

Comparative analysis of machine learning models for concentration level prediction in e-learning

Abdessamad Chanaa¹, Fatima-ezzahra Ait Bennacer², Nour-eddine El Faddouli¹

¹RIME Team, MASI Laboratory, E3S Research Center, Department of Computer Science, Mohammed V University in Rabat, Rabat, Morocco

²ELITES Laboratory, Department of Computer Science, Faculty of Sciences, Chouaib Doukkali University, El Jadida, Morocco

Article Info

Article history:

Received Dec 12, 2025

Revised Jul 23, 2026

Accepted Sep 4, 2026

Keywords:

Classification

Concentration level

eXtreme gradient boosting

Machine learning

Online assessments

ABSTRACT

Concentration level plays an important role while accomplishing cognitive tasks. Although there are several methods for predicting concentration levels, the majority of the methods now in use for assessing concentration levels rely on labor-intensive and time-consuming human coding. In this work, learners' performance and involvement in online assessments are utilized to categorize the concentration level using a publicly accessible dataset. The goal of the project is to identify the best machine learning (ML) model for concentration level prediction. Six distinct ML classifiers are employed for comparative analysis in this context. When compared to alternative classifiers for the assessment of concentration level, the eXtreme gradient boosting (XGBoost) model has the best precision (80.87%). These findings are highly promising and have implications for the future development of automated concentration level analytic tools for online learning.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Abdessamad Chanaa

RIME Team, MASI Laboratory, E3S Research Center, Department of Computer Science

Mohammadia School of Engineers, Mohammed V University in Rabat, Rabat, Morocco

Email: abdessamad.chanaa@gmail.com

1. INTRODUCTION

For many years, attention has been a fascinating topic for scholars. The capacity to engage both attentional and executive processes while working toward a certain objective is known as concentration. Concentration is the mental state in which a person's entire senses and thoughts are focused on a single topic, and it is something that can be learned [1]. Generally speaking, attention is defined as one's limited intellectual capacity [2], [3]. Another way to think about concentration is as the mechanism that encodes linguistic input, maintains its activity in short-term memory, and retrieves it from long-term memory [4]. Learning and education are closely related to focus and attention [5]. During the learning phase, students' motivation and focus are crucial since the more motivated and focused they are, the better their learning outcomes will be.

The lack of connection between professors and students is one of the primary drawbacks of e-learning, despite its benefits. It is clear that online learning cannot take the place of some parts of education, such as interacting with instructors and studying with classmates. These drawbacks frequently result in inefficient instruction, which causes learning loss in a large number of learners [6]. Concentration is a fundamental and undeniable factor in online education. Effective and adequate evaluation of the learners' focus is essential for providing meaningful feedback to learners and tutors and thus the evolution of efficient and personalized smart

tutoring systems.

Artificial intelligence (AI), most precisely machine learning (ML), attained remarkable progress in solving complex data representation and automatically learn and analyse classification models, especially in education [7]. From this angle, the objective of this research is to automatically examine and detect the learner's concentration level in online assessments and develop a methodology to identify factors that influence the learner's concentration level during the learning process using the ML approach.

The structure of the paper is arranged as follows: consistent literature is reviewed in section 2. Section 3 explores the proposed ML approach for e-learners assessments. We exhibit the experiment, the analysis, results discussion, and limitations of this work are presented in section 4. Finally, section 5 presents the strengths of our methodology and future works.

2. LITERATURE REVIEW

Recently, there has been an growing interest in research within areas related to e-learning with ML approaches. Analyzing and detecting learners' concentration level within learning process is not an exception. In this direction, numerous attempts have been made to use e-learners' behavior to gauge their level of concentration. Jeong *et al.* [8] used artificial recurrent neural networks (RNN), gated recurrent units (GRU), and long short-term memory (LSTM) to predict the concentration of learners based on facial landmark extracted from video features. Li *et al.* [9] used support vector machine (SVM) algorithm on data collected by mouse and webcam to predict learners' concentration level. Hasnine *et al.* [10] utilized convolutional neural network (CNN) to extract six types of basic emotions, they were employed to assess students' attention in a virtual classroom. Chanaa and Faddouli [11] used the concentration level as decisive factor to predict cognitive level via neural network (NN).

Vijayakumar *et al.* [12] used the dimension reduction unsupervised ML approach for facial feeling acknowledgment by understudying facial expression, eye stare checking and head movements. Villegas-Ch *et al.* [13] used cameras to monitor the attention level of students in real time during online classes based on standardized tests, behavioral observation, and assessment of academic achievement. Lokare and Netak [14] sought to find an effective ML model to predict concentration levels using brain signals, recorded through Electroencephalogram. In the same direction, Al-Nafjan and Aldayel [15] developed and implement a real-time attention detection system for detecting concentration level using physiological data via random forest (RF) classifier using electroencephalography (EEG) signals in online classroom. Kawamura and Murase [16] focused on learner's facial reaction to tutor's behavior aimed to attract attention using Jaccard coefficients and SVM ML algorithm. Through e-learning photos, Shamika *et al.* [17] discovered the importance of emotions such as fear, surprise, happiness and calm to determine a student's concentration level using CNN. Sharma *et al.* [18], [19] used a laptop's built-in webcam to capture real-time information on students' eye movements (eye tracking) and facial expressions in order to assess their concentration levels using a CNN. This approach also provides valuable insights into students' emotional responses during lessons, helping teachers better understand their level of engagement in the classroom. By gaining a deeper understanding of these emotional cues, educators can adapt their teaching strategies and offer more effective support to learners [20].

The presented literature discusses the adoption of ML algorithms to detect and analyse learners' concentration level. However, those methods are generally based on image processing and electroencephalography; therefore, those methods are hard to be generalized into any e-learning platform, since only a few learning management systems (LMS) can provide those technical tools in online education. Also, processing real time images or spontaneous electrical activity of the brain is very complex tasks and time-consuming. We propose a new method that analyzes learners' behavioral activities in the context of online assessment (more easier data to process, and very available on most of the LMS) that takes into consideration epistemological emotions like confusion that directly affect concentration level in the learning process.

3. MACHINE LEARNING APPROACH

Our approach is to establish a ML model that considers a variety of epistemic and cognitive aspects in the context of online evaluation. An outline of our proposed approach is shown in Figure 1. Our method is divided into three steps: the extraction and aggregation of variables/features, ML and predictive approaches, and making adjustments by determining the degree of concentration of learners. We will go into further depth about each step separately in the next subsections. In order to determine each learner's precise concentration level and develop an appropriate pedagogical proposal for each learner in the system, we first define our variables/features

that could determine the learner's concentration level during the learning process. Next, the ML model will categorize learners into various classes based on the identified features.

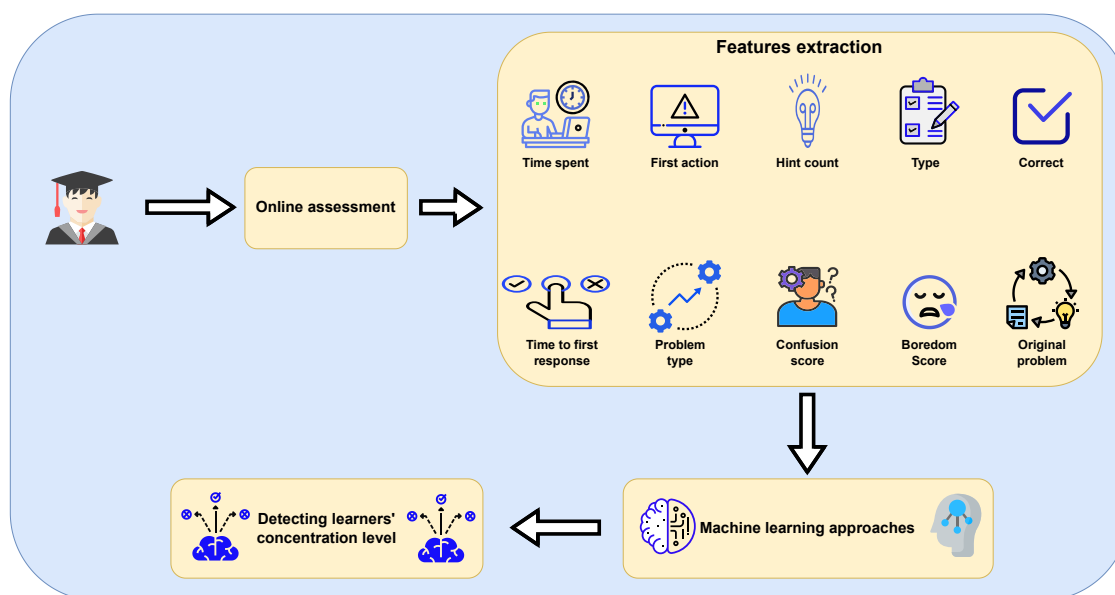


Figure 1. Methodology steps

3.1. Extraction and aggregation of variables/features

It was challenging to carefully choose and compile appropriate elements that may help define the learner's concentration level because our approach is focused on online learning assignments. The primary obstacle in this study problem is the improper categorization of the online assignment questions and the incorrect weighting of each assignment, category, and subject in the final choice to score learners. Furthermore, we are unable to accurately identify learners based on the assignments' yield or score. As a result, we suggest a mix of affective and cognitive characteristics that influence each learner's concentration level when they are studying a certain learning topic. These variables are defined as follows:

- Time spent:** this is the amount of time that passes between the beginning of the task and the final answer submission. Determining whether the learner takes a long time or a short time to solve the task is a crucial variable. In general, students who maintain high levels of concentration are more likely to complete their problem assignments faster than those who do not. However, this is not a general rule. We can identify learners who struggle with concentration and attempt to complete assignments quickly. So the time spent is very important feature but does not give a full image about learners' concentration level.
- Time to first response:** this feature represents the student's first response time. Similar to the time spent, this feature provides a good overview of the concentration level, but not enough on its own. The time to first response may be equal to the time spent if the student had a good answer from the first try without any hints or false attempts. This is also a beneficial factor for measuring the focus during the learning process.
- Hint count:** represents the number of hints utilized in the task. There are several tips that each student may apply to help them solve the assignment challenge. **Correct:** this feature represents whether the student answered correctly on the first attempt, used up all of the credit from too many hints/attempts, or gave a wrong response on the first try.
- Problem type:** it displays the problem that the learner tries to solve; different problem types require different types of concentration and a cognitive way to solve. This feature improves the model's excitability. Four problem types exist:
 - Choose: radio buttons for multiple choice.
 - Algebra: string evaluated using math (text box).
 - Fill in: basic string-compared response (text box).
 - Open response: keeps track of students' responses, but they are always marked as correct.

- e. Original problem: it indicates whether the problem is the main problem or a scaffolding problem.
- f. First action: exhibit the learner's intuition and cognitive level are determined by the first action. The options are to try an answer/attempt or ask for a hint.
- g. Type: this feature indicates the type of head section of the issue set. Every issue set is usually one of the following three:
 - Linear: each activity is finished by the learner in the specified sequence.
 - Random: every student finishes all assignments and problems, but each learner's issue order is different to other learners.
 - Mastery: this follows a random sequence where the student must demonstrate mastery of a problem by correctly answering a specific number of questions in a row before proceeding. Based on this categorization, we may ascertain if the learner is intellectually involved and focused to answer the issue in a preset or right order, or whether they attempt to solve the problem at random with no cognitive engagement or concentration level.
- h. Confusion score: the emotional and cognitive systems clearly function independently, yet they nonetheless influence each other. It is believed that confusion is both an emotional and an absolute cognitive state [21]. The learner's confusion score during problem-solving is directly correlated with the concentration level. A student with a low concentration level is often the one with a elevated confusion measure (confused learner). As a result, the confusion score is an extra element that is crucial for figuring out learners' focus.
- i. Boredom score: according to the cognitive theory of boredom, boredom arises when a person finds it difficult or impossible to interact intellectually with an item. According to the most well-known variation of cognitive theory, boredom is intimately linked to focus difficulties or to the misallocation of attention [22]. Based on this definition, boredom is a cognitive state linked directly to focus and concentration, especially in the learning process. Generally, learners with a high boredom score/level have a low concentration level. This proves that boredom score in the learning process is a decisive factor in determining the learner's concentration level.

3.2. Machine learning and predictive approach

It is important to keep in mind that data pre-processing, a simple but essential step that guarantees data reliability for better analysis, must be completed prior to data analysis. This stage, called as noise reduction, filters the data by eliminating unnecessary and duplicate elements. The act of collecting, creating, and presenting data in order to get valuable information, spot learning trends, and predict future behavior is known as data analysis. This is an essential part of the ML model. Building an accurate prediction system throughout the courseware's several learning periods necessitates processing and representing the data at each learning phase. For example, statistical measures are used to understand the data. Feature aggregations and correlations are used to minimize dimension, while ML techniques are used for prediction.

ML is a crucial component of AI. "The field of study that gives computers the ability to learn without being explicitly programmed" is the definition of ML [23]. It is an example-based learning system. Through observations and self-training, ML aims to enable a computer to do certain tasks in novel contexts. E-learning is one of the many facets of daily life where ML is becoming more and more prevalent. Reinforcement learning, unsupervised learning and supervised learning are the three primary categories of ML. While reinforcement learning produces a sequence of judgments over time to maximize performance, unsupervised learning assesses input X without guidance from output Y. In supervised learning, the output data Y is predicted using input data X.

Following the development of data vectors that show each learner's cognitive state over time. We want to carry out a supervised ML study to identify concentration level classes of learners in order to classify data according to qualities and features. More precisely, we want to group objects into "classes" according to the highest degree of compatibility within each category/class. In order to predict the output data Y, we construct a model based on the input vector X of the characteristics shown in the previous section based on the supervised learning technique. There are two classes of learners according concentration level: high concentration level and low concentration level. Since it is unlikely that the best classification model will be discovered beforehand, we first carried out a guided research to examine a number of well-known models in order to select the best one. We plan to analyze a range of ML models and select the best model based on the accuracy/precision of the results.

3.3. Determining the concentration level

After the concentration level of each student was extracted. The supervisory process can be facilitated by the provided ML classification model. The quality of choices and learning options may really be much enhanced by including the ML model into the e-learning platform. The suggested approach's design tackles the issue of learners' concentration levels and maintaining focus during learning process and their capacity to anticipate them earlier through online assignments. The concept of analyzing and forecasting a learner's concentration level might provide insight into the learner's concentration, and therefore their motivation involvement throughout the process.

4. EXPERIMENT

In order to verify the suggested methodology. We oversee the experiment in this section to provide specifics about our methodology and confirm its viability. The used data, the solution's experiment settings, and the results and discussion are the three subsections that make up this section. It is significant to note that the dataset (after processing) are accessible to the general public in https://github.com/Abdessamad139/Concentration_ML.

4.1. Data

We use the dataset from the ASSISTments data for the 2012–2013 academic year [24]. In order to align the dataset with our approach, we selected, cleaned, and aggregated pertinent attributes. We selected “addition and subtraction of fractions” as the learning topic in order to assess students' level of concentration on this particular concept in the online assessment. There are 146992 entries from various learners in the systems in the final dataset that was prepared. Two groups were labeled/assigned to the dataset: low concentration level (represented by 0) and high concentration level (represented by 1). Examples of items from the final dataset are shown in Table 1.

Table 1. Examples of entries from dataset

Id_student	Correct	Hint	Time	First response	Type	First action	Confusion score	Boredom score	Label
Id_8	Incorrect	3	158.07	143.55	Random	Hint	0	0.088	1
Id_438	Correct	0	89.27	89.27	Linear	Attempt	0	0.442	0

4.2. Experiment settings

Pandas and Numpy libraries were used in conjunction with the Python 3.10 programming language to carry out the experiment. After pre-processing, cleaning, and duplicate removal, 146849 unique entries remained for ML algorithm analysis. According to the 70:30 rule, the training and test sets are randomly divided into 102751 and 44098 entries, respectively. The open-source ML library Scikit-learn was utilized for the data manipulation and ML classification tasks [25]. The experiment was conducted in Google Colab, in T4-GPU runtime with 15 GigaBytes GPU and 12 Giga-Bytes RAM support setting <https://colab.research.google.com/>.

SVM [26], logistic regression (LR) [27], Naïve Bayesian (NB) [28], eXtreme gradient boosting (XGBoost) [29], RF [30], and NN [31] were the potential ML models that we compared. We used four different metrics to evaluate the efficacy of the possible ML models: precision, recall, accuracy, and F1-score. F1-score provides a balanced measure of a model's accuracy in identifying pertinent instances. It is the harmonic mean of recall and precision.

4.3. Results and discussions

The implementation's goal is to train the suggested model and assess its correctness and efficiency. The efficacy of several ML algorithms is compared in this section. Table 2 lists the best outcomes from the dataset that was examined. We see that the XGBoost algorithm performs the best out of all ML approaches with a precision of 80.87% and a F1-score 84.58%. That's a relative improvement of 0.11% over NN, 1.84% over NB and 1.03% over SVM. In classification-based problems, the classification result of 84.58% precision and f1-score over 80% is fairly significant. This illustrates how well our model detects learners' concentration levels and how well XGBoost perform in comparison to other ML.

Table 2. Examples of entries from dataset

	Precision	Accuracy	Recall	F1-score
SVM	0.7179	0.7177	0.9993	0.8355
LR	0.7436	0.7199	0.9305	0.8266
NB	0.7243	0.7112	0.9647	0.8274
XGBoost	0.8087	0.7680	0.8865	0.8458
RF	0.7962	0.7203	0.8201	0.8080
NN	0.7937	0.7619	0.9027	0.8447

The strong performance of XGBoost can be attributed to its nature as an ensemble learning method based on gradient boosting, which combines multiple weak learners (typically decision trees) to produce a more robust predictive model. Ensemble approaches such as XGBoost are particularly effective because they reduce variance and bias while capturing complex nonlinear relationships within the data. Interactions like high hint count+high confusion score or long time spent+incorrect response are very difficult for simple models to capture. As a result, XGBoost is better at capturing very complex relationships effectively.

Although the SVM achieved the highest recall (0.9993) among all models, its overall performance remains lower when considering other evaluation metrics such as precision (0.7179) and F1-score (0.8355). Recall measures the model's ability to correctly identify positive instances, meaning that the SVM successfully detects nearly all true positive cases (learners with high concentration level). However, high recall alone does not necessarily indicate superior performance. A model can achieve high recall while simultaneously producing a large number of false positives (non-concentrated learners predicted incorrectly). This is reflected in the lower precision value of the SVM model compared with XGBoost.

Previous studies have relied on image processing and EEG to measure learners' concentration, but these methods are hard to generalize across learning systems and often miss key factors. In contrast, our approach (to the best of our knowledge) is the first to integrate behavioral (e.g., time spent and first action), cognitive (problem type), and affective (confusion and boredom scores) factors to capture concentration during learning. We use tree-based models like XGBoost, which handle heterogeneous data without strict distributional assumptions and effectively capture feature interactions, improving predictive performance. For instance, identifying that a learner with high confusion but many hints may still succeed. This multi-factor, tree-based method provides a more comprehensive and generalizable assessment of learner concentration than classical ML approaches.

Overall, the results suggest that ensemble-based approaches such as XGBoost are more effective in capturing the underlying patterns of special dataset fields (such as education), leading to improved predictive performance compared to traditional ML methods.

When students are unable to maintain the necessary concentration level, educators or instructional designers who determine the learner's cognitive level prior to learning beginning can administer them and determine the best solution or intervention to help them overcome their learning challenges. By efficiently automating logical tasks involving logical decision-making that avoid the necessity of costly training, this approach aims to streamline the entire learning process. The ability of XGboost approach to automate judgments has been the focus of recent research in the field of cognitive analysis. This opens up a number of research topics that have already been covered in the literature, such as personalized learning object suggestions and categorical analysis of summative assessments and learning objects recommendation systems.

By leveraging behavioral, cognitive, and affective features, the proposed approach enables more accurate identification of students' engagement and focus. This can support the development of adaptive learning systems that provide timely interventions, personalized feedback, and improved learning experiences. The outcomes are really encouraging. Through an online learning evaluation, we can currently anticipate each learner's concentration level with a precision of 80.87%. This can assist the instructional designers and instructors in creating more effective learning materials for every type of student (adaptive learning). Additionally, students who struggle with attention could benefit from adaptive tutoring that matches their level of focus, such as shortening tasks, increasing interactivity, or switching up instructional modalities (video to text or vice versa). Conversely, students who have a high level of focus may also benefit from a recommendation system that presents them with challenging assignments or more complex tasks. Also, by placing students in homogeneous collaborative groups that correspond to their degree of focus, we may also improve their learning acquisition and focus. For tutors and the educational system, keeping students' attention and focus is both a problem and an incentive. Predicting students' level of focus can be crucial, since it is a cognitive state that directly affects epistemic emotion like confusion, boredom and frustration; consequently maximizing learners' attention span

can help reduce the number of students dropping out of online courses and improve their motivation through learning process.

4.4. Limitations

This study has several limitations. First, the dataset is limited to a specific e-learning context (online assessment), which may affect the generalizability of the results to other learning environments/context. Second, the selected features, although informative, may not capture all relevant aspects of learner behavior, such as prior knowledge or external factors. Finally, model performance may be influenced by data imbalance or noise inherent in educational datasets.

5. CONCLUSION

The paper aims to classify learners according to their degrees of concentration while they are interacting in an online learning assessment. Our approach intends to predict learners' concentration levels throughout the learning process, particularly when completing tasks such as answering questions in a certain order/logic, making attempts and requesting hints. Six ML techniques have been used to examine learners' focus in online learning assessment. The XGBoost classifier is found to provide the best precision, at about 80.87%. The suggested XGBoost-based model approach facilitates effective forecasting of students' levels of focus while completing learning assessment activities. To increase accuracy, this work could be expanded by incorporating a deep learning approach that utilises a broader range of more complex cognitive and affective features/variables. Also, future research will concentrate on incorporating this model into an adaptive learning systems that enable the system to provide more effective learning materials for every learner. This allows us to assess and analyse learners' cognitive level and compartmental behaviour while they are studying.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Chanaa Abdessamad	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			
Fatima-ezzahra Ait Bennacer		✓	✓			✓	✓	✓	✓	✓	✓			
Nour-eddine El Faddouli	✓			✓		✓			✓			✓	✓	

C : Conceptualization

M : Methodology

So : Software

Va : Validation

Fo : Formal Analysis

I : Investigation

R : Resources

D : Data Curation

O : Writing - Original Draft

E : Writing - Review & Editing

Vi : Visualization

Su : Supervision

P : Project Administration

Fu : Funding Acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data supporting the findings of this study are publicly available in the following GitHub repository: https://github.com/Abdessamad139/Concentration_ML.

REFERENCES

- [1] S. Kumar, "An innovative method to enhance interaction during lecture sessions," *Advances in Physiology Education*, vol. 27, no. 1, pp. 20–25, 2003, doi: 10.1152/advan.00043.2001.

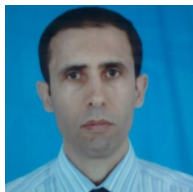
- [2] K. Watson, "Thinking, Fast and Slow. New York, NY: Farrar, Straus and Giroux. 499 pages," *Canadian Journal of Program Evaluation*, vol. 26, no. 2, pp. 111–113, 2011, doi: 10.3138/cjpe.26.010.
- [3] B. Bruya and Y.-Y. Tang, "Is attention really effort? revisiting daniel kahneman's influential 1973 book attention and effort," *Frontiers in Psychology*, vol. 9, 2018, doi: 10.3389/fpsyg.2018.01133.
- [4] P. Robinson, "Attention and Memory during SLA," *The handbook of second language acquisition*, pp. 631–678, 2003, doi: 10.1002/9780470756492.ch19.
- [5] T. Al'Omairi and H. Al Balushi, "The influence of paying attention in classroom on students' academic achievement in terms of their comprehension and recall ability," in *Proceedings of INTCESS15-2nd International Conference on Education and Social Sciences*, 2015, pp. 2–4.
- [6] C. Kalenzi, D. Back, and M. Yim, "The future of online education: Lessons from south korea," *World Economic Forum*, vol. 13, 2020.
- [7] S. Rizwan, C. K. Nee, and S. Garfan, "Identifying the Factors Affecting Student Academic Performance and Engagement Prediction in MOOC Using Deep Learning: A Systematic Literature Review," *IEEE Access*, vol. 13, pp. 18952–18982, 2025, doi: 10.1109/ACCESS.2025.3533915.
- [8] Y.-S. Jeong, M.-S. Joo, and N.-W. Cho, "A study on evaluation of e-learners' concentration by using machine learning," *Journal of Korea Society of Digital Industry and Information Management*, vol. 18, no. 4, pp. 67–75, 2022, doi: 10.1007/s11227-022-04804-w.
- [9] J. Li, G. Ngai, H. V. Leong, and S. C. F. Chan, "Multimodal human attention detection for reading from facial expression, eye gaze, and mouse dynamics," *ACM SIGAPP Applied Computing Review*, vol. 16, no. 3, p. 37–49, Nov. 2016, doi: 10.1145/3015297.3015301.
- [10] M. N. Hasnine, H. T. Bui, T. T. T. Tran, H. T. Nguyen, G. Akçapınar, and H. Ueda, "Students' emotion extraction and visualization for engagement detection in online learning," *Procedia Computer Science*, vol. 192, pp. 3423–3431, 2021, doi: 10.1016/j.procs.2021.09.115.
- [11] A. Chanaa and N.-e. El Faddouli, "A cognitive level evaluation method based on machine learning approach and Bloom of taxonomy for online assessments," *Journal of Education and Learning (EduLearn)*, vol. 18, no. 2, pp. 553–560, 2024, doi: 10.11591/edulearn.v18i2.20948.
- [12] S. Vijayakumar, K. R. Sai, and B. S. Reddy, "Concentration level of a learner using facial expressions on e-learning platform," *Advances in Data Science and Management*, 2022, pp. 575–584, doi: 10.1007/978-981-16-5685-9_56.
- [13] W. Villegas-Ch., J. García-Ortiz, I. Urbina-Camacho, and A. Mera-Navarrete, "Proposal for a system for the identification of the concentration of students who attend online educational models," *Computers*, vol. 12, no. 4, 2023, doi: 10.3390/computers12040074.
- [14] V. T. Lokare and L. D. Netak, "Concentration level prediction system for the students based on physiological measures using the eeg device," *Intelligent Human Computer Interaction*, 2021, pp. 24–33, doi: 10.1007/978-3-030-68449-5_3.
- [15] A. Al-Naffjan and M. Aldayel, "Predict students' attention in online learning using EEG data," *Sustainability*, vol. 14, no. 11, 2022, doi: 10.3390/su14116553.
- [16] R. Kawamura and K. Murase, "Concentration estimation in e-learning based on learner's facial reaction to teacher's action," in *IUI '20 Companion: Companion Proceedings of the 25th International Conference on Intelligent User Interfaces*, 2020, pp. 103–104, doi: 10.1145/3379336.3381487.
- [17] U. B. P. Shamika, W. A. C. Weerakoon, P. K. P. G. Panduwawala, and K. A. P. Dilanka, "Student concentration level monitoring system based on deep convolutional neural network," in *2021 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, Colombo, Sri Lanka, 2021, pp. 119–123, doi: 10.1109/SCSE53661.2021.9568328.
- [18] P. Sharma, M. Esengönül, S. R. Khanal, T. T. Khanal, V. Filipe, and M. J. C. S. Reis, "Student concentration evaluation index in an e-learning context using facial emotion analysis," *Technology and Innovation in Learning, Teaching and Education*, 2019, pp. 529–538, doi: 10.1007/978-3-030-20954-4_40.
- [19] P. Sharma *et al.*, "Student engagement detection using emotion analysis, eye tracking and head movement with machine learning," *Technology and Innovation in Learning, Teaching and Education*, 2022, pp. 52–68, doi: 10.1007/978-3-031-22918-3_5.
- [20] F. R. Chan, F. Annas, Y. E. Yuspita, and G. Darmawati, "Implementation of convolutional neural networks (CNN) in an emotion detection system for measuring learning concentration levels," *Knowbase: International Journal of Knowledge in Database*, vol. 4, no. 1, pp. 49–61, Aug. 2024, doi: 10.30983/knowbase.v4i1.8429.
- [21] G. L. Clore and J. R. Huntsinger, "How emotions inform judgment and regulate thought," *Trends in Cognitive Sciences*, vol. 11, no. 9, pp. 393–399, 2007, doi: 10.1016/j.tics.2007.08.005.
- [22] A. Elpidorou, "Boredom and cognitive engagement: A functional theory of boredom," *Review of Philosophy and Psychology*, vol. 14, no. 3, pp. 959–988, 2023, doi: 10.1007/s13164-021-00599-6.
- [23] A. L. Samuel, "Some Studies in Machine Learning Using the Game of Checkers. II—Recent Progress," in *IBM Journal of Research and Development*, vol. 11, no. 6, pp. 601–617, Nov. 1967, doi: 10.1147/rd.116.0601.
- [24] M. Feng, N. Heffernan, and K. Koedinger, "Addressing the assessment challenge with an online system that tutors as it assesses," *User modeling and user-adapted interaction*, vol. 19, no. 3, pp. 243–266, 2009, doi: 10.1007/s11257-009-9063-7.
- [25] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
- [26] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995, doi: 10.1007/BF00994018.
- [27] D. R. Cox, "The regression analysis of binary sequences," in *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 20, no. 2, pp. 215–232, 1958, doi: 10.1111/j.2517-6161.1958.tb00292.x.
- [28] D. D. Lewis, "Naive (bayes) at forty: The independence assumption in information retrieval," in *European Conference on Machine Learning*, 1998, pp. 4–15, doi: 10.1007/BFb0026666.
- [29] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 785–794, doi: 10.1145/2939672.2939785.
- [30] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.
- [31] J. E. Dayhoff, "Neural network architectures: an introduction, Van Nostrand Reinhold Co., 1990.

BIOGRAPHIES OF AUTHORS

Abdessamad Chanaa    is Ph.D. on recommendation systems at Mohammadia School of Engineers (EMI), Mohammed V University in Rabat, Morocco. He is also a computer science engineer from the National School of Applied Science of Marrakech. His interests lie in the fields of deep learning, neural network, and educational engineering. He can be contacted at email: abdessamad.chanaa@gmail.com.



Fatima-ezzahra Ait Bennacer    holds a Ph.D. in Computer Sciences from the University of Chouaib Doukkali, Faculty of Sciences of El Jadida, Morocco. Her research focuses on digital transformation in higher education institutions, with particular emphasis on quality management systems, knowledge management and artificial intelligence. Her interests also include the design and implementation of digital platforms that enhance institutional performance and support innovation within research ecosystems. She can be contacted at email: aitbennacer.f@ucd.ac.ma.



Nour-eddine El Faddouli    is full professor at Mohammadia School of Engineers (EMI), Mohammed V University in Rabat, Morocco. He is a professor of Computer Science in Mohammadia School of Engineers since 1996. In addition, he is director of many thesis projects and supervisor of final year graduate projects in computer science and big data fields. His research interests are: e-learning, big data, pedagogy and education, assessment, higher education, computational linguistics, teaching, and learning. He can be contacted at email: elfaddouli@emi.ac.ma.