education sciences

MDPI

Article

Competency Learning by Machine Learning-Based Data Analysis with Electroencephalography Signals

Javier M. Antelis, Myriam Alanis-Espinosa, Omar Mendoza-Montoya, Pedro Cervantes-Lozano and Luis G. Hernandez-Rojas *

Tecnologico de Monterrey, School of Engineering and Sciences, Monterrey 64700, Nuevo León, Mexico; mauricio.antelis@tec.mx (J.M.A.); myriam.alanis@tec.mx (M.A.-E.); omendoza83@tec.mx (O.M.-M.); pedro.cervantes@tec.mx (P.C.-L.)

* Correspondence: luisg.hernandez@tec.mx

Abstract

Data analysis and machine learning have become essential cross-disciplinary skills for engineering students and professionals. Traditionally, these topics are taught through lectures or online courses using pre-existing datasets, which limits the opportunity to engage with the full cycle of data analysis and machine learning, including data collection, preparation, and contextualization of the application field. To address this, we designed and implemented a learning activity that involves students in every step of the learning process. This activity includes multiple stages where students conduct experiments to record their own electroencephalographic (EEG) signals and use these signals to learn data analysis and machine learning techniques. The purpose is to actively involve students, making them active participants in their learning process. This activity was implemented in six courses across four engineering careers during the 2023 and 2024 academic years. To validate its effectiveness, we measured improvements in grades and self-reported motivation using the MUSIC model inventory. The results indicate a positive development of competencies and high levels of motivation and appreciation among students for the concepts of data analysis and machine learning.

Keywords: learning activity; electroencephalography; experiments; data collection; learning gain; student motivation

check for updates

Academic Editor: Eleanor Dommett

Received: 31 May 2025
Revised: 18 July 2025
Accepted: 22 July 2025
Published: 25 July 2025

Citation: Antelis, J. M., Alanis-Espinosa, M., Mendoza-Montoya, O., Cervantes-Lozano, P., & Hernandez-Rojas, L. G. (2025). Competency Learning by Machine Learning-Based Data Analysis with Electroencephalography Signals. *Education Sciences*, *15*(8), 957. https://doi.org/10.3390/educsci15080957

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Data analysis and artificial intelligence methods are increasingly indispensable skills across many types of industries, as they are essential for boosting the development of new services and products, fueling decision-making, driving innovation, and steering strategic planning (Iansiti & Lakhani, 2020; Jagatheesaperumal et al., 2022). Therefore, higher education institutions are challenged to develop innovative educational solutions that foster these cross-disciplinary competencies in their students, preparing professionals to proficiently and confidently navigate an emerging future with unknown challenges in their professional lives (Kuleto et al., 2021; Luan et al., 2020). Traditionally, data analysis and machine learning competencies are taught to freshman, sophomore, junior, and senior college students depending on the career and the target level of specialization. For instance, descriptive statistics and linear regression topics are taught to freshmen and sophomore students in all engineering careers (Balady & Taylor, 2024; Legaki et al., 2020), while more advanced topics such as digital signal processing and classification models are taught to

junior and senior students in computing and robotics careers (Anna Åkerfeldt & Petersen, 2024; Terboven et al., 2020). In general, prerequisites are first- and second-year mathematics and statistics, along with a programming language. For more specialized courses, additional prerequisites include linear algebra, differential equations and numerical analysis. Regardless of the academic year, data analysis and machine learning courses primarily focus on theoretical concepts accompanied by practical activities using a programming language and previously recorded and pre-processed datasets (Bin Tan & Cutumisu, 2024).

This conventional approach can hinder the development of practical problem-solving abilities and reduce intrinsic motivation, as students may not fully grasp the relevance or application of the concepts which often limit students' engagement with the complete data lifecycle, from collection and preparation to the contextualization of real-world applications. In addition, students are also expected to develop skills to design and execute experiments in unexpected application fields, as required in their professional lives, often without having had the opportunity to do so in their courses. Developing the competency to design and implement real-life solutions based on courses with solid theoretical components, and crucially, using data collected and prepared by others rather than by the students themselves, is a highly challenging pedagogical task.

The development of competencies in engineering students has been investigated in various domains, such as in a multidisciplinary project-based learning method for a mechanism course (Guajardo-Cuéllar et al., 2022), in a classical product development project versus a challenge-based project to enhance innovation skills (Guido Charosky & Bragós, 2022), in the fabrication of a boomerang device using analytical and computational tools (Guajardo-Cuéllar et al., 2020), and with mixed reality for control engineering (Navarro-Gutiérrez et al., 2023). Several studies have also proposed learning activities and reported best practices to develop data analysis and machine learning skills in university students (Padilla & Campos, 2023; Tsai, 2024). A method to integrate data analysis and machine learning in a learning management system was also proposed to improve student learning in an online education model (Villegas-Ch et al., 2020). Other efforts advocate for scholars teaching data science-related topics to possess, in addition to theoretical expertise, experience in solving real-world problems, collecting data, and considering ethical implications (DeMasi et al., 2020; Hicks & Irizarry, 2018). In addition, contemporary educational strategies aim to enhance student engagement and motivation. For instance, participatory teaching methods in hands-on courses have been shown to positively impact learning motivation by allowing students to actively participate in content planning and evaluation (Ma, 2023). Similarly, gamification has been explored as an approach to increase motivation and reduce cognitive load, thereby improving learning outcomes in various settings (Baah et al., 2024). Furthermore, the application of machine learning models to predict student success in online courses also underscores the increasing integration of ML in education itself, validating the relevance of these competencies within pedagogical research (Arévalo-Cordovilla & Peña, 2024).

Despite significant efforts by educators and the widespread availability of online courses and tools for teaching data analysis and machine learning models (Alam & Mohanty, 2023; Ismail et al., 2023; Thomas Donoghue & Ellis, 2021), these common teaching methodologies often do not incorporate domain-specific contextual information relevant to the application area. Also, these approaches contribute significantly to enhancing educational experiences; many still rely on structured or pre-processed datasets, limiting the authentic experience of managing raw, noisy data from real-world experiments. This drastically limits students' ability to understand and leverage the relationship between theory and practice, which is essential for developing the competencies needed to address real-world situations demanded by the labor market. Consequently, the challenge lies in

fostering an educational environment that bridges the gap between theoretical knowledge and practical, hands-on experience, ensuring that students are active participants in their learning journey. Research indicates that active engagement significantly enhances learning outcomes and performance, highlighting the need for innovative pedagogical strategies that promote deep interaction with the subject matter (Oliver & McNeil, 2021; Savonen et al., 2024). Hence, to promote the development of knowledge, skills and attitudes related to data analysis and machine learning, we designed and implemented a learning activity that requires students' active participation in the entire process of data recording, processing, analysis, and application of machine learning methods. Motivated by this, we propose a learning activity that aims to develop data analysis and machine learning concepts, skills, and attitudes in engineering students from diverse careers through hands-on experiences. The goal is to actively involve students by making them active participants in the learning process, using their own data.

TECNOLOGICO DE MONTERREY, a top-ranked higher education and research institution in Mexico, offers a new teaching model based on competency development, which fosters learning and knowledge discovery through solutions to real challenges proposed by external and independent educational partners. This educational model called TEC21 (ITESM, 2018; Olivares Olivares et al., 2021), allows the development of skills, attitudes and values to address real problems through the creation of a natural link between theory and practical situations. This model aligns with current trends of higher education institutions that promote moving from the objective-based learning model towards a competency-based learning model (Gooding, 2023; Thompson & Harrison, 2023). This new teaching trend aims for students to develop skills and attitudes focused on understanding and bridging the gap between theory and practice. This is precisely what needs to be encouraged in the teaching of data analysis and machine learning for engineering students.

Aligned with our educational model, this paper presents the implementation of a practical learning activity consisting of the collection and visualization of students' electroencephalographic (EEG) brain signals in real settings, the application of signal processing techniques, and the use of machine learning methods to decode the gathered data. The use of non-invasive brain signals was motivated by the strong appeal and interest it generates in students regardless of their career, its user-friendliness for conducting experiments, and its suitability for exploring and understanding data analysis and machine learning concepts and tools.The activity was implemented in six courses across four engineering careers over two consecutive academic years. The learning gain was used to measure the grade change brought about by the activity, and the MUSIC model inventory was used to quantify the students' perceptions of empowerment, usefulness, success, interest, and caring generated by the activity. In addition, contextual examinations were applied to spark interest and to measure students' pre- and post-activity experience related to the learning activity's topics. This practical activity fostered learning by doing, which in turn enhanced the development of data analysis and machine learning competencies in the courses where it was applied, and increased students' motivation and appreciation for these topics.

Our work distinguishes itself by proposing and implementing a novel learning activity that immerses engineering students in the entire cycle of data analysis and machine learning by having them collect their own electroencephalographic (EEG) signals through hands-on experiments. This distinctive element ensures that students are not merely processing data but are intimately involved in its generation and contextualization, providing an understanding of data characteristics and their implications for analysis. Unlike studies focusing solely on pre-existing data or simulated environments, our methodology directly addresses the critical gap in developing the competency to design and execute experiments in unexpected application fields, a skill highly valued in professional life.

2. Materials and Methods

2.1. Learning Activity

The learning activity consists of the following four stages to be described in the subsequent subsections.

2.1.1. Stage 1: EEG Acquisition and Visualization Experiment

In this stage students are first instructed on how to use the hardware and software components for acquisition and visualization of electroencephalographic (EEG) brain signals. Students are then organized into groups and instructed to connect all hardware components and make all necessary software configurations to acquire and visualize EEG signals on their own. For this, one of the students adopts the role of experimental subject (from whom the EEG signals will be recorded) while the rest adopt the role of experimenters (responsible for executing the experiment).

With the signals displayed in real time, all students are instructed to recognize noise-free EEG signals in a relaxed state with open and closed eyes, as well as EEG signals contaminated by artifacts such as blinking, eye movement, jaw clenching and body movement. This practical activity is used to introduce and describe technical concepts such as continuous-time and discrete-time signals, analog-to-digital conversion, sampling frequency, Nyquist frequency, frequency band, and digital filtering, among others.

2.1.2. Stage 2: Evoked Potentials Experiment

The purpose of this stage is to carry out EEG experiments to develop data recording skills, and to obtain their own dataset of EEG signals that will be used in the next stages for data analysis and machine learning purposes. Once again, students are organized into teams where one member (different from the one in the first stage) is assigned as the subject and the rest are assigned as experimenters. Experimenters first configure the hardware and software to acquire and display the EEG signals (as previously carried out in stage 1) from the subject. The experimenters then start a graphical user interface (GUI) that will guide the subject through the tasks to be performed while the EEG signals are recorded. The GUI presents a classic P300 evoked potential experimental protocol (see Chailloux Peguero et al. (2020); Picton (1992) for a detailed description of the P300 paradigm). In this protocol, the subject is seated in front of a computer screen wearing the EEG system and is instructed to follow the instructions presented in the GUI. The GUI displays five symbols: four arrows located at the left, right, top, and bottom of the screen and a "STOP" traffic sign located in the center.

The students then initiate the experiment, which is executed as follows. One of the symbols is randomly highlighted with a blue background for a couple of seconds, indicating the target symbol that the subject has to focus on. The symbols are then randomly highlighted, one by one, by superimposing a yellow smiling cartoon face on them. The subject is requested to mentally count each time the target symbol is highlighted while ignoring instances when the other non-target symbols are being highlighted. This experiment lasts almost 5 min and at the end, the students will obtain a data file with the raw EEG data along with a marker signal that indicates the exact time instants the target and non-target symbols were highlighted.

2.1.3. Stage 3: EEG Data Preparation and Processing

In this stage, students are instructed to work with the previously collected data using software for data analysis such as Python 3.10or Matlab R2022b. Students can work individually or in pairs. The first task is to upload and plot the raw data using code they develop, while being guided by the professor. They inspect the EEG signals and the marker

signal that indicates the time instant of the presentation of the target and non-target symbols. This visual analysis is used to relate the experiment with the recorded data. They then proceed to build a code to extract segments of EEG signals with respect to the presentation of the happy face superimposed in the symbols. These data segments are extracted with a duration of one second and are labeled as target or non-target depending on whether the subject was or was not attending to, as encoded in the marker signal. Note that this task promotes the acquisition of skills and technical details about data preparation and manipulation, which are not typically performed when data is pre-recorded and prepared by someone else.

In the second task, students are instructed to apply data analysis techniques such as descriptive statistical analyses, and signal processing techniques such as band-pass filtering, resampling and artifact-contaminated segment identification and rejection. Technical details about these data analysis and signal processing methods are discovered by and/or suggested to the students before they are asked to implement and apply them. During the application of these techniques, students are always asked to identify and understand their effects by comparing the signals before and after the application of the method. Finally, students are asked to apply the ensemble averaging technique to compute the evoked responses (Chailloux Peguero et al., 2020; Delijorge et al., 2020). This technique is separately applied for the segments of target and non-target stimulation, independently for each EEG channel.

2.1.4. Stage 4: Machine Learning Model

This final stage aims to develop students' knowledge and skills in feature extraction techniques, classification models, training and testing of the classification models, as well as methodologies and metrics for the assessment of machine learning models. As described in stage 3, students can work individually or in pairs using software for data analysis. They use the pre-processed and cleaned EEG data segments previously extracted for the target and non-target conditions.

The first task here is to build their own code to extract (compute) features or attributes. Since the experiment is based on evoked potentials (as described in stage 2), this task focuses on extracting temporal features (Torres-García et al., 2022). Students are instructed and guided to perform downsampling and to append the data across all EEG channels. This is repeated for all data segments and as a result, they construct and organize a feature matrix with columns representing the features. They are also guided to build a class label vector indicating whether each feature vector belongs to a non-target condition (defined as class 0) or a target condition (defined as class 1). Hence, students learn by doing via various important concepts such as feature vector, dimension, class label, feature selection, and feature transformation, among others. Notice that all of these machine learning concepts are derived from and related to the experiment, and therefore, students learn to directly associate these concepts with real-life applications.

In the next task, students learn to use a classification algorithm to distinguish between target and non-target conditions using a feature vector extracted from an EEG data segment as input. They are instructed to split the feature matrix and its associated class label vector into two parts, one for training purposes and the other for testing purposes. They are requested to explore and select a classification model suitable for discriminating between target and non-target situations. Technical details of the classifiers are explored and discovered. Through coding, they initialize the selected classification model and fit its parameters using the training data (this constitutes the model's training). The testing data is fed into the tuned classifier to generate predictions. Finally, students discover performance metrics to assess the effectiveness of the classification model in recognizing target and

non-target EEG data segments. They are also encouraged to propose graphical ways to present results. Again, this hands-on learning procedure allows students to discover more machine learning concepts. As a closing remark, a discussion about how to deploy the machine learning model is held.

Figure 1 depicts the description of the four stages of the learning activity. Notice that all four stages involve hands-on practical tasks performed by the students, exploration and discovery of technical concepts, and computational and programming tasks, guided by the professor. Moreover, the four stages are time-flexible and can be adapted for implementation in one or more teaching sessions depending on the career, course, competencies, learning goals, thematic content, or any other relevant aspects. For example, more emphasis can be placed on stage 4 (machine learning model) for data science students, or on stage 2 (evoked potential experiment) for biomedical students. In other words, the number of sessions and the duration of the learning activity might vary depending on the particular characteristics and needs specific to the career and the course topics.

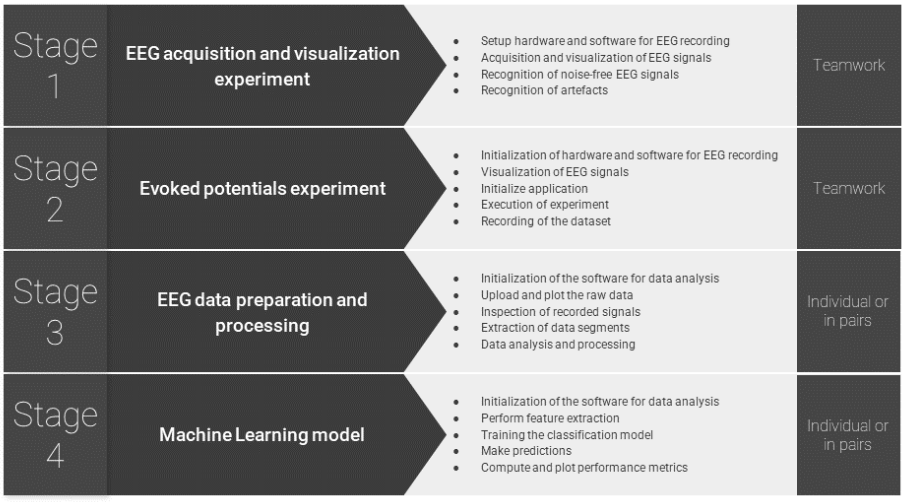


Figure 1. Illustration of the four stages of the proposed learning activity designed and implemented to develop data analysis and machine learning knowledge, skills, and attitudes in engineering students.

2.2. Validation Instruments

To assess the impact of the learning activity, two instruments were employed: (1) the learning gain (LG), measured by comparing grades before and after the activity; and (2) the dimensions of the MUSIC model of academic motivation inventory. Details regarding the application of these two instruments are described below.

2.2.1. Learning Gain (LG) Based on Academic Performance

The first validation instrument aimed to quantify the learning gain (LG), defined as the change in academic performance, i.e., grades of the students posterior to versus prior to the implementation of the learning activity. To accomplish this, an examination was administered to the students to assess their performance before (*Pre*) and after (*Post*) the initiation of the stages. This validation instrument, based on the academic performance, has been successfully applied in previous studies to assess improvements achieved through various learning methods and tools (Hartikainen et al., 2019; Vermunt et al., 2018).

In our work, the inquiries in the *Pre* and *Post* examinations consisted of several closed-ended questions directly related to the topics of the course and the learning activity. The professor responsible for a course proposed the examination questions, while other professors reviewed them and offered suggestions for improvement. Thus, examinations were independently designed and administered for each course. This was performed

to consider the specific concepts and terminology of each course, in spite of all courses addressing similar data analysis and machine learning topics, and to develop competencies with similar characteristics. Note that the academic performance or grade achieved by the student (either *Pre* or *Post*), represents a generic measurement across all courses and careers, thus enabling overall evaluations and fair comparisons.

The number of questions ranged from sixteen to twenty; nonetheless, seven questions were the same across all the courses where the learning activity was implemented. These seven common questions, with their corresponding response options and their correct responses (highlighted in bold and italic), were as follows:

Q1: Electroencephalography is a method of recording changes in blood flow in the brain.

R1: True/**False**

Q2: The sampling frequency of a signal is the number of samples of the signal.

R2: True/**False**

Q3: What sampling frequency of EEG signals is sufficient to be able to analyze delta, alpha, beta and gamma rhythms?

R3: 16 Hz/30 Hz/40 Hz/80 Hz/**256 Hz**

Q4: A typical amplitude range of EEG signals is:

R4: −100 to 100 mV/**−100 to 100 μV**/−1 to 1 μV/−1 to 1 A −1 to 1 μA

Q5: It is not source of artifacts in EEG signals:

R5: Blinking/Head movement/Chew/**Cognitive load**/Bad electrode placement

Q6: Where could the reference and ground electrodes be placed for the acquisition of EEG signals?

R6: Heart and elbow/Earlobe and neck/Wrist and right foot/**Head and earlobe**/Head and neck

Q7: If a signal was recorded at a sample rate of 1200 Hz and there are 4s of that signal, how many samples does the signal have in total?

R7: 300/600/1200/2400/**4800**

Note that these questions were translated here to English but they were administered in Spanish since all students were native Spanish speakers. The examinations were provided through an online form. We hypothesized that the scores would be low in the *Pre* examination and they would significantly increase in the *Post* examination as a consequence of the learning activity the students were involved in.

2.2.2. Motivation Based on the MUSIC Model Inventory (MMI)

As a second validation instrument, we assessed students' perceptions of the motivational climate through the college student version of the MUSIC Model of Academic Motivation Inventory (*MMI*) (Jones, 2009, 2012, 2018, 2019). This self-report instrument consists of 26 items that are easily completed on a six-level scale ranging from strongly disagree to strongly agree (the response options for each item are 1 = Strongly Disagree; 2 = Disagree; 3 = Somewhat Disagree; 4 = Somewhat Agree; 5 = Agree; 6 = Strongly Agree. These items are used to compute five scores that assess students' motivational perceptions in five dimensions: empowerment, usefulness, success, interest, and caring. These dimensions measure respectively the degree to which a student believes or perceives the following: (i) She/he is in control of the learning environment; (ii) The coursework or the learning activities are useful to her/his future; (iii) She/he can succeed at the coursework or the learning activities; (iv) The methods and tasks in the coursework or the learning

activities are interesting and enjoyable; (v) The instructor cares about the student's success and well-being during the coursework and learning activities.

The *MMI* has been extensively used to assess students' perceptions about the motivational climate after coursework or learning activities in different learning environments and situations (Amato-Henderson & Sticklen, 2022), whose validations have been well established in health science courses (Jones et al., 2023; Tehmina Gladman & Ali, 2020), English and other second-language courses, and student pharmacists, among others (Jones, 2020; Jones et al., 2023; Pace et al., 2016). In the present work, all students that decided to be part of the research by signing the informed consent form and that responded to the *Pre* and *Post* examinations were also requested to respond to the *MMI* items provided through an online form.

2.3. Participants and Courses

Students from several courses and undergraduate programs at the School of Engineering and Sciences (EIC) at TECNOLOGICO DE MONTERREY participated in this study. All students enrolled in the courses were informed about the goals of the research in a session prior to the implementation of the learning activity; also, they were informed that their participation was completely voluntary. No academic grade or penalty was assigned for accepting or declining to participate in the research. Students accepted participation by signing a written informed consent form that detailed the goals and procedures for the corresponding learning activity.

Table 1 presents the information where the learning activity was implemented, including academic year, career, course ID, semester in which the course was taught, number of students enrolled in the course, and the number of participants who signed the consent form and completed all procedures of the learning activity. Note that the course ID encodes the name of the course itself as follows: BI2010B for design and development in neuroengineering; MA3001B for development of mathematical engineering projects; TC3002B for development of advanced data science applications; and TE3003B for integration of robotics and intelligent systems.

Table 1. Summary of the academic year, career, course, the semester in which the course was taught, the number of students enrolled, and the number of students that were finally part of the study. List of involved careers: IMD: Biomedical Engineering; IDM: Data Science and Mathematics Engineering; ITC: Computer Technology Engineering; IRS: Robotics and Digital Systems Engineering. (*) For the BI2010B course in 2023, the data from the *Post* examination was inadvertently lost, which prevented us from computing the learning gain (*LG*) metric for that specific cohort.

Academic Year	Career	Course ID	Semester	Enrolled Students	Participant Students
2023	IMD	BI2010B	6	59	0 *
	IDM	MA3001B	8	16	16
	ITC	TC3002B	8	60	29
2024	IMD	BI2010B	6	47	46
	ITC	TC3002B	8	91	76
	IRS	TE3003B	8	19	18
Total				292	185

Despite being different courses and careers, all courses have similar elements and characteristics in the topics of data analysis and machine learning. The list below shows the topics and/or competencies most closely associated with data analysis and machine learning within each course. That is, other topics and/or competencies are also declared in

the course; however, we do not show them because they might not be directly related to data analysis and machine learning.

- BI2010B: Design and development in neuroengineering.
  - Competency: Recording and analysis of neural signals.
- MA3001B: Development of mathematical engineering projects.
  - Competency: Data analysis of natural phenomena.
- TC3002B: Development of advanced data science applications.
  - Competency: Implementation of computational algorithms.
- TE3003B: Integration of robotics and intelligent systems.
  - Competency: Signal processing and data analysis.

2.4. Experimental Procedure

2.4.1. Description of the Implementations

The implementations were carried out in the 2023 and 2024 academic years. The number of sessions and the duration of each stage of the learning activity (see Figure 1) were decided by the professor responsible for the course. However, the same course in the same or in different academic years consisted of the same sessions and duration (see Table 1). A description of the goals, the research procedures, and an invitation to participate were presented to the students in the first session of the learning activity. Students who accepted to be part of the study then signed the written informed consent form. Then, all students were invited to answer several contextual questions as described in Section 2.4.2 along with an examination to assess their knowledge (*Pre* examination). They were informed to respond individually and honestly. Subsequently, the four stages of the learning activity were implemented following the descriptions and procedures presented in Section 2.1. Immediately after all the activities were completed, the corresponding examination to assess their knowledge (*Post* examination) and questions of the MUSIC survey were administered.

2.4.2. Context Questions

To examine students' perceived baseline level in data recording in general, and in EEG data recording specifically, contextual questions were asked both before (*Pre*) and after (*Post*) the implementation of the learning activity. These questions focused entirely on their previous experiences with data recording and were designed to increase students' interest and engagement in the learning activity. Consequently, three questions were created and administered to all students across all courses and careers where the learning activity was implemented.

The questions with their corresponding response options are as follows:

C1: Have you previously carried out experiments to obtain your own data with which you studied and applied data analysis and machine learning techniques?

R1: Yes/No

C2: From 0 to 5, where 0 is "not important" and 5 is "very important", how essential do you consider the data recording process to be for data analysis and machine learning?

R2: 0/1/2/3/4/5

C3: Have you participated and/or performed experiments to acquire and record electroencephalogram (EEG) signals?

R3: Yes/No

Note that there are no correct or incorrect responses, and students were informed about this prior to the examination. These contextual questions were administered via an online form along with the assessment questions.

2.5. Data Analysis and Statistical Tests

We only considered data from students who signed the informed consent form and fully completed the *Pre* and *Post* examinations, contextual questions and *MMI* items. That is, data from participants who did not complete all these elements were discarded and not used in the remainder of this study. Student IDs were only used to match the *Pre* and *Post* responses with the *MMI* scores. Right after this, student ID information was discarded to maintain full anonymity of the data. The data from all participants were merged into a resulting dataset to be subsequently subjected to the following analyses.

For the context questions, we computed the proportion of responses (for instance, the proportion of Yes and No, in the case of questions with this type of response) before and after the learning activity. Note that there are no correct or incorrect responses, and therefore these results are used to establish how students' personal experiences evolved as a result of the learning activity. For the examination questions, the percentage of correct and incorrect responses was computed separately for each question in the *Pre* and *Post* examinations. The improvement percentage for each question was simply computed as the difference in the percentage of correct responses between *Post* and *Pre*. The total grade for each student, before and after the learning activity, was computed to obtain the distribution of grades in the *Pre* and *Post* examinations. These analyses are essential to quantify the learning gain and to evaluate their statistical significance. In the case of the MUSIC inventory, the responses to the 26 items were used to compute the scores of the five dimensions (empowerment, usefulness, success, interest, and caring) that assess students' motivational perceptions.

To examine significant differences between the distributions of grades before and after the learning activity, given their ordinal nature, the Wilcoxon rank-sum test was employed. All statistical tests were performed at a confidence level of $\alpha = 0.01$.

2.6. Learning Activity Deployment

2.6.1. Hardware

All activities involving the acquisition of the EEG signals in our implementations utilized the Unicorn Hybrid Black EEG system (from the manufacturer g.tec medical engineering GmbH, Schiedlberg, Austria). We decided to select this system because of its wireless capabilities, wearable properties, easy-to-use aspects, and significantly driver-oriented capability to provide high-quality signals (Pontifex & Coffman, 2023).

Note however that any other EEG recording system can be used. The Unicorn Hybrid Black system is a 24-bit amplifier that digitizes brain signals at a sampling rate of 250 Hz from the eight scalp locations FZ, C3, CZ, C4, PZ, PO7, OZ, and PO8 according to the international EEG 10-20 system. Reference and ground electrodes are fixed on the left and right mastoids of the participants using disposable electrodes. In the implementation of the learning activities, each work team was provided with one of these systems to carry out their own practical activities as described below. The use of the EEG recording system was learned by the students in stage 1 of the learning activity.

2.6.2. Software

To guide the students during the recording and visualization of the EEG signals, we used an in-house software implemented in C++, which included the graphical user interface (GUI) for the P300 experiments. This software has been previously used in our lab in several

brain–computer interface (BCI) experiments for movement recovery and rehabilitation based on P300 and motor imagery (Hernandez-Rojas et al., 2022; Peguero et al., 2023).

Importantly, the software included the recording of the marker signal, in addition to the raw EEG signals, which indicates the exact time of stimulus presentation for target and non-target symbols. This marker signal is essential for the practical activities of the learning activity because it is used by the students to obtain the epochs of EEG signals. The use of software and the GUI is learned by the students in stages 1 and 2 of the learning activity.

2.6.3. EEG Preprocessing

The technical details of the EEG data preparation and processing methods that students learn and apply in stage 3 are as follows. The initial step in processing the EEG signals involves filtering to remove low- and high-frequency artifacts, which are typically caused by eye movements, muscle activity, or environmental noise. Specifically, a 6th-order digital Butterworth band-pass filter with cutoff frequencies from 1 Hz to 20 Hz is proposed to the students. This frequency band has been widely used in the literature for the analysis of event-related potentials (ERPs), particularly the P300 component, as it retains the low-frequency content critical for the detection of this potential while effectively attenuating slower drifts and high-frequency noise such as electromyographic activity (Chailloux Peguero et al., 2020; Delijorge et al., 2020).

Although some studies suggest using lower cutoff frequencies (e.g., 0.5 Hz), we found that using a 1 Hz lower bound preserves sufficient information for reliable P300 detection, as confirmed by empirical results in prior implementations of this activity. Students were also encouraged to experiment with alternative filter settings to assess the impact on signal quality and classification performance. Additionally, a Notch filter centered at 60 Hz may be applied to suppress power line interference. However, students were informed that the 60 Hz component lies outside the band of interest defined by the band-pass filter, and thus, its removal has a negligible impact on the ERP signal or the outcomes of the P300 detection process. In addition to filtering, another important pre-processing aspect is EEG referencing. In this activity, the EEG signals were originally referenced to mastoid electrodes, which is a standard and widely accepted reference in ERP research, particularly for visual paradigms involving the P300 component. Given the limited number of electrodes (8 channels) and their focused spatial distribution, we did not apply common average referencing (CAR), as it could introduce distortions in ERP morphology rather than improve signal quality. However, more advanced students were encouraged to explore the impact of alternative referencing schemes, such as CAR, as part of the open-ended exploration suggested in stage 3.

2.6.4. Epoching and Rejection of Noisy Windows

After the band-pass filtering step, the EEG signals are segmented into epochs associated with the target and non-target conditions. This segmentation is performed based on the time markers provided by the user interface of the application used to implement the experimental protocol. Each epoch is time-locked to the stimulus onset and spans a window in which the evoked P300 response is expected to occur. All resulting epochs are then analyzed to assess signal quality and detect the presence of artifacts. To determine whether an epoch is contaminated by noise, students are instructed to compute, for each electrode, two metrics: (i) the peak-to-peak voltage, defined as $v_i = max(x_i) - min(x_i)$, and (ii) the standard deviation as $\sigma_i = \sqrt{\frac{1}{N_s-1}\sum_{t=1}^{N_s}(x_i - \mu_i)^2}$, where x_i is the band-pass-filtered signal of the i-th electrode, $i = 1, 2, \cdots, 8$, N_s is the total number of samples in x_i, and μ_i is the average of x_i. Therefore, an epoch is considered as noisy if in at least one electrode the

following conditions are fulfilled: first, the peak-to-peak voltage is greater than 100 μV, and second, the standard deviation is greater than 50 μV.

2.6.5. ERP Computation and Classification

After selecting the artifact-free epochs, the EEG signals for the target and non-target conditions are averaged separately to obtain the corresponding event-related potentials (ERPs). Students are encouraged to inspect the averaged waveforms for each channel to identify the presence of an evoked response, particularly the P300 component, which is typically visible in the target condition approximately 300 milliseconds after stimulus onset. In addition to ERP averaging, students implement a linear Support Vector Machine (SVM) classifier using features extracted from the clean epochs. The model is evaluated using k-fold cross-validation (k = 5) to estimate its ability to distinguish between target and non-target. Through this process, students examine whether the classifier achieves performance above chance level, linking signal quality and ERP presence with classification success.

All these EEG signal processing and analysis methods are implemented in Python 3.10 or Matlab R2022b depending on the analysis software required to be used in each career and course where the learning activity was implemented. Note however that, as indicated in the description of stage 3 of the learning activity, students can also discover and apply other methods for EEG signal analysis.

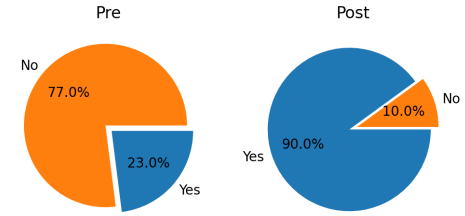
3. Results

3.1. Implementations

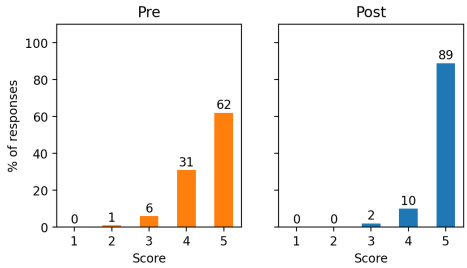
The implementation summary, including profiles of the 185 students who completed all participation prerequisites (consent forms, *Pre* and *Post* examinations, and the *MMI* survey), is presented in Table 1. The results presented below contemplate these participants' profiles.

3.2. Contextualization

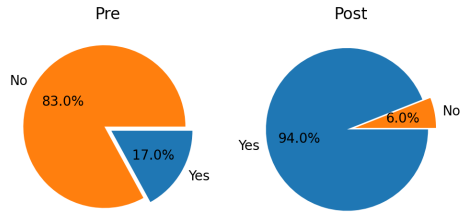
Figure 2 shows the results for all participants for the three context questions in the *Pre* (left panels) and *Post* (right panels) examinations. First, 23% of the students had carried out experiments and activities to collect (any) data before the application of the learning activity and this percentage increased to 90% after it (question C1). This suggests that almost all students had their first practical experience in data collection through this learning activity. It is unclear why a small proportion of students (10%) reported no such experience even after the learning activity, as shown in Figure 2a (right pie chart). Second, regarding the appreciation of the relevance of the data collection process for data analysis and machine learning (question C2), 62% of the students considered it as very important (score of 5) before the learning activity, which impressively increases to 89% after it. On the other hand, 7% of the students first ranked the data recording process with low importance (scores of 1, 2 and 3), which dropped to 2% in the *Post* examination. Finally, the previous experiences in EEG experiments (question C3) increased from 17% to 94% in the *Pre* and *Post*, respectively. The small proportion of students with EEG experience before the application of the learning activity is attributed to those who reported participation in EEG-based brain–computer interface (BCI) research experiments. Similarly, the reasons for the 6% of participants who reported no EEG experience after the learning activity remain unclear (as shown in Figure 2b, right pie chart). These results reveal a generally positive perception of and personal connection to the importance of data collection, which was further enhanced after the learning activity.



(**a**) Context question C1.



(**b**) Context question C2.



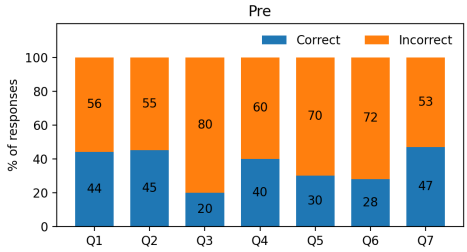
(**c**) Context question C3.

Figure 2. Across-all-participants results of the three context questions (C1, C2 and C3) in the *Pre* (**left** panels) and *Post* (**right** panels) examinations.

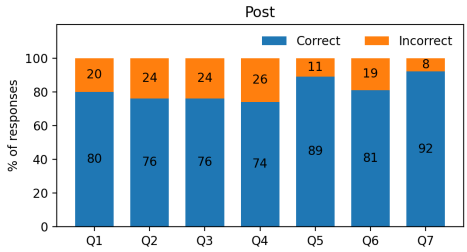
3.3. Learning Gain (LG)

The percentage of correct and incorrect responses to each question is presented in Figure 3. In the *Pre* examination, the proportion of correct answers varied from 20% to 47% for all questions (see Figure 3a), which increased in the *Post* examination, reaching between 74% and 92% of correct responses (see Figure 3b). Indeed, the percentage of improvement in each question was consistently positive, ranging from 31% for question Q2 to 59% for question Q5 (see Figure 3c). The lowest improvements were obtained in questions Q1, Q2, and Q4 (less than 35%), since a high percentage of these questions were answered correctly in the *Pre* exam, giving little room for improvement in the *Post* exam. Conversely, for the remaining questions, the improvement in the proportion of correct responses exceeded 50%. In general terms, the large improvement in the percentage of correct responses suggests a positive learning effect induced by the learning activity.

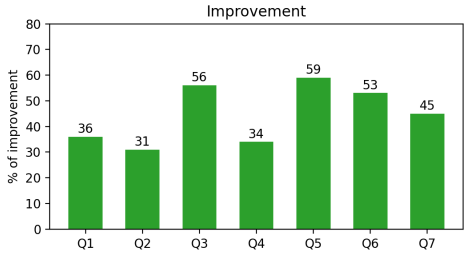
The grade, on a scale of 0–100, was then computed for each student, and the distributions for all participants are presented in Figure 4 for the *Pre* and *Post* examinations. Significant differences were found between the medians of the grade distributions before (*Pre* examination) and after (*Post* examination) the application of the learning activity ($p < 0.01$, Wilcoxon rank sum test). Indeed, the average grade was 36.47 ± 24.79 and 81.17 ± 21.13 in the *Pre* and *Post* exams, respectively, showing that on average, the grade was 44.70 points greater after the implementation than before. This shows a significant increase in the grades as a result of the application of the proposed learning activity.



(**a**) *Pre* examination.



(**b**) *Post* examination.



(**c**) Improvement.

Figure 3. Across-all-participants percentage of correct (blue) and incorrect (orange) responses for each question (Q1 to Q7) in the (**a**) *Pre* and (**b**) *Post* examinations. (**c**) Percentage of improvement in each question.

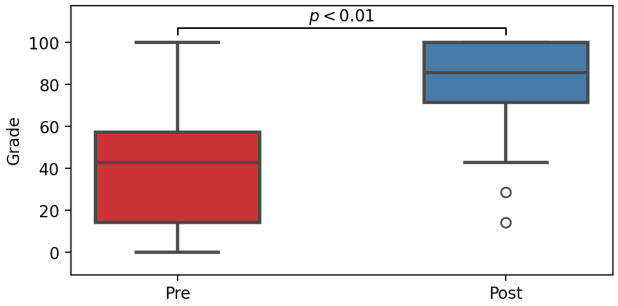


Figure 4. Distribution of grades of all participants in the *Pre* (boxplot in red) and *Post* (boxplot in blue) application of the learning activity.

3.4. Motivation (MT)

The distributions of the five dimensions of the MUSIC motivational inventory are presented in Figure 5. All five motivational scores were high, with median values around 5 and only a few outliers below 3. On average, these measures for empowerment, usefulness, success, interest, and caring were 4.99 $\pm$ 0.86, 5.09 $\pm$ 0.87, 5.19 $\pm$ 0.84, 5.07 $\pm$ 0.86, and 5.43 $\pm$ 0.75, respectively. Accordingly, students perceived the learning activity as pleasant and felt it provided them with control over their learning environment. Moreover, it was valuable for their academic development and fostered an environment of success. It is important to highlight that these results indicate that students actively participated and were motivated in the classes.

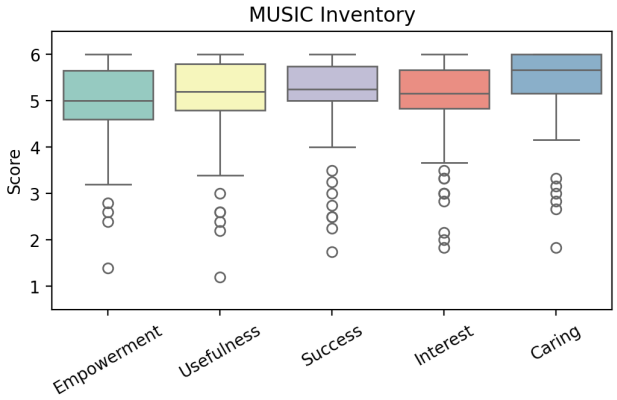


Figure 5. Distribution of scores for the five dimensions of the MUSIC inventory applied to all participants after the implementation of the learning activity.

4. Discussion

We proposed and implemented a learning activity designed to encourage and develop data analysis and machine learning knowledge, skills, and attitudes in college engineering students across diverse careers. The growing demand for professionals with robust training and experience in these competencies is a critical need in many industries (Smaldone et al., 2022). Additionally, these competencies are also considered critical for individuals ranging from the K-12 population to life-long learners (Sanusi et al., 2022; Srinivasan, 2022). Therefore, it is necessary for students and young professionals across various disciplines to understand and effectively use these methods and techniques. Given their responsibility in preparing students for the labor market, higher education institutions play an essential role in developing data analysis and machine learning competencies.

In our learning-by-doing methodology, we chose to conduct experiments involving the recording of students' own electroencephalogram (EEG) brain signals given the interest and engagement it generates and its suitability for exploring and discovering data analysis and machine learning concepts. The use of EEG devices in learning scenarios has been previously explored in medical education. For instance, they have been used to assess the efficacy of the flipped classroom model with neurology residency students (Novroski & Correll, 2018) and to investigate the effectiveness of combining problem-based learning (PBL) with case-based learning (CBL) in teaching clinical EEG interpretation to residents and refresher physicians (Li et al., 2024). Other studies have also enhanced knowledge and skills in medical students through EEG experiments (Nascimento et al., 2021). However, our learning activity, utilizing an EEG apparatus, differs in its focus, as it does not aim to develop medical skills. Instead, we use the device to collect students' own data in hands-on situations for non-medical training. Specifically, we aim to develop competencies in students from various engineering disciplines, such as Biomedical Engineering (measurement analysis and modeling), Data Science and Mathematics Engineering (pattern recognition and AI), Computer Technology Engineering (implementation of computational algorithms), and Robotics and Digital Systems Engineering (signal processing and data analysis).

The learning activity consisted of four stages: experiments for acquisition and visualization of brain signals, experiments to record brain signals in real situations, preparation and processing of the collected data, and application of machine learning models using the collected brain signals. This cross-disciplinary active learning activity aimed to motivate and engage students in their own learning process, promoting learning by doing, knowledge discovery, problem-solving, teamwork, and group discussions. The activity was implemented with nearly 300 students from four different engineering disciplines. Nevertheless, the results reported in this work are based on data from a total of 185 students who

completed all procedures, including signing the informed consent form and completing both the *Pre* and *Post* examinations. To assess the impact of the activity, we employed two validation instruments: the learning gain, determined by comparing students' grades before and after the activity, and the self-reported measures of empowerment, usefulness, success, interest, and caring, as calculated from inquiries using the MUSIC model inventory.

Considering the context questions, an increase in students' personal experience in the subject area was observed. First, it is notable that the proportion of students with hands-on experience collecting data increased from 23% before the learning activity to 90% after (Context question C1, Figure 2a), which indicates the effectiveness of carrying out practical activities. The 10% of students who did not report this experience warrant further analysis to determine specific reasons (e.g., absence of that session or lack of interest in practical activities), which should be addressed in the future.

Second, a significant increase was observed in students' appreciation of the importance of data collection for data analysis and machine learning (Context question C2, Figure 2b). Before the activity, 62% of students rated data collection very important, a percentage that rose to 89% afterward. Additionally, low importance rates decreased from 7% to 2%, revealing a trend towards a better personal appreciation of the importance of data collection.

Finally, there was a substantial increase in students' experience with EEG experiments, rising from 17% before the activity to 94% afterward (Context question C3, Figure 2c). This increase demonstrates the effectiveness of the learning activity in providing practical EEG experiences to students. The initial non-zero, low percentage of students (17%) with prior EEG experiences is because some of them had participated in EEG-based brain–computer interface research before. The remaining 6% who reported no EEG experience post-activity possibly indicates a misunderstanding of the question, or that the student, in fact, did not participate (again, absence of that session). This is critical and should be explored further to ensure that all students participate.

In line with our hypothesis, we found a significant increase in the learning gain associated with the application of the learning activity. This was observed individually for each question, with the percentage of improvement ranging from 30% in the lower case to 60% in the best case (see Figure 3c). Additionally, there was a significant improvement in students' grades, with the average grade rising from 36.47 ± 24.79 in the *Pre* examination to 81.17 ± 21.13 in the *Post* examination (see Figure 4), resulting in an average increase of 44.70 points. Indeed, the statistical analysis showed a significant improvement in students' grades following the learning activity. These findings indicate that the learning activity had a substantial positive impact on students' academic performance (consistent with the context questions), demonstrating its effectiveness in enhancing their understanding and application of data analysis and machine learning concepts.

On the other hand, the results of the MUSIC test revealed high motivation and appreciation for the learning activity. Empowerment, usefulness, success, interest, and caring dimensions reached average and median values around 5 points. This indicates that students found the learning activity enjoyable and felt it provided them control over their learning environment. Additionally, they perceived the activity as valuable for their academic development and felt it fostered an environment of success. Similar to the context questions, few students self-reported low motivational scores (as indicated by the atypical values in the distributions presented in Figure 5). These cases require further analysis, and a contingency plan should be formulated to identify and address these situations. However, our findings indicate that this active, experimental approach significantly improved student grades and fostered high levels of self-reported motivation, curiosity, and satisfaction, particularly concerning the relevance and engagement aspects of the learning experience.

The integration of real-world data collection, processing, and application appears to be a powerful pedagogical tool for complex, interdisciplinary subjects. Overall, this study highlights the effectiveness of the learning activity in motivating students for learning.

To our knowledge, this is the first study that formally introduces the use of an EEG brain signal recording device as part of an active learning methodology for developing data analysis and machine learning competencies in engineering (non-medical) students. While many studies explore innovative teaching methods for data science and machine learning, a significant number still rely on curated datasets or theoretical instruction. For example, studies on using machine learning to predict student success in online programming courses highlight the application of ML in education, but often do not involve students in the full data lifecycle themselves. Similarly, research into gamification and learning pathways focuses on enhancing engagement and motivation through structured digital environments, but may not incorporate the hands-on collection and processing of raw, experimental data that is central to our approach. Our unique contribution lies in directly embedding the real-world data acquisition of EEG signals into the learning process, thereby providing a more authentic and comprehensive experience of the data analysis pipeline. This contrasts with approaches where data is simply provided, pushing students beyond mere algorithmic application to a deeper understanding of data source, quality, and contextualization, a critical skill set often underdeveloped in traditional curricula.

Nonetheless, several aspects require further consideration or improvement for future implementations. First, a control group was not included. Hence, it would be advisable to assess the true effect of our educational innovation in the experimental group by comparing the results of the validation instruments with those obtained from groups of students who did not participate in the implementation. In this case, the students in the control group would not participate in stages 1 and 2 of the learning activity. Second, as mentioned above, there remains a small proportion of students who reported no experience with the collection of EEG brain signals and low motivational scores. We did not find associations between these factors, and a more detailed and personalized follow-up needs to be incorporated to ensure that all students are engaged in the activity. Third, it would be interesting to investigate to what extent the proposed educational innovation enhances the development of the same competencies in students from different engineering disciplines. For instance, Biomedical Engineering students may be more inclined and receptive to performing EEG experiments compared to Computer and Data Sciences students, who might have less interest in hands-on experimentation and field experiences. Despite these limitations, the pedagogical framework underpinning this learning activity demonstrates strong potential for generalizability. The core principle of active, experiential learning, where students engage with the full data analysis lifecycle through real data collection, is highly adaptable. This approach could be transferred to other STEM disciplines by simply changing the type of data collected (e.g., environmental sensor data for civil engineering, financial data for business analytics). The emphasis on critical thinking, problem-solving, and interdisciplinary application of data science and machine learning is universally valuable. The findings on enhanced motivation and perceived relevance align with studies on learning pathway systems and gamification, suggesting that integrating engaging elements and practical relevance can broadly benefit student learning outcomes, regardless of the specific data modality.

5. Conclusions

The proposed active learning activity, based on the collection and processing of students' own brain signals to develop data analysis and machine learning competencies, has proven to be effective in enhancing students' knowledge, skills, attitudes, and motivation.

By guiding students to conduct their own experiments and to record and process their own data before exploring data analysis and machine learning methods, this activity addresses the limitations inherent in using pre-recorded data, which is typical of traditional teaching methods with instructors or video lectures. Therefore, this activity promotes a shift from the traditional passive learning model, where students listen to an expert (either face-to-face or via video lectures) and use existing datasets recorded by others, toward a collaborative environment that fosters higher-order thinking and empowers students as owners of their learning process.

The learning activity comprises four stages, adaptable to various teaching sessions of differing lengths and formats, depending on the specific career, course, learning goals, thematic content, or other factors. The learning activity was implemented in six different courses across four engineering careers and facilitated the development of competencies for students in Biomedical Engineering (measurement analysis and modeling), Data Science and Mathematics Engineering (pattern recognition, natural language and AI), Computer Technology Engineering (implementation of computational algorithms), and Robotics and Digital Systems Engineering (signal processing and data analysis). The positive outcomes, demonstrated by significant learning gain and high self-reported empowerment, usefulness, success, interest, and caring, indicate that this hands-on, experiential learning approach can significantly improve students' understanding and appreciation of data analysis and machine learning concepts.

This study has significant broader impacts on computing education and beyond. Firstly, it offers a replicable model for teaching complex computational skills by fostering a deeper, more intuitive understanding of data science and machine learning concepts through active engagement rather than passive reception. Finally, the positive influence on student motivation and appreciation for data analysis suggests that similar active learning methodologies can contribute to a more engaging and effective learning environment in higher education, potentially leading to increased student retention and success in challenging STEM fields. Our findings align with broader educational trends emphasizing active learning and technology integration for enhanced pedagogical effectiveness, demonstrating a pathway to cultivate crucial competencies for the future workforce.

Author Contributions: Conceptualization, J.M.A. and L.G.H.-R.; methodology, J.M.A., O.M.-M., L.G.H.-R., M.A.-E. and P.C.-L.; software, J.M.A., O.M.-M. and L.G.H.-R.; validation, J.M.A., O.M.-M., L.G.H.-R., M.A.-E. and P.C.-L.; formal analysis, J.M.A., O.M.-M., L.G.H.-R., M.A.-E. and P.C.-L.; visualization, J.M.A., O.M.-M., L.G.H.-R., M.A.-E. and P.C.-L.; writing—original draft preparation, all authors. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Institute for the Future of Education, Tecnologico de Monterrey, Mexico, under NOVUS Grant N22-301.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of Tecnologico de Monterrey (CA-EIC-2407-01).

Informed Consent Statement: Written informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The dataset and recordings are available upon request to the corresponding author due to the institution's privacy guidelines regarding information and data collected from its students.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

Alam, A., & Mohanty, A. (2023). Educational technology: Exploring the convergence of technology and pedagogy through mobility, interactivity, AI, and learning tools. *Cogent Engineering*, *10*(2), 2283282. [CrossRef]

Amato-Henderson, S., & Sticklen, J. (2022, October 8–11). *Work in progress: Utilizing the MUSIC instrument to gauge progress in first-year engineering students*. 2022 IEEE Frontiers in Education Conference (FIE) (pp. 1–6), Uppsala, Sweden.

Anna Åkerfeldt, S. K., & Petersen, P. (2024). A research review of computational thinking and programming in education. *Technology, Pedagogy and Education*, *33*(3), 375–390. [CrossRef]

Arévalo-Cordovilla, F. E., & Peña, M. (2024). Comparative analysis of machine learning models for predicting student success in online programming courses: A study based on LMS data and external factors. *Mathematics*, *12*(20), 3272. [CrossRef]

Baah, C., Govender, I., & Subramaniam, P. R. (2024). Enhancing learning engagement: A study on gamification's influence on motivation and cognitive load. *Education Sciences*, *14*(10), 1115. [CrossRef]

Balady, S., & Taylor, C. (2024). Students' expert-like attitudes in calculus and introductory computer science courses with active-learning pedagogy. *Computer Science Education*, *34*(1), 37–67. [CrossRef]

Bin Tan, H.-Y. J., & Cutumisu, M. (2024). The applications of machine learning in computational thinking assessments: A scoping review. *Computer Science Education*, *34*(2), 193–221. [CrossRef]

Chailloux Peguero, J. D., Mendoza-Montoya, O., & Antelis, J. M. (2020). Single-option P300-BCI performance is affected by visual stimulation conditions. *Sensors*, *20*(24), 7198. [CrossRef] [PubMed]

Delijorge, J., Mendoza-Montoya, O., Gordillo, J. L., Caraza, R., Martinez, H. R., & Antelis, J. M. (2020). Evaluation of a P300-based brain-machine interface for a robotic hand-orthosis control. *Frontiers in Neuroscience*, *14*, 589659. [CrossRef] [PubMed]

DeMasi, O., Paxton, A., & Koy, K. (2020). Ad hoc efforts for advancing data science education. *PLoS Computational Biology*, *16*(5), e1007695. [CrossRef] [PubMed]

Gooding, B. (2023). Competency-based education: Transforming learning experiences. *Instructure Journal*, *4*(3), 45–57.

Guajardo-Cuéllar, A., Cárdenas, D., & Lepe, M. (2020). Fabrication of a boomerang using computational tools to develop engineering competences. *International Journal of Mechanical Engineering and Robotics Research*, *9*(10), 1406–1410. [CrossRef]

Guajardo-Cuéllar, A., Vázquez, C. R., & Navarro Gutiérrez, M. (2022). Developing competencies in a mechanism course using a project-based learning methodology in a multidisciplinary environment. *Education Sciences*, *12*(3), 160. [CrossRef]

Guido Charosky, K. P., Hassi, L., & Bragós, R. (2022). Developing innovation competences in engineering students: A comparison of two approaches. *European Journal of Engineering Education*, *47*(2), 353–372. [CrossRef]

Hartikainen, S., Rintala, H., Pylväs, L., & Nokelainen, P. (2019). The concept of active learning and the measurement of learning outcomes: A review of research in engineering higher education. *Education Sciences*, *9*(4), 276. [CrossRef]

Hernandez-Rojas, L. G., Cantillo-Negrete, J., Mendoza-Montoya, O., Carino-Escobar, R. I., Leyva-Martinez, I., Aguirre-Guemez, A. V., Barrera-Ortiz, A., Carrillo-Mora, P., & Antelis, J. M. (2022). Brain-computer interface controlled functional electrical stimulation: Evaluation with healthy subjects and spinal cord injury patients. *IEEE Access*, *10*, 46834–46852. [CrossRef]

Hicks, S., & Irizarry, R. (2018). A guide to teaching data science. *American Statistician*, *72*(4), 382–391. [CrossRef] [PubMed]

Iansiti, M., & Lakhani, K. R. (2020). Competing in the age of AI. *Harvard Business Review*, *98*(1), 60–67.

Ismail, A., Mutalib, S., & Haron, H. (2023). Data science technology course: The design, assessment and computing environment perspectives. *Education and Information Technologies*, *28*, 10209–10234. [CrossRef] [PubMed]

ITESM. (2018). *Tec21 educational model.* Available online: https://internationalfaculty.tec.mx/sites/g/files/vgjovo1851/files/tec21-model.pdf (accessed on 27 May 2024).

Jagatheesaperumal, S. K., Rahouti, M., Ahmad, K., Al-Fuqaha, A., & Guizani, M. (2022). The duo of artificial intelligence and big data for Industry 4.0: Applications, techniques, challenges, and future research directions. *IEEE Internet of Things Journal*, *9*(15), 12861–12885. [CrossRef]

Jones, B. D. (2009). Motivating students to engage in learning: The MUSIC model of academic motivation. *The International Journal of Teaching and Learning in Higher Education*, *21*, 272–285.

Jones, B. D. (2012). *User guide for assessing the components of the MUSIC model of motivation.* Available online: www.theMUSICmodel.com (accessed on 1 October 2022).

Jones, B. D. (2018). *Motivating students by design: Practical strategies for professors*. CreateSpace Independent Publishing Platform.

Jones, B. D. (2019). Testing the MUSIC model of motivation theory: Relationships between students' perceptions, engagement, and overall ratings. *The Canadian Journal for the Scholarship of Teaching and Learning*, *10*(3), n3. [CrossRef]

Jones, B. D. (2020). Engaging second language learners using the MUSIC model of motivation. *Frontiers in Psychology*, *11*, 1204. [CrossRef] [PubMed]

Jones, B. D., Khajavy, G. H., Li, M., Mohamed, H. E., & Reilly, P. (2023). Examining the cross-cultural validity of the MUSIC model of academic motivation inventory in English language courses. *Sage Open*, *13*(1), 21582440231156583. [CrossRef]

Jones, B. D., Wilkins, J. L. M., Schram, A. B., Gladman, T., Kenwright, D., & Lucio-Ramirez, C. A. (2023). Validating a measure of motivational climate in health science courses. *BMC Medical Education*, *23*, 548. [CrossRef] [PubMed]

Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O. M., Păun, D., & Mihoreanu, L. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. *Sustainability*, *13*(18), 10424. [CrossRef]

Legaki, N.-Z., Xi, N., Hamari, J., Karpouzis, K., & Assimakopoulos, V. (2020). The effect of challenge-based gamification on learning: An experiment in the context of statistics education. *International Journal of Human-Computer Studies*, *144*, 102496. [CrossRef] [PubMed]

Li, F., Luo, J., & Zhang, H. (2024). The application of problem-based learning combined with case-based learning in EEG teaching. *Journal of Medical Education and Curricular Development*, *11*, 23821205241252277. [CrossRef] [PubMed]

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J. H., Ogata, H., Baltes, J., Guerra, R., Li, P., & Tsai, C.-C. (2020). Challenges and future directions of big data and artificial intelligence in education. *Frontiers in Psychology*, *11*, 580820. [CrossRef] [PubMed]

Ma, Y.-C. (2023). Using participatory teaching in hands-on courses: Exploring the influence of teaching cases on learning motivation. *Education Sciences*, *13*(6), 547. [CrossRef]

Nascimento, F. A., Maheshwari, A., Chu, J., & Gavvala, J. R. (2021). EEG education in neurology residency: Background knowledge and focal challenges. *Epileptic Disorders*, *22*, 769–774. [CrossRef] [PubMed]

Navarro-Gutiérrez, M., Vázquez, C. R., Guajardo-Cuéllar, A., & Yungaicela-Naula, N. M. (2023, June 19–22). *A mixed reality laboratory for developing competencies in control engineering*. 9th International Conference on Higher Education Advances (HEAd'23), Valencia, Spain. Available online: https://api.semanticscholar.org/CorpusID:259720081 (accessed on 1 February 2023).

Novroski, A.-R., & Correll, C. (2018). Flipped classroom: Applications in teaching EEG. *Neurology*, *90*(15), P3.020. [CrossRef]

Olivares Olivares, S. L., López Islas, J. R., Pineda Garín, M. J., Rodríguez Chapa, J. A., Aguayo Hernández, C. H., & Peña Ortega, L. O. (2021). *Modelo educativo tec21: Retos para una vivencia que transforma*. Available online: https://search.ebscohost.com/login.aspx?direct=true&db=ir01477a&AN=ritec.11285.639177&lang=es&site=eds-live&scope=site (accessed on 27 May 2024).

Oliver, J. C., & McNeil, T. (2021). Undergraduate data science degrees emphasize computer science and statistics but fall short in ethics training and domain-specific context. *PeerJ Computer Science*, *7*, e441. [CrossRef] [PubMed]

Pace, A. C., Ham, A.-J. L., Poole, T. M., & Wahaib, K. L. (2016). Validation of the MUSIC® model of academic motivation inventory for use with student pharmacists. *Currents in Pharmacy Teaching and Learning*, *8*(5), 589–597. [CrossRef]

Padilla, E., & Campos, E. (2023, January 16–18). *Students' understanding of the mean through technology-mediated analysis of real-life data*. 2023 Future of Educational Innovation-Workshop Series Data in Action (pp. 1–8), Monterrey, Mexico. [CrossRef]

Peguero, J. D. C., Hernández-Rojas, L. G., Mendoza-Montoya, O., Caraza, R., & Antelis, J. M. (2023). SVEP detection assessment by combining visual stimuli paradigms and no-training detection methods. *Frontiers in Neuroscience*, *17*, 1142892. [CrossRef]

Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, *9*(4), 100480. [CrossRef] [PubMed]

Pontifex, M. B., & Coffman, C. A. (2023). Validation of the g.tec Unicorn Hybrid Black wireless EEG system. *Psychophysiology*, *60*(9), e14320. [CrossRef] [PubMed]

Sanusi, I. T., Olaleye, S. A., Agbo, F. J., & Chiu, T. K. (2022). The role of learners' competencies in artificial intelligence education. *Computers and Education: Artificial Intelligence*, *3*, 100098. [CrossRef]

Savonen, C., Wright, C., Hoffman, A., Humphries, E., Cox, K., Tan, F., & Leek, J. (2024). Motivation, inclusivity, and realism should drive data science education. *F1000Research*, *12*, 1240. [CrossRef] [PubMed]

Smaldone, F., Ippolito, A., Lagger, J., & Pellicano, M. (2022). Employability skills: Profiling data scientists in the digital labour market. *European Management Journal*, *40*(5), 671–684. [CrossRef]

Srinivasan, V. (2022). AI & learning: A preferred future. *Computers and Education: Artificial Intelligence*, *3*, 100062. [CrossRef]

Tehmina Gladman, S. G., & Ali, A. (2020). MUSIC® for medical students: Confirming the reliability and validity of a multi-factorial measure of academic motivation for medical education. *Teaching and Learning in Medicine*, *32*(5), 494–507. [CrossRef] [PubMed]

Terboven, C., Miller, J., Wienke, S., & Müller, M. S. (2020). Self-paced learning in HPC lab courses. *The Journal of Computational Science Education*, *11*, 61–67. [CrossRef] [PubMed]

Thomas Donoghue, B. V., & Ellis, S. E. (2021). Teaching creative and practical data science at scale. *Journal of Statistics and Data Science Education*, *29*(Suppl. 1), S27–S39. [CrossRef]

Thompson, J. R., & Harrison, P. L. (2023). Transitioning to competency-based education: Institutional perspectives and practices. *International Journal of Educational Reform*, *32*(1), 45–63. [CrossRef]

Torres-García, A. A., Mendoza-Montoya, O., Molinas, M., Antelis, J. M., Moctezuma, L. A., & Hernández-Del-Toro, T. (2022). Chapter 4—Pre-processing and feature extraction. In A. A. Torres-García, C. A. Reyes-García, L. Villaseñor-Pineda, & O. Mendoza-Montoya (Eds.), *Biosignal processing and classification using computational learning and intelligence* (pp. 59–91). Academic Press. Available online: https://www.sciencedirect.com/science/article/pii/B9780128201251000142 (accessed on 1 April 2024). [CrossRef]

Tsai, Y.-C. (2024). Empowering students through active learning in educational big data analytics. *Smart Learning Environments*, *11*(1), 14. [CrossRef]

Vermunt, J. D., Ilie, S., & Vignoles, A. (2018). Building the foundations for measuring learning gain in higher education: A conceptual framework and measurement instrument. *Higher Education Pedagogies*, *3*(1), 266–301. [CrossRef]

Villegas-Ch, W., Román-Cañizares, M., & Palacios-Pacheco, X. (2020). Improvement of an online education model with the integration of machine learning and data analysis in an LMS. *Applied Sciences*, *10*(15), 5371. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.