POLITECNICO DI TORINO

Master of Science in Engineering and Management

Specialization in Innovation and Entrepreneurship

Artificial Intelligence in Higher Education: Critical Success Factors and Business Model Innovation. The Case of *Speeding* in Italian STEM Education

Supervisor

Prof. Marco CANTAMESSA

Candidate

Giuseppe CANTALE AEO

Academic Year
2025/2026

ABSTRACT

This research examines the role of artificial intelligence in supporting exam preparation in Italian STEM degrees, where high dropout and delay rates are concentrated around a few “gatekeeper” courses such as Calculus 1–2, Physics, and Linear Algebra. It asks whether vertically specialised, exam-specific AI tutoring platforms can provide effective learning support and achieve sustainable unit economics in a context where general purpose chatbots are already widely used and largely free.

Theoretically, the work reviews the evolution of Intelligent Tutoring Systems from rule-based architectures to data-driven and LLM-based tutors, summarising evidence on learning gains, engagement, equity, and risks such as accuracy issues, cognitive offloading, academic integrity, and the constraints introduced by emerging regulation (e.g. the EU AI Act). It then discusses medium-term trends in AI in education (2026–2030), including adaptive learning and multimodal systems, the commoditisation of general LLMs, and the resulting need for verticalization, robust data governance, and compliance-by-design in EdTech business models.

Empirically, the thesis adopts the case study of Speeding, an Italian EdTech startup that pivoted from live tutoring courses to an AI-first platform called *Pepe*, focused on Calculus and related STEM exams. Drawing on a survey of 251 STEM students, a follow-up survey of 24 “power users” and internal product analytics, it tests three hypotheses concerning: (H1) the pedagogical effectiveness of a vertical AI tutor compared to textbooks, videos, private tutoring and generic chatbots; (H2) students’ willingness to pay for such a tutor; and (H3) retention patterns and unit economics in an exam-centric usage context.

Findings show strong product market fit on the learning side: the targeted courses match students' most difficult exams. Nearly all respondents already use AI to study but report concerns about correctness, generic explanations and syllabus misalignment. However, intensive users of Speeding report higher perceived understanding, clarity of step-by-step explanations and a sense of "learning the method" rather than merely copying solutions. At the same time, stated willingness to pay averages around 5€ per month and early monetisation data reveal strong price sensitivity, seasonality and one-shot usage limited to the duration of a single exam.

The thesis concludes that a vertically specialised AI tutor like *Pepe* can deliver credible, cost-efficient learning support in STEM most tough exams, but that current configurations are only pre-sustainable economically. Achieving full sustainability will require business model innovation (e.g. lower-priced or exam-specific plans, hybrid AI-plus-human tiers, broader subject coverage) and deliberate pedagogical design aimed at cultivating AI-enabled problem-solving skills rather than simply optimising short-term exam performance.

Table of Contents

Table of figures

1. THE CONTEXT

University dropout rates in Italy remain stubbornly high, particularly in STEM fields where approximately 30% of enrolled students do not complete their degree and an additional 20% are significantly delayed (do not complete the degree within its standard duration). Quantitative subjects such as Calculus, Physics, Chemistry, Geometry and Linear Algebra emerge as gatekeeper exams that block progression and contribute disproportionately to attrition: as shown by a survey of 251 STEM students conducted as part of this thesis, 59.4% identify Calculus 1-2 as their most difficult exam, followed by Physics 1-2 (40.2%) and Linear Algebra (35.9%). These phenomena are especially pronounced during the first years of university, where foundational STEM exams often act as critical bottlenecks. Typically, failure points are exacerbated by fragmented preparation methods (textbooks, YouTube videos, generic online resources) and by the high cost and limited availability of human tutoring, leaving many students stuck without effective support.

AI-based tutoring platforms offer a potential solution to this problem, combining personalisation, accessibility and scalability in ways that traditional methods cannot match. Large language models (LLMs) such as ChatGPT, Gemini or Copilot can generate step-by-step explanations, targeted exercises and adaptive feedback at marginal cost close to zero, potentially reaching hundreds of thousands of students who would never afford private tutors. However, the rapid diffusion of general-purpose AI tools also creates new challenges: 98.4% of surveyed students already use AI for studying, but 66.5% worry about the correctness of answers, 62.2% report calculation errors, and 45% find explanations too generic for their specific syllabus. The central question is whether verticalised, exam-specific AI tutors can deliver superior learning outcomes and sustainable business models, or whether they will merely commoditise as "yet another chatbot". So, the objective of this thesis is to analyze the role of artificial intelligence in supporting university students during exam preparation, with a specific focus on STEM disciplines. The research aims to evaluate whether AI-based tutoring systems can provide effective learning support, generate sustainable engagement, and represent a viable complement or alternative to traditional study resources.

This thesis examines these issues through the case study of *Speeding*, an Italian EdTech start-up that pivoted from live tutoring courses to an AI-based platform called *Pepe* in July 2025. *Pepe* focuses primarily on Calculus 1 and related STEM exams, offering step-by-step solutions, practice quizzes, flashacards and progress tracking powered by customised LLMs. The full empirical details of the case are presented in Chapter 4.

The thesis aim to answer to this question: "*Can a vertically specialised AI tutoring platform such as Pepe deliver effective learning support for university STEM students while achieving sustainable unit economics in the current landscape of AI adoption?* " To answer this question, the thesis tests three working hypotheses (H1–H3) related to pedagogical effectiveness, willingness to pay and retention/unit economics, using a mix of surveys (251 general students, 24 power users) and internal platform analytics.

The remainder of the thesis is structured as follows:

- **Chapter 2** reviews the state of the art on AI in higher education, distinguishing between general-purpose chatbots and specialised tutoring systems (Intelligent Tutoring Systems, or ITS), and discussing empirical evidence on learning outcomes, student engagement and skill development.

- **Chapter 3** analyses future scenarios for AI-driven tutoring business models, considering trends such as the commoditisation of LLMs, the role of verticalisation, pricing challenges and the need to foster “AI-enabled” competences in a world where generic tools are ubiquitous.
- **Chapter 4** presents the empirical case study of the startup *Speeding*, including business model evolution, experimental validation of H1–H3 through surveys and analytics, and a detailed assessment of critical success factors and sustainability.
- **Chapter 5** synthesises the findings, discusses their implications for EdTech strategy and outlines directions for future research and product development.

2. ARTIFICIAL INTELLIGENCE IN EDUCATION: STATE OF THE ART

This chapter reviews the state of the art on AI in education with a specific focus on what exists and what works, particularly in the domain of tutoring and learning support.

The review concentrates on the evolution of Intelligent Tutoring Systems (ITS) toward LLM-based tutoring, the mechanisms through which AI may improve learning, the empirical evidence on effectiveness, and the eventually limitations and risks such as over-reliance, academic integrity, and equity.

2.1 Intelligent Tutoring System (ITS)

The landscape of Artificial Intelligence in Education (AIED) has undergone a profound transformation over the last decade. A bibliometric analysis of 1,830 articles[1] published between 2010 and 2019 identifies "Intelligent Tutoring Systems" (ITS) as one of the most persistent and central research themes in the field.

Historically, these systems relied on rule-based architectures to model learner behavior. To this extent, Intelligent Tutoring Systems[2] (ITS) emerged as educational technologies designed to deliver individualized instruction by modeling aspects of the learner (e.g., knowledge state, misconceptions) and providing adaptive feedback. These tools have been a cornerstone of computer-assisted instruction since the 1980s and recent

[1] https://link.springer.com/article/10.1007/s40593-021-00244-4
[2] https://arxiv.org/html/2507.18882v1

trends indicate a shift toward data-intensive approaches, utilizing Educational Data Mining (EDM) and Learning Analytics (LA) to create more dynamic and responsive environments.

As shown in Figure 1.1, Intelligent Tutoring Systems represent the most recurrent keyword in AIED research between 2010 and 2019, confirming their centrality in the field. Adjacent themes such as learning analytics, educational data mining and machine learning highlight the progressive integration of data-driven approaches into educational technologies. The prominence of natural language processing anticipates the later emergence of LLM-based tutoring systems.

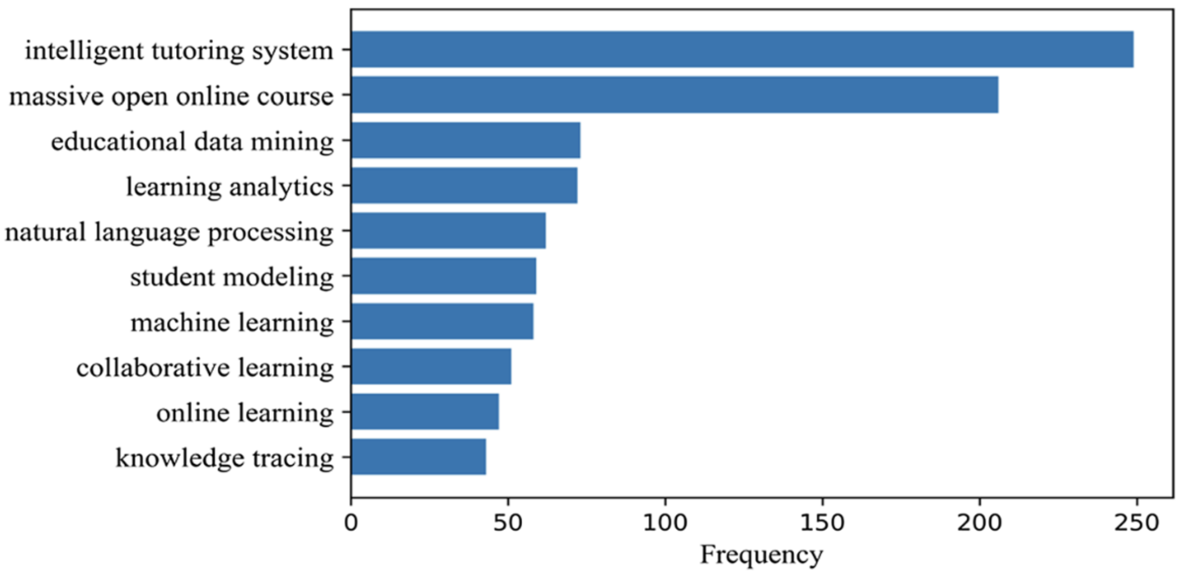

Fig. 1 Top ten most frequent keywords in Artificial Intelligence in Education (2010–2019)

2.1.1 Evolution and core mechanism of ITS

Traditional rule-based ITS relied on handcrafted knowledge bases and expert systems to model domain knowledge (e.g. algebra rules) and student skills, delivering personalised feedback through predefined decision trees and Bayesian knowledge tracing. While effective in controlled domains such as arithmetic and basic geometry, these systems struggled to scale to complex, open-ended subjects like university-level calculus, where the space of possible student misconceptions and solution paths is combinatorially explosive. Over time, ITS research expanded into more data-driven approaches, integrating machine learning and learner analytics to better estimate student mastery and personalize instructional sequences.

A key reason ITS gained attention is that its core promise mirrors that of human tutoring, in which consistent and personalized feedbacks are the huge values. Comparative evidence suggests that tutoring systems can be substantially more effective than no-tutoring baselines, though usually not matching the full impact of high-quality human tutoring.

2.1.2 Switch from rule-based ITS to LLM-based Tutoring

Modern ITS and adaptive learning systems typically operate on a "closed loop" architecture[3]: this mechanism continuously collects learner data, evaluates progress against a target model, and delivers customized feedback or content adjustments in real-time. In this context, the integration of Machine Learning algorithms has accelerated dramatically with the advent of large language models (LLMs)[4] such as GPT-4, Gemini and Claude. More advanced systems are now incorporating Deep Learning (DL) and neural networks to detect learning styles automatically and predict academic performance based on complex behavioral traces. For instance, recent chatbot-based tutoring systems allow for non-linear

[3] https://www.mdpi.com/2227-7102/13/12/1216
[4] https://www.mdpi.com/2227-7102/15/3/343

progression rather than simply grading a final answer: the system can offer hints, explanation videos, or easier sub-problems when a student gets stuck, fostering a more supportive learning scaffold[5]. Algorithms such as K-Nearest Neighbors, Decision Trees, and Bayesian networks are frequently employed to classify learner knowledge levels and recommend appropriate remediation strategies.

Hence, the rise of large language models (LLMs) introduced a qualitatively different paradigm: instead of relying primarily on explicit domain rules, LLM-based systems can support tutoring interactions through natural language dialogue, flexible explanations, and contextual adaptation[6]. This shift to LLM-based tutoring represents a paradigm change, because generative models pretrained on vast corpora of mathematical texts, textbooks and solved problems can now infer reasoning chains and adapt explanations to individual student inputs without explicit programming. Systems like LPITutor[7] combine LLMs with Retrieval-Augmented Generation (RAG) to retrieve syllabus-specific content and fine-tune prompts for particular exams, achieving flexibility that rule-based systems lacked. This evolution enables deployment of sophisticated tutoring at scale, but also raises questions about the reliability of generative outputs in high-stakes domains like STEM exam preparation.

Eventually, this shift increases conversational usability, but it also introduces new challenges, such as hallucinations, and difficulties in guaranteeing pedagogically sound reasoning in every interaction. In generative AI, hallucinations means that the model produces an output that

---

[5] https://www.ceeol.com/search/article-detail?id=1226718

[6] https://www.sciencedirect.com/science/article/abs/pii/S1041608023000195

[7] https://peerj.com/articles/cs-2991/

sounds fluent and confident, but is not grounded in the input data or in verified knowledge. In other words, it can invent facts, missing details and most of the times gives a well-structured explanation that still ends in a wrong result; sometimes, the latter case occurs when AI try to solve advanced mathematics problems.

Recent educational discussions therefore emphasize not only what LLMs can do, but also how they should be designed and governed to support learning rather than shortcut it.

Table 1. Conceptual comparison between rule-based Intelligent Tutoring Systems and LLM-based tutoring architectures

Rule-based ITS	LLM-based Tutoring
Expert rules	Pre-trained models
Closed domain	Open domain
High control	High flexibility
Limited scalability	Massive scalability

2.1.3 Evidence of effectiveness

The deployment of AI and ML in adaptive learning has demonstrated tangible benefits for student engagement and retention. By analyzing individual strengths and weaknesses, these systems can optimize learning paths, often leading to improved test scores and academic performance[8]. For example, the use of adaptive remediation strategies, where the system identifies specific error types and suggests targeted practice, has been shown to enhance the efficiency of the learning process.

[8] https://www.mdpi.com/2227-7102/13/12/1216

Meta analyses of ITS [9] effectiveness show moderately large learning gains across K–12 and higher education, with average effect sizes of Hedges' g = 0.66 for rule-based systems and emerging evidence of similar or better results for LLM-based tutors. In STEM subjects, where procedural fluency is critical, ITS have proven particularly effective for algebra and early calculus, reducing error rates by 20–30% compared to traditional instruction.

Benchmark studies comparing LLMs to human tutors provide mixed but encouraging results: ChatGPT-4[10] passes university admissions tests in mathematics with scores in the top 20–30% percentile[11], but struggles with advanced proof-based problems where reasoning chains exceed context windows or require domain-specific intuition.

In controlled experiments, students using ChatGPT for homework assistance showed statistically significant improvements in procedural skills (e.g. integration techniques) but required additional scaffolding to transfer these skills to novel problems. Retention effects are also positive: ITS users exhibit 10–15% higher long-term retention of mathematical procedures, thanks to spaced repetition and adaptive practice.

These findings suggest that LLM-based tutors are already viable for procedural STEM content like calculus exercises[12], but their effectiveness depends on combining generative power with vertical knowledge grounding (e.g. syllabus alignment, error detection) to overcome hallucinations and superficial explanations. Importantly, the evidence also suggests nuance:

[9] https://www.researchgate.net/publication/264545127_Intelligent_Tutoring_Systems_and_Learning_Outcomes_A_Meta-Analysis
[10] https://www.ceeol.com/search/article-detail?id=1360495
[11] https://www.iejme.com/article/chatgpts-performance-in-university-admissions-tests-in-mathematics-15517
[12] https://www.ceeol.com/search/article-detail?id=1226718

- ITS tends to be particularly effective when it supports active problem solving and step-based guidance;
- learning gains depend on whether the system encourages learners to engage in reasoning rather than passive copying.

2.2 AI Application in higher education

Generative AI has rapidly integrated into student workflows. Recent data indicates that 95.6% of surveyed students use AI technologies in their academic activities, with virtual assistants (like ChatGPT) being the most dominant tool, utilized by 88.2% of respondents. Students primarily use these tools for information retrieval, task management, and clarifying complex concepts, effectively using them as personalized academic assistants[13]. Scholarly discussions emphasize that these tools are "here to stay" and that education systems must manage adoption through thoughtful pedagogy and policy rather than denial.

In an internal suvery conducted by *Speeding*, 98.4% STEM university students[14] reporting that they have used LLMs for studying at least occasionally. Usage patterns fall into three main categories: explanations of complex concepts (e.g. "explain the chain rule in calculus"), exercise solving (e.g. "solve this integration problem step-by-step"), and summarisation (e.g. "summarise these lecture notes on limits"). These tools excel at providing instant, natural-language responses that reduce cognitive friction compared to textbooks or forums, but students frequently report limitations[15]: 66.5% worry about the correctness of

[13] https://www.mdpi.com/2227-7102/15/3/343
[14] Speeding's internal analytics
[15] Speeding's internal analytics

answers, 62.2% encounter calculation errors, and 45% find explanations too generic for specific syllabi.

2.2.1 Tutoring and Homework help

Among AI applications, tutoring and homework support is one of the most direct extensions of classical ITS. The educational value here depends strongly on interaction design (e.g., prompting learners to attempt steps, verifying understanding, and encouraging reflection), because students use conversational systems to ask for explanations, hints, and step-by-step guidance.

Tools like "Mentor Académico Inteligente" (MAI)[16], a personalized GPT-based virtual tutor, have been successfully used to resolve academic queries, generate study materials (such as summaries and diagrams), and assist with time management. In a study of MAI, 59.7% of students reported being "very satisfied" with the tool, citing its ability to generate quality content and clarify complex topics.

[16] https://posthumanism.co.uk/jp/article/view/1729

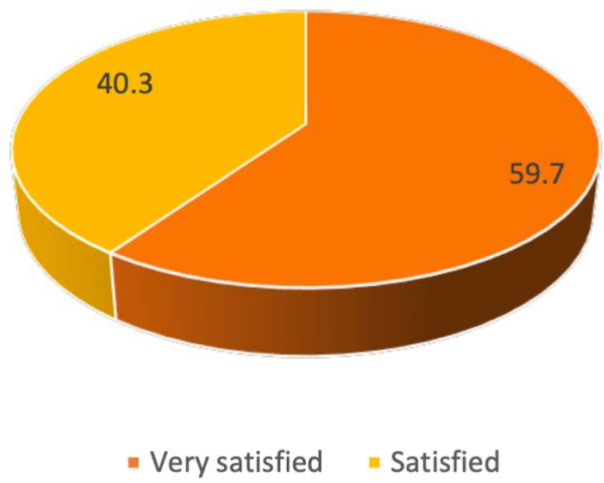

Fig. 2 Percentage of students satisfied and very satisfied about the tool's effectiveness (https://posthumanism.co.uk/jp/article/view/1729)

Conversely, in the domain of mathematics, the performance of AI solvers is mixed. A study testing the "AI Math Solver" on non-standard "Kangaroo" competition problems revealed a success rate of just under 30% for both English and Serbian tasks[17]. While capable of handling standard queries, these tools struggle significantly with problems requiring visual interpretation or multimodal reasoning, often failing to extract necessary data from diagrams or misunderstanding geometric constraints. Specific use cases are evolving beyond simple text generation: in the MAI study, students utilized the AI to generate quizzes, study guides, and schedules and, in more specialized contexts, students are using AR/VR tools like "Reality Labo" to create location-based games for tourism courses, blending technical skills with content creation. Additionally, generative bots are being used to provide personal feedback and create social learning experiences. However, empirical observations of chatbot tutoring reveal that students often use these tools to bypass the learning

[17] https://casopisi.junis.ni.ac.rs/index.php/FUTeachLearnTeachEd/article/view/13642

process: analysis of solution times showed that some students confirm correct answers in less than 2 seconds, suggesting they are "gaming" the system rather than solving the problems[18].

Table 2. Description of tasks and activities performed by MAI tool

Main Tasks of MAI	Description	Common Activities
Resolving Academic Queries	MAI acts as an immediate consultation tool for students, providing clear and contextualized explanations.	-Quick responses -Personalized explanations -Topic expansions
Generating Study Materials	MAI produces educational content tailored to each student's needs.	- Summaries - Diagrams and concept maps - Personalized quizzes - Study guides
Planning and Time Management	MAI functions as a virtual academic assistant, helping students organize tasks and optimize their time.	- Schedules - Reminders - Time management tips - Schedule adjustments
Formative assessment	MAI plays a significant role in formative assessment, providing continuous feedback to help students identify areas for improvement before final evaluations.	- Quizzes and exercises - Immediate feedback - Improvement suggestions - Mock exams

2.2.2 Assessment and administrative AI

Beyond tutoring, AI is reshaping administrative and assessment landscapes. In China, the adoption of AI in Higher Education Institutions (HEIs) has been linked to improvements in Organizational Performance (OP), particularly when mediated by Knowledge Management (KM) practices[19]. AI tools in these contexts are used to automate administrative tasks, such as enrollment management and resource allocation, thereby reducing inefficiencies. Furthermore, AI-driven platforms facilitate the organization of academic resources and streamline decision-making

[18] https://www.ceeol.com/search/article-detail?id=1226718
[19] https://www.igi-global.com/article/adoption-of-artificial-intelligence-and-organizational-performance-in-higher-education-institutions-in-china/388758

processes by utilizing predictive analytics to plan for institutional needs. Furthermore, AI in higher education also appears in learning analytics, early-warning systems, and administrative processes. This thesis focuses on AI tutoring for learning support and these institutional applications are considered secondary in this review. So, they are referenced mainly to contextualize broader AI adoption trends.

Systematic reviews note that AI research in higher education has historically been fragmented and often not fully aligned with educators' practical needs, reinforcing the importance of studying concrete tutoring use cases.

2.3 Pedagogical Impact

The impact of AI on learning outcomes is generally perceived as positive. In a survey of university students, 82.4% believed that using AI contributed to enhancing their academic performance, primarily by optimizing study time and facilitating quick access to resources[20].

In the context of the MAI tool, 71.8% of students found the AI's explanations useful for understanding complex topics[21].

Engagement,during study, is further amplified through "immersive learning" technologies that combine AI with AR and VR: these environments can create empathy-driven engagement by placing students in realistic global scenarios (e.g. experiencing the implications of climate change firsthand), which deepens the connection between the student and the content.[22]

---

[20] https://www.mdpi.com/2227-7102/15/3/343
[21] https://www.ceeol.com/search/article-detail?id=1360495
[22] https://aisberg.unibg.it/retrieve/93d826ed-5cdf-4c3f-9193-24f7ca0cdd80/CERLIS%20vol12_completo.pdf

UNESCO's guidance[23] highlights the need for human-centered implementation and safeguards, particularly in education and research contexts. Products like Khanmigo[24], which integrates GPT-4 with Khan Academy content, illustrate an intermediate approach: by grounding LLMs in curated educational materials, it reduces hallucinations while preserving generative flexibility and achieving higher satisfaction in procedural math tasks. However, even advanced implementations struggle with proof-based reasoning and motivational scaffolding, areas where human tutors retain advantages.

2.3.1 Equity, accessibility and student engagement

AI tools offer significant potential for inclusivity. As we said ealier, immersive learning technologies allow for tailored environments that accommodate learners with specific needs, such as sensory impairments or learning disabilities[25]. However, the digital divide remains a critical barrier. In the study of the MAI tool, 16% of students reported connectivity issues or a lack of adequate devices as a primary difficulty in using the AI tutor, highlighting that access to these pedagogical benefits is contingent upon reliable infrastructure[26].

The literature suggests that AI enhances motivation primarily through personalization, gamification, and the creation of immersive states of concentration. In the context of Extended Reality (XR), AI-driven environments are designed to cultivate "flow" (a state of deep concentration where a learner is fully engaged). This occurs when the AI

[23] https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research
[24] https://aclanthology.org/2025.emnlp-main.885.pdf
[25] https://aisberg.unibg.it/retrieve/93d826ed-5cdf-4c3f-9193-24f7ca0cdd80/CERLIS%20vol12_completo.pdf
[26] https://www.ceeol.com/search/article-detail?id=1360495

balances the student's skill level with the challenge level of the task. To this extent, the integration of game design elements (points, badges, leaderboards) in AI systems is a proven method to enhance both intrinsic and extrinsic motivation as well.
Adaptive e-learning platforms use these AI/ML algorithms to create interactive experiences that increase student engagement, leading to improved learning outcomes and motivation: by analyzing learner data, these systems optimize learning paths, ensuring that students are neither bored by easy tasks nor overwhelmed by difficult ones, which is crucial for maintaining focus and engagement[27].
In addition, AI tools are increasingly used to create safe learning environments where the fear of failure or social embarrassment is minimized. In language learning, speaking anxiety is a significant barrier ad and VR applications like *Immerse* use AI to provide a layer of anonymity through avatars. This reduces the cognitive load and social anxiety students often feel when speaking a foreign language in front of peers, allowing them to practice fluency and confidence in a low-stakes environment.[28]
In mathematics tutoring, AI chatbots have been designed to remove the pressure associated with traditional scoring. A study of an AI math chatbot[29] found that by presenting only the correct solution for review after an attempt, without a punitive grading system, the tool reduced the pressure that leads to random guessing. This approach encourages students to ask for hints and engaging with the material rather than just surviving the test.

[27] https://www.mdpi.com/2227-7102/13/12/1216
[28] https://aisberg.unibg.it/retrieve/93d826ed-5cdf-4c3f-9193-24f7ca0cdd80/CERLIS%20vol12_completo.pdf
[29] https://www.ceeol.com/search/article-detail?id=1226718

However, it is important to underline that AI can also cause anxiety: a study at the National University of Science and Technology POLITEHNICA Bucharest [30] found that while AI is helpful, nearly half (48.2%) of students expressed anxiety regarding the accuracy of AI-generated content, fearing they might learn incorrect information

2.4 Risks, ethics, and limitation

The widespread use of generative AI in higher education introduces significant risks and limitations. One of the most relevant risks for AI tutoring is that students may outsource too much cognitive effort to the system. The cognitive offloading literature shows that people routinely rely on external tools to reduce mental demands, which can help performance but may shift how individuals approach tasks and self-regulation. Applied to AI tutoring, this implies a key design requirement: systems should scaffold learning without replacing it.

As will be discussed in the empirical case study (Section 4.4), early evidence from student surveys confirms this risk: a non-negligible share of users expresses concern about over-reliance, suggesting that system design must actively scaffold autonomy. This also reflects a broader debate in the literature, where AI tutors risk to promote superficial shortcut learning if they provide complete solutions, without forcing students to engage with intermediate steps.

Moreover, generative AI raises direct concerns about plagiarism, cheating, and the validity of traditional assessments and so, academic integrity is another flashpoint. AI tools make it trivial to generate homework solutions

[30] https://www.mdpi.com/2227-7102/15/3/343

or even exam essays, raising questions about assessment design in an AI-pervasive world. Academic work [31]on ChatGPT in higher education highlights the tension between legitimate educational support and misuse, and discusses strategies such as redesigning assessment, increasing transparency, and clarifying acceptable use policies

In addition, a major limitation of current AI applications is accuracy, particularly with non-standard or visual tasks. The "AI Math Solver" frequently failed to interpret visual data in math problems correctly, sometimes providing hallucinated solutions or failing to model key constraints of a problem (e.g. assuming item prices were not distinct whole numbers)[32]. In student surveys, nearly half (48.2%) expressed concern about receiving incorrect or imprecise answers from AI tools[33].

Then, there is a growing fear that excessive dependence on AI may hinder the development of critical thinking. Students report that while AI saves time, it can reduce their willingness to ask questions or engage deeply with material, leading to a passive learning style[34]. The phenomenon of students using AI to simply confirm answers without attempting to solve them is a documented behavior that threatens genuine skill acquisition[35].

Furthermore, issues of academic integrity are prevalent; the ease of generating text and code raises plagiarism concerns, necessitating clear ethical guidelines.

The regulatory landscape is tightening in response to these risks: the EU AI Act classifies AI systems intended for use in education, specifically those used for assessing students or determining access to institutions, as

[31] https://www.tandfonline.com/doi/full/10.1080/14703297.2023.2190148
[32] https://casopisi.junis.ni.ac.rs/index.php/FUTeachLearnTeachEd/article/view/13642
[33] https://www.mdpi.com/2227-7102/15/3/343
[34] https://www.mdpi.com/2227-7102/15/3/343
[35] https://ejournal.upi.edu/index.php/IJOMR/article/view/60882

"high-risk" [36]. This classification imposes strict obligations on providers, including the establishment of risk management systems, high-quality data governance to prevent bias, and mandatory human oversight to ensure that AI tools do not perpetuate discrimination or exploit vulnerable groups.

2.5 Synthesis

The state of the art in AIED demonstrates a transition from experimental pilots to integrated tools that support tutoring, assessment, and administration. While adaptive systems and generative AI offer personalized support and high satisfaction, significant challenges regarding accuracy, equity, and cognitive offloading remain. The tension between the efficiency of AI-driven "survival English" or math solving and the pedagogical need for "productive struggle" is evident. These dynamics, technological promise, pedagogical risk, and unresolved equity questions, set the agenda for Chapter 3, which analyses how EdTech business models and regulation are evolving in response.

[36] https://www.wsgr.com/a/web/qrkz1SnNzWw6nk7B3oAyDa/10-things-you-should-know-about-the-eu-artificial-intelligence-act_v2.pdf

3. WHERE IS AI IN EDUCATION GOING? TRENDS 2026-2030

This chapter analyzes the trajectory of Artificial Intelligence in Education (AIED) as it moves from experimental pilot programs to critical institutional infrastructure. It explores the convergence of technological advancements, evolving business models, and the stringent regulatory frameworks that will define the educational landscape between 2026 and 2030.

3.1 From pilot to infrastructure

In higher education, the integration of AI, is shifting from isolated applications to systemic infrastructure. This transition is driven by the need for operational efficiency and personalized learning at scale.
The education technology (EdTech) market is entering a phase of accelerated expansion, driven by digital delivery, data-intensive learning tools, and the integration of AI into study workflows. Recent market estimates place global EdTech at around $163.5B in 2024, projected to reach $348.4B by 2030 (CAGR 13.3%, from 2026 to 2030)[37].

[37] https://www.grandviewresearch.com/industry-analysis/education-technology-market

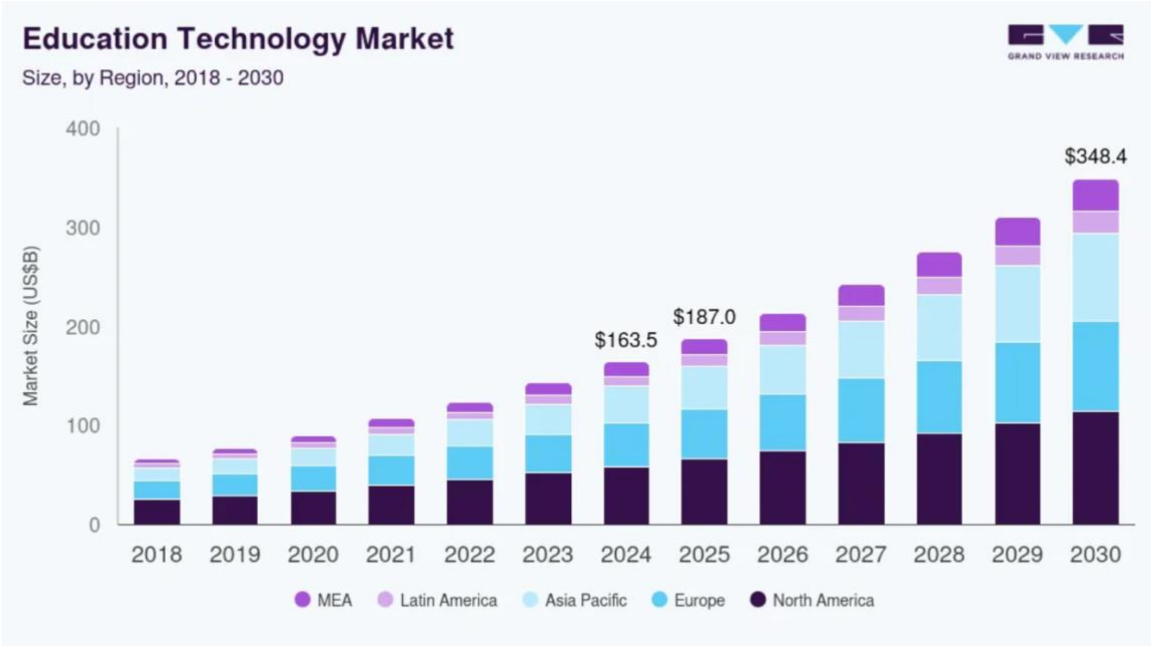

Fig. 3 Global EdTech Market size (in billion of dollars)
(https://www.grandviewresearch.com/industry-analysis/education-technology-market)

Adoption is already mainstream. Survey evidence suggests that student use of generative AI has shifted from experimentation to routine. For example, the UK HEPI/Kortext Student Generative AI Survey (2025) reports that 92% of students used AI tools, up from 66% in 2024, and 88% used generative AI for assessments[38].

However, successful scaling requires more than just software acquisition: it necessitates robust Organizational Capacity (OC) and Managerial Capacity (MC). Research indicates that an institution's ability to adopt AI is contingent on its structural and financial resources and the strategic leadership required to align technology with institutional goals. For instance, in China, government initiatives like the "Next Generation Artificial Intelligence Development Plan" have accelerated this infrastructure shift, demonstrating that robust policy frameworks and

[38] https://www.hepi.ac.uk/reports/student-generative-ai-survey-2025/

substantial investment are prerequisites for moving beyond pilot phases[39] Despite the momentum, significant barriers remain. These include inadequate technological infrastructure, insufficient faculty training, and deep-seated resistance to change. A major impediment to the commercial development of AI-based solutions is the so-called "cold-start problem," where systems lack sufficient initial data on learners to provide effective personalization immediately[40]. Furthermore, the lack of data sovereignty and the complexity of integrating new AI tools with legacy Learning Management Systems (LMS) continue to hinder widespread infrastructural adoption[41]

3.2 Technology trends: from chatbot to adaptive learning

The technological trend in Education Technology is shifting from static, text-based chatbots toward adaptive systems that personalize learning paths and multimodal environments that integrate vision, voice, and immersive reality. The literature landscape shows that ITS (intelligent tutoring systems) remain a central focus of AI in education, with adjacent growth around personalization, deep learning, and analytics-driven online learning ecosystems[42]. Current generative AI tools (like ChatGPT) function primarily as virtual assistants that answer questions, generate summaries, and provide immediate feedback. However, the trend is moving toward integrating these chatbots into "closed loop" architectures. The idea behind is: collect learner data → model the learner → adapt content/sequence/feedback → collect more data.

[39] https://www.igi-global.com/article/adoption-of-artificial-intelligence-and-organizational-performance-in-higher-education-institutions-in-china/388758
[40] https://www.mdpi.com/2227-7102/13/12/1216
[41] https://ieeexplore.ieee.org/abstract/document/9997887
[42] https://link.springer.com/article/10.1007/s40593-021-00244-4

This transition to adaptive tutoring is driven by the integration of specific Machine Learning (ML) algorithms that go beyond simple text generation: algorithms like Ant Colony Optimization (ACO) and Genetic Algorithms are used to generate optimal learning paths for different groups of students, dynamically adjusting the sequence of materials based on performance. The adaptive-learning literature shows platforms combining multiple ML families to do different jobs: prediction, clustering, recommendation, sequencing, and common patterns include:
- Clustering patterns (techniques like K-means) that categorize learners into groups (e.g., beginner, intermediate) and identify learning styles to tailor content delivery.
- Deep learning models such as Long Short-Term Memory (LSTM) networks, are employed to predict learner performance and identify at-risk students early.
- Reinforcement learning (e.g. Q-learning) to optimize learning paths using implicit feedback.

These modern system don't just look at final scores, but they ingest LMS logs, online interactions, assessments, digital resources, and build learner profiles (strengths, weaknesses, preferences, gaps) that get updated over time. Finally, this is what enables real personalization toward users.

To this aim, adaptive systems are increasingly expected to justify recommendations so learners and instructors can trust them (and so users can debug them when they do something dumb). The adaptive-learning review explicitly highlights growing emphasis on explainable AI for transparency, trust, accountability, and ethical use[43].

---

[43] https://www.mdpi.com/2227-7102/13/12/1216

3.2.1 AI, AR/VR, and multimodal learning

Adaptive learning work points to more language-capable interactions (chatbots, voice assistants, natural-language assessment/feedback) as a way to deliver personalization in a more usable format. Future trends point toward multimodal learning, where AI systems process not just text, but also visual and auditory data. In this context, the future of pedagogy lies in the convergence of Augmented Reality (AR), Virtual Reality (VR), and AI as well. This synthesis creates "immersive learning" environments that allow for just-in-time information delivery, in which information appears exactly when the learner needs it in-context (e.g., when they interact with an object or scenario). For example, AI-driven VR can place students in realistic global scenarios , thereby deepening the connection between the student and the content. In language learning, VR applications like *Mondly VR* and *Immerse* utilize AI to create realistic social interactions, allowing students to practice survival English in low-stakes, simulated environments (e.g., checking into a hotel or ordering food). XR language-learning examples also highlight the *practical* multimodality: spatial audio, spoken interaction with AI interlocutors, realistic scenarios (hotel, shops), and avatar-based interaction. The reasearch notes that immersion/presence can be high, but real systems still fail in annoying human ways (e.g. glitches or the AI avatar not processing user input).[44] Regarding STEM context, current AI math solvers still struggle significantly with multimodal reasoning, tasks that require synthesizing text and images. For example, in the Kangaroo math competition, an AI tool failed to solve geometry problems because it misidentified visual

[44] https://aisberg.unibg.it/retrieve/93d826ed-5cdf-4c3f-9193-24f7ca0cdd80/CERLIS%20vol12_completo.pdf

segments in the diagrams, achieving a success rate of less than 30% on non-standard tasks. The trend is to improve these vision combined with language models to handle STEM subjects more effectively[45].

Fig. 4 Benefits in immersive learning
(https://aisberg.unibg.it/retrieve/93d826ed-5cdf-4c3f-9193-24f7ca0cdd80/CERLIS%20vol12_completo.pdf)

3.3 Adaption patterns, business model and value capture

The EdTech industry is transitioning toward data-driven business models, specifically Software as a Service (SaaS) and Data Infrastructure as a Service (DIaaS). These models allow EdTech firms to co-create value with educational institutions by hosting software and designing curricula. However, a significant tension exists regarding data sovereignty: many EdTech firms struggle to fully capture value because they lack access to the granular data needed to train their algorithms effectively, often operating in "low-data models" where the consumer retains the data[46].
The adoption of AI is positively related to Organizational Performance in Higher Education Institutions (HEIs) and this relationship is significantly

---

[45] https://casopisi.junis.ni.ac.rs/index.php/FUTeachLearnTeachEd/article/view/13642
[46] https://ieeexplore.ieee.org/abstract/document/9997887

mediated by Knowledge Management. AI tools enhance the efficiency of knowledge acquisition, storage, and sharing, which in turn improves teaching quality, research productivity, and administrative decision-making,. Therefore, the value capture for universities lies not just in the technology itself, but in how AI facilitates a culture of knowledge sharing and innovation[47]

Eventually, as general-purpose LLMs become widely available, AI tutoring products face a commoditisation threat. Sustainable differentiation increasingly depends on *verticalisation* (domain constraints, syllabus grounding, verification, pedagogy-by-design) and on proprietary data loops (learning analytics, progress tracking, retention systems).

On the business model adoption side, the dominant models remain:

B2C direct-to-student (freemium, subscription, usage-based credits): fastest adoption, hardest monetisation due to price sensitivity.

B2B2C / university partnerships (licensing, bundled access through institutions): slower sales cycles, potentially stronger retention and defensibility.

Hybrid models combining AI tutoring with periodic human, able to differentiate the experience.

EdTech business model research also highlights systemic frictions, since education decision-makers are institutions and educators. For this reason, data sovereignty and ambiguity can block analytics-driven models, which are critical to personalize user's experience. Therefore, many providers still underutilize data-driven value creation compared to other sectors.

---

[47] https://www.igi-global.com/article/adoption-of-artificial-intelligence-and-organizational-performance-in-higher-education-institutions-in-china/388758

3.4 Regulation and Governance: The AI Act

In Europe, the AI trend trends will be heavily defined by the EU Artificial Intelligence Act (AI Act)[48]. This legislation classifies AI systems used in education (specifically those for assessing students, determining access to institutions, or assigning students to schools) as "high-risk". Providers of high-risk AI systems must adhere to strict obligations, including:

1- Risk Management Systems: Continuous iteration to identify and mitigate risks to health, safety, and fundamental rights.
2- Data Governance: Ensuring training, validation, and testing data are of high quality to prevent bias and discrimination.
3- Human Oversight: Systems must be designed to allow human reviewers to override or stop the AI ("stop button"), ensuring that machines do not have the final say in critical educational trajectories

Having taken these rules, product strategy in higher education must consider compliance-by-design (risk management, transparency, data governance, incident response) becomes part of the business model, not an afterthought.

Beyond safety, ethical governance must address the black box nature of AI algorithms. There is a pressing need for Explainable AI (XAI) to ensure that students and teachers understand why an algorithm made a specific recommendation. Furthermore, the rise of Generative AI has sparked legal battles over copyright infringement, as Large Language Models (LLMs) are often trained on copyrighted texts without permission. This has led to

[48] https://www.wsgr.com/a/web/qrkz1SnNzWw6nk7B3oAyDa/10-things-you-should-know-about-the-eu-artificial-intelligence-act_v2.pdf

lawsuits by authors and demands for licensing frameworks to protect creative and intellectual property in the AI era[49].

Fig. 5 Implementation timeline of the EU AI Act.

3.5 Scenario synthesis

By 2030, the world with AI will likely be replaced by a regulated, blended model. Across market, technology, and policy trends, research converges on a practical synthesis: sustainable AI tutoring in higher education requires trustable tutoring, vertical grounding, adaptive progression and a credible value capture mechanism that matches student willingness-to-pay and retention dynamics.

Successful company and institutions will be those that possess the managerial capacity to navigate the complex regulatory requirements of the AI Act while leveraging the knowledge management systems to translate AI data into improved organizational performance. The primary challenge will shift from access to technology to the integrity of the

[49] https://aisberg.unibg.it/retrieve/93d826ed-5cdf-4c3f-9193-24f7ca0cdd80/CERLIS%20vol12_completo.pdf

learning process, ensuring that AI remains a tool for scaffolding rather than a crutch for cognitive offloading.

As the empirical case discussed in Chapter 4 will show, these structural tensions between adoption and monetisation are not theoretical, since they are already being experienced by early-stage EdTech start-up operating in this space.

4. APPLICATION: THE SPEEDING CASE

4.1 Case context

Speeding is an Italian EdTech start-up founded in 2023 with the explicit goal of reducing university dropout in STEM degrees by helping students overcome gatekeeper quantitative exams. In its first phase (from October 2022 to June 2025), the company operated as a high-touch course provider, delivering live preparatory classes for Calculus and related subjects to more than 400 students with excellent satisfaction scores (4.6/5 on Trustpilot). While this model validated the existence of strong demand for structured support in difficult STEM exams, it also proved difficult to scale due to the need for human instructors and limited capacity per cohort.

In July 2025, *Speeding* pivoted from live courses to an AI-first platform centred on *Pepe*, a large-language-model-based tutor for mathematics and related subjects. The idea was to encode the pedagogical approach tested in live courses (step-by-step explanations, exam-oriented exercises, and continuous feedback) into an automated system that could serve hundreds or thousands of students simultaneously, at a fraction of the marginal cost of human tutoring. The initial scope of *Pepe* focused on Calculus 1 and 2, Linear Algebra and introductory Physics, reflecting the results of preliminary interviews and the later survey showing that 59.4% of STEM students consider Calculus 1 and 2 their most difficult exam, followed by Physics 1-2 and Linear Algebra.

Pepe is accessible via web interface and allows students to upload exercises, lecture slides or photos of handwritten problems, and to receive detailed, step-by-step solutions generated by an LLM guided by custom prompts and domain-specific constraints. The system can also generate additional practice questions, quizzes and short summaries, and is being extended towards adaptive study plans, progress monitoring and community features. Between July and December 2025 the platform operated under a freemium model, offering generous free usage while collecting data on user behaviour, engagement and learning perceptions. A paid model based on credit packs and optional subscription was introduced only in January 2026.

As of the time of analysis, the platform counts approximately 800 registered users, mainly enrolled in engineering and scientific degree programs across several Italian universities. While the current implementation deliberately focuses on STEM disciplines, the underlying technological architecture is designed to be extensible to other academic fields. This vertical-first approach allows for a controlled evaluation of learning effectiveness in domains where structured reasoning and feedback are essential, making *Speeding* a relevant case study for analyzing the adoption and sustainability of AI-driven tutoring systems in higher education.

4.2 Business model evolution

The development of *Speeding,* initially called *"Speeding Corsi",* has been characterized by a progressive evolution of its business model, driven by scalability constraints and by the opportunities enabled by recent advances in artificial intelligence.

In the live-course phase, the company followed a traditional B2C model: students paid for intensive preparatory courses in Calculus and related exams, delivered in small groups and supported by human tutors. This approach produced strong learning outcomes and positive word-of-mouth, but unit economics were constrained by the high marginal cost of instructor time and by the difficulty of scaling beyond a limited number of parallel classes. From the perspective of the thesis, this phase serves as a baseline representing "pre-AI" tutoring: effective but labour-intensive and accessible only to students who can afford private courses.

With the launch of *Pepe* in July 2025, the model shifted towards a vertical B2C AI tutoring platform, where the core of the value proposition became "an AI tutor specialised in your hardest exam, available 24/7, at a fraction of the cost of private lessons", reflecting broader trends in AI adoption in higher education discussed in Chapter 3. During the freemium period, *Speeding* focused on acquisition and engagement, managing to reach its maximum peak 279 monthly active users. These users registered a 68% upload material conversion (meaning that student desire personalization during the study session) and deep usage in roughly 38% of sessions, while maintaining an infrastructure cost of about 0.94€ per MAU. In this phase, *Pepe* was treated as an investment in product–market fit and data collection.

Starting in January 2026, *Speeding* introduced a hybrid monetisation model combining pay-per-use credit packs (priced around 1.99€ for small bundles) and monthly subscriptions positioned initially in the 5€/month range. Early results, analysed in Section 4.4, show that only 13 distinct customers made 18 payments in the first three months, revealing strong price sensitivity and confirming survey evidence that the average willingness to pay is closer to 3-5€/month than to the original target price. At the same time, the general survey indicates significant interest in a potential hybrid tier combining AI tutoring with occasional human

sessions: 21.5% of students would use such a tutor frequently, 14.7% occasionally, and 48.6% "depending on price", which opens a path towards higher-value offerings for students facing high-stakes exams.
In summary, *Speeding*'s business model has evolved from a human-intensive tutoring provider to an AI-first vertical platform that aims to balance accessibility (low marginal cost and freemium access), pedagogical effectiveness (exam-specific step-by-step support) and economic sustainability (through credits, subscriptions and potentially hybrid AI-human services). The following sections of this chapter will show that while the pedagogical and cost foundations of this model are solid, further work is needed on pricing, product scope and long-term value capture to reach full sustainability.

4.3 Experimental validation: does it work?

This section try to empirically testing whether the *Speeding* model is effective and economically viable in the current context of AI adoption in higher education. Building on the trends discussed in Chapters 2 and 3, three working hypotheses are validated through platform analytics (internal data of the company), two structured surveys and direct user interviews.

H1 – Pedagogical effectiveness: LLM-based AI tutoring can provide effective adaptive learning support, comparable to or better than traditional alternatives such as private tutoring, YouTube videos, or generic chatbots like ChatGPT.

H2 – Willingness to pay: An accessible price point around 10€/month is sufficient to sustain a B2C model, direct to students in this case.

H3 – Retention and unit economics: A vertical focus on gatekeeper STEM exams (especially Calculus 1-2) combined with personalization and continuous usage should generate retention levels compatible with positive unit economics.

To test these hypotheses, the analysis combines three data sources:

- **1.** A large-scale survey of 251 STEM students about difficult exams, study habits, and willingness to pay for AI tutors;
- **2.** A focused survey of 25 "power users" (at the moment) of *Pepe* with ≥ 5 sessions, capturing perceived effectiveness and behavioural intention; and
- **3.** Product analytics and paid-user data from the *Pepe* platform between July 2025 and January 2026.

4.3.1 Data sources and sample

The first dataset used in the experimental analysis consists of a structured survey administered to university students enrolled in STEM degree programs. The survey collected 251 valid responses and was distributed through university WhatsApp groups, ensuring rapid dissemination among active student communities.

Respondents mainly came from the bachelors of three Italian universities: University of Catania, University of Florence, and University of Bologna, providing a heterogeneous yet coherent sample of students exposed to similar academic challenges in STEM disciplines. The sample is strongly skewed towards Engineering (76.9%), with the remainder mostly in Economics/Management (22.7%), which matches the initial go-to-market focus of *Speeding*.

The questionnaire was fully anonymous, and no personal or identifying information was collected. Participation was voluntary, and respondents were informed that the data would be used exclusively for academic research purposes. The survey aimed to investigate students' study habits, perceived difficulties in exam preparation, current use of artificial intelligence tools in learning activities, what they dislike about existing AI solutions, and how much they would be willing to pay for a specialised AI tutor. The dataset provides a broad contextual framework to understand the role of AI in higher education from the students' perspective, independently from the specific use of the *Speeding* platform.

4.3.2 Results – AI usage and study habits among stem students (H1)

The survey results confirm that exam preparation in STEM disciplines is widely perceived as challenging. A significant share of respondents reported difficulties related not only to time management, but also to understanding theoretical concepts and solving complex exercises. This finding supports the idea that traditional study resources often fail to provide sufficient personalized guidance, especially in procedural and problem-solving intensive subjects.

When asked which exams they find most difficult, 59.4% of respondents mention Calculus 1-2, 40.2% indicate Physics 1-2, and 35.9% cite Linear Algebra and Geometry. These three courses clearly emerge as gatekeeper exams that block progression and contribute to dropout, and they are precisely the subjects covered by *Pepe*'s initial content roadmap. Moreover, survey evidence confirms that *Speeding* is targeting a real and concentrated pain point in Italian STEM education.

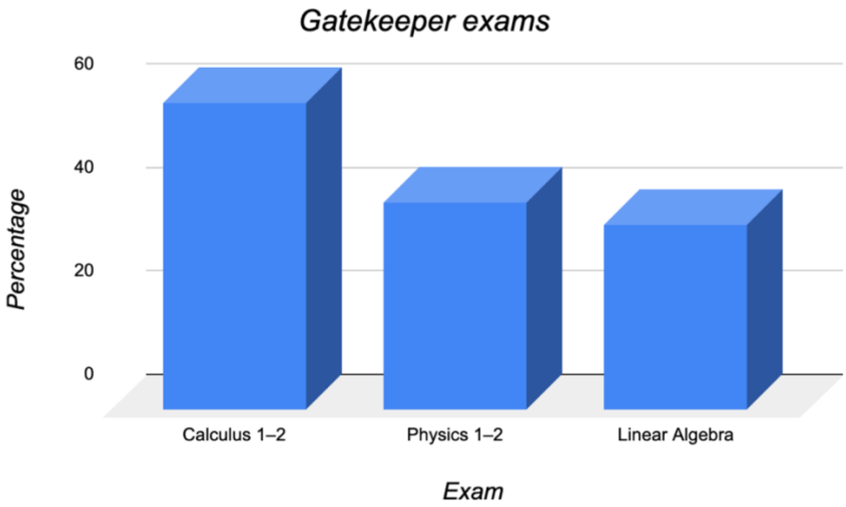

Fig. 6 Most difficult STEM "gatekeeper" exams according to surveyed students (n = 251)

A large proportion of respondents declared that they already use artificial intelligence tools as part of their study routine. This indicates that AI adoption in higher education is not a future hypothesis but an ongoing phenomenon, particularly among students seeking immediate feedback and clarification during individual study sessions.

Despite the widespread adoption of AI-based tools, many students reported that existing solutions do not fully address their learning needs. Generic AI systems and video-based resources are often perceived as useful for quick explanations but insufficient for structured, step-by-step problem solving and personalized learning paths. This gap highlights the limitations of horizontal tools when applied to complex academic contexts.

In terms of preparation strategies, students already employ a multi-tool mix. The majority rely on traditional materials such as textbooks and lecture slides (82.5%), but a very large share also uses AI tools (76.1%) and YouTube or other video resources (55.4%) to cope with difficult topics. Only 8% report taking private lessons, and 12.7% use university tutoring services, which suggests that high-touch human support is either too expensive or not sufficiently accessible at scale. This combination of high difficulty in core quantitative exams and reliance on fragmented, often generic, tools motivates the need for a more structured and specialised AI-driven tutoring solution such as *Speeding*, and its AI-based product.

Although almost all respondents have already experimented with AI for studying (98.4% report having used tools such as *ChatGPT*, *Gemini* or *Copilot* at least once), satisfaction with these generic solutions is far from perfect. The survey reveals three dominant pain points:

1- Trust and correctness**:** 66.5% state that they "do not know whether the answers are correct", indicating a lack of reliable verification and transparency in the reasoning process.

2- Mathematical accuracy: 62.2% complain that AI "makes errors in calculations or formulas", which is critical in subjects such as calculus and linear algebra.

3- Lack of contextualisation: 45% describe explanations as "too generic and not specific to my course", pointing to a gap between general-purpose large language models and the local, exam-oriented syllabi of Italian universities.

Additional, but less frequent, issues include responses that are too long or complicated (14.3%), lack of feedback on learning progress (16.7%), and even feelings of guilt when using AI for studying (10.8%). These findings align with the design choices behind *Pepe*: vertical content aligned with specific exams, step-by-step solutions with explicit intermediate steps, and an emphasis on checking rather than blindly trusting AI outputs.

Eventually, data show a clear openness toward more structured AI-based tutoring solutions. Many respondents expressed interest in tools capable of guiding them through exercises in a personalized and interactive way, suggesting the presence of unmet demand for vertical AI tutors specifically designed for academic problem solving
At the same time, students expressed ambivalent attitudes toward the long-term impact of AI on learning. While many respondents believe that AI can enhance understanding and confidence, a non-negligible share expressed concerns about potential over-reliance on AI tools and reduced autonomy. This dual perception reflects an ongoing transformation in the meaning of learning "in a world with AI," where support and dependency coexist.

4.3.3 Willingness to pay for specialised AI Tutoring (H2)

The same survey asks students how much they would be willing to pay for an AI tutor that "explains exercises step-by-step, is specific to their subject, and adapts to their study path". Responses reveal a heterogeneous but price-sensitive demand. Overall, 36.7% of respondents would use such a tool only if it were free, while 63.3% declare some willingness to pay. Among paying students, 38.2% indicate a price range of 3-5€/month, 13.1% are willing to pay 6-10€/month, 4.4% would pay 11-15€/month, and 4.8% are ready to spend more than 15€/month.
By assigning mid-point values to each price bracket (4€, 8€, 13€, 20€ respectively), the average stated willingness to pay is approximately 4.11 €/month. Only 22.3% of students fall into the segment willing to pay at least 6€/month, which is the range of *Pepe*'s current pricing (around 10 €/month in the initial freemium-to-paid model). This result partially

supports H2: there is a non-negligible paying segment, but the size of the market at 10€/month is smaller than initially assumed, and pricing strategy becomes a critical variable for business model sustainability.

In addition to AI-only support, *Speeding* is exploring a hybrid model where students can book targeted sessions with human tutors who have access to their progress data on the platform. The survey tests demand for this option by asking whether respondents would use such a human tutor. Results indicate substantial latent interest: 21.5% would "use it often", 14.7% would "use it occasionally", and 48.6% answer "maybe, depending on the price". Only 10.4% explicitly prefer to rely solely on AI, and 4.8% state that they would rather go to their trusted private tutor. These figures suggest that a premium hybrid tier (combining *Pepe* as always available AI tutor with occasional human interventions) could be a viable upsell path, especially for students facing high-stakes exams such as Calculus 1-2.

Finally, the survey assesses activation potential by offering respondents free access to *Pepe*. 61.4% of students ask to "receive access to *Pepe*" and voluntarily provide their email or phone contact, while 38.6% decline the offer. This high opt-in rate confirms that the value proposition of a specialised AI tutor resonates strongly with the target audience.

However, later sections of this chapter show that actual conversion from interest to sustained usage and paid subscriptions is significantly lower, constrained by pricing, exam seasonality, and the "use-it-only-until-I-pass" behaviour typical of exam-centric products. In other words, H2 and H3 are only partially validated: students perceive clear value in a specialised AI tutor and are willing to try it, but the proportion of users ready to pay 10

€/month and to stay active across multiple exam cycles is limited, which has direct implications for unit economics and growth strategy.

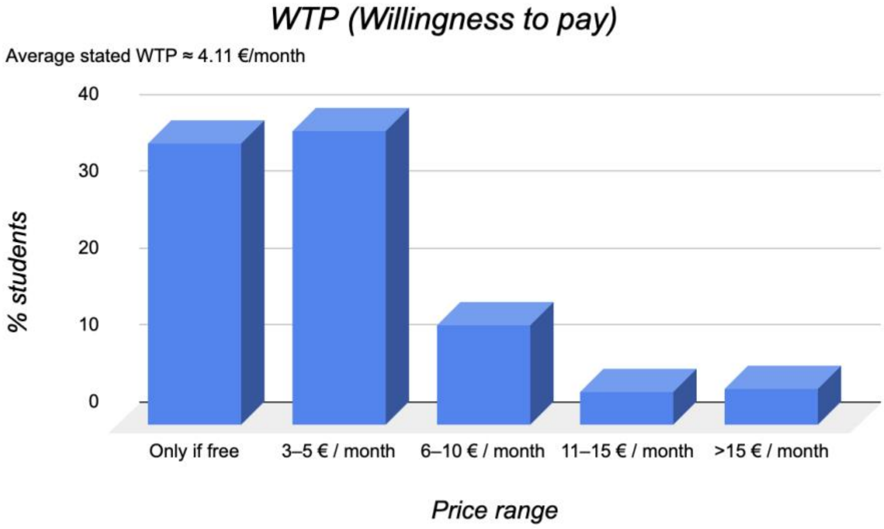

Fig. 7 Distribution of students' willingness to pay for a specialised AI tutoring platform (n = 251)

4.3.4 Retention and unit economics (H3)

This subsection analyses the first three months of monetisation on *Speeding*, combining paid user transactions with aggregate platform metrics in order to assess the economic dimension of the model and whether vertical focus and high engagement can translate into sustainable unit economics.

Speeding introduced a paid model on 6 January 2026, after operating as a free to use service since July 2025. In the monetisation period up to March 2026, the platform recorded 18 valid payments, made by 13 distinct paying customers, generating a total revenue of 85.10€. Transactions are split between pay-per-use credit packs and subscriptions: 12 payments correspond to credit purchases, while 6 correspond to subscription plans. Importantly, 29.28€ of total revenue derives from renewals, indicating the emergence of recurring revenue behaviour rather than purely one shot transactions.

The revenue distribution remains relatively concentrated, with a subset of highly engaged users responsible for repeated purchases and subscription renewals, while several users test the paid features through low value micro transactions (e.g. 1,99€ credit packs). This pattern is consistent with the willingness to pay results from the general survey, where most students indicate a price range of 3-5€ per month rather than the initially envisioned 10€ subscription model.

To contextualise these numbers, it is necessary to compare them with broader platform usage metrics. In the period 16 November-16 December 2025, *Speeding* counted 279 monthly active users (MAU), who generated 11.9k messages and consumed 182 million tokens, with a relatively low infrastructure cost per MAU thanks to efficient prompt design and caching mechanisms. Assuming a similar order of magnitude in active users during the monetisation phase, 13 paying customers still represent a conversion rate in the low single digits, confirming that only a fraction of engaged users currently transition to paid plans.

This limited conversion is coherent with survey findings: although 63.3% of students declare some willingness to pay for a specialised AI tutor, only 22.3% are willing to spend at least 6€ per month, and many *Speeding* users explicitly state that they would stop paying once they have passed the target exam. In other words, short term retention observed among power users does not automatically translate into long term recurring revenue, because the perceived value of the product remains strongly tied to a specific blocking exam rather than to the entire degree programme.

From a cost perspective, early unit economics remain structurally favourable. Total API costs of approximately 261.6€ over one month, combined with deep usage (37.9% of sessions contain more than 10 messages), imply a variable cost per active user below 1€. This confirms the theoretical advantage of LLM based tutoring: once the system architecture is established, marginal cost scales efficiently with usage intensity.

Revenue, however, still does not cover variable costs. With 85.10€ in total revenue and an estimated MAU base of approximately 270 users, implied ARPU (Average Revenue Per User) is around 0.30€ per MAU, which remains below the estimated variable cost per MAU (~0.94€), implying a negative contribution margin in the short term. Nevertheless, approximately one third of total revenue, signals an important structural shift compared to the very first monetisation month, suggesting that recurring behaviour is beginning to emerge.

Table 3. Summary of the monetisation metrics

Metric	Value
Total Payments	18
Paying Customers	13
Total Revenue	€85.10
Revenue from Renewals	€29.28
Renewal Share of Revenue	34%
Credits (Pay-per-use)	12
Subscriptions	6
Estimated MAU	~279
Estimated ARPU	~€0.30
Variable Cost per MAU	~€0.94

The data therefore lead to a partial validation of H3. On the one hand, the vertical focus on Calculus and related STEM exams, combined with personalised step by step explanations, clearly generates high engagement and satisfaction among a subset of users, as confirmed by the power user survey and by the share of deep sessions. On the other hand, conversion rates remain limited and average spending aligns more closely with the 3-5€ per month bracket indicated by the general survey than with the initially targeted 10€ subscription price.

In practical terms, *Speeding*'s data suggest that an AI tutor like *Pepe* can deliver pedagogical value at low marginal cost, while early signs of recurring revenue indicate improving retention dynamics. However, the business model still requires adaptation through pricing experimentation (e.g. lower monthly fees or exam specific bundles), broader subject coverage to mitigate seasonality, or premium hybrid features (such as human tutoring sessions) in order to reach full economic sustainability. These implications are further discussed in Section 4.5

4.4 Result and analysis

While Section 4.3 presented the results of the two surveys at an aggregate level, this section deepens the analysis through the follow-up survey administered to *Speeding*'s power users (students who engaged with *Pepe* in at least five sessions). The goal is to move beyond stated preferences and capture perceived effectiveness, emotional impact, and behavioural intentions from those who have directly experienced the platform. Section 4.4.1 focuses on H1 (pedagogical effectiveness and transferable skills), while Section 4.4.2 addresses H3 (retention and willingness to pay among existing users).

4.4.1 Perceived effectiveness, impact of confidence, transferable skills – linked to H1

Overall, survey results indicate a high level of pedagogical effectiveness. On a 1-5 Likert scale, users report an average score of 4.25 for the statement "*Pepe* helps me understand exercises better than before", with 22 out of 24 respondents choosing values 4 or 5. Similarly, the clarity of explanations is rated 4.42 on average, and half of the sample gives the maximum score of 5 to the item "*Pepe's* explanations are clear and help me work through the steps on my own".

When asked to compare *Pepe* with alternative tools, users perceive a strong advantage over generic AI and video resources. In terms of exam preparation, *Pepe* scores 4.33 on average against ChatGPT and other AI tools and 4.00 against YouTube or other online videos (1–5 scale, where 3 = "same usefulness"). The comparison with private tutoring is more balanced (mean 3.25), suggesting that human one-to-one support still retains marginal advantages but at a significantly higher price point.

Qualitative comments reinforce this picture: users frequently cite "step-by-step explanations", "exam-specific language" and "correction with targeted exercises" as the most valuable features of *Pepe*.

These findings support H1: for students who actually adopt the tool and use it intensively, LLM-based tutoring can provide an effective and credible alternative to traditional resources, especially when compared to generic chatbots and asynchronous content such as videos.

Beyond comprehension, *Pepe* appears to have a positive but more moderate impact on students' emotional experience of studying. The average score for "Using *Pepe* I feel more confident in my exam preparation" is 3.88, with the distribution skewed towards 4 and 5 but with a non-negligible minority choosing 2 or 3. The item "*Pepe* helps me reduce anxiety and mental blocks while studying" scores slightly lower, at 3.58 on average, with responses spread across the full 1–5 range. This suggests that while *Pepe* is effective in clarifying content and procedures, emotional support and anxiety management are only partially addressed, leaving room for future features such as progress tracking, reassurance messages, or integration with human tutors for high-stakes moments.

A central question for this thesis is whether AI-based tutoring fosters transferable problem-solving skills or merely facilitates "shortcut learning". On this dimension, results are encouraging but nuanced. The statement "Using *Pepe* I feel I really understand the steps, not just copy a solution" receives an average score of 4.33, with 21 out of 24 users selecting 4 or 5. This indicates that most power users perceive the interaction with *Pepe* as an aid to understanding rather than a pure answer-generator. When explicitly asked whether "Using AI for studying helps me develop skills that will also be useful at work (reasoning, problem solving, etc.)", the mean score drops to 3.71. The distribution is more dispersed: while 13 users select 4 or 5, 10 choose 3 and one assigns the minimum score. This pattern suggests that a substantial share of students believes that AI-supported study contributes to long-term skill development, but this perception is not yet universal.

Interestingly, concerns about loss of autonomy are present but not dominant. The item "I am afraid that using AI too much could make me less autonomous in studying" has an average of 3.04, with responses almost evenly spread across the scale. This ambivalence reflects the tension highlighted in Chapter 3: students are simultaneously attracted by the efficiency gains provided by AI and worried about over-reliance and superficial learning.

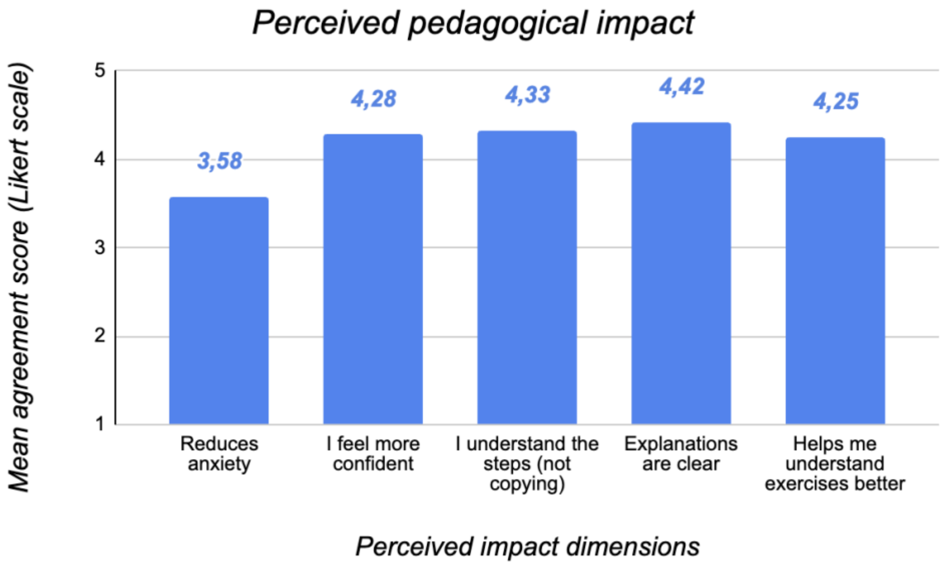

Fig. 8 Perceived pedagogical impact among Speeding's power users (Likert scale)

For *Speeding*, this underlines the importance of designing *Pepe* explicitly as a learning companion that forces users to engage with intermediate steps and encourages them to attempt exercises independently, rather than as a mere solution provider.

Taken together, these results provide empirical support to the claim that *Pepe* can contribute to "learning in a world with AI" rather than simply optimising performance in a pre-AI assessment regime, although further work is needed to strengthen the link between AI-assisted study and long-term professional competences.

4.4.2 Usage intensity, retention and willingness to pay among power users – linked to H3

Complementing the effectiveness data presented in Section 4.4.1, this subsection examines how power users actually engage with *Pepe* over time, how often they use it, whether they intend to continue, how much they are willing to pay, and the development of transferable skills in an AI-rich learning context.

From a behavioural perspective, power users exhibit high engagement with the product: 18 out of 24 respondents report using *Pepe* at least once a week, and 18 of them use it 3-4 times per week or daily. Only one user declares having stopped using the tool, and another reports using it "rarely (less than once a week)". This confirms that, for a subset of students, *Pepe* becomes an integral part of their study routine during the preparation phase for difficult exams.

Retention intention is also strong but not absolute. On a 1–5 scale, the average response to "How likely is it that you will continue to use *Pepe* for other exams?" is 3.96, with 18 out of 24 respondents choosing 4 or 5. Net Promoter Score is very high: the mean likelihood of recommending *Pepe* to a classmate is 8.83/10, with 10 users giving a score of 10 and only one detractor (score ≤ 6). These indicators suggest that, conditional on initial adoption and satisfaction, both loyalty and word-of-mouth potential are considerable.

At the same time, the reasons given for potential churn highlight a structural limitation of exam-centric products. The most frequent answer to the question "What is the main reason why you might stop using *Pepe*?" is "I won't need it anymore once I have passed the exam", followed by "Lack of time / change in exam period". This confirms that many students perceive *Pepe* primarily as a transactional tool to overcome a specific exam, rather than as a long-term study companion across the whole degree. As discussed later in this chapter, this has direct implications for H3: while verticalization and personalization are sufficient to achieve high short-term retention during exam preparation, sustaining engagement across multiple semesters requires either broader subject coverage or additional features that create longitudinal value (e.g. planning, tracking, cross-course knowledge).

The survey also explores willingness to pay among users who already know and appreciate *Pepe*. Results are more favourable than in the general student sample but still reveal sensitivity to price. When asked "If *Pepe* became paid, how much would you be willing to pay per month?", 8 respondents select "up to 5€/month", 2 choose "up to 10€/month" and 1 chooses "up to 15€/month". Five respondents explicitly state that they would not be willing to pay, and another five answer "maybe, it depends on the period". These figures indicate that existing satisfied users have a higher WTP than the general, but they still confirm that a 10€/month price point addresses only a subset of the potential market. For H3, this means that even if retention during exam periods is strong, the proportion of users converting to paid plans will be constrained by affordability considerations, pushing *Speeding* to experiment with alternative pricing schemes (lower monthly price, credit-based system, exam-specific packs).

4.5 Critical success factors

The empirical evidence from surveys and platform analytics suggests that the viability of *Speeding*'s business model depends on a small set of critical success factors (CSFs). These are not isolated features, but interdependent conditions: some are already strongly validated by data (product market fit and cost structure), while others emerge as constraints that must be addressed to reach full sustainbaility (pricing, seasonality, perceived long-term value). These five CSFs identified below bridge the empirical evidence from Section 4.3 and 4.4, with the strategic discussion in Chapter 5.

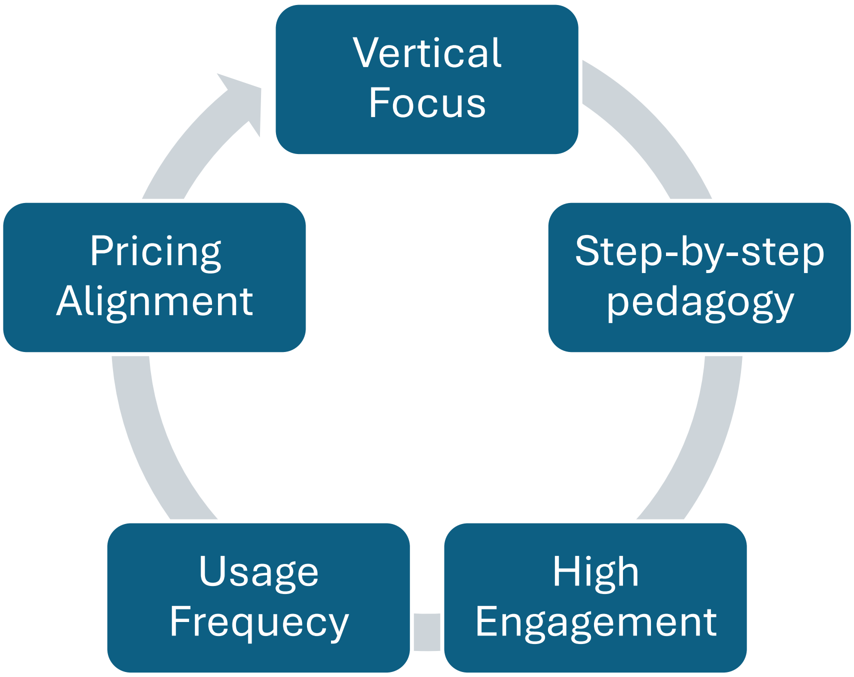

Fig. 9 Critical success factors for sustainable AI tutoring

4.5.1 Vertical focus on "gatekeeper" STEM exams

A first critical success factor is the decision to focus *Speeding* on a narrow set of highly problematic exams, rather than trying to cover the entire university curriculum from day one.

As established in Section 4.3.2, the survey confirms that Calculus 1-2, Physics 1-2 and Linear Algebra are the three dominant gatekeeper exams. This aligns between product scope and student pain points, since these three courses concentrate a disproportionate share of failures and delays. For this reasons, they coincide with the subjects initially addressed by *Pepe's* content roadmap.

This alignment between perceived difficulty and product coverage generates a clear problem solution fit: users recognise *Pepe* as a tool designed exactly for the exams that block their progression, rather than as a generic AI assistant. In a crowded landscape where almost all students already use general-purpose AI tools (98.4% of survey respondents report having used ChatGPT, Gemini or similar for studying), such verticalization is essential to differentiate *Speeding* and to justify any form of payment.

4.5.2 Step-by-step pedagogy and trust in the AI

A second CSF is the pedagogical design of *Pepe* as a step-by-step tutor rather than as a simple answer generator. The survey of 251 students reveals three dominant pain points with generic AI tools: 66.5% "do not know whether the answers are correct", 62.2% report "errors in calculations or formulas", and 45% complain that explanations are "too generic and not specific to my course". These issues undermine trust and make it risky to rely on large language models for high-stakes exams in mathematics and physics.

Among *Pepe's* power users, the data indicate that this design choice is effective. On a 1–5 scale, respondents give an average score of 4.25 to the statement "*Pepe* helps me understand exercises better than before" and 4.42 to "*Pepe's* explanations are clear and help me work through the steps on my own", with the vast majority selecting values 4 or 5. When compared directly to alternatives, *Pepe* scores 4.33 against ChatGPT and other AI tools and 4.00 against YouTube in terms of usefulness for exam preparation, whereas the advantage over private tutoring is more modest (3.25).

Moreover, users tend to perceive *Pepe* as supporting genuine understanding rather than mere copying: the item "Using *Pepe* I feel I really understand the steps, not just copy a solution" receives an average rating of 4.33. This is strategically important, because it positions *Speeding* not only as a tool for passing exams in the current system, but also as a platform that can help students build robust problem-solving habits in an AI-rich environment, an aspect emphasised in the literature on AI literacy and 21st-century skills.

4.5.3 Engagement depth and favourable cost structure

From an operational standpoint, a third CSF is the combination of deep engagement and low marginal cost per user. A significant portion of interactions is intense, since 37.9% of all sessions contain at least 10 messages, indicating that users often work through full exercises or problem sets rather than asking isolated questions.

Despite this depth of usage, the variable infrastructure cost remains modest, at approximately 0.94€ per MAU for the month. From the unit economics perspective, it means that the cost side of the model is structurally favourable: once the platform and content are in place, the system can handle high volumes of tutoring interactions without incurring the labour costs associated with human tutors.

In the terminology of Chapter 3, *Pepe* already exhibits the cost characteristics of a scalable AI-native tutoring platform, where the main bottleneck is not infrastructure but the ability to convert engagement into revenue and to smooth demand over time.

4.5.4 Monetisation constraint: pricing, seasonality and “one shot usage”

At the same time, monetisation evidence highlights structural constraints that are as relevant as the strengths identified above. Survey results clearly indicate that willingness to pay remains below the initially assumed threshold: the average stated WTP for a specialised AI tutor is approximately 4.11€ per month, with over one third of students willing to adopt the service only if free and less than one quarter prepared to pay at least 6€ per month. Even among highly engaged power users, the majority position themselves within the “up to 5€ per month” range, with only a limited minority expressing comfort with higher subscription tiers. This confirms that price sensitivity is not limited to marginal users but extends to the core segment.

Monetisation dynamics reflect this constraint. Although recurring revenue has begun to emerge, signalling early signs of retention beyond one-shot transactions, total revenue remains modest relative to the active user base. The gap between expressed interest in a specialised AI tutor and effective paid conversion suggests that pricing architecture, perceived value alignment, and timing of purchase decisions remain central strategic variables. The issue is therefore not demand for the product per se, but the monetisation of that demand at the current price structure.

In addition, usage patterns display pronounced seasonality. Both survey responses and internal cohort analysis indicate that many students perceive the platform primarily as a tool to overcome a specific blocking exam, rather than as a longitudinal study companion across the entire degree programme. Weekly activity tends to decline after exam sessions, reinforcing the interpretation that demand is concentrated around assessment windows rather than evenly distributed throughout the academic year.

Taken together, these elements imply that H3 remains only partially validated. Vertical focus and personalisation generate strong engagement and favourable cost metrics, and early renewal behaviour suggests improving retention dynamics. However, exam-centric usage and persistent price sensitivity limit conversion and recurring revenue potential. Achieving positive unit economics will therefore require refinement of the pricing model, expansion of subject coverage to smooth seasonality, or the introduction of higher-value hybrid configurations capable of increasing perceived long-term utility.

4.5.5 Future-proofing: from exam preparation to AI-enabled competence

A final critical success factor, still emerging but strategically important, concerns how *Pepe* contributes to skill development in a “world with AI”. The power-user survey provides early evidence that students do not perceive *Pepe* merely as a shortcut: as noted, most users feel they understand the steps rather than copying solutions, and about half of them agree that using AI to study helps them develop reasoning and problem-solving skills that will also be useful in their future work.

At the same time, the average score on this latter item (3.71/5) and the mixed responses regarding fear of losing autonomy (mean 3.04/5) show that this dimension is not yet fully realised. The literature on AI in education emphasises that long-term value will depend on helping students learn how to use AI critically and productively, rather than simply replacing manual calculation or memorisation. For *Speeding*, this implies that future iterations of *Pepe* should make the development of transferable competences (such as problem decomposition, verification of AI outputs, and comparison of multiple solution paths) an explicit design goal.

In summary, the analysis identifies five main critical success factors for the *Speeding* case:

1- Tight vertical focus on high-impact STEM gatekeeper exams (strongly validated).
2- Step-by-step, exam-specific pedagogy that restores trust in AI explanations (strongly validated).
3- High engagement depth combined with very low marginal costs per user (validated).
4- Significant constraints in monetisation due to price sensitivity, seasonality and exam-centric usage (validated, implies need for strategic adjustment).
5- An emerging but still incomplete role in building AI-enabled, transferable problem-solving skills (partially supported, key area for future development).

These factors provide the bridge between the empirical case study and the broader discussion in Chapter 5 about which configurations of AI tutoring, pricing and product scope are most likely to yield sustainable and future-proof business models in higher education.

4.6 Sustainability assessment

This section synthesises the findings of Sections 4.3-4.5 into a unified sustainability assessment. It evaluates the *Pepe* business model along two dimensions, pedagogical effectiveness and unit economics. The purpose is to provide a clear, evidence-based foundation for the strategic implications discussed in Chapter 5.

4.6.1 Pedagogical sustainability

From a learning perspective, the data indicate that *Pepe*
delivers consistently high value to its core users. As detailed in Sections 4.4 and 4.5.2, power users consistently rate *Pepe*'s pedagogical effectiveness between 4.25 and 4.42/5, with a Net Promoter Score of 8.83/10 and strong perception of genuine understanding rather than shortcut learning. On this basis, the pedagogical side of the model appears sustainable: there is strong evidence that the product improves perceived understanding and satisfaction, and nothing in the cost structure prevents further scaling of tutoring interactions.

4.6.2 Unit economics sustainability

Economic sustainability represents the most delicate dimension of the current business model. While engagement metrics and qualitative feedback confirm strong pedagogical value, revenue generation remains structurally weaker than cost scalability.
The first months of monetisation data show initial revenue growth and the emergence of subscription renewals, indicating that a subset of users is willing to move beyond one-off micro-transactions. However, overall revenue density per active user remains insufficient to offset variable

infrastructure costs. While the technological architecture scales efficiently, the monetisation does not yet scale at the same pace and this imbalance is consistent with behavioural evidence discussed earlier. The platform is primarily perceived as an exam-solving tool rather than as a continuous academic companion and, as a consequence, demand is concentrated around specific assessment periods and therefore weakens once the target exam is passed. At the same time, willingness to pay remains clustered in lower price brackets, limiting average revenue per user even among satisfied students.

The core issue is therefore not product relevance, nor cost structure, but the episodic nature of demand combined with price sensitivity. Engagement is strong, but temporally concentrated. Monetisation exists, but lacks depth and continuity.

Taken together, these elements suggest that the current configuration is economically viable yet not fully sustainable: the cost side validates the scalability of an LLM-based tutoring model but the revenue side still requires strategic refinement. Economic viability will depend on increasing revenue per active user through better alignment between pricing logic, product scope, and user behaviour.

Potential directions include designing monetisation mechanisms that match exam cycles more precisely, broadening subject coverage to reduce seasonality effects, and strengthening the perceived long-term value of the platform beyond single-exam preparation. The path to sustainability does not require technological reinvention, but a more coherent integration between pedagogy, timing, and pricing.

4.6.3 Overall sustainability

In conclusion, the sustainability assessment for the *Speeding* case can be summarised as follows: pedagogically sustainable, economically pre-sustainable. *Pepe* already demonstrates that an AI-native tutoring platform can deliver high-quality support at low marginal cost and with strong user satisfaction, but the current combination of pricing, product scope and exam-centric usage does not yet yield positive unit economics. The remainder of the research discusses how these insights can inform the design of more robust and future-proof AI business models in higher education.

5. CONCLUSION AND STRATEGIC IMPLICATION

This thesis set out to answer a specific question: Can a vertically specialized AI tutoring platform such as *Speeding* deliver effective learning support for university STEM students while achieving sustainable unit economics in the current landscape of AI adoption? Through a combination of literature review, empirical case study, surveys of 251 STEM students and 24 power users, and internal platform analytics from *Speeding*, the research provides a nuanced answer: yes, but not yet in its current configuration. The findings confirm that AI-based tutoring can address a real and concentrated pain point in this sector, delivering pedagogical value that students perceive as superior to generic chatbots and comparable to human tutoring at a fraction of the cost. However, the path to economic sustainability requires deliberate business model innovation, pricing adjustments, and a shift in product strategy from exam-centric survival tools toward platforms that cultivate long-term AI-enabled problem-solving competences.

This concluding chapter synthesizes the main findings, discusses their implications for EdTech strategy and policy, acknowledges the limitations of the study, and outlines directions for future research and product development.

5.1 Validation of research hypothesis

The thesis tested three working hypotheses through mixed-methods empirical analysis:

H1 – Pedagogical Effectiveness: LLM-based AI tutoring can provide effective adaptive learning support, comparable to or better than traditional alternatives such as private tutoring, YouTube videos, or generic chatbots.

This hypothesis is strongly supported and these results are consistent with the literature on Intelligent Tutoring Systems, which highlights the importance of adaptive feedback and scaffolding. The full numerical evidence is presented in Section 4.4.1, where we note that students reported improved clarity, better understanding of intermediate steps, and increased confidence when approaching complex exercises.

H2 – Willingness to Pay: An accessible price point around 10€/month is sufficient to sustain a direct-to-student B2C model.

Results related to H2 are more complex. While students express appreciation for the pedagogical value of the platform, their stated willingness to pay remains relatively low. This reflects a structural tension in the current AI landscape, where students compare specialised educational platforms not only to human tutors, but also to free general-purpose AI tools.

H3 – Retention and Unit Economics: A vertical focus on "gatekeeper" STEM exams combined with personalization and continuous usage should generate retention levels compatible with positive unit economics.

This hypothesis is partially validated: short-term engagement is strong, but long-term retention and economic sustainability remain constrained by exam-centric usage patterns and seasonality. Since retention is linked to academic cycles rather than platform engagement dynamics alone, sustainability depends not only on cost structure but on product positioning, subject expansion, and frequency of use beyond single-exam preparation.

Overall, the findings suggest that vertically specialised AI tutoring platforms can generate meaningful learning value, but their long-term economic sustainability requires careful alignment between pedagogical differentiation and business model innovation.

5.2 Theoretical implication

From a theoretical perspective, this study contributes to the ongoing debate on AI in higher education in three main ways.

First, it confirms that LLM-based tutoring systems can replicate some of the core mechanisms traditionally associated with Intelligent Tutoring Systems, particularly step-by-step feedback and personalised explanation. However, it also highlights that the effectiveness of AI tutoring is not solely dependent on the underlying model, but on the design layer, such as structure, prompting architecture, exam-specific adaptation, and pedagogical framing.

Second, the research highlights a growing distinction between generic AI assistance and vertically integrated AI learning environments. While general-purpose chatbots provide flexibility, specialised systems can create contextual relevance, domain-specific guidance, and consistent pedagogical scaffolding. This distinction may become increasingly important as AI tools become commoditised.

Third, the study reinforces the idea that AI in education cannot be evaluated exclusively through learning outcomes or exclusively through economic metrics. Pedagogical effectiveness and economic viability are interdependent. An AI tool that improves learning but fails to monetise may not survive long enough to scale. Conversely, a profitable tool that does not improve learning outcomes risks reinforcing superficial engagement or academic shortcut behaviours.

The case of *Speeding* demonstrates that the strategic positioning of AI in higher education requires simultaneous consideration of instructional design, behavioural economics, and platform sustainability.

5.3 Managerial implication for EdTech firms

For AI-driven educational startups such as *Speeding*, the findings suggest several strategic implications.

1. Vertical specialisation as differentiation
   In a context where general AI tools are widely accessible and often free, differentiation must come from contextual depth rather than technological novelty. Exam-specific content, structured reasoning paths, and alignment with local curricula can justify premium positioning.
2. Beyond exam-centric usage
   Seasonal usage tied to exam preparation constrains recurring revenue. Platforms should explore complementary features such as long-term skill development, cross-course integration, or AI-supported metacognitive training to increase usage frequency.
3. Pricing Innovation
   Traditional subscription models may not fully align with student expectations in a freemium AI environment. Alternative models such as pay-per-exam bundles, micro-credits, or institutional partnerships may better match perceived value. Evidence from adjacent markets (language learning, coding bootcamps) suggests that outcome-linked

pricing (e.g. "pass-or-refund" guarantees, may further reduce perceived risk and increase conversion among price-sensitive segments).

4. From assistance to skill formation
   A key long-term opportunity lies in shifting from "help me solve this exercise" toward "help me develop AI-enhanced reasoning skills." This aligns with the broader transformation of labour markets and may increase both perceived value and institutional interest.

5.4 Limitation of the study

While this thesis provides valuable insights into the viability of vertical AI tutoring platforms, several limitations should be acknowledged

1- Limited sample size and geographic scope

The primary survey includes 251 respondents, and the power-user survey includes only 24 participants, all from Italian universities (primarily Catania, Florence, and Bologna). This sample is not representative of the global higher education landscape, and findings may not generalize to other national contexts with different exam structures, tuition models, or AI adoption patterns. Future research should validate these findings across larger, more diverse samples.

2- Short observation period for monetisation data

The thesis analyses only the first month of paid operations (January 2026), which may not capture longer-term trends in conversion, retention, or revenue stabilisation. Seasonality effects, cohort maturation, and pricing adjustments could significantly alter unit economics over subsequent quarters. Longitudinal studies tracking multiple exam cycles and cohorts are needed to assess true sustainability.

3- Self-reported data and selection bias

Survey responses rely on self-reported perceptions of effectiveness, willingness to pay, and skill development, which may be subject to social desirability bias or inaccurate self-assessment. Additionally, the power-user sample is inherently biased toward students who found value in *Pepe* and continued using it, potentially overstating effectiveness relative to the broader population. Future studies should complement surveys with objective learning outcome measures (e.g. exam grades, time-to-degree completion) and controlled experiments comparing AI-tutored versus non-AI-tutored students.

4- Single-case study design

The empirical analysis focuses exclusively on *Speeding*, an Italian EdTech startup targeting STEM most tough exams. While this provides depth and context, it limits generalisability to other AI tutoring platforms with different business models, pedagogical approaches, or target markets. Comparative case studies of multiple platforms (e.g. Khanmigo, Duolingo Max, Chegg AI) would strengthen external validity.

5- **Lack of long-term skill transfer evidence**

The thesis finds preliminary evidence that power users perceive *Pepe* as supporting genuine understanding rather than shortcut learning, but it cannot yet demonstrate whether this translates into durable skill development, workplace performance, or retention of knowledge beyond exam periods. Longitudinal research tracking graduates into professional contexts is needed to validate claims about AI-enabled competence development.

Finally, the rapid evolution of AI technologies means that competitive and regulatory conditions may shift significantly in the coming years, potentially altering the strategic landscape for educational AI platforms.

5.5 Future research direction

Future research could expand along several dimensions:

- Conducting longitudinal studies to assess actual performance improvement linked to AI tutoring.
- Comparing vertically specialised AI tutors with general-purpose AI tools in controlled experimental settings.
- Investigating the long-term cognitive effects of AI-assisted study habits.
- Exploring institutional adoption models and B2B integration of AI tutoring within university ecosystems.
- Additionally, further inquiry is needed into how AI tools shape metacognition, autonomy, and transferable problem-solving skills.

5.6 AI, competence, and the "World with AI"

Beyond the immediate empirical findings, this thesis also raises a broader strategic and philosophical question: does AI merely help students succeed within traditional educational frameworks, or does it reshape the very meaning of competence?

Currently, AI tutoring platforms such as *Speeding* primarily support students in navigating existing assessment structures. In this sense, AI acts as an accelerator within a pre-AI institutional model. However, as AI becomes embedded in professional environments, the relevant competences of graduates may evolve.

The challenge for educational AI is therefore twofold:

1. Support students in mastering traditional academic content.
2. Simultaneously foster the ability to operate effectively "in a world with AI."

This may require explicit design for AI collaboration skills, critical evaluation of machine-generated outputs, and reflective understanding of when and how to rely on automated systems.

If educational AI remains confined to optimising exam performance, it risks reinforcing short-term instrumental learning. If instead it evolves toward competence amplification and cognitive partnership, it may become a transformative infrastructure for higher education.

5.7 Final remarks

Artificial Intelligence is rapidly reshaping the educational landscape and this thesis demonstrates that vertically specialised AI tutoring platforms can generate significant perceived pedagogical value and operate with scalable cost structures. However, economic sustainability depends on strategic differentiation and behavioural alignment in a market increasingly saturated with free AI alternatives.

The case of *Speeding* illustrates both the opportunity and the tension inherent in AI-driven education: technology can dramatically increase accessibility and personalisation, yet its value must be clearly articulated in a context where AI assistance is becoming ubiquitous.

Ultimately, the future of AI in higher education will depend not only on model performance, but on thoughtful integration between pedagogy, economics, and long-term human competence development.

Bibliography & Sitography

[1] https://www.wsgr.com/a/web/qrkz1SnNzWw6nk7B3oAyDa/10-things-you-should-know-about-the-eu-artificial-intelligence-act_v2.pdf

[2] https://www.ceeol.com/search/article-detail?id=1226718

[3] https://www.ceeol.com/search/article-detail?id=1360495

[4] https://aisberg.unibg.it/retrieve/93d826ed-5cdf-4c3f-9193-24f7ca0cdd80/CERLIS%20vol12_completo.pdf

[5] https://casopisi.junis.ni.ac.rs/index.php/FUTeachLearnTeachEd/article/view/13642

[6] https://posthumanism.co.uk/jp/article/view/1729

[7] https://www.mdpi.com/2227-7102/13/12/1216

[8] https://www.mdpi.com/journal/education

[9] https://www.sciencedirect.com/science/article/abs/pii/S1041608023000195

[10] https://arxiv.org/html/2507.18882v1

[11] https://www.mdpi.com/2227-7102/15/3/343

[12] https://www.researchgate.net/publication/264545127_Intelligent_Tutoring_Systems_and_Learning_Outcomes_A_Meta-Analysis/link/53e427b00cf21cc29fc8eecc/download

[13] https://link.springer.com/article/10.1007/s40593-021-00244-4

[14] https://link.springer.com/article/10.1007/s40593-021-00244-4