The impact of using a supervised learning system on student performance in a humanities and social sciences course

BAHLOUL Farouk [1,*]
University Abderrahmane Mira of Bejaïa, (Algeria), farouk.bahloul@univ-bejaia.dz

Received: 02/01/2025 **Accepted:** 25/07/2026 **Published**: 30/09/2026

Abstract: This study examines the impact of integrating a locally hosted supervised learning system into the teaching of a data analysis module within an Information and Communication Sciences undergraduate programme at the University of Béjaïa. Adopting a descriptive, design-based methodology, the research targets the validation of an adaptive pedagogical instrument across one lecture group and six tutorial groups per cohort. Student performance is operationalized as a process-based, multi-dimensional construct monitored through five interaction indicators — reading time, GSE query logs, chatbot interactions, MCQ scores, and graded exercise results — aggregated via a backend dashboard. The system's architecture draws on Self-Regulated Learning theory (Zimmerman, 2000), Vygotsky's Zone of Proximal Development (1978). Preliminary correlational analysis reveals a positive association between MCQ performance and the conceptual specificity of student queries, while passive reading time alone does not predict academic outcomes. These findings suggest that AI-driven adaptive environments measurably improve students' cognitive engagement and self-regulatory behavior in heterogeneous learner contexts.

Keywords: adaptive learning systems; supervised machine learning; process-based assessment; self-regulated learning; natural language processing; higher education.

ملخص: تتناول هذه الدراسة أثر دمج نظام تعلم آلي خاضع للإشراف في تدريس مقياس تحليل البيانات لطلبة السنة الثانية علوم الإعلام والاتصال بجامعة بجاية. يعتمد البحث منهجية وصفية طولية قائمة على التصميم، تستهدف التحقق من صلاحية أداة بيداغوجية تكيفية عبر محاضرة واحدة وستة أفواج أعمال موجهة في كل دفعة. يُشغَّل الأداء الأكاديمي بوصفه مفهوماً متعدد الأبعاد يُرصد عبر خمسة مؤشرات تفاعلية — مدة القراءة، واستعلامات محرك البحث، وتفاعلات الشاتبوت، ودرجات الاختبارات والتمارين — تُجمع آلياً من لوحة بيانات خلفية. يستند تصميم النظام إلى نموذج التعلم المنظم ذاتياً، ومنطقة النمو الوشيك لفيغوتسكي. تكشف القراءة الأولية للبيانات عن ارتباط إيجابي بين درجات الاختبارات ودقة استعلامات الطلاب، في حين لا تبرز مدة القراءة منفردةً متنبئاً داً بالتحصيل. تدعم هذه النتائج فاعلية البيئات التعليمية التكيفية في تحسين الانخراط المعرفي لدى الطلاب.

الكلمات المفتاحية :أنظمة التعلم التكيفي؛ التعلم الآلي الخاضع للإشراف؛ التقييم القائم على المسار؛ التعلم المنظم ذاتياً؛ معالجة اللغة الطبيعية؛ التعليم العالي

* Corresponding author, e-mail:

1- Introduction

The advent of artificial intelligence has profoundly changed practice in certain aspects of our cultural, technological and scientific life. The horizons opened by the use of AI in education, more especially in higher education (Wang et al., 2024), are beyond doubt; a significant number of establishments and universities around the world have joined this movement which continues to grow. the scale of incessant discoveries flourishing at a breakneck pace. Even the most natural and rudimentary tasks are entrusted to this new form of representing the world (Eager et al., 2023).

Algeria is not part of this movement; the Ministry of Higher Education and Scientific Research (MESRS) has made the integration of AI a primary objective in its strategy to revitalize the sector. During the last global conference on the level of AI governance in July 2024, the supervisory minister affirmed the need for the development and integration of AI in various sectors, a prelude to a new competitive technological and innovative strategy[1].

The purpose of this contribution is to examine the impact of integrating generative intelligence and a supervised learning system into data analysis teaching for second-year students in the Information and Communication Sciences program at the University of Bejaia. we had already worked on the integration of a conversational chatbot in the data analysis module, we noted that our approach needed to be supplemented by more intensive teaching of the integrated model.

We implemented scripted online teaching containing a chatbot as well as supervised learning of the module data. The process was trained to make student learning optimal in various activity formats. A corpus of search queries was analysed in order to follow the evolution and results of the introduction of such a technique.

This presentation will first outline a definition and a look at the fields of AI. Secondly, we will discuss the importance and implementation of such teaching and finally conclude with a practical examination of the approach taken and its impact on the progression of the module's teaching.[2]

1.1- Artificial intelligence, a definition

Artificial intelligence could be defined as a form of simulation of human intelligence by machines (McCarthy, 2007), this simulation includes the triptych: learning (acquisition), reasoning (logical, rule-based) as well as self-correction (through a test and error system, among others). It is customary to distinguish between so-called weak AI (whose main task can be described as specific: SIRI, ALEXA, the recommendation systems of YouTube, Netflix, Amazon, etc. to name only these few examples) from so-called strong AI which is designed to perform tasks in a manner analogous to human intelligence (this type of AI faithfully takes up the commonly used and accepted definition of the subject, where the AI is endowed with a capacity for understanding and doubling of consciousness assimilated to human intelligence).

More formally, we can define AI as referring to systems that exhibit intelligent behaviour by analysing their environment and taking action – with some degree of autonomy – to achieve specific goals. (Briganti, 2023). Some authors even call for the need to give meaning (Villani et al., 2018) to this notion which continues to mutate and is tinged with a certain temporal polysemy, others emphasize the notion of mastery as well as the natural limits of the notion (Fetzer & Fetzer, 1990, p. 4).

It would be vain, even illusory, to want to give a definitive definition of AI (Chiu, 2024) ; we will see in the following sections that the expansion of the fields of AI gives more meaning to attempts to define it.

1.2- Fields and types of artificial intelligence

In this section of our presentation, we will discuss the categories that underpin AI as well as its fields. You should know that this terminology is not clear-cut and acquired by all specialists in the field (Southworth et al., 2023; Tayan et al., 2024). A brief overview of the issue will also allow us to mark the path and introduce the next section which constitutes the essential part of our research approach.

To take the previous definition and clarify it, we will simply say that AI can be divided into two broad categories: Machine Learning and Deep Learning. Machine Learning manages and facilitates the learning of machines from data in order to improve their performance. Deep Learning, anchored in the philosophy of AI, is inspired by the human brain process, its neural networks, to create other artificial ones in order to process and interpret massive quantities of information (big data among others).

The following figure attempts to resume by explaining the types of artificial intelligence in a very simplified way. This presentation follows a division, the most common in this field of research, namely that operated between the capabilities of AI and its functionalities..

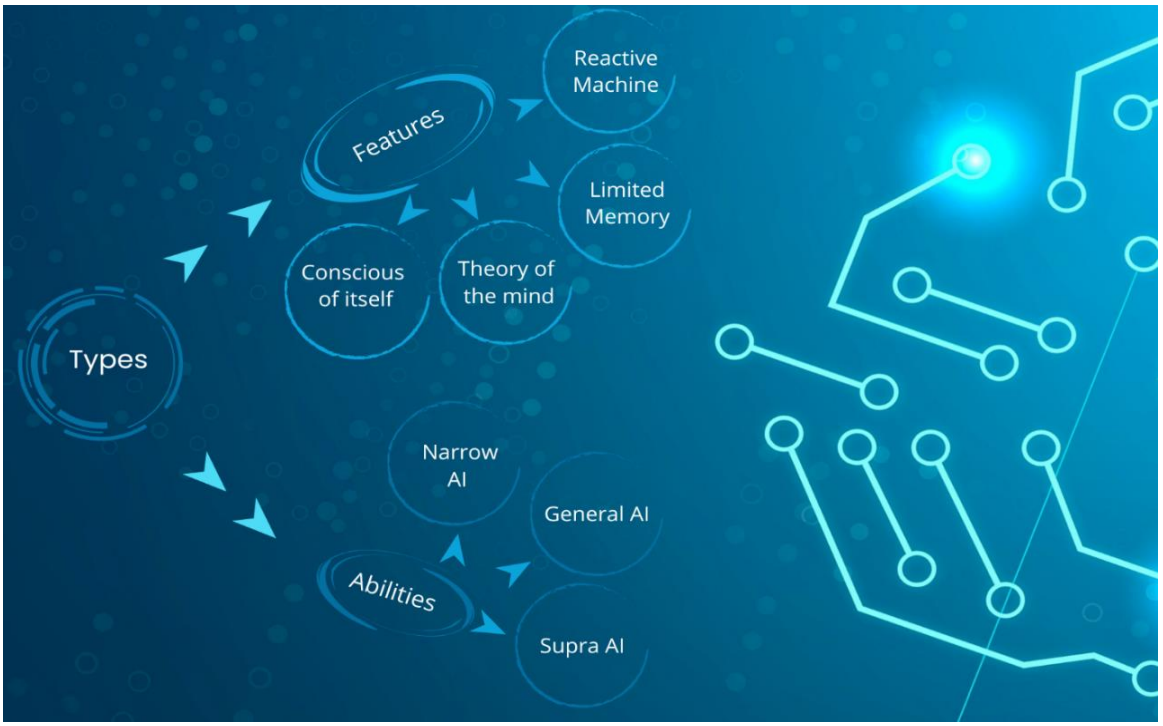

Figure (1) Types of artificial intelligence. Source: Author

Explanation. Taken by capacity, we can divide AI into three subtypes: Narrow AI (also called Weak), General AI (Strong) and finally Super-Artificial AI (Supra). It is the version that we currently know, trained for specific tasks, and others that humans cannot necessarily perform, this type of AI requires human intervention. The other two types of AI have no formal existence, they are only theoretical. Strong AI is therefore intended to be a theoretical concept which, unlike the first type, does not require human intervention for its training. On a more elaborate scale, we find the last type, Supra AI, also called super artificial, will be able to feel emotions, to develop cognitive mechanisms. In its functional aspect, AI can be subdivided into four types (it is this point that will rightly allow us to discuss the notion of the domain of AI). In this subdivision we find reactive AI which is specialized in very specific tasks through the exploitation of data with the aim of producing optimal results. Limited memory AI just exploits past information, it improves with training and uses this internalized memory for tasks as intelligent as reactive AI. The last two subtypes are

theoretical models, the AI of the mind will be the one that will give an interpretation to the thoughts as well as the feelings of man in a personalized way. Finally, the AI that is aware of itself will react by itself for itself and others, it will be able to develop its own emotions.

We can now more easily understand the fields of AI from the categories mentioned above, it is mainly about: Machine Learning therefore teaches machines to decide based on the data they have. Deep Learning, for its part, makes it possible to make voice and image recognition more targeted and optimized using neural networks.

Natural Language Processing (NLP) is another field of AI which allows, as its name suggests, to understand human language to perform various tasks. Finally, Symbolic AI simulates human reasoning to solve problems.

1.3- Presentation of the project approach

The idea of the project is to offer an interactive and enriching learning experience for students, in this sense there have been writings which have addressed this theme (Chan & Hu, 2023; Dai et al., 2023; Eager et al., 2023), our questioning is part of a more technical concern which is that of adapting the teaching to a group of students who for the most part consider the module as an initiation, because they come from literary or philosophical or language classes.

The project is structured around a gradual assimilation of the notions of statistics and data analysis in human and social sciences , supported by a chatbot that has a dual role: deep learning from the data inserted by ourselves (mainly and exclusively on the presentation and analysis of data) as well as another learning based this time on the interaction between the chatbot and the work of the students on the website dedicated to the module (and this via the activities and other means available on the module website). The module website is available in two languages: French and English and also has an audio description for the visually impaired[3].

1.3.1- Brief description of the data presentation and analysis module

The following lines, written in accordance with Order No. 903 of July 20, 2023 (setting the teaching program for obtaining a Bachelor's degree in Communication) present, in turn, the module, its teaching, the prerequisites. By clicking on each accordion, you will find more details.

The data presentation and analysis module is part of the methodological teaching unit, as such, it is of practical orientation. The data presentation and analysis module require the student to master a set of knowledge whose objective is to be able to provide an analysis based on both theoretical reasoning [logical and epistemological] and analytical reasoning [Statistics, Data Science]. The Module is at the intersection of several scientific disciplines, it takes advantage of statistical sciences to determine the appropriate terminology for the analysis. The Module is also based, to a large extent, on ***datalogy*** for the analysis of typologies of relationships between variables included in the framework of a quantitative data analysis. Finally, the module is also partly based on the methodological teaching of the first year, in fact, data analysis requires a certain skill that echoes a battery of knowledge. The content of the Module also involves considerations related to writing skills and abilities. One of the aims of the module is to be able to assist the student in mastering the methods and techniques for composing analysis reports (quantitative, qualitative or mixed) by obeying the most commonly used bibliographic standards (styles). [We will present in the following sections the objectives, prerequisites and content of the subject].

1.3.2- The objectives of introducing AI into the teaching of the module

In this study, student performance is conceptualized as a process-based, multi-dimensional construct rather than a terminal outcome measure. Following (Boud & Falchikov, 2007) framework of formative and process-oriented assessment, performance is not reduced to a single summative score but is operationalized through a set of observable behavioral and cognitive indicators generated by each student’s interaction with the learning environment over time.

Specifically, the following five indicators collectively constitute the operational definition of academic performance adopted in this research:

- scores obtained on multiple-choice questionnaires (MCQs) embedded within each lesson unit;
- completion rates and quality of responses to graded exercises, which vary in difficulty according to each student's interaction profile.
- reading time logged per lesson section, used as a proxy for attentional engagement with course content.
- search queries submitted through the integrated Google Search Engine (GSE) add-on, analyzed as indicators of autonomous knowledge-seeking behavior.
- chatbot interaction patterns, which reflect the nature and depth of students' conceptual difficulties.

These indicators are not used in isolation but are aggregated and cross-referenced within the backend dashboard, enabling the system to generate a dynamic and individualized performance profile for each registered student. A preliminary correlational analysis, discussed in Section 3, reveals a statistically meaningful association between MCQ scores and the thematic content of GSE and chatbot queries, while passive reading time alone does not significantly predict MCQ outcomes — a finding consistent with the distinction between surface and deep learning strategies (Marton & Säljö, 1976).

1.3.3. Theoretical Grounding of the Pedagogical Device

While the preceding sections have described the technical architecture of the supervised learning system, it is essential to situate this architecture within the established theoretical frameworks of cognitive and educational psychology from which it draws, whether explicitly or by structural analogy.

Self-Regulated Learning (SRL)

The system's design is primarily informed by Zimmerman (2000) model of Self-Regulated Learning, which posits that effective learning is characterized by the learner's capacity to set goals, monitor progress, and adapt strategies in response to feedback. The supervised learning environment operationalizes these three regulatory functions by: (a) orienting each student toward personalized learning objectives derived from their interaction profile; (b) providing continuous feedback through the chatbot and the dashboard's performance indicators; and (c) adapting the content and difficulty of exercises as a function of each student's demonstrated competency level. In this sense, the system acts as an external scaffolding structure that progressively internalizes the regulatory cycle described by Zimmerman.

Zone of Proximal Development (ZPD)

The adaptive difficulty mechanism of the chatbot is grounded in Vygotsky's (1978) concept of the Zone of Proximal Development, which defines the space between what a learner can achieve independently and what can be accomplished with appropriate guidance. By analyzing each student's query patterns and MCQ results, the system continuously recalibrates the nature of its interventions so as to position each learner at the frontier between mastered and emerging competencies. Exercises are neither repetitions of already-acquired knowledge nor unreachable challenges; they are generated to occupy precisely the productive zone of supported learning.

Taken together, these two theoretical frameworks converge on a common pedagogical principle: learning is not a uniform, linear process, but a situated, adaptive, and iterative one that benefits from responsive and personalized instructional support. The supervised learning system described in this study represents a practical instantiation of this principle in a real institutional context.

1.3.4- Tools and technologies used in the development of the ChatBot

We now move on to the description of the tools used in the development of the data presentation and analysis course as well as how it is structured. In a second step, we will

explain how we proceeded to integrate the two areas of AI into our teaching, the advantages and limitations of this integration will be discussed in the following section, a reminder of the main results is the subject of the last section of this work.

The content of the web pages. Each teaching is delivered in a homogeneous manner. The course has Three Blocks containing a total of eleven lessons (one lesson represents at least one course session and four at most), we have divided each lesson into fourteen axes which are listed below:

- Introduction and Summary of the Block
- Objectives of the session
- Concepts and themes to be covered during the session
- Presentation (oral) of the lesson
- The content of the lesson
- Summary
- Bibliography of the lesson
- Summary Questions
- MCQs
- Course Sheets & TD
- To go further
- On the Module app
- Download the Flipbook
- Discussion forum

Writing the content of web pages. In the layout of the teachings arranged in the form of a website, we used the following technologies and programs:

- HyperText Markup Language (HTML): HTML is the mark-up language used to create and structure the content of a web page. It allows you to define elements such as headings, paragraphs, images, links, lists, and much more. Each element is structured using tags, which allows browsers to understand how to display the content;
- Cascading Style Sheets (CSS): CSS is a style sheet language used to control the visual appearance of HTML elements. It allows you to define styles such as colours, fonts, spacing, borders, and animations. CSS separates the structure (HTML) from the presentation, which makes it easier to lay out and design websites. With CSS, a site can be made aesthetically pleasing and consistent;
- Bootstrap Framework: Bootstrap is a CSS Framework developed to facilitate the development of responsive and modern websites. It includes a collection of pre-designed components (such as buttons, forms, layout grids, and many others) and allows developers to quickly create elegant and adaptable user interfaces. Bootstrap is based on a grid system, which makes it easy to organize elements on different screen sizes;
- JavaScript: JavaScript is a client-side programming language, mainly used to add interactive features to web pages. Unlike HTML and CSS, which structure and style content, JavaScript allows HTML and CSS elements to be dynamically manipulated based on user actions (such as mouse clicks, scrolling, or data entry);
- Various libraries: of which we briefly describe the most important ones:
- jQuery: JavaScript library facilitating DOM manipulation, animations, and AJAX requests;
- Popper.js: JavaScript library to manage tooltips and pop-up elements in a precise way;
- Bootstrap Icons: Collection of icons designed to integrate perfectly with Bootstrap;
- Font Awesome: Library of icons to add custom symbols and logos to websites;
- footnote.js: JavaScript library to create interactive footnotes on web pages;
- MathJax: Tool to display mathematical notations in web pages using LaTeX, Math ML or Ascii Math;

– Artyom.js: JavaScript library to integrate speech recognition and synthesis into web applications;
– Video.js: Library to easily integrate customizable HTML5 video players;
– Google Fonts: Collection of free-to-use web fonts, with direct loading via URLs;
– jsPanel: Framework to create floating windows (like modals, tooltips, etc.) with various customization options;
– APlayer: Custom HTML5 audio player, offering an elegant design and intuitive controls for websites.

We have also, for the needs of collecting information and data from which our chatbot will feed, included a google search engine at the top of each course page.

The artificial intelligence part. The artificial intelligence part took a lot of time, the initial project started in 2019, with the release of the very limited version of GPT-2, we then wrote most of the content of the Course using Java Script for the interactivity it offered, however with the release of GPT-3.5, we reconsidered a lot of the content of this work, we then made changes that can be explained in the following lines:

Writing code with python. We used the most popular Python language in writing the code used to set up individual files for each student enrolled in the course, to write part of the chatbot as well as learning from student requests. In addition, Python is the data processing language in the module. We integrated an online compiler, in this case TRINKET, thanks to its simplicity and its large community. Python offers a multitude of libraries for machine learning (such as TensorFlow, Keras, PyTorch), data processing (Pandas, Numpy) and scientific computing (SciPy) integrated directly into the online compiler.

Integration of wit.ai, meta ai and botpress: wit.ai allowed us to delegate the learning operation to a ready-made machine, we were able to recover the botpress API to be able to store the data in a secure and reliable server, and above all free, in order to avoid possible costs that would have slowed down our progress.

1.3.5- Benefits of the AI-assisted learning model introduction approach

The advantages of the approach are to be included in the objectives assigned to the module, we take up in this section the objectives assigned to the teaching as they appear in the section: ***Presentation*** available on the module teaching site[4].

The teaching of the module aims at objectives that we can divide into categories:

Introducing students to analytical reasoning in the HSS

Through this course material, we want to support the student in the process of encountering data. Officially included in the methodological teaching unit, the data presentation and analysis module is an opportunity for work on concrete information. We will direct the knowledge contained in this module towards its technical aspect, data analysis is a matter of calculations and graphic representations;

Preparing the student to conduct a field survey

The content of the support emphasizes the obligation to collect data. The three axes that make up this support go from conceptualization (definition in variables) to the analysis and interpretation of the survey data. This exercise cannot be carried out without mastering the fundamentals of survey work. With the development of computer tools for data analysis, we have also reserved a significant part for work on the Python language;

Familiarize the student with the principles of the logic of statistics for the SHS

In data analysis and representation, a large amount of knowledge from statistics is involved. Data analysis is thus largely based on those of statistics, we considered it important to give more space to the logic of mathematical formulas, their construction and their meaning. We support the idea that learning a tool, a language, computer data analysis requires mastery of the theoretical notions that structure this conversion between theory and computerized application;

Developing the student's data analysis mechanisms for the survey

Any data analysis is a new experience in itself. Data analysis takes a relatively long time given the scope of the research being conducted. One of the teaching objectives is to

forge in students a mindset that questions data systematically and continuously, otherwise this analysis work will turn into a succession of steps to be validated;

Allow the student to write a report based on the survey data

This objective is the logical result of the previous ones. Indeed, data analysis, however relevant it may be, must be well communicated, transmitted using rules of "analysis" and writing. In this teaching, we will focus on providing students with the basics of information dissemination work.

2- Method and Tools:

2.1.Nature and Epistemological Positioning of the Research

The present study adopts a descriptive-longitudinal research design, situated within the tradition of design-based research (DBR) as theorized by Collins, Joseph and Bielaczyc (2004) and further developed by Reeves (2006). Design-based research is characterized by the iterative development, implementation, and evaluation of educational interventions in authentic instructional settings, with the explicit aim of producing both practical improvements and theoretical insights. This framework is particularly appropriate for research that seeks to validate a pedagogical instrument, as is the case here, rather than to test causal hypotheses through controlled experimentation.

Accordingly, the present study does not operate within a hypothetico-deductive paradigm; it does not require a control group, nor does it seek to establish a causal relationship between the supervised learning system and student outcomes through randomized group comparison. Its purpose is to document the operational conditions of the instrument, to analyze the behavioral traces generated by student–system interaction, and to assess the relevance and internal coherence of the pedagogical device as a whole. The objective of the approach, already reported by a number of authors (Khan, 2023; Memarian et al., 2023; Michel-Villarreal et al., 2023) is to assess the students continuously on the material of the module taught.

2.2. Study Population

The study is conducted on the full cohort of second-year undergraduate students enrolled in the Information and Communication Sciences (ICS) program at the University of Abderrahmane Mira, Bejaïa, Algeria. For each academic year under observation, the cohort consists of one lecture group and six tutorial groups, representing the standard institutional structure for this program.

Students were enrolled in the Data Presentation and Analysis module, a compulsory methodological unit within the ICS curriculum, as defined by Order No. 903 of 20 July 2023. The student population is predominantly composed of learners from literary, philosophical, and linguistic secondary school tracks, for whom quantitative and statistical reasoning constitutes a genuinely new disciplinary challenge. This characteristic feature of the population directly motivated the design of an adaptive, AI-assisted learning environment capable of accommodating a wide range of prior mathematical competencies.

Each student was registered individually via a structured enrolment form, which collected the following information: unique student registration number, language proficiency level, and self-assessed competency in mathematics. This data constituted the first layer of personalized input fed into the system’s backend.

2.3. System Development and Deployment Timeline

The supervised learning model at the core of this study was developed and trained on a locally hosted infrastructure. The training corpus was composed of data extracted from course materials, past examination sessions, and student interaction logs accumulated during the initial deployment phase. The system was trained and iteratively refined over the period 2019–2021, constituting a two-year pilot phase during which the model’s parameters were progressively adjusted to the specific content of the Data Presentation and Analysis module and to the learning profiles of the target student population.

Following this training phase, the system was operationally deployed for regular instructional use. The results reported in this study draw on data collected during the post-training deployment period, during which the chatbot and its associated supervised learning mechanisms functioned as an integral component of the module's pedagogical environment.

2.4. Data Collection Protocol

Data collection is organized around five distinct interaction flows, each corresponding to a specific dimension of student engagement with the learning platform. These flows are automatically captured and aggregated in the backend dashboard, which constitutes the primary analytical instrument of the study.

- Registration data: collected via Google Forms and imported into the backend as the baseline student profile, including registration number, language level, and mathematical background.
- Reading time data: a time-tracking script logs the duration of each student's engagement with each lesson section, generating a measure of attentional behavior that is subsequently transformed into a thematic word cloud and fed into the chatbot's learning module (Wit.ai).
- MCQ and exercise results: scores from embedded multiple-choice questionnaires and personalized exercises are stored in the system's flash memory and used to identify conceptual gaps at the individual level, which in turn orient the chatbot's subsequent interactions with each student.
- Search query logs: all queries submitted through the integrated Google Search Engine add-on are automatically retrieved and inserted into the chatbot's learning corpus, enriching its capacity to respond to students' evolving informational needs.
- Chatbot interaction logs and forum data: requests submitted to the chatbot, as well as comments extracted from the Disqus discussion forum, constitute an additional source of linguistic data used to refine the system's deep learning through Wit.ai.

2.5. Analytical Approach

The analytical approach combines two complementary methods. At the quantitative level, the dashboard aggregates performance indicators for each student individually and by tutorial group, enabling both longitudinal monitoring and cross-sectional comparison. A correlational analysis was subsequently conducted to examine the relationships between the five performance indicators defined in Section 1.3.2.

At the qualitative and computational level, a corpus of student queries — comprising all requests submitted to the chatbot, and keywords entered in the GSE add-on — was analyzed using scikit-learn, a Python library for machine learning and natural language processing. This corpus-based analysis made it possible to identify recurring thematic categories within student queries and to classify learners according to their apparent level of engagement with the module content.

3- Results and Discussion:

In this section, devoted to the discussion of the results obtained by introducing the AI-assisted learning procedure, we will first start with a reminder of the approach in the form of a figure tracing the process in its entirety, then we will discuss the elements that make up the process. Finally, we will illustrate our point by examining a practical case.

The integration process consists of two environments: a Frontend environment (what the students and the teacher see) and a Backend environment (what the teacher sees in terms of a Dashboard).

The procedure for collecting user data goes through four stages, the scraping thus considered is fed by these said sources in order to offer a targeted experience for each of the students registered for the course. The first stage consists of a registration using a Google form, the latter is received in the back end and treated as the first source of information on the registered student, the form contains a very important element which is the unique registration number specific to each student, it also contains other information (telephone number,

language level and skills in mathematics, ...), it can be consulted by clicking on this link. Once this first operation is carried out, the student will be able to download the module application, available in Android and iOS, it allows the student to interact with the teacher as well as to revise at the time of the exam, it also allows him to stay in touch by receiving notifications, one of the advantages of this application being its ability to offer an environment where the student can carry out his exercises during the tutorial sessions. The use of the site, third step, consists in recovering from the user's interaction with the course the following elements: the number of sections consulted, the keywords searched in the GSE, the concepts and questions searched in the chatbot and other interaction statistics. The last phase of the process consists in exploiting the data from the comments extracted from the Disqus discussion forum whose web application allows to recover the comments to be able to insert them into Wit.ai. On the backend side, once the information from each phase is retrieved, the chatbot and its learning will offer guided learning systems, the teacher will have a dashboard at his disposal to follow the personal developments of each student and this by registration number, first name, the details of each phase as well as the orientations offered by the installed chatbot.

This figure represents the chatbot's learning system with the possibilities it offers by developing from the different requests issued by users:

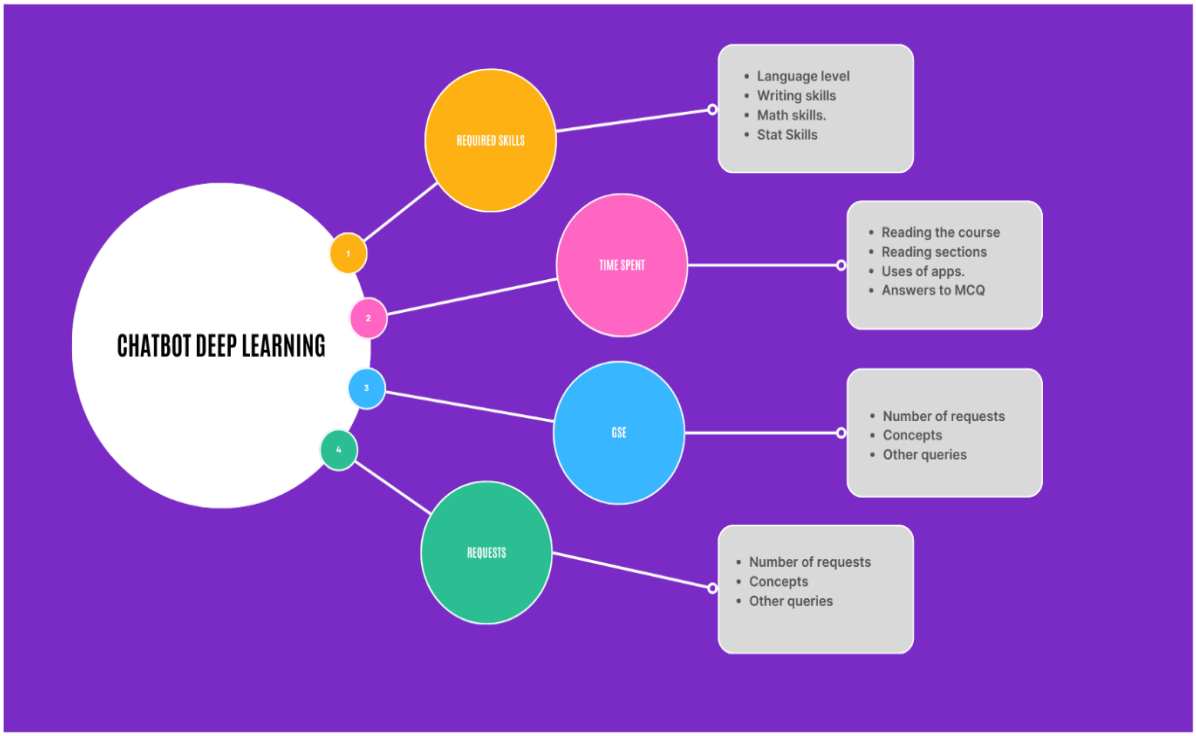

Figure (2) Backend learning and deep learning

In order to assess the relationships between the five performance indicators collected through the platform, a preliminary quantitative analysis was conducted on the aggregated dashboard data. The analysis focuses on two central questions: (1) the degree of association between MCQ scores and the thematic content of student queries submitted to the GSE and the chatbot; and (2) the predictive value of passive reading time with respect to MCQ outcomes.

The results of this analysis reveal a positive and statistically meaningful correlation between MCQ performance and the specificity of search and chatbot queries. Students who submitted queries directly related to the conceptual content of the module — as identified through the scikit-learn thematic classification — consistently obtained higher MCQ scores than those whose queries fell into the second category (less targeted searches). This finding is

consistent with deep learning theory, which associates intentional, concept-directed information-seeking with stronger conceptual understanding, and with recent empirical work on AI-driven teaching assistants in higher education (Zhang & Xiao, 2024).

Conversely, reading time alone — the total duration logged per lesson section — does not emerge as a significant predictor of MCQ outcomes. This result suggests that time-on-task, taken as an undifferentiated measure, is insufficient to characterize the quality of student engagement with course content. It is the nature of the cognitive activity accompanying reading, as reflected in subsequent query behavior, that proves more informative. This distinction between quantitative exposure and qualitative engagement is a well-documented finding in the learning sciences and reinforces the validity of a multi-indicator approach to performance measurement.

These quantitative results are preliminary in nature and reflect a single academic cohort. Their primary purpose is to demonstrate the analytical potential of the data collected through the platform and to identify directions for more systematic statistical investigation in subsequent research phases.

what we learn from corpus analysis

The corpus is made up of all the requests made by the students to the chatbot, key words, concepts or sentences searched in the Google search engine. We were thus able to collect a large number of verbatim reports which we subsequently analyzed using scikitlearn, a Python library for data analysis. The main lessons learned from this analysis are summarized in the following lines:

Firstly, we can divide the corpus into three categories of query register: first come the queries which tend to be confused with the main interest of the tool, this type of query has a progressive tendency, as evidenced by the hours spent in reading the course. To this end we agree with the constant made by Xia & Zhang who affirm that: *AI-driven teaching assistants can automatically grade assignments, answer common student queries, and even provide feedback. This not only ensures consistent and unbiased grading but also allows educators to focus on more complex student interactions and queries. The incorporation of artificial intelligence (AI) in higher education, especially through intelligent teaching assistants, is a burgeoning trend* (Zhang & Xiao, 2024).

The second group is made up of a set of less targeted searches and queries. We observed this through the time spent reading the related concepts as well as the exercises, MCQ and end-of-course assignments covered there. There have been recent attempts to define a way of designing artificial intelligence-assisted teaching (Sandu & Gide, 2019) for these latter authors as for others, AI (materialized by Chatbots) can help to extend learning capacity, an observation nuanced by another group of authors who see in AI-assisted teaching a way to distance certain learners from others (Neumann et al., 2023; Popenici, 2023).

4- Conclusion

The introduction of supervised learning in the teaching of data presentation and analysis, specifically designed for second-year undergraduate students in CIS, as well as the integration of a chatbot, made it possible to precisely target students' skills needs. This system also highlighted potential gaps, identified through the requests collected by the GSE and the chatbot.

In addition, information collected from various sources, such as MCQs, exercises and interactions with the course site, is used to further refine supervised learning and offer a more personalized educational path. Each piece of data collected plays a key role in adjusting the content, thus strengthening students' ability to progress in a learning environment adapted to their specific needs.

We also believe that such a methodology, which combines supervised learning and diversified data collection, could be applied to various educational content. A similar approach would allow for even more targeted monitoring for each pathway, with particular attention to students’ individual skills. This is a flexible and scalable model, capable of

meeting the demands of many other areas of education that would benefit from greater personalisation of learning.

Endnotes

[1] There is also a dynamic of introducing artificial intelligence into university programs as well as the creation of schools and institutes specialized in the field since 2019: Creation of SKAILAB at the University of Skikda in 2020, Creation of the national higher school of artificial intelligence (ENSIA) in 2021, creation of the Scientific Council for artificial intelligence in June 2023.

[2] The full work can be found on the website dedicated to teaching the module at the following address : https://elearning.univ-bejaia.dz/pluginfile.php/1133055/mod_resource/content/7/2.%20Forword.html#gsc.tab=0

[3] French website : https://elearning.univ-bejaia.dz/pluginfile.php/1132896/mod_resource/content/6/1.%20Avant-propos.html#gsc.tab=0
English website: https://elearning.univ-bejaia.dz/pluginfile.php/1133055/mod_resource/content/7/2.%20Forword.html#gsc.tab=0

[4] On this link, information related to prerequisites, content, teaching and assessment methods are also available:
https://elearning.univ-bejaia.dz/pluginfile.php/1133142/mod_resource/content/3/3.%20Fiche%20technique %20 et% 20informations %20sur%20le%20module.html#gsc.tab=0

References

Boud, D., & Falchikov, N. (2007). *Rethinking assessment in higher education: Learning for the longer term*. Routledge.

Briganti, G. (2023). Intelligence artificielle: une introduction pour les cliniciens. *Revue des Maladies Respiratoires*, *40*(4), 308-313.

Cardon, D., Cointet, J.-P., & Mazières, A. (2018). La revanche des neurones: L'invention des machines inductives et la controverse de l'intelligence artificielle. *Réseaux*(5), 173-220.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *20*(1), 43.

Chiu, T. K. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and artificial intelligence*, *6*, 100197.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences*, 15-42.

Copeland, B. J., & Proudfoot, D. (2000). What Turing did after he invented the universal Turing machine. *Journal of Logic, Language and Information*, *9*, 491-509.

Cordeschi, R. (2007). AI turns fifty: revisiting its origins. *Applied Artificial Intelligence*, *21*(4-5), 259-279.

Dai, Y., Liu, A., & Lim, C. P. J. P. C. (2023). Reconceptualizing ChatGPT and generative AI as a student-driven innovation in higher education. *119*, 84-90.

Eager, B., Brunton, R. J. J. o. U. T., & Practice, L. (2023). Prompting higher education towards AI-augmented teaching and learning practice. *20*(5).

Fetzer, J. H., & Fetzer, J. H. (1990). *What is artificial intelligence?* Springer.

Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the national academy of sciences*, *79*(8), 2554-2558.

Jensen, L. X., Buhl, A., Sharma, A., & Bearman, M. J. H. E. (2024). Generative AI and higher education: a review of claims from the first months of ChatGPT. 1-17.

Khan, M. (2023). *Ethics of Assessment in Higher Education–an Analysis of AI and Contemporary Teaching* (2516-2314).

Kuhn, T. S., & Meyer, L. (1983). *La structure des révolutions scientifiques* (Vol. 2). Flammarion Paris.

Machinery, C. (1950). Computing machinery and intelligence-AM Turing. *Mind*, *59*(236), 433.

Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I. Outcome and process. *British Journal of Educational Psychology*, *46*(1), 4-11.

McCarthy, J. (2007). What is artificial intelligence.

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the dartmouth summer research project on artificial intelligence, august 31, 1955. *AI magazine*, *27*(4), 12-12.

Memarian, B., Doleck, T. J. C., & Intelligence, E. A. (2023). Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI), and higher education: A systematic review. 100152.

Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. J. E. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *13*(9), 856.

Neumann, M., Rauschenberger, M., & Schön, E.-M. (2023). "We need to talk about ChatGPT": The future of AI and higher education. 2023 IEEE/ACM 5th International Workshop on Software Engineering Education for the Next Generation (SEENG),

Popenici, S. (2023). The critique of AI as a foundation for judicious use in higher education. *Journal of Applied LearningTeaching*, *6*(2).

Reevs, T. (2006). *Design research from a technology perspective* . Routledge.

Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, *65*(6), 386.

Russell, S., & Norvig, P. (2010). *Intelligence artificielle: Avec plus de 500 exercices*. Pearson Education France.

Sampson, J. R. (1976). Adaptation in natural and artificial systems (John H. Holland). In: Society for Industrial and Applied Mathematics.

Sandu, N., & Gide, E. (2019). Adoption of AI-Chatbots to enhance student learning experience in higher education in India. 2019 18th international conference on information technology based higher education and training (ITHET),

Southworth, J., Migliaccio, K., Glover, J., Reed, D., McCarty, C., Brendemuhl, J., . . . Intelligence, E. A. (2023). Developing a model for AI Across the curriculum: Transforming the higher education landscape via innovation in AI literacy. *4*, 100127.

Tayan, O., Hassan, A., Khankan, K., & Askool, S. J. M. L. w. A. (2024). Considerations for adapting higher education technology courses for AI large language models: A critical review of the impact of ChatGPT. *15*, 100513.

Teuscher, C. (2002). *Turing's connectionism: an investigation of neural network architectures*. Springer Science & Business Media.

Turing, A. (1936). On computable numbers, with an application to the Entscheidungs problem. *Proceedings of the London Mathematical Society Series/2 (42)*, 230-242.

Turing, A. (1948). Intelligent machinery: a report. *London: National Physical Laboratory*, 27.

Vygotsky, L. S. (1978). *Mind in Society: Development of Higher Psychological Processes.* Harvard University Press.

Villani, C., Bonnet, Y., Berthet, C., Levin, F., Schoenauer, M., Cornut, A. C., & Rondepierre, B. (2018). *Donner un sens à l'intelligence artificielle: pour une stratégie nationale et européenne*. Conseil national du numérique.

Wang, H., Dang, A., Wu, Z., Mac, S. J. C., & Intelligence, E. A. (2024). Generative AI in higher education: Seeing ChatGPT through universities' policies, resources, and guidelines. *7*, 100326.

Zhang, Y., & Xiao, M. (2024). Exploring Teachers' Perspectives and Strategies on Student Assessment Using Generative AI in Chinese Universities. *International Conference on Emerging Technologies*, (pp. 403-407).

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. Dans M. Boekaerts, P. R. Pintrich, & M. Zeidner, *Handbook of self-regulation* (pp. 13-39).

How to cite this article by the APA style:

BAHLOUL, Farouk (2026). The impact of using a supervised learning system on student performance in a humanities and social sciences course. *Journal of Psychological and Educational Sciences.* 12 (3). Algeria: El-Oued University.. 257-270