12-15-2024

Fostering and Assessing Authentic Learning with “Intelligent Assessor”: Preliminary Report on Building a Generative AI Tool to Teach SQL

Raja Sooriamurthi
Carnegie Mellon University, raja@cmu.edu

Xiaoying Tu
Carnegie Mellon University, xiaoying@cmu.edu

TREO
Technology, Research, Education, Opinion

Fostering and Assessing Authentic Learning with "Intelligent Assessor"

Preliminary Report on Building a Generative AI Tool to Teach SQL

Raja Sooriamurthi (raja@cmu.edu); Xiaoying Tu (xiaoying@cmu.edu)

The ultimate goal of teaching is authentic learning, where the student has meta-knowledge about what they have learned—nuances, variations, and implications—and are able work with their knowledge at the higher end of Bloom's taxonomy (analyze, evaluate, create). An effective way of assessing authentic learning is for an instructor to have a dialog with the student probing subtler issues, alternative approaches, and implications of what the student has learned. With the advent of generative AI (genAI), we now have technology that can potentially assist with 1-1 personalized learning that scales to a large number of students.

In this TREO talk, we discuss our preliminary experience with building a genAI/LLM powered "Intelligent Assessor" for conducting 1-1 dialogs with students by automatically generating targeted questions and evaluating responses to assess students' genuine understanding. Our research goals are twofold: (i) our current focus is to develop a tool to assess authentic learning (ii) subsequently we intend to extend our tool to foster authentic learning. We explore the potential of such genAI techniques for assessment in the context of teaching how to write SQL queries. Because of the restricted pedagogical context of SQL query formulation it is well suited for piloting this novel pedagogical approach and tool.

The prototype Intelligent Assessor system leverages an LLM e.g., GPT-4o or Llama3.1, seeded with the details of the 3 step heuristic process of formulating any SQL query that we developed and successfully tried over years of teaching database courses. We also fed the system SQL style guidelines, common mistakes made by students, and a non-exhaustive list of sample questions it may ask. We fine-tuned its behavior so that, when taking the input of the dataset / table schema, the question prompt, and the SQL query written by the student, the systems is able to assess the correctness and appropriateness of the SQL constructs used in the student's query. Without directly revealing the potential improvements, it can iteratively generate questions that probe the use of key query constructs, analyze the student's responses, provide feedback, and ask follow-up questions until the student has demonstrated a clear, in-depth understanding of those constructs. A small-scale pilot is currently underway with a larger-scale experiment planned during the upcoming offerings of the database courses at both undergraduate and graduate levels to measure the effect of such Intelligent-Assessor-led dialogs on students' learning outcomes, attitudes, engagement among other factors compared to human-assessor-led dialogs and no dialogs at all.

Our project distinguishes itself among prior studies on the use of AI in education by directly assessing the degree of authentic learning underlying the final product rather than just evaluating the product itself. Furthermore, our proposed application is conversational, fully leveraging one of ChatGPT's most distinctive features to provide a "human touch" in a pedagogically safe context which is critical to the success of AI integration in education.