InsPIRE Insight

Improving Course Recommendation Systems with Explainable AI: LLM-Based Frameworks and Evaluations

By

- Jiawei Li (School of Electrical and Electronic Engineering)
- Qianru Lyu (National Institute of Education)
- Wei Qiu (Centre for the Applications of Teaching and Learning Analytics for Students (ATLAS), InsPIRE)
- Andy W. H. Khong (School of Electrical and Electronic Engineering; Lee Kong Chian School of Medicine)

Nanyang Technological University
Singapore

KEY IMPLICATIONS

- Structured LLM prompts that integrate course content with social science theories (such as Relevance Theory) significantly enhance the clarity and quality of AI-generated explanations.
- Course descriptions that focus on specific topics and content are more effective in aiding student decision-making than those that rely solely on abstract learning outcomes.
- Advanced prompting strategies allow AI models to generate high-quality, personalized explanations for students across all academic performance levels, including those who are struggling.

BACKGROUND

Deep learning-based systems are frequently used to recommend courses in higher education, but they often lack interpretability, functioning as "black boxes" that fail to explain the rationale behind their suggestions. This lack of transparency limits their practical utility, as students and educators may struggle to trust or act upon the recommendations. Explainable Artificial Intelligence (XAI) is essential in this context to provide insights that guide decision-making and improve student engagement. While Large Language Models (LLMs) show potential for generating these explanations, there is a need to optimize their design specifically for educational contexts.

FOCUS OF INITIATIVE

This study aimed to improve the interpretability of course recommendation systems by developing and evaluating a modular framework that uses LLMs to generate post-hoc explanations. The initiative focused on identifying the most effective prompt designs by comparing four model variations. These models varied in their prompting patterns (Persona vs. Cognitive Verifier), the type of course descriptions used (Outcome-based vs. Content-based), and whether they incorporated social science theories like Relevance Theory.

KEY OUTCOMES

- **Superior Model Performance:** The model which combined the "Cognitive Verifier" prompt pattern with content-based descriptions and Relevance Theory consistently outperformed all other models in Clarity, Effectiveness, Relevance, and Specificity.
- **Content Over Outcomes:** Explanations based on content-based descriptions (detailing specific topics) were found to be more actionable and effective than those based on outcome-based descriptions, which were often too abstract.
- **Theory-Driven Clarity:** Integrating Relevance Theory helped the AI explicitly connect a student's prior performance to the recommended course, significantly improving the clarity and logical flow of the explanations.
- **Inclusivity:** The optimized model demonstrated robust performance across diverse student profiles—top, average, and struggling performers—indicating its ability to provide tailored support regardless of a student's academic standing.

SIGNIFICANCE OF OUTCOMES

- **Implications for practice:** To enhance AI-driven support, universities should ensure course catalogs include detailed, topic-level content rather than just learning outcomes, as this data enables AI to generate more specific and useful explanations.
- **Implications for policy and future initiatives:** Institutions should develop guidelines for "Prompt Engineering" in educational tools, ensuring that AI implementations are grounded in pedagogical theories to maximize their effectiveness.
- **Proposed follow-up activities:**
  1. Develop multimodal explanations that include visual aids (e.g., knowledge graphs) to further improve comprehension.
  2. Integrate quantitative data, such as predicted grades, into the explanations to provide deeper insights and build greater trust in the system.

PARTICIPANTS/SCOPE

The recommendation model was trained on data from 2,795 engineering students. The qualitative evaluation involved three engineering course instructors with varying levels of experience (a course coordinator, a senior lecturer, and a junior lecturer) who evaluated 120 explanations generated for six representative student profiles.

METHODOLOGY/APPROACH

The study utilized a two-stage framework: a BERT-based model first generated course recommendations based on student history and course embeddings, followed by an LLM (Gemini 1.5 Flash) that generated explanations. The four explanation models were evaluated using a survey where experts rated them on a 6-point Likert scale across four criteria: Clarity, Effectiveness, Relevance, and Specificity.

References

Devlin, J. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
Hilton, D. J. (1991). A conversational model of causal explanation. European Review of Social Psychology, 2(1), 51-81.
White, J., et al. (2023). A prompt pattern catalog to enhance prompt engineering with ChatGPT. arXiv preprint arXiv:2302.11382.

Link to full article: https://doi.org/10.5281/zenodo.15870185
Contact Jiawei Li at jiawei009@e.ntu.edu.sg or Wei Qui at qiuwei@ntu.edu.sg for more information about the initiative.