Title: Use of Collaborative Annotation Technology with Tutor AI in Computer Science Education

Student Name: Mark P. McCormack
Supervisor's Name: John G. Keating

Abstract:

Recent developments in the literature suggest that Large Language Model (LLM) systems, such as ChatGPT, are becoming increasingly more effective in the realm of computer science education (CSE). While most studies focus on their problem solving capabilities, new studies are emerging looking at how these and other Machine Learning (ML) models may be used in augmentation of students learning such as automated assessment, discussing assignments with students and engaging in peer-programming exercises.

In this study, we look at how effective these Large Language Model systems can be in acting as tutors for students, and whether these tools help or hinder the students quality of learning. To do this, we provide this tool to a class of Higher Education Computer Science Masters' students. We divide the students into three groups. The control group will complete an assessment activity without the AI tutor, experiment group one will complete the activity with the AI tutor and experiment group two will complete the activity with the AI tutor who also has direct access to auxiliary resources relevant the assignment. The students using the AI tutor will use the Mistral 7B model for this initial experiment. For the auxiliary resources, we select relevant materials from the Open Education Resources (OER) Commons webpage, as this information would be accessible and usable by every educator.

We look to analyse the students summative grade in the assignment, gather their opinions using the ASPECT questionnaire[5] and finally conduct a follow-up test some weeks later to assess if the knowledge was retained. From this, we can draw conclusions as to if this AI tutor improves students' quality of learning and if augmenting it with auxiliary relevant materials provides better feedback. Regarding future work, our study provides us preliminary results for several areas. Firstly, we can experiment with if different AI models (GPT-3, Mixtral, Claude etc.) are better or worse for particular assignments or activities. Also, should the students benefit from using AI with auxiliary resources, we can look to see what types of resources are the most relevant for the students.