Autonomous Motivational Tutor Agent

1 Abstract

While Artificial Intelligence is changing education and triggering a great growth of agentic educational systems[1], current Intelligent Tutoring Systems (ITS) mainly focus on cognitive improvements, often ignoring students' emotions and mental health[1]. This lack of emotional understanding limits the long-term effectiveness of human-AI educational interactions. This proposal introduces a new "Sustainable AI Tutor" designed to offer personalized academic support while also tracking students' mental well-being. Therefore, we propose an architecture that includes five key parts: **(1)** an Affective Processing Unit (APU) to understand emotional signals; **(2)** a Hybrid Memory Store (HMS) that merges Knowledge Graphs with Vector Databases for effective Retrieval-Augmented Generation; **(3)** a Multi-Agent Deliberation Layer (MADL) that uses an "Inner Parliament" approach for realistic decision-making; **(4)** a Predictive Analytics Pipeline (PAP) for real-time risk evaluation; and **(5)** a Human Escalation Module (HEM) to safely direct high-risk situations to human experts. The proposed system effectively limits the LLM Conversational Core within safe and educational limits, ensuring that academic help does not risk psychological safety. By connecting personalized learning with emotional computing, this study offers a practical framework for safe, compassionate, and sustainable use of AI in education.

2 Introduction and background

2.1 Background Information

As AI is developing at a rapid pace, it has been widely integrated into the educational system. However, students who are learning in the real world are actually struck by anxiety about their future, current academic ability and feeling distant from the educational process.[1] Therefore, the current educational AI system must be beyond a simple task-oriented, but rather self-adjustable and provide students with non-cognitive skills: resilience and motivation.

Research strongly supports the application of conversational AI agents that use dialogue, empathetic responses, and behavior change techniques to build positive student-agent relationships.[2] In specialized populations, such as students with ADHD, four weeks of chatbot-aided training demonstrated superior outcomes in attention and emotional regulation compared to control groups.[2] This evidence establishes that the effectiveness of AI interventions lies in their capacity to actively train self-help skills and address the psychological roots of anxiety, thereby enhancing student resilience and persistence.

2.2 Theoretical Foundation

2.2.1 Self-Determination Theory (SDT)

The design of motivational AI agents must be grounded in psychological theory. Self-Determination Theory (SDT), developed by Deci and Ryan (2012), points out that human motivation is sustained by the satisfaction of three fundamental psychological needs: autonomy, competence, and relatedness.[5] SDT provides a clear framework for engineering AI interventions that foster intrinsic drive and persistence.

(1) **Competence Enhancement:** AI agents are good at supporting competence through adaptive strategies. By analyzing student work and retrieving common misconceptions, modern LLM can have feedback loops that are timely and specific.[3] Furthermore, agents can provide detailed explanations and examples, gradually shifting to hints and questioning as the student demonstrates mastery, thereby enhancing students' competence. [3]
(2) **Autonomy Support:** AI supports autonomy by moving beyond the traditional "one-size-fits-all" approach.[6] Agents personalize the learning experience by changing the teaching materials to adapt individual needs.[6] By addressing the need for autonomy, the system transforms learning from an obligation into an exploration, thereby supporting students' motivation.[6]
(3) **Relatedness through Empathy:** Relatedness requires the agent to be perceived as a supportive partner. Agents fulfill this by giving affective and empathetic dialogue to the students..[2] This capability is critical for achieving psychological outcomes, as evidence suggests that agents aiding emotional control are effective in improving emotional regulation skills.[2]

A recent study confirmed the mechanistic importance of SDT, demonstrating that quality of AI-driven feedback exerts direct effects on engagement and persistence.[8] Crucially, this feedback also influenced outcomes indirectly through intrinsic motivation and learner autonomy.[8] Thus, these researches confirm that SDT, as the fundamental theory for the AI-driven educational system is crucial and can be effective.

2.2.2 Integrating Growth Mindset

To build resilience, AI systems must build a growth mindset. Research suggests that AI feedback should be designed with autonomy-supportive prompts and competence-affirming messages to activate this growth mindset.[8] Furthermore, fostering persistence often requires addressing feelings of inadequacy, such as the fixed mindset that can lead to underperformance, particularly in non-inclusive learning environments.[9]

The approach to develop the growth mindset is affective computing.[10] Affective computing allows AI systems to recognize, interpret, and respond to human emotions, which allows the AI response to be more emotional and affectionate.[10] This capability allows the AI agent to detect motivational decline—such as frustration, anxiety, or early signs of a fixed mindset—enabling a prompt, relatedness-focused intervention consistent with SDT principles. This transition from logical task execution to emotional understanding is fundamental to next-generation psycho-educational agents.[10]

3 Project Goal and Solution

3.1 Personalized Tutoring

3.1.1 Goal 1 and Methodology

This educational system aimed to provide students with tailored tutoring using a retrieval-augmented generation (RAG), Knowledge Graphs, along with recommendation algorithms.

(1) **Hybrid Memory Architecture**: The convergence of LLMs and Knowledge Graphs (KGs) provides this solution. KG enables extracting structured knowledge from unstructured text with speed and efficiency.[12] This new efficiency makes KGs scalable for enterprise-level applications, including education.[12] Specifically, frameworks use Knowledge Graph Completion (KGC) by retrieving two types of knowledge—analogical and subgraph knowledge—to enhance the LLM's logical reasoning ability.[14] By integrating a chain-of-thought (CoT) prompting strategy, the model is guided to filter and rerank candidate entities based on this supplementary knowledge, which reduces omissions and incorrect responses typical of unconstrained LLMs.[14]

(2) **Retrieval-augmented Generation**: Using a retrieval-augmented generation, we connect the Large Language Models to the knowledge graphs. While standard RAG architectures rely on vector databases to perform searches, they often struggle to capture the complex relational structures, hierarchical dependencies, and multi-hop connections inherent in domain-specific data[13]. By integrating a KG, the system can leverage structured, human-curated knowledge to provide a more transparent and logically grounded context for the LLM.

(3) **Motivational Feedback**: We use self-determination driven feedback loops. The feedback would initially be detailed and comprehensive, followed by a fading into hints or brief introduction to promote autonomous study and resilience.

3.1.2 Goal 2 and Methodology

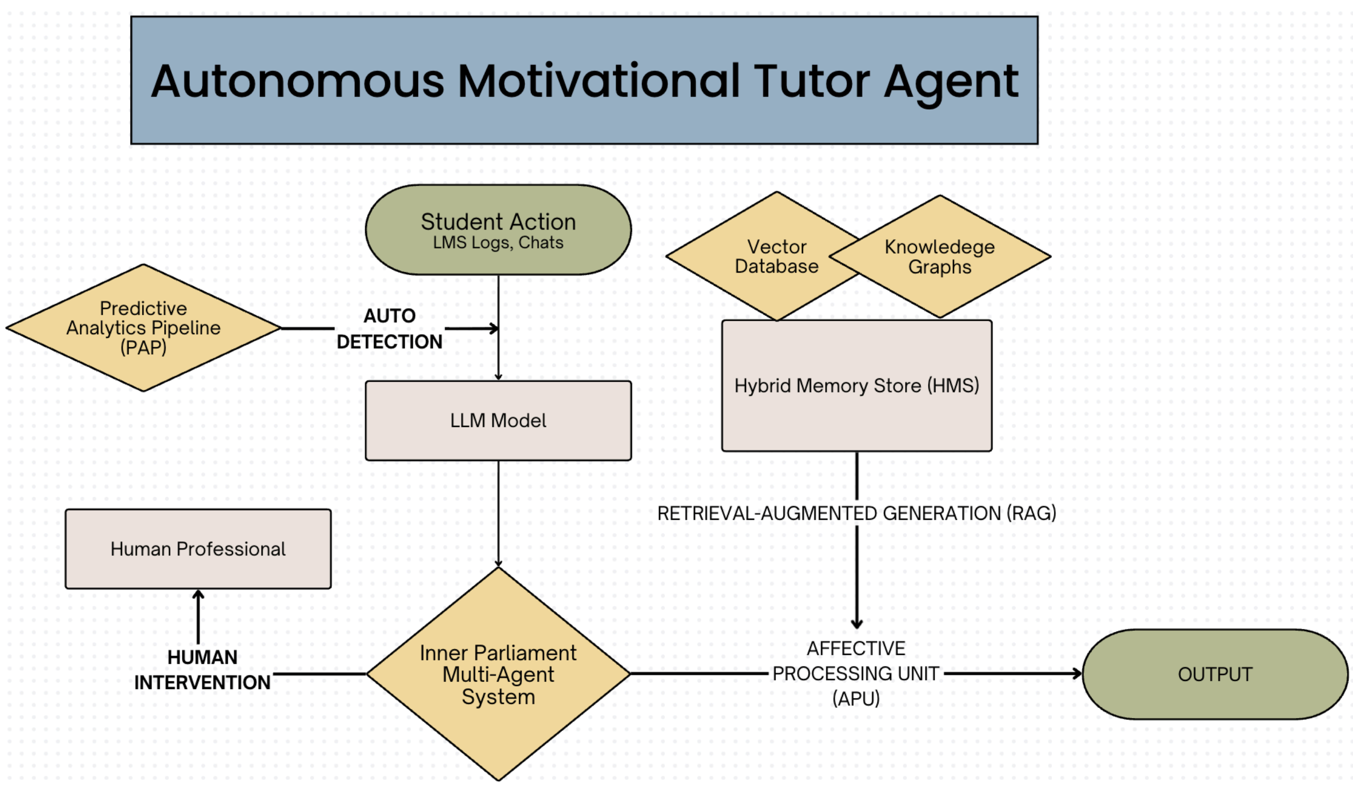
In this goal we want to achieve a real-time AI-driven emotional predictive analysis.

(1) **Data Pipeline**: We would develop a realtime data pipeline based on the streaming data to process attendance, assignments, and interaction with AI tutors.[17]
(2) **Predictive Modeling:** We deploy machine learning models for the early identification of at-risk students.[18]
(3) **Intervention:** Implement SHapley Additive exPlanations (SHAP)-based explainability for the predictive models.[18] Any risk analysis will be categorized into different sources. This ensures that interventions are targeted.[18]
(4) **Affective Monitoring**: Integrate an Affective Processing Unit (APU) utilizing NLP/ASA techniques to detect subtle linguistic markers (e.g., reduced spontaneity, emotional tone shifts) indicative of internal motivational decline or a shift towards learned helplessness.[10]

3.1.3 Goal 3 and Methodology

In this goal, we want to build a context-aware coaching AI agent.
(1) **State Detection:** The APU continually monitors the student's emotional state. Upon detection of distress[10], the agent will shift from instructive to empathetic, supportive interaction, using established emotional coaching strategies.[2]
(2) **Psychological Authenticity:** We build a Multi-Agentic system to give coaching responses.[15] There should be an internal conflict among agents, allowing the agentic system to provide coaching that addresses the student's internal struggles related to competence and resilience.[16]

3.2 Flow chart of System Architecture



4 Implementation Plan

4.1 Phase 1: Foundation, Governance, and Ethics

The initial phase focuses on establishing the ethical and theoretical guardrails for high-stakes psycho-educational interventions.

- Oversee compliance with the current AI law frameworks privacy related laws
- Define the structured knowledge schema for the student profile, incorporating Self-Determination Theory (SDT) dimensions—autonomy, competence, relatedness, etc.
- Develop informed consent and privacy protocols to protect data safety and privacy.

4.2 Phase 2: Technical Architecture and Module Development

This phase involves building the core engines of the system.

- Hybrid Memory Store (HMS) Construction: Utilize LLMs to extract structured triplets from university handbooks and academic data to create the initial Knowledge Graph.
- Predictive Analytics Pipeline (PAP): Develop real-time data processing from Learning Management Systems (LMS). Implement XGBoost and Deep Neural Network models for early at-risk detection.
- Explainable AI (XAI) Integration: Integrate SHAP-based model to ensure all risk flags are transparent and based on specific behavioral markers.
- Affective Processing Unit (APU): Develop natural language processing (NLP) pipelines to identify linguistic markers of motivational decline and emotional distress.

Phase 3: System Integration and Pilot Testing

The system is integrated into existing infrastructure for small-scale testing under strict supervision.

- LMS Integration: Deploy the system within the campus learning system (e.g., Canvas) for data exchange and deep linking.
- Human Escalation Module (HEM) Validation: Conduct stress tests of crisis detection protocols (suicide risk, severe anxiety) to ensure accuracy and immediate handoff to campus counseling services.

Phase 4: Full Deployment and Long-Term Evaluation

The final phase scales the system campus-wide while monitoring its impact on student outcomes.

- Campus-Wide Rollout: Expand system access to all students, supported by a 24/7 care team for Tier 3 escalations.

- Monitoring Hybrid Workflows: Evaluate the impact on professional staff workloads. Use the agent to automate psycho-educational report drafts, aiming to save practitioners up to 6 hours per week.
- Impact Assessment: Use Structural Equation Modeling to measure how AI feedback influences engagement through the dual pathways of intrinsic motivation and autonomy.
- Longitudinal Research & Bias Audits: Initiate a multi-semester study to track long-term effects on resilience and critical thinking. Conduct regular bias audits to ensure the system serves diverse student groups equitably.

5 Reference

[1] "From digital disruption to mental health: the impact of AI-induced educational anxiety on teacher well-being in the era of smart education," *PubMed Central*, accessed Dec. 13, 2025. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC12625245/

[2] L. D. Lucía, P. Cecchetto, and M. Busso, "Well-Being Technologies and Positive Psychology Strategies for Training Metacognition, Emotional Intelligence and Motivation Meta-Skills in Clinical Populations: A Systematic Review," *MDPI*, vol. 6, no. 1, 2025. [Online]. Available: https://www.mdpi.com/2624-8611/6/1/19

[3] A. Kumar, R. Singh, and L. Zhang, "AI-Powered Educational Agents: Opportunities, Innovations, and Ethical Challenges," *MDPI*, vol. 16, no. 6, 2025. [Online]. Available: https://www.mdpi.com/2078-2489/16/6/469

[4] J. Balkin, "AI Agents and Democratic Resilience," *Knight First Amendment Institute*, accessed Dec. 13, 2025. [Online]. Available: https://knightcolumbia.org/content/ai-agents-and-democratic-resilience

[5] H. Li and M. Chen, "The human touch in AI: optimizing language learning through self-determination theory and teacher scaffolding," *Frontiers in Psychology*, vol. 16, 2025. [Online]. Available: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2025.1568239/full

[6] "How AI is changing student motivation and engagement," *SchoolAI Blog*, accessed Dec. 13, 2025. [Online]. Available: https://schoolai.com/blog/ai-changing-student-motivation-engagement-classrooms

[7] S. N. Tan and A. Rahman, "The Impact of Artificial Intelligence on Specialized Learning Motivation in Higher Education: A Conceptual Paper," *ResearchGate*, 2025. [Online]. Available: https://www.researchgate.net/publication/394101560_THE_IMPACT_OF_ARTIFICIAL_INTELLIGENCE_ON_SPECIALIZED_LEARNING_MOTIVATION_IN_HIGHER_EDUCATION_A_CONCEPTUAL_PAPER

[8] J. Y. Park, "The role of AI-driven feedback in fostering growth mindset and engagement: A self-determination theory perspective," *ResearchGate*, 2025. [Online].

Available:
https://www.researchgate.net/publication/396155363_The_role_of_AI-driven_feedback_in_fostering_growth_mindset_and_engagement_A_self-determination_theory_perspective

[9] Y. Xia, C. Chiu, and S. Lee, "A Self-determination Theory (SDT) Design Approach for Inclusive and Diverse Artificial Intelligence (AI) Education," *selfdeterminationtheory.org*, Jan. 2023. [Online]. Available: https://selfdeterminationtheory.org/wp-content/uploads/2023/01/Manuscript_XiaChiuEtAl_SDTDesignApproach.pdf

[10] A. Rossi and E. Martínez, "Affective Computing and Emotional Data: Challenges and Implications in Privacy Regulations, The AI Act, and Ethics in Large Language Models," *arXiv preprint arXiv:2509.20153v2*, 2025. [Online]. Available: https://arxiv.org/html/2509.20153v2

[11] P. Navarro and D. Xu, "Integrating artificial intelligence to assess emotions in learning environments: a systematic literature review," *Frontiers in Psychology*, vol. 15, 2024. [Online]. Available: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1387089/full

[12] C. Branzan, "From LLMs to Knowledge Graphs: Building Production-Ready Graph Systems in 2025," *Medium*, accessed Dec. 13, 2025. [Online]. Available: https://medium.com/@claudiubranzan/from-llms-to-knowledge-graphs-building-production-ready-graph-systems-in-2025-2b4aff1ec99a

[13] M. Iqbal, F. Nguyen, and L. Zhao, "Supporting Student Decisions on Learning Recommendations: An LLM-Based Chatbot with Knowledge Graph Contextualization for Conversational Explainability and Mentoring," *arXiv preprint arXiv:2401.08517v3*, 2024. [Online]. Available: https://arxiv.org/html/2401.08517v3

[14] J. Kim and S. Patel, "KLR-KGC: Knowledge-Guided LLM Reasoning for Knowledge Graph Completion," *MDPI Electronics*, vol. 13, no. 24, 2025. [Online]. Available: https://www.mdpi.com/2079-9292/13/24/5037

[15] Q. Huang, "A Multi-Agent Psychological Simulation System for Human Behavior Modeling," *arXiv preprint arXiv:2511.02606v1*, 2025. [Online]. Available: https://arxiv.org/html/2511.02606v1

[16] R. S. Mehta and T. Wong, "Which Type of Students Can LLMs Act? Investigating Authentic Simulation with Graph-based Human–AI Collaborative System," *arXiv preprint arXiv:2502.11678v4*, 2025. [Online]. Available: https://arxiv.org/html/2502.11678v4

[17] "Predictive Analytics for Student Success: AI-driven Early Warning Systems and Intervention Strategies for Educational Risk Management," *Upubscience Publisher*, accessed Dec. 13, 2025. [Online]. Available: http://www.upubscience.com/upload/20250808150514.pdf

[18] "A Data Analytics-Based System for Proactive Student Performance Assessment,"

Preprints.org, accessed Dec. 13, 2025. [Online]. Available: https://www.preprints.org/manuscript/202510.1809