediction of university students' cognitive performance in the classroom. *Brain Sciences, 11*(6), 698

Vesin, B., Mangaroska, K., & Giannakos, M. (2018). Learning in smart environments: user-centered design and analytics of an adaptive learning system. *Smart Learning Environments, 5*, 1-21.

Xu, J., & Zhong, B. (2018). Review on portable EEG technology in educational research. *Computers in Human Behavior, 81*, 340-349.