Revista Brasileira de Informática na Educação – RBIE
Brazilian Journal of Computers in Education
(ISSN online: 2317-6121; print: 1414-5685)
http://br-ie.org/pub/index.php/rbie

Submission: 07/08/2020; Camera ready: 21/05/2021; 1st round notif.: 10/12/2020; Edition review: 22/05/2021; New version: 25/01/2021; Available online: 13/06/2021; 2nd round notif.: 11/03/2021; Published: 13/06/2021;

On the joint use of Artificial Intelligence and Brain-Imaging Techniques in Technology-enhanced Learning Environments: A Systematic Literature Review

Kamilla Tenório
Federal University of Alagoas
kktas@ic.ufal.br

Jário Santos
University of São Paulo
jariojj@usp.br

Victor Accete
Federal University of Alagoas
victoraccete@ic.ufal.br

Sterfanno Remigio
Federal University of Alagoas
ssrc@ic.ufal.br

Alan Pedro da Silva
Federal University of Alagoas
alanpedro@ic.ufal.br

Diego Dermeval
Federal University of Alagoas
diego.matos@famed.ufal.br

Ig Ibert Bittencourt
Federal University of Alagoas
ig.ibert@ic.ufal.br

Leonardo Brandão Marques
Federal University of Alagoas
leonardo.marques@cedu.ufal.br

Abstract

Recent meta-analysis and literature reviews support that adaptive learning systems are components of effective instruction. These exciting results are motivating researchers to explore new technologies that provide relevant students' information to promote a better-personalized experience for students to achieve better learning outcomes in technology-enhanced learning environments. A new trend is related to the studies that use brain-imaging techniques to provide relevant students' information for educational systems, aiming to enable an enhanced personalized experience. Some of these studies are making use of artificial intelligence to provide real-time monitoring of students' cognitive phenomena supplied by brain-imaging techniques such as electroencephalography and functional magnetic resonance imaging. Therefore, considering the relevance of the application of artificial intelligence in studies that use brain-imaging techniques combined with technology-enhanced learning environments and the lack of a current understanding of how these techniques have been used in this context, we present a systematic literature review (SLR) that aims to explore which artificial intelligence algorithms have been adopted, what are their purposes in studies that apply brain-imaging techniques in educational technologies and which were the results reported in these studies related to the use of artificial intelligence algorithms. The systematic literature review was conducted according to the recommendation of a well-accepted guideline to perform a rigorous review of the current literature. The search was conducted in seven academic databases in January 2020 and resulted in a total of 6089 studies that was reduced to 20 studies for the final analysis.

Keywords: *Brain-imaging Techniques; Technology-enhanced Learning Environments; Artificial Intelligence Algorithms.*

Cite as: Tenório, K., . . . & Marques, L. B. (2021). On the joint use of Artificial Intelligence and Brain-Imaging Techniques in Technology-enhanced Learning Environments: A Systematic Literature Review. Brazilian Journal of Computers in Education (Revista Brasileira de Informática na Educação – RBIE), 29, 502-518. DOI: 10.5753/RBIE.2021.29.0.502.

1 Introduction

Researchers consider adaptive teaching – the adaptation of instruction by teachers to facilitate students' learning in classrooms (Smale-Jacobse et al., 2019) – a component of effective instruction (Parsons et al., 2018; Hattie, 2008). Due to the rapid development of information and communication technology (ICT), the personalizing of learning has also become possible in the context of technology-enhanced learning context (Xie et al., 2019; Klašnja-Milićević et al., 2011; Akbulut & Cardak, 2012; Normadhi et al., 2019). Examples of educational technologies that provide personalized teaching based on users' inputs are adaptive learning systems and intelligent tutoring systems (Naik & Kamat, 2015; Phobun & Vicheanpanya, 2010; Deunk et al., 2018). Considering the positive results pointed out in the literature concerning the effects of these systems (Fang et al., 2018; Ma et al., 2014; Kulik & Fletcher, 2016; Steenbergen-Hu & Cooper, 2013; VanLehn, 2011; Verdú et al., 2008; Steenbergen-Hu & Cooper, 2014), some researchers are motivated to investigate new technologies to provide a better-personalized experience for students.

A new trend is to use brain-imaging techniques such as electroencephalography (EEG), functional near-infrared spectroscopy (fNIRS), and functional magnetic resonance imaging (fMRI) to provide relevant students' information (e.g., level of attention, emotion, learning state, motivation, and so on) for adaptive learning systems in order to enable an enhanced personalized experience (Y. Li et al., 2011; C.-C. Wang & Hsu, 2014; Bauer et al., 2019; Ghiani et al., 2015; Mailhot et al., 2018). Therefore, adaptive learning environments could rely on direct pieces of evidence of students' current cognitive phenomena (Spüler et al., 2016), instead of depending only on learners' explicit interaction behavior for adaptation such as correctness of responses and reaction times (Käser et al., 2013).

Classification of signals supplied by brain-imaging techniques might play a vital role (Babiker et al., 2019) to provide personalized experiences in educational systems. Based on it, a relevant number of studies are jointly using artificial intelligence algorithms and brain-imaging techniques in the adaptive learning context to present real-time adaptive learning based on students' current cognitive phenomena (Babiker et al., 2019; Y. Li et al., 2011; Kang et al., 2015). For example, a recent study improves the recognition rate of emotions in intelligent tutoring systems by means of a genetic algorithm for optimization of hyper-parameters in a CNN using data from an EEG-based brain-computer interface (Zatarain Cabada et al., 2019).

Therefore, considering the relevance of the application of artificial intelligence algorithms to classify students' information retrieved from brain-imaging techniques in technology-enhanced learning environments and the lack of secondary studies investigating this field, we conducted a systematic literature review, which is a review of primary studies that aim to identify what evidence is available in a specific area (Kitchenham & Charters, 2007). Based on it, this systematic literature review aims to explore which artificial intelligence algorithms have been used in studies that apply brain-imaging techniques in educational technologies, their purposes, and the results obtained in these studies related to the use of AI algorithms.

We structure the remainder of this paper as follows. In the Method section, we depict the systematic literature review planning and execution. In the Results section, we describe the results obtained for each research question after the analysis process. Finally, in the Discussion and Conclusion section, we discuss the results found, and present the concluding remarks of this work.

2 Method

This section aims to detail the systematic literature review method followed by this paper. We developed this review study according to the guidelines proposed by Kitchenham & Charters (2007).

2.1 Planning

In the planning phase, we defined the systematic literature review's objective, research questions, search string, academic databases and inclusion/exclusion criteria.

2.1.1 Objective

This systematic literature review aims to explore, analyze studies in the literature that jointly use artificial intelligence and brain-imaging techniques (e.g. EEG, fNIRS, fMRI) in the technology-enhanced learning context.

2.1.2 Research Questions

Based on the previously explained objective, the two following research questions are targeted in this article:

RQ1: Which artificial intelligence algorithms have been used, and what are their purpose in studies that applied brain-imaging techniques in educational technologies?

RQ2: Which were the results reported in the studies related to the use of artificial intelligence algorithms?

2.1.3 Search String

Considering the purpose of this systematic literature review, we defined the search string based on specialists in educational technologies and psychology fields and based on related works (Santos et al., 2018). The search string is compounded by keywords related to the three following domains: (1) neuroscience; (2) brain-imaging techniques; (3) educational technologies, as we can see in Table 1. We defined the search string by grouping the three-domain keywords with the logic operator "AND".

2.1.4 Academic Databases

Seven academic databases were chosen for the search strategy, as follows:

ACM Digital Library, Engineering Village, IEEE Xplore, Pubmed, Scopus, Springer Link, Web of Science.

The academic databases were chosen based on experts' recommendations through meetings and based on the sources chosen in recent systematic reviews conducted in the technology-enhanced learning domain (Indriasari et al., 2020; Kumar & Chand, 2019; Bano et al., 2018).

Table 1: Keywords of the Search String .

Id	Domain	Keywords
1	neuroscience	"memor*" OR "neuro*" OR "cognit*"
2	brain-imaging techniques	"magnetic resonance imaging" OR "functional magnetic resonance imaging" OR "computed tomography" OR "positron emission tomography" OR "single positron emission tomography" OR "electroencephalography" OR "magnetoencephalography" OR "near infrared spectroscopy" OR "functional near infrared spectroscopy" OR "diffuse optical imaging" OR "event-related optical signal" OR "brain-computer interface"
3	educational technologies	"educational software platform" OR "computers in education" OR "informatics in education" OR "technology in education" OR "educative software" OR "educational software" OR "educational system" OR "learning management system" OR "online education" OR "educational environment" OR "learning environment" OR "virtual learning environment" OR "artificial intelligence in education" OR "artificial intelligence for education" OR "web-based learning" OR "e-learning" OR "electronic learning" OR "m-learning" OR "mobile learning" OR "t-learning" OR "transformative learning" OR "internet-based learning" OR "web-based education" OR "semantic web-based education" OR "semantic web and education" OR "semantic web for education" OR "semantic web-based education" OR "collaborative learning" OR "cooperative learning" OR "collaborative networked learning" OR "collaborative learning in virtual worlds" OR "adaptive hypermedia" OR "adaptive educational systems" OR "hypermedia-based education" OR "intelligent tutoring system" OR "intelligent educational systems" OR "intelligent tutor" OR "distance education" OR "distance learning" OR "*MOOC" OR "massive open online courses" OR "web-based online courses" OR "web-based courses" OR "internet conducted courses" OR "educative game" OR "game-based learn" OR "game-based learning" OR "educational game" OR "game-based education" OR "serious game"

2.1.5 Study Selection Criteria

We define in Table 2 the selection criteria used to determine the inclusion or exclusion of the retrieved studies.

2.2 Procedure

We conducted the search in the seven academic databases in January 2020 and resulted in 6089 articles. The studies returned in the databases were automatically downloaded and inserted into, and organized in the Parsifal tool. First, in the selection process, three researchers reviewed titles, keywords, and abstracts. We excluded the articles that were not related to the scope according to the inclusion and exclusion criteria, remaining 103 studies. In the next phase, we downloaded the remaining 103 studies, and we performed a detailed analysis of the full articles to verify if the remaining studies were under the pre-defined inclusion and exclusion criteria. After this phase, 20 studies remained, and we included them in the systematic literature review. It is essential to highlight that when there was a case of disagreement about accepting a study by the three researchers during the selection phase, we consulted the fourth researcher. When the fourth researcher agreed the article was in the review's scope, we included it in the SLR. The summary of this process can be visualized through Fig. 1.

Table 2: Study selection criteria.

Inclusion criteria	Exclusion criteria
1. Primary Studies	1. Non-English written studies
2. Studies published before December 2019	2. Gray literature
3. Peer-reviewed studies	3. Short-papers and Posters
	4.Duplicated studies (only one copy of each study was included)
	5.Redundant studies of same authorship
	6.Studies not accessible
	7.Studies that do not use brain-imaging techniques in the technology-enhanced learning context
	8. Studies that do not apply artificial intelligence algorithms

2.3 Extraction

We fully analyzed the studies selected in the analysis process. Afterwards, relevant data were extracted from each study according to the extraction form validated by experts through meetings. We extracted the following information from the articles:

1. Authors, year, title, publication source.
2. Artificial intelligence algorithms used to support studies that use brain-imaging techniques in educational technologies.
3. The studies' objective concerning the use of the chosen artificial intelligence algorithm.
4. Results reported in the included studies related to the use of AI algorithms.

3 Results

The objective of this section is to present the results obtained for each research question after the analysis process.

3.1 General Information

3.1.1 Year of Publication

The studies included in this systematic literature review were published between 2010 and 2019. According to Fig. 2, 2018 and 2019 (5 and 4 articles published, respectively) were the years with the largest number of publications of studies, showing an increasing interest of researchers in this research field. Moreover, considering the total number of publications and the year of the first publication (2010), we can perceive that the research field is in its early stages.

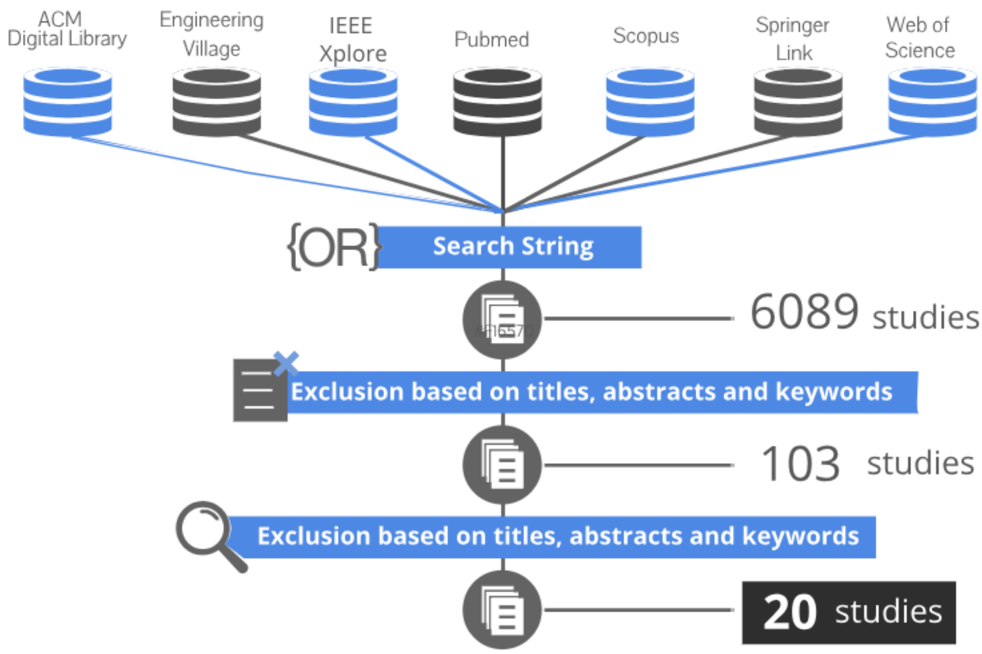

Figure 1: Search and selection process summary.

3.1.2 *Type of Publication*

The type of publication of the included studies was classified in: conference paper, journal publication, or book chapter. In Fig. 3, it is possible to visualize that most of the studies were conference papers (12 out of 20 studies), followed by journal publications (7 out of 20 studies).

3.2 RQ1

The first research question explores which artificial intelligence algorithms have been used and their purposes in studies that applied brain-imaging techniques in educational technologies. In Table 3, we present an overview of the results obtained concerning this research question. Our results point out that most of the papers adopted AI algorithms aiming to classify students' attention levels (4 out of 20 studies), confusing levels (4 out of 20 studies), and emotions (4 out of 20 studies) during their interaction with educational technologies. It is also important to note that the most recent works that aimed to classify students' attention levels, confusing levels, and emotions were the ones that achieved the highest accuracies.

Moreover, our results point out that one of the most popular artificial intelligence algorithms was the support vector machine (7 out of 20 studies), used to classify students' emotions (2 out of 7 studies), students' true and false memories (1 out of 7 studies), students' mental effort (1 out of 7 studies), students' known and unknown words (1 out of 7 studies), students' situational interest (1 out of 7 studies), and students' cognitive load (1 out of 7 studies). Another popular artificial intelligence technique adopted was the neural network (8 out of 20 studies), which is divided into multilayer perceptron (4 out of 8 studies), used to classify students' emotions, attention levels,

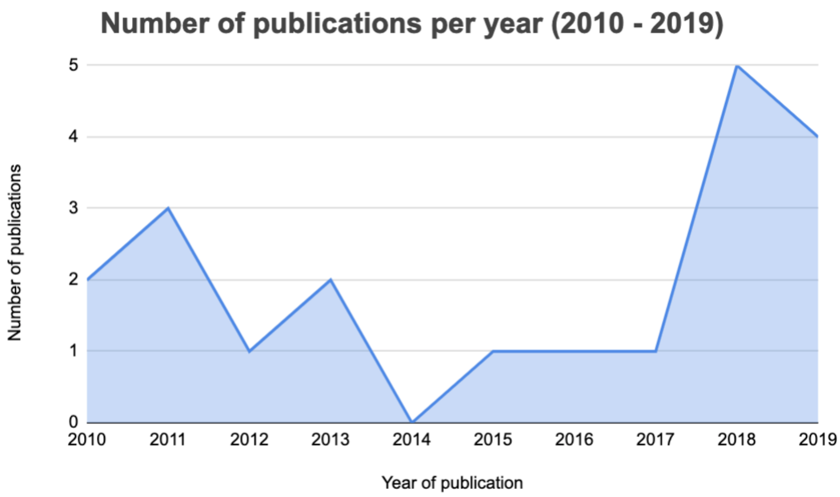

Figure 2: Publications per year of the included articles in the systematic literature review.

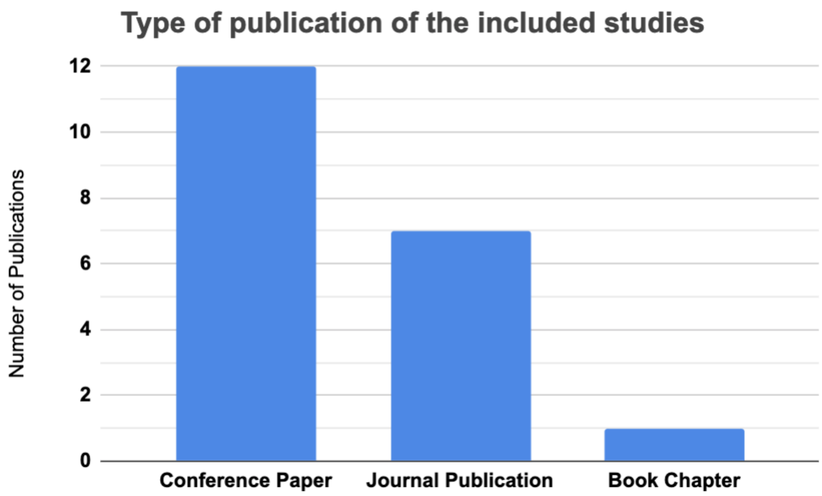

Figure 3: Type of publication of the included articles in the systematic literature review.

error, working memory; into a convolutional neural network (3 out of 8 studies), used to classify students' confusion level, emotions, error; and into bidirectional LSTM recurrent neural networks (1 out of 8 studies), used to classify students' confusion level.

3.3 RQ2

Considering that this systematic literature review's main objective is to investigate the use of artificial intelligence in studies that apply brain-imaging techniques in the technology-enhanced adaptive learning context, this research question analyzes the results reported in the studies concerning the application of AI algorithms. As such, we considered the reported results of the accuracy or F-score metrics to measure how well the algorithms are regarding their application in this context.

As previously reported in RQ1, most of the articles aimed to use artificial intelligence al-

Table 3: Objectives, Artificial Intelligence Algorithms and Results.

Objective	Algorithms	Results	Studies
Classify students' attention levels	K-nearest neighbors - KNN	Accuracy: 57.03% / Accuracy: 51.9%	Y. Li et al. (2011) X. Li et al. (2010)
	Correlation-based feature selection - CFS and K-nearest-neighbor - KNN	Accuracy: $80.84 \pm 3.0\%$	Hu et al. (2018)
	Correlation-based feature selection - CFS and Naive Bayes	Accuracy: $78.00 \pm 5.1\%$	Hu et al. (2018)
	Correlation-based feature selection - CFS and SMO	Accuracy: $78.04 \pm 2.3\%$	Hu et al. (2018)
	Correlation-based feature selection - CFS and Random Forest	Accuracy: $78.18 \pm 5.1\%$	Hu et al. (2018)
	Naive Bayes	Accuracy: 37%	X. Li et al. (2010)
	Linear support vector classifier - SVC	Accuracy: 90%	Mohamed et al. (2020)
	Neural network - NN	Accuracy: 86%	Mohamed et al. (2020)
Classify students' true and false memories	Support vector machines - SVM with RBF kernel	Accuracy: 97%	Bamatraf et al. (2016)
Classify students' confusion level	GP Function	Accuracy: 89.16%	Tahmassebi et al. (2018)
	Bidirectional LSTM Recurrent Neural Networks	Accuracy: 73.3%	Ni et al. (2017)
	RNN-LSTM	Accuracy: 69.0%	Ni et al. (2017)
	Gaussian Naïve Bayes	Accuracy: 57%	H. Wang et al. (2013)
	Convolutional Neural Network - CNN	Accuracy: 91.04%	Zhou et al. (2019)
Classify students' hemispherical synchronization state	One Rule classifier	Accuracy: 99%	Kaszuba & Kostek (2012)
	LADTree	Accuracy: 99%	Kaszuba & Kostek (2012)
	Logistic Model Tree - LMT	Accuracy: 99%	Kaszuba & Kostek (2012)
Classify students' emotions	Support vector machine - SVM with RBF kernel	F-score: 91.49%	Gruenewald et al. (2018)
	Support vector machine - SVM	F-score: 65%	Azcarraga & Suarez (2013)
	Multilayer Perceptron - MLP	F-score: 92%	Azcarraga & Suarez (2013)
	Convolutional Neural Network - CNN and Genetic algorithms GA	Accuracy: 82%	Zatarain Cabada et al. (2019)
Classify students' mental effort	Decision tree	Accuracy: 95.6%	Lin & Kao (2018)
	Support vector machine - SVM with RBF kernel	Accuracy: 91.6%	Lin & Kao (2018)
	Artificial neural network - ANN	Accuracy: 93.1%	Lin & Kao (2018)
Classify students' mental workload behavior	Gaussian Process Regression - GPR	Accuracy: 91%	Chaouachi et al. (2011)
Classify students' known and unkown words	Support vector machine - SVM with RBF kernel	Accuracy: 80.16%	Kang et al. (2015)
Classify students' mental state	Hidden Markov	Accuracy: 83%	Fincham et al. (2010)
Detect students' error	Convolutional Neural Network - CNN	Accuracy: 94.16%	Pinto et al. (2019)
	MultiLayer Perceptron - MLP	Accuracy: 82.50%	Pinto et al. (2019)
Classify students' situational interest	Support vector machine - SVM	Accuracy: 93.3%	Babiker et al. (2019)
	K-nearest-neighbor - KNN	Accuracy: 87.5%	Babiker et al. (2019)
Classify students' cognitive load	Support vector machine - SVM	F-score: 74.3% / F-score: 76.4%	Das et al. (2013)
	Component based Fuzzy c-Means - CFCM	F-score: 76.7% / F-score: 77.9%	Das et al. (2013)
	Traditional Fuzzy c-Means - FCM	F-score: 65.7% / F-score: 68.8%	Das et al. (2013)
Classify students' working memory	Linear support vector classifier - SVC	Accuracy: 85%	Mohamed et al. (2020)
	Neural network - NN	Accuracy: 87%	Mohamed et al. (2020)

gorithms to classify students' attention levels through the data acquired from brain-imaging techniques while using some educational technology. As seen in Table 3, researchers used different algorithms to perform this task. According to the study of Mohamed et al. (2020), the algorithms that achieved the best performance in classifying students' attention levels was the linear support vector classifier - SVC (reported accuracy: 90%), followed by Multilayer Perceptron - MLP (reported accuracy: 86%). Hu et al. (2018) also note that the combination of correlation-based feature selection (CFS) and k-nearest-neighbor (KNN) algorithms demonstrated an accuracy of $80.84 \pm 3.0\%$.

Classifying students' confusion level is another aim of some studies. As seen in Table 3, in Zhou et al. (2019), the convolutional neural network - CNN obtained an accuracy of 91.04% to classify students' confusion level. Moreover, in Tahmassebi et al. (2018), GP Function presented an accuracy of 89% to classify students' confusion level while using some educational technology. Concerning the classification of students' emotions, Gruenewald et al. (2018) reported an F-score of 91.49% using Support vector machine - SVM with RBF kernel, as seen in Table 3. In Azcarraga & Suarez (2013), Multilayer Perceptron - MLP presented an F-score of 88% to classify students' emotions.

As shown in Table 3, several algorithms are used, starting from probabilistic models, such as Naive Bayes, logical models of grouping/classification, such as KNN and SVM, and models based on deep learning, such as MLP and LSTM. We could observe that studies that present a limited training set have satisfactory results with the traditional Artificial Intelligence algorithm such as KNN. Although this model achieves results without the need for training and adjustment based on various distance measures such as Euclidean or cosine, the model has a disadvantage linked to large data sets. On the other hand, although the most recent deep learning techniques have better accuracy when trained with big data sets, they need a long training, validation, and testing for their due convergence.

3.4 Overview of the Studies

This subsection aims to present an overview of each study included in this systematic literature review. Each overview provides the following information: the AI algorithms adopted, the objective of the AI algorithms in the study's context, the tasks students were performing while the data were collected, the brain-imaging techniques used, and the results achieved by the algorithms.

In Y. Li et al. (2011), KNN was chosen for real-time attention level recognition. The data was collected from an EEG-based brain-computer interface (BCI) while the learners performed several tasks throughout seven sessions. The experiment achieved the highest accuracy in the seventh session, achieving 65% accuracy on one subject and 57% average accuracy across all subjects with response delay times shorter than 300ms.

In Pinto et al. (2019), MLP with Resilient Backpropagation and a neural network with convolutional and LSTM layers were chosen to detect errors from brain signals. The data was collected from an EEG device while the learners played a game with intentional bugs and talked to a robot which also presented erroneous behavior. The CNN and LSTM neural network's results were much better than the MLP's results, achieving an accuracy of 94.16% compared to the MLP's 82.50% accuracy, which, as the authors say, "clearly demonstrated that temporal information is important for signal analysis".

Bamatraf et al. (2016) chose SVM with RBF kernel to classify (predict) true and false memories, both short term memories (STM) and long term memories (LTM). The data was collected by an EEG device while the learners watched learning material; half the learners watched in 2D, and the other half watched the same material in 3D. An EEG device was also used during the memory recall tests for both STM (after 30 minutes) and LTM (after two months). According to the authors, the average prediction accuracy of the SVM with RBF kernel was 97% for both 2D and 3D.

In Tahmassebi et al. (2018), GP Classifier was chosen to classify confusion in a binary way (confused or not confused). The data was collected from an EEG device while students were watching MOOC videos. The best model was a GP Classification model that outperformed the others and yielded an accuracy of 89.16%. Moreover, Hu et al. (2018) combined correlation-based feature selection (CFS) with KNN, Naive Bayes, SMO and Random Forests for attention recognition. The data was collected from an EEG device while subjects performed various tasks given randomly in an e-learning environment. The test results with cross-validation yielded an average accuracy of 80.84 $\pm$ 3.0% across test subjects for the CFS and KNN combination; 78.00 $\pm$ 5.1% for the CFS and Naive Bayes combination; 78.04 $\pm$ 2.3% for the CFS and SMO combination; and 78.18 $\pm$ 5.1% for the CFS and Random Forests combination.

Gruenewald et al. (2018) tried SVM, MLP, CNN, LSTM and hybrid models combining CNN and LSTM for emotion classification. For the data collection, this work introduced a new biomedical multi-sensor platform, collecting a variety of physiological data: temperature, electroencephalography (EEG), electrooculography (EOG), galvanic skin response (GSR), heart rate and blood oxygen saturation. Results tested using 5-fold cross-validation showed that SVM with RBF kernel approach with hand-crafted features achieved better results than the ANNs approaches, yielding an average F1-score of 91.49% across the folds.

In Zhou et al. (2019) was chosen an ANN to classify learners' confusion. The data was collected from an EEG device; the models were trained on the data of one standardized cognitive test paradigm (Raven's test) and tested on the data of real tasks in gameplay (Sokoban Game). As for the results, the ANN developed by this work yielded an accuracy of 91.04%.

Ni et al. (2017) applied bidirectional LSTM to classify students' confusion, in addition to SVM, KNN, CNN, Deep Belief Network and RNN. The data was collected by an EEG device while students watched online course videos, and results showed that the most important feature to detect confusion is the gamma 1 wave of EEG signal. The bidirectional LSTM yielded the best average accuracy across 5 folds of cross-validation, 73.3%, followed by RNN-LSTM, with an average accuracy of 69.0%.

In Kang et al. (2015) was used SVM to classify students' confusion to predict if the learner sees an unknown word to her/him in real-time. This paper used eyeball movement data and brain waves data. Eyeball movement data were collected by an eye tracker device, and brain waves data was collected from an EEG device while the learners performed various tasks related to language study. The SVM model yielded an accuracy of 80.16% on the test set.

Babiker et al. (2019) applied SVM and KNN to classify learners' interest. The data was collected using an EEG device while watching lectures. Across 10 folds of cross-validation, SVM achieved a high accuracy of 93.3% and 87.5% for two datasets using features from four EEG channels, and KNN achieved an accuracy of 87.5% and 86.7% using the same datasets using a single EEG channel. Kaszuba & Kostek (2012) chose One Rule, LADTree, and Logistic Model Tree (LMT) to classify learners' hemispherical synchronization state. The data used was collected from an EEG device while students were performing several tests. The test results performed with the cross-validation procedure showed high efficiency for each of the chosen methods, presenting accuracies of 99% for all three classifiers chosen.

In Mohamed et al. (2020) was used neural network (NN) and linear support vector classifiers (SVC) to detect learners' focused attention and working memory. The data used was collected from an EEG device while students were undergoing a cognitive assessment battery. For the detection of focused attention, the SVC yielded an accuracy of 90%, and NN yielded an accuracy of 86%. For working memory, the SVC obtained an accuracy of 85%, and NN obtained an accuracy of 87%. Zatarain Cabada et al. (2019) adopted a convolutional neural network (CNN) optimized using genetic algorithms (GA) to recognize emotion in intelligent tutoring systems. The data used was collected from an EEG device. This approach obtained an 82% accuracy rate.

X. Li et al. (2010) applied k-Nearest-Neighbor and Naive Bayes to classify students' attention in e-learning systems. The data used was collected from an EEG device while the subjects were studying different learning materials. If applied to the same feature group, the best algorithm was KNN(K=5) classifier, with an average of 44.45%. With analysis of using same Features with Different Classifiers, the KNN The average for the Naive Bayes classifier was 37.0%. The average for the KNN (K = 3) classifier was 51.9%.

Lin & Kao (2018) adopted decision tree, SVM, and ANN to classify students' mental effort in e-learning contexts. The data used was collected from EEG devices while students were watching online videos. The decision tree obtained the highest average accuracy, which was 95,6%, followed by ANN, which was 93,1%, and SVM, which was 91,6%.

In Chaouachi et al. (2011) was used Gaussian Process Regression to predict students' mental workload when performing different cognitive tasks in Intelligent Systems. The data used was collected from EEG devices while students were performing different tasks such as trigonometry and logic activities. A mean accuracy rate of 91% across all the participants was reached. Azcarraga & Suarez (2013) used a multi-layered perceptron (MLP) and support vector machine (SVM) to predict academic emotions (confidence, excitement, frustration, and interest) during learning sessions. The data used was collected from EEG devices and mouse click information while students were solving algebra equations. The overall performance for all emotions in terms of average F-measure was 92% for MLP and 65% was for SVM.

Das et al. (2013) used Support Vector Machine (SVM), Component-based Fuzzy c-Means (CFCM), and traditional Fuzzy c-Means (FCM) to classify the level of cognitive load on an individual for a given stimulus. The data used was collected from an EEG device while students were performing different tasks: typing phrases using two different on-screen layouts and doing a logical reasoning test. On-Screen Keyboard tasks, the overall performance in terms of average F-measure was 74.3% for SVM, 76.7% for CFCM and 65.7% for FCM. On Logical Reasoning tasks, the overall performance in terms of average F-measure was 76.4% for SVM, 77.9% for CFCM and 68.8% for FCM.

In C.-C. Wang & Hsu (2014) Gaussian Naive Bayes classifiers were trained to detect when the student is confused while watching the course material. The data used was collected from an EEG device while the students watched a 2-minute video. Student-specific classifiers achieved a classification accuracy of 57%. Fincham et al. (2010) used a hidden Markov algorithm to predict students' mental states during problem-solving episodes. The data used was collected from Functional magnetic resonance imaging (fMRI) while students worked with a tutoring system that taught an algebra isomorph. In terms of predicting what state a student was in during any 2 second period, the algorithm achieved 83% accuracy on the test data.

4 Discussion and Conclusion

This systematic literature review presents the results obtained from the analysis of 20 studies that jointly used artificial intelligence algorithms and brain-imaging techniques in the technology-enhanced learning context. The included studies were published between 2010 and 2019 in conferences, journals, and books. The main objective of this study was to analyze and synthesize what is currently done in the literature in this research field. Therefore, this article's two research questions focused on reporting which artificial intelligence algorithms were used in the included studies, which were their purposes and the results obtained in these papers through AI. In the present study, it is possible to distinguish the algorithms used in the context of technology-enhanced learning to achieve different objectives that range from logic, probabilistic models to more sophisticated mathematical models such as Deep Learning. This SLR reports exciting results, discussions, and comparisons between different algorithms to help this emerging area.

The first research question pointed out that support vector machines (SVM) is the most popular artificial intelligence algorithms used in the studies that use brain-imaging techniques in educational technologies. An explanation for this favoritism is that SVM is robust, versatile, and easy to implement and presents many free resources and toolboxes, explaining its numerous implementations in the literature (Babiker et al., 2019). However, although SVM demonstrates an excellent performance to classify students' cognitive phenomena during the learning process, the results point out the importance of exploring a wide range of algorithms to verify which provides the best outcomes for different purposes. For example, although SVM achieved an excellent performance to classify students' cognitive load, traditional fuzzy c-means (FCM) reported better results. This result could be an incentive for future works to explore artificial intelligence algorithms such as SVM, KNN, MPL, and other potentially effective ones.

It is also important to highlight a trend detected in some included articles. Some studies investigated the combination of data retrieved from brain-imaging techniques with data retrieved from other devices to produce more robust data sets in order to better classify students' cognitive phenomena in the technology-enhanced learning context. For example, studies that aimed to detect students' emotions combined data related to students' brain activities with data related to facial recognition (retrieved with camera (Zatarain Cabada et al., 2019)) and related to behavior in the system (retrieved with mouse (Azcarraga & Suarez, 2013)).

The use of artificial intelligence and brain-imaging techniques in educational technologies is a promising approach, considering that the most recent published studies in the field presented a considerable performance improvement in the results achieved by using AI algorithms compared to the first published articles. For example, as expected, studies whose purpose is to use artificial intelligence algorithms to classify students' attention level during the learning process published in 2018 and 2020 (Hu et al., 2018; Mohamed et al., 2020) presented a significant performance improvement compared to studies with the same objective published in 2010 and 2011 (X. Li et al., 2010; Y. Li et al., 2011). Similarly, the same phenomenon occurred with studies that aimed to classify students' confusion level, cognitive load, and emotions, showing that the research field is becoming more mature.

Finally, the results reported in this study show promising results related to the use of brain-imaging techniques-based devices in the technology-enhanced learning context. More specifically, this study points out the effectiveness of using different artificial intelligence algorithms to classify students' cognitive phenomena during the learning process in order to provide personalized learning in educational environments (e.g., Massive Open Online Courses, intelligent tutoring systems), based on the data collected using these brain-imaging technologies. Therefore, it is possible to identify significant results by using traditional classification algorithms to create customized feedback automatically, related in contexts of synchronous or asynchronous learning.

5 Theoretical Agenda

In general, the systematic literature review's results are encouraging. Most of the included papers reported exciting results concerning artificial intelligence algorithms to classify students' cognitive phenomena during the learning process. These results demonstrate the feasibility of jointly using artificial intelligence algorithms and brain-imaging techniques in the technology-enhanced learning context to provide a personalized and better learning experience for students. However, this is a new field and needs to be advanced in the next few years. Therefore, based on our study results, we propose a theoretical plan to promote the area.

1. **Create standardized data sets**: Despite recent advances in deep learning-based models,

there is an apparent use of traditional algorithms that may be justified by computational complexity and the convergence rate of neural networks. Taking EEG into consideration, it would be further necessary to map which attributes and variables may be relevant to study and create models based on deep learning.

2. **Compare algorithms using standardized data sets:** A low number of studies compare different algorithms using standardized data sets for the same purpose. Therefore, it would also be necessary to compare the different algorithms on standardized data sets to find the best algorithms for specific purposes in terms of time, accuracy, and quality of results.

3. **Comparison between techniques with unique data sets**: Understanding the models created and traditional classification techniques for expansion into more sophisticated deep learning models can be more easily accomplished with comparisons and availability of data sets for replication and adaptation of algorithms, and mechanisms for transfer learning, commonly used in image processing with large neural networks.

4. **Diversify studies' goals:** Most of the studies used artificial intelligence algorithms to classify students' attention levels, confusing levels, and emotions during their interaction with educational technologies. However, there is a low number of studies exploring the use of artificial intelligence algorithms to classify students' mental workload behavior, known and unknown words, mental state, and true and false memories, for example. Therefore, future studies could better explore other artificial intelligence algorithms' use to these and other goals poorly or not explored yet, such as anxiety, concentration, interest, and relaxation.

Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

References

Akbulut, Y., & Cardak, C. S. (2012). Adaptive educational hypermedia accommodating learning styles: A content analysis of publications from 2000 to 2011. *Computers and Education*, *58*(2), 835–842. doi: 10.1016/j.compedu.2011.10.008 [GS Search]

Azcarraga, J., & Suarez, M. T. (2013). Recognizing Student Emotions Using Brainwaves and Mouse Behavior Data. *Int. J. Distance Educ. Technol.*, *11*(2), 1–15. doi: 10.4018/jdet.2013040101 [GS Search]

Babiker, A., Faye, I., Mumtaz, W., Malik, A. S., & Sato, H. (2019). EEG in classroom: EMD features to detect situational interest of students during learning. *Multimedia Tools and Applications*, *78*(12), 16261–16281. doi: 10.1007/s11042-018-7016-z [GS Search]

Bamatraf, S., Hussain, M., Aboalsamh, H., Qazi, E.-U.-H., Malik, A. S., Amin, H. U., ... Imran, H. M. (2016). A System for True and False Memory Prediction Based on 2D and 3D Educational Contents and EEG Brain Signals. *Computational Intelligence and Neuroscience*, *2016*, 8491046. doi: 10.1155/2016/8491046 [GS Search]

Bano, M., Zowghi, D., Kearney, M., Schuck, S., & Aubusson, P. (2018). Mobile learning for science and mathematics school education: A systematic review of empirical evidence. *Computers and Education*, *121*, 30–58. doi: 10.1016/j.compedu.2018.02.006 [GS Search]

Bauer, M., Bräuer, C., Schuldt, J., Niemann, M., & Krömker, H. (2019). Application of Wearable Technology for the Acquisition of Learning Motivation in an Adaptive E-Learning Platform. In T. Z. Ahram (Ed.), *Advances in human factors in wearable technologies and game design* (pp. 29–40). Campinas: Springer, Cham. doi: 10.1007/978-3-319-94619-1_4 [GS Search]

Chaouachi, M., Jraidi, I., & Frasson, C. (2011). Modeling Mental Workload Using EEG Features for Intelligent Systems. In J. A. Konstan, R. Conejo, J. L. Marzo, & N. Oliver (Eds.), *User modeling, adaption and personalization* (pp. 50–61). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-22362-4_5 [GS Search]

Das, D., Chatterjee, D., & Sinha, A. (2013). Unsupervised approach for measurement of cognitive load using EEG signals. In *13th ieee international conference on bioinformatics and bioengineering, ieee bibe 2013.* doi: 10.1109/BIBE.2013.6701686 [GS Search]

Deunk, M. I., Smale-Jacobse, A. E., de Boer, H., Doolaard, S., & Bosker, R. J. (2018). Effective differentiation Practices:A systematic review and meta-analysis of studies on the cognitive effects of differentiation practices in primary education. *Educational Research Review*, *24*, 31–54. doi: 10.1016/j.edurev.2018.02.002 [GS Search]

Fang, Y., Ren, Z., Hu, X., & Graesser, A. C. (2018). A meta-analysis of the effectiveness of ALEKS on learning. *Educational Psychology*, 1278–1292. doi: 10.1080/01443410.2018.1495829 [GS Search]

Fincham, J. M., Anderson, J. R., Betts, S., & Ferris, J. L. (2010). Using Neural Imaging and Cognitive Modeling to Infer Mental States while Using an Intelligent Tutoring System. In *Educational data mining 2010 - 3rd international conference on educational data mining* (pp. 51–60). doi: 10.1184/R1/6618986.v1 [GS Search]

Ghiani, G., Manca, M., & Paternò, F. (2015). Dynamic User Interface Adaptation Driven by Physiological Parameters to Support Learning. In *Proceedings of the 7th acm sigchi symposium on engineering interactive computing systems* (pp. 158–163). New York, NY, USA: ACM. doi: 10.1145/2774225.2775081 [GS Search]

Gruenewald, A., Kroenert, D., Poehler, J., Brueck, R., Li, F., Littau, J., ... Niehaves, B. (2018). Biomedical Data Acquisition and Processing to Recognize Emotions for Affective Learning. In *2018 ieee 18th international conference on bioinformatics and bioengineering (bibe)* (pp. 126–132). doi: 10.1109/BIBE.2018.00031 [GS Search]

Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge Taylor & Francis Group. doi: 10.4324/9780203887332 [GS Search]

Hu, B., Li, X., Sun, S., & Ratcliffe, M. (2018). Attention Recognition in EEG-Based Affective Learning Research Using CFS+KNN Algorithm. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *15*(1), 38–45. doi: 10.1109/TCBB.2016.2616395 [GS Search]

Indriasari, T. D., Luxton-Reilly, A., & Denny, P. (2020). Gamification of student peer review in education: A systematic literature review. *Education and Information Technologies*, *25*(6), 5205–5234. doi: 10.1007/s10639-020-10228-x [GS Search]

Kang, J.-S., Ojha, A., & Lee, M. (2015). Development of Intelligent Learning Tool for Improving Foreign Language Skills Based on EEG and Eye Tracker. In *Proceedings of the 3rd international conference on human-agent interaction* (pp. 121–126). New York, NY, USA: ACM. doi: 10.1145/2814940.2814951 [GS Search]

Käser, T., Baschera, G. M., Busetto, A. G., Klingler, S., Solenthaler, B., Buhmann, J. M., & Gross, M. (2013). Towards a Framework for Modelling Engagement Dynamics in Multiple Learning Domains. In *International journal of artificial intelligence in education* (Vol. 22, pp. 59–83). Springer New York LLC. doi: 10.3233/JAI-130026 [GS Search]

Kaszuba, K., & Kostek, B. (2012). Employing a Biofeedback Method Based on Hemispheric Synchronization in Effective Learning. In Z. S. Hippe, J. L. Kulikowski, & T. Mroczek (Eds.), *Human – computer systems interaction: Backgrounds and applications 2: Part 2* (pp. 295–309). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-23172-8_20 [GS Search]

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. In (Vol. 2). Keele University and Durham University Joint Report. [GS Search]

Klašnja-Milićević, A., Vesin, B., Ivanović, M., & Budimac, Z. (2011). E-Learning personalization based on hybrid recommendation strategy and learning style identification. *Computers and Education*, *56*(3), 885–899. doi: 10.1016/j.compedu.2010.11.001 [GS Search]

Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of Intelligent Tutoring Systems: A Meta-Analytic Review. *Review of Educational Research*, *86*(1), 42–78. doi: 10.3102/0034654315581420 [GS Search]

Kumar, B. A., & Chand, S. S. (2019). Mobile learning adoption: A systematic review. In *Education and information technologies* (Vol. 24, pp. 471–487). Springer New York LLC. doi: 10.1007/s10639-018-9783-6 [GS Search]

Li, X., Zhao, Q., Hu, B., Liu, L., Peng, H., Qi, Y., . . . Liu, Q. (2010). Improve Affective Learning with EEG Approach. *Computing and Informatics*, *29*, 557–570. [GS Search]

Li, Y., Li, X., Ratcliffe, M., Liu, L., Qi, Y., & Liu, Q. (2011). A Real-time EEG-based BCI System for Attention Recognition in Ubiquitous Environment. In *Proceedings of 2011 international workshop on ubiquitous affective awareness and intelligent interaction* (pp. 33–40). New York, NY, USA: ACM. doi: 10.1145/2030092.2030099 [GS Search]

Lin, F.-R., & Kao, C.-M. (2018). Mental effort detection using EEG data in E-learning contexts. *Computers & Education*, *122*, 63–79. doi: 10.1016/j.compedu.2018.03.020 [GS Search]

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, *106*(4), 901–918. doi: 10.1037/a0037123 [GS Search]

Mailhot, T., Lavoie, P., Maheu-Cadotte, M.-A., Fontaine, G., Cournoyer, A., Côté, J., ... Cossette, S. (2018). Using a Wireless Electroencephalography Device to Evaluate E-Health and E-Learning Interventions. *Nursing Research*, *67*, 43–48. doi: 10.1097/NNR.0000000000000260 [GS Search]

Mohamed, Z., Halaby, M. E., Said, T., Shawky, D., & Badawi, A. (2020). Facilitating Classroom Orchestration Using EEG to Detect the Cognitive States of Learners. In *Advances in intelligent systems and computing* (Vol. 921, pp. 209–217). Springer Verlag. doi: 10.1007/978-3-030-14118-9_21 [GS Search]

Naik, V., & Kamat, V. (2015). Adaptive and Gamified Learning Environment (AGLE). In *2015 ieee seventh international conference on technology for education (t4e)* (pp. 7–14). doi: 10.1109/T4E.2015.23 [GS Search]

Ni, Z., Yuksel, A. C., Ni, X., Mandel, M. I., & Xie, L. (2017). Confused or Not Confused?: Disentangling Brain Activity from EEG Data Using Bidirectional LSTM Recurrent Neural Networks. In *Proceedings of the 8th acm international conference on bioinformatics, computational biology,and health informatics* (pp. 241–246). New York, NY, USA: ACM. doi: 10.1145/3107411.3107513 [GS Search]

Normadhi, N. B. A., Shuib, L., Nasir, H. N. A., Bimba, A., Idris, N., & Balakrishnan, V. (2019). Identification of personal traits in adaptive learning environment: Systematic literature review. *Computers and Education*, *130*, 168–190. doi: 10.1016/j.compedu.2018.11.005 [GS Search]

Parsons, S. A., Vaughn, M., Scales, R. Q., Gallagher, M. A., Parsons, A. W., Davis, S. G., ... Allen, M. (2018). Teachers' Instructional Adaptations: A Research Synthesis. *Review of Educational Research*, *88*(2), 205–242. doi: 10.3102/0034654317743198 [GS Search]

Phobun, P., & Vicheanpanya, J. (2010). Adaptive intelligent tutoring systems for e-learning systems. *Procedia - Social and Behavioral Sciences*, *2*(2), 4064–4069. doi: 10.1016/j.sbspro.2010.03.641 [GS Search]

Pinto, A., Nardari, G., Mijam, M., Morya, E., & Romero, R. (2019). A Serious Game to Build a Database for ErrP Signal Recognition. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 11507 LNCS, pp. 186–197). Springer Verlag. doi: 10.1007/978-3-030-20518-8_16 [GS Search]

Santos, W. O. d., Bittencourt, I. I., Isotani, S., Dermeval, D., Marques, L. B., & Silveira, I. F. (2018). Flow Theory to Promote Learning in Educational Systems: Is it Really Relevant? *Revista Brasileira de Informática na Educação*, *29–59*(02), 29. doi: 10.5753/rbie.2018.26.02.29 [GS Search]

Smale-Jacobse, A. E., Meijer, A., Helms-Lorenz, M., & Maulana, R. (2019). Differentiated Instruction in Secondary Education: A Systematic Review of Research Evidence. *Frontiers in Psychology*, *10*, 23–66. doi: 10.3389/fpsyg.2019.02366 [GS Search]

Spüler, M., Walter, C., Rosenstiel, W., Gerjets, P., Moeller, K., & Klein, E. (2016). EEG-based prediction of cognitive workload induced by arithmetic: a step towards online adaptation in numerical learning. *ZDM*, *48*(3), 267–278. doi: 10.1007/s11858-015-0754-8 [GS Search]

Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology*, *105*(4), 970–987. doi: 10.1037/a0032447 [GS Search]

Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology*, *106*(2), 331–347. doi: 10.1037/a0034752 [GS Search]

Tahmassebi, A., Gandomi, A., & Meyer-Base, A. (2018). An Evolutionary Online Framework for MOOC Performance Using EEG Data. In *Ieee congress on evolutionary computation* (pp. 1–8). doi: 10.1109/CEC.2018.8477862 [GS Search]

VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems, Educational Psychologist. *Educational Psychologist*, *46*(4), 197–221. doi: 10.1080/00461520.2011.611369 [GS Search]

Verdú, E., Regueras, L. M., Verdú, M. J., De Castro, J. P., & Pérez, M. (2008). An Analysis of the Research on Adaptive Learning: The Next Generation of e-Learning. *WSEAS Trans. Info. Sci. and App.*, *5*(6), 859–868. [GS Search]

Wang, C.-C., & Hsu, M.-C. (2014). An exploratory study using inexpensive electroencephalography (EEG) to understand flow experience in computer-based instruction. *Information & Management*, *51*(7), 912–923. doi: 10.1016/j.im.2014.05.010 [GS Search]

Wang, H., Li, Y., Hu, X., Yang, Y., Meng, Z., & Chang, K.-M. (2013). Using EEG to improve massive open online courses feedback interaction. In *Ceur workshop proceedings* (Vol. 1009, pp. 59–66). [GS Search]

Xie, H., Chu, H. C., Hwang, G. J., & Wang, C. C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers and Education*, *140*, 103599. doi: 10.1016/j.compedu.2019.103599 [GS Search]

Zatarain Cabada, R., Rodriguez Rangel, H., Barron Estrada, M. L., & Cardenas Lopez, H. M. (2019, 10). Hyperparameter optimization in CNN for learning-centered emotion recognition for intelligent tutoring systems. *Soft Computing*, 1–10. doi: 10.1007/s00500-019-04387-4 [GS Search]

Zhou, Y., Xu, T., Li, S., & Shi, R. (2019). Beyond engagement: an EEG-based methodology for assessing user's confusion in an educational game. *Universal Access in the Information Society*, *18*(3), 551–563. doi: 10.1007/s10209-019-00678-7 [GS Search]