Proceedings of Machine Learning Research 273:255–259, 2025 iRAISE Workshop at AAAI 2025

Enhancing Learning Outcomes within a Large-Scale Online Learning System through AI-Powered Feedback

Aylin Ozturk aylino@andrew.cmu.edu
Robin Schmucker rschmuck@cs.cmu.edu
Tom Mitchell tom.mitchell@cs.cmu.edu
Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

Building on prior research, which demonstrated the effectiveness of a learning analytics-based feedback system in improving learner engagement and learning outcomes, this study addresses scalability challenges by automating feedback authoring using generative AI. Focusing on critical issues in distance education, including limited academic support, social isolation, and reduced learner motivation, we design and evaluate an AI-powered feedback system within a large-scale online learning environment. This study utilizes input data comprising learners' online learning environment interactions, learning material engagement patterns, academic performance metrics, behavioral indicators, and demographic characteristics. The system generates AI-powered personalized feedback interventions based on the ARCS-V Motivation Model, Self-Regulated Learning principles, and Nudge Theory as its primary outputs. To assess the system's effectiveness, more than $30,000$ learners at a large distance education university will be randomly assigned to experimental and control groups. Preliminary work demonstrated the system's readiness for a pilot evaluation. The next steps include assessing the system's impact on diverse learner subgroups and refining system design based on user feedback. The study aims to advance our understanding of how AI-powered, personalized feedback influences self-regulated learning, motivation, and learning outcomes in online environments.

Keywords: feedback, online learning, generative AI in education, learning analytics

1. Introduction

Online learning environments (OLEs) face persistent challenges, including limited academic support, social isolation, and declining learner motivation. These issues significantly impact academic performance and retention, particularly in large, diverse learner populations (Mubarak et al., 2020; Wang et al., 2019). Feedback, a critical mechanism for fostering engagement and learning (Iraj et al., 2021; Winstone et al., 2017), often fails to meet personalized needs at scale (Cavalcanti et al., 2023). Emerging technologies like generative artificial intelligence (GenAI) and AI-powered learning analytics (LA) offer scalable solutions, enabling adaptive and personalized interventions (Dai et al., 2024).

Our prior research demonstrated the effectiveness of instructor-written, LA-based feedback in improving engagement and learning outcomes (Ozturk, 2022). However, the manual feedback process limits scalability, making it highly resource-intensive for large-scale implementation. This motivates the integration of generative AI, which enables the automated creation of personalized feedback messages at scale. Thus, this study aims to design and evaluate an AI-powered feedback system tailored for learners in large-scale OLEs. The system is grounded in the ARCS-V Motivation Model (Keller, 1987), Self-Regulated Learning principles (Zimmerman, 2000), and Nudge Theory (Thaler and Sunstein, 2008).

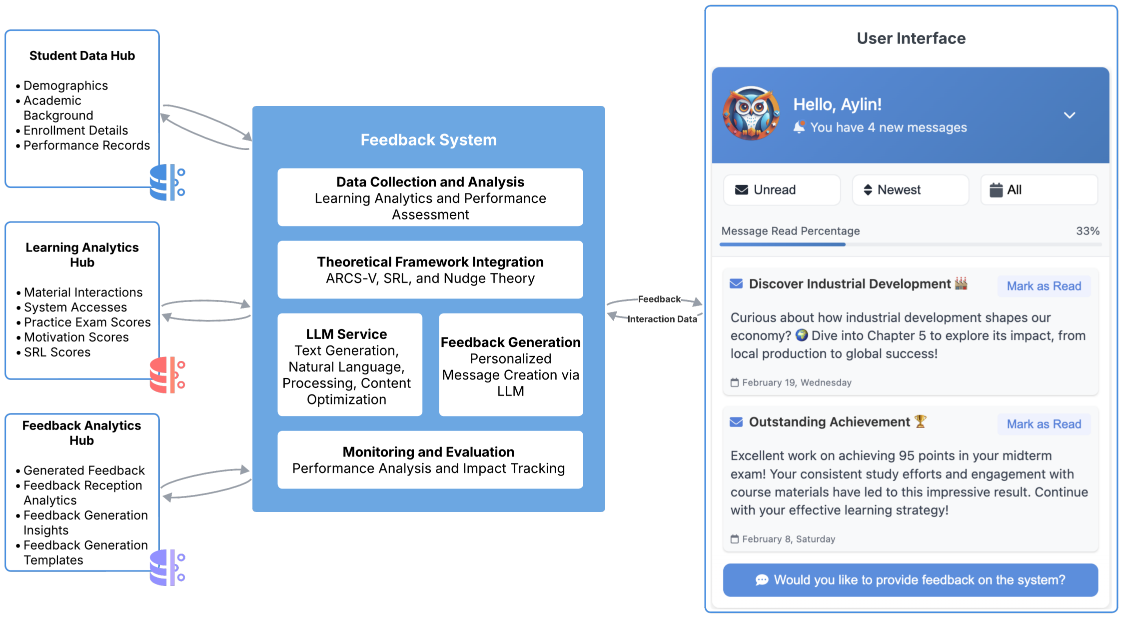

Figure 1: AI-Enhanced Educational Feedback System Architecture

2. Study Design and Methodology

The research is being conducted within the Anadolu University Open Education System in Türkiye, which serves over one million learners across its distance education programs. The research is driven by the central question: *How can an AI-powered feedback system enhance learner performance, engagement, self-regulation, and motivation within an online learning environment?*

The study integrates design-based research with an experimental research design. Primary data sources comprise demographic information, academic metrics, and interaction logs. Quantitative data from validated scales and questionnaires measure cognitive, affective, and behavioral dimensions, while qualitative insights from semi-structured interviews and focus groups provide a deeper understanding of learner experiences. All data collection adheres to ethical standards with robust data protection protocols. The design-based research follows five iterative phases (Wang and Hannafin, 2005):

1. Analysis: Expert interviews and focus groups have been conducted to identify design principles for the effective feedback system.

2. Design: Data structures and feedback mechanisms have been designed and implemented to address the identified needs and principles.

3. Development: Figure 1 illustrates the core architectural framework of the developed system. Data collected from three primary data hubs is processed within the central feedback system to manage learning analytics and performance assessment processes. The system integrates theoretical frameworks into message generation through the LLM service while maintaining continuous interaction with the performance analysis and impact tracking module. This structure forms the foundation for our learner-centered and data-driven feedback mechanism.

Building upon this architecture, generative AI technologies have been integrated into the system to generate personalized textual feedback using large language models, and the feedback system has been fully integrated into the learning management system (LMS). The system leverages advanced LLMs such as GPT (OpenAI, n.d.), LLaMA (Meta, n.d.), and Claude (Anthropic, n.d.) for their effectiveness in education-focused tasks and contextually appropriate feedback generation. To ensure reliability and fairness, the system has implemented a comprehensive validation pipeline combining automated checks, human review, and fairness audits (IBM, 2024). Additionally, statistical fairness metrics and tools such as Microsoft Fairlearn have been integrated for robust bias detection and mitigation (Bird et al., 2020).

4. Implementation: The system has been prepared for deployment in the upcoming semester's online course involving 30,000 online learners, randomly assigned to experimental and control groups. The experimental group will receive personalized feedback, while the control group will follow standard instructions. To enhance trust in AI-generated feedback, the implementation phase includes a mechanism for integrating human oversight, where instructors or moderators can review and validate a subset of AI-generated feedback. This ensures the quality, transparency, and reliability of the feedback process.

5. Evaluation: The effectiveness of feedback is being assessed through multiple quantitative measures. Learner engagement is measured via LMS log data, including access timestamps, login frequency, and content interactions. Academic performance is evaluated through quiz scores and exam results, while motivation and self-regulation skills are assessed using pre- and post-intervention surveys and learning analytics data. Interaction data and performance metrics within the LMS are analyzed using statistical methods, including independent samples t-tests, to quantify the impact of the intervention by comparing learning outcomes between the experimental and control groups. To further explore subgroup differences, the causal forest algorithm (Wager and Athey, 2018) is being used to identify learner characteristics that influence the effectiveness of the intervention.

3. Expected Findings and Implications

This study aims to provide empirical, data-driven insights into the impact of AI-powered feedback on engagement, academic performance, and self-regulation skills for learners within a large-scale online education environment. Initial findings demonstrate the potential of generative AI to address persistent challenges in online learning, such as limited support and declining motivation while enhancing scalability and access to personalized support. The current implementation at a large-scale distance education university is positioned to benefit over one million learners annually. This research contributes to the growing body of knowledge on AI in education by providing actionable insights for integrating adaptive analytics into diverse educational contexts while upholding ethical AI principles.

Acknowledgments

This study was supported by TUBITAK (The Scientific and Technological Research Council of Türkiye) under the 2219 - International Postdoctoral Research Fellowship Program for Turkish Citizens.

References

Anthropic. Meet claude, n.d. URL https://www.anthropic.com/claude. Accessed: 2025-02-15.

Sarah Bird, Miroslav Dudík, Richard Edgar, Brandon Horn, Roman Lutz, Vanessa Milan, Mehrnoosh Sameki, Hanna Wallach, and Kathleen Walker. Fairlearn: A toolkit for assessing and improving fairness in AI. Technical report, Microsoft, September 2020. Version: September 22, 2020.

A. P. Cavalcanti, R. F. Mello, D. Gašević, and F. Freitas. Towards explainable prediction feedback messages using bert. *International Journal of Artificial Intelligence in Education*, 2023. ISSN 1560-4292. doi: 10.1007/s40593-023-00375-w.

W. Dai, Y.-S. Tsai, J. Lin, A. Aldino, H. Jin, T. Li, D. Gašević, and G. Chen. Assessing the proficiency of large language models in automatic feedback generation: An evaluation study. *Computers and Education: Artificial Intelligence*, 7:100299, 2024. ISSN 2666-920X. doi: 10.1016/j.caeai.2024.100299.

IBM. Techniques for avoiding undesirable output in foundation models, November 2024. URL https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/fm-hallucinations.html. IBM watsonx.ai Documentation.

H. Iraj, A. Fudge, H. Khan, M. Faulkner, A. Pardo, and V. Kovanović. Narrowing the feedback gap: Examining student engagement with personalized and actionable feedback messages. *Journal of Learning Analytics*, 8(3):101–116, 2021. doi: 10.18608/jla.2021.7184.

J. M. Keller. Development and use of the arcs model of instructional design. *Journal of Instructional Development*, 10(3):2–10, 1987. doi: 10.1007/BF02905780.

Meta. Llama: Large language model meta ai, n.d. URL https://www.llama.com/. Accessed: 2025-02-15.

A. A. Mubarak, H. Cao, and W. Zhang. Prediction of students' early dropout based on their interaction logs in online learning environment. *Interactive Learning Environments*, 30(8):1414–1433, 2020. doi: 10.1080/10494820.2020.1727529.

OpenAI. Gpt-4: Openai's most advanced system, n.d. URL https://openai.com/index/gpt-4/. Accessed: 2025-02-15.

Aylin Ozturk. *Modeling Learners' Behavioral Patterns and Profiles, Predicting the Academic Performance and Investigating the Effects of a Dashboard in Open and Distance Learning*. PhD thesis, Anadolu University, 2022.

R. H. Thaler and C. R. Sunstein. *Nudge: Improving Decisions About Health, Wealth, and Happiness*. Springer, 2008.

S. Wager and S. Athey. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242, 2018. doi: 10.1080/01621459.2017.1319839.

F. Wang and M. J. Hannafin. Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, 53:5–23, 2005. doi: 10.1007/BF02504682.

W. Wang, L. Guo, L. He, and Y. J. Wu. Effects of social-interactive engagement on the dropout ratio in online learning: Insights from mooc. *Behaviour & Information Technology*, 38(6):621–636, 2019. doi: 10.1080/0144929X.2018.1549595.

N. E. Winstone, R. A. Nash, J. Rowntree, and M. Parker. 'it'd be useful, but i wouldn't use it': Barriers to university students' feedback seeking and recipience. *Studies in Higher Education*, 42(11):2026–2041, 2017. doi: 10.1080/03075079.2015.1130032.

B. J. Zimmerman. Attaining self-regulation: A social cognitive perspective. In Monique Boekaerts, Paul R. Pintrich, and Moshe Zeidner, editors, *Handbook of Self-Regulation*, pages 13–39. Academic Press, 2000. ISBN 9780121098902. doi: 10.1016/B978-012109890-2/50031-7.