Embracing chatbots in teaching: improving students' R programming skills by enhancing responsible use of large language models

Summary

Large language models (LLMs), such as chatGPT, are here to stay and teachers cannot entirely prevent students from using them to perform assignments related to their academic curriculum. LLMs present multiple educational challenges, such as the risk for plagiarism and the use and spread of misinformation. However, LLMs also offer diverse educational opportunities. LLMs can for example provide personalised answers to students' queries, which may support them in completing difficult assignments. Hence, it is the teachers' duty to inform students about the limitations of LLMs, but also to foster the responsible use of these digital tools. A particular skill that is increasingly important in academia is programming. LLMs offer significant advantages for programming as they can support in writing, debugging, and optimising scripts. The proposed teaching innovation therefore aims to teach students how to responsibly use LLMs to enhance their programming skills. This innovation will be implemented in the course "health economic modelling" of the Master Health Sciences in which students have to develop a health economic model using the statistical software R.

Relevance of the project

In November 2022, OpenAI launched chatGPT (chat Generative Pre-trained Transformer) worldwide. In a couple of months, millions of individuals across the globe, amongst other students, have used chatGPT. ChatGPT is a large language model (LLM) powered by natural language processing methods, and trained on extensive internet content, which can provide human-like answers to users' prompts. Besides, chatGPT is able to generate text such as essays, course outlines, and scripts for programming languages. ChatGPT is also able to provide feedback and suggestions to improve users' texts (or script of a programming language). The use of chatGPT, and LLMs in general, consequently opens opportunities for developing innovative educational designs, e.g. to create personalised content and feedback for students[1,2]. However, their use raises issues relating to privacy and copyrights, and the potential for spread of misinformation[1,3].

LLMs are here to stay and will also impact education. However, it is not desirable to let this technology loose without informing students about its limitations. Using LLMs responsibly requires students to be aware of the potential pitfalls of LLMs and to critically evaluate their outputs. Hence, the proposed teaching innovation aims to teach students to leverage the potential of LLMs and to create awareness about the limitations and potential threats of LLMs. Therefore this teaching innovation fits perfectly within the University of Twente's educational vision which aims that students *"develop a broader range of '21st century' skills"* and "*understand and anticipate on current trends and developments".*

The teaching innovation will be implemented in the "health economic modelling" course of the Master Health Sciences during which students have to develop a health economic model using the statistical software R. This course has been selected because teachers witnessed that learning R programming as well as troubleshooting erroneous scripts was challenging for these students, who, in general, have limited programming experience. Hence, LLMs may be a useful tool for to improve their skills[4].

[1] Kasneci, Enkelejda, et al. "ChatGPT for good? On opportunities and challenges of large language models for education." Learning and Individual Differences 103 (2023): 102274.
[2] Qadir, Junaid. "Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education." (2022).
[3] Dwivedi, Yogesh K., et al. "'"So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy." International Journal of Information Management 71 (2023): 102642.
[4] MacNeil, Stephen, et al. "Generating diverse code explanations using the gpt-3 large language model." Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 2. 2022.

Problem statement

The problem that we will address is twofold. First, Health Sciences master students have difficulties with developing a health economic model using the statistical software R. This is explained by their lack of hands-on experience in writing R scripts and by their lack of problem-solving skills to troubleshoot errors in scripts. Second, LLMs are not (yet) implemented in teaching, but it cannot be prevented that students will use LLMs to perform different teaching-related tasks. This creates a risk that LLMs will be misused by students. Hence, teachers have a duty to educate students on how to responsibly use LLMs.

Proposed solution

Our solution to improve students' programming skills is to teach students how LLMs can be used responsibly when developing a health economic model in R. We will use the brain-based learning principles as described below when developing and implementing our teaching activities. Additionally, we will use the findings of and our experiences gained during the project to develop a generic framework for integrating LLMs in other non-health-economic courses using a programming language.

Theoretical background

LLMs in teaching

Due to the novelty of LLMs, empirical evidence on their use in teaching is limited. The scientific literature on this topic is currently dominated by opinion papers and pre-prints and contains a handful reports of experiences of actually using LLMs in teaching. However, these opinions and first experiences provide useful insights in the potential added value of using LLMs in education, and especially for enhancing students' programming skills.

LLMs increasingly receive attention in supporting students' learning of programming languages due to the ability of LLMs (in helping) to generate, comment, and troubleshoot lines of scripts. For instance, LLMs have already been implemented in a programming course to generate explanations of scripts to enhance the students' understanding of the script[5]. Additionally, LLMs have been shown to improve the understandability of error messages provided by programming software[6]. These capabilities of LLMs are useful to enhance students' learning because a) high quality explanations of scripts foster students' understanding of these scripts, and b) LLMs may support troubleshooting erroneous scripts, which is considered one of the main barriers in learning a programming language[6].

Even though first experiences indicate that LLMs likely have a positive effect on improving students' programming skills, challenges remain concerning the use of LLMs. These challenges entail issues relating to plagiarism, licensing, and over-reliance on results generated by LLMs. In particular the latter poses a threat towards learning good coding practices[7]**.** Hence, teachers should discuss these challenges with students to foster the responsible use of LLMs during their education and their professional career. To the best of our knowledge, there is currently no evidence on using LLMs for developing health economic models.

[5] MacNeil, Stephen, et al. "Experiences from using code explanations generated by large language models in a web software development e-book." Proceedings of the 54th ACM Technical Symposium on Computer Science Education V. 1. 2023.
[6] Leinonen, Juho, et al. "Using Large Language Models to Enhance Programming Error Messages." arXiv preprint arXiv:2210.11630 (2022).
[7] Becker, Brett A., et al. "Programming Is Hard--Or at Least It Used to Be: Educational Opportunities And Challenges of AI Code Generation." arXiv preprint arXiv:2212.01020 (2022).

Brain-based learning theory

The use of brain-based learning theories, such as the SCARF model, enhances the development of effective teaching activities. SCARF involves five domains of human social experience: Status, Certainty, Autonomy, Relatedness and Fairness. Through this project we aim to improve the "Autonomy" of Health Sciences students. More specifically, these students currently find programming particularly difficult and mostly rely on extensive support of teachers or fellow students with more advanced programming skills. The use of LLMs is expected to increase their perception that they are able to write and troubleshoot scripts by themselves, thereby increasing their sense of autonomy. The proposed teaching innovation will be designed in a brain-friendly way by using activating brain principles as defined by the Brain Based Learning Institute. We will for instance "build upon" previous students' knowledge to enhance their learning, as they are most likely familiar with LLMs (although not for programming), and because they have some experience with R programming from previous courses.

Expected impact

Through this teaching innovation, we aim to a) increase students' programming skills and their ability to troubleshoot in case they encounter errors in their scripts, b) increase students' awareness about the opportunities offered by LLMs for programming, as well as about their limitations, and c) enhance students' critical thinking skills concerning the output generated by LLMs.

Implementation plan

We will incorporate the use of LLMs into the master course "Health Economic Modeling". As part of this course, students develop a health economic model using R by adapting and extending a predefined template. In this modeling assignment, the students are provided with a series of steps to guide them through this process. During previous years, we recommend students to use the built-in help function in R and to search Google to find answers to their programming questions. However, these resources often proved insufficient, as these do not provide personalised answers to their questions. In contrast, LLMs like chatGPT, do provide such personalised answers. Therefore, we will advise students to use LLMs for performing this assignment after having introduced them to the capabilities and limitations of LLMs to support them in programming. The steps to implement this teaching innovation are described below. Steps 1 to 5 will be implemented within a single tutorial session.

1. Introduce students to the assignment and to the predefined health economic model template.
2. Ask students to solve five to ten hands-on exercises in R within fifteen minutes using the built-in help function in R and Google, but without using LLMs. Before performing the exercises, we will ask the students to rate their current programming skills (scale 1-10). After working on the exercises, we will ask them to indicate how many exercises they were able to solve and how complex they found these exercises (scale 1-10).
3. Demonstrate students the potential use of LLMs for programming to:
    a) Generate scripts, e.g. ask a LLM how to add three columns to an existing dataframe.
    b) Troubleshoot, e.g. ask a LLM why the script generated under a) does not work.
    c) Optimise scripts, e.g. ask a LLM to reduce execution times of previously written scripts.
4. Illustrate how NOT to use LLMs, for instance by letting LLMs generate or review an entire R script, or letting LLMs generate scripts and use it without carefully checking it. We will show a number of examples to illustrate these risks, to stimulate proper use of LLMs and to maximise the students' learning process.
5. Ask students to solve five to ten exercises in R (comparable to those in step 2) within fifteen minutes, using a LLM (if they wish to). Afterwards, ask them 1) how many exercises they were

able to solve, and 2) how complex they found these exercises (scale 1-10), and 3) whether and how using LLMs affected their confidence for successfully completing the exercises.
6. Let students work on the assignment, and provide support when necessary.
7. At the start of the next tutorial, discuss the most frequently-asked questions, and give advice on how LLMs could be used to solve these questions.
8. Repeat steps 6 and 7 during each tutorial session.
9. During a tutorial sessions towards the end of the course, we will let students review each other's code in pairs, both with and without using LLMs. The students will be asked to:
    a) Review each other's code, which increases their ability to understand code, and provides insights into alternative ways of solving the same problems.
    b) Challenge them to optimise each other's code as much as possible through a small competition. The group who reduces the execution time of another group's health economic model the most will win the competition and receive a gift (note that the running time of a code can be easily determined in R).
10. Repeat step 5 during the last tutorial.
11. Students finalise the assignment and hand it in. Students will be asked to indicate whether they used a LLM to solve each question of the assignment. Additionally, we will ask them to provide the outputs they obtained from the LLMs, and how they integrated it in their own health economic model. The use a LLM to complete the assignment will not affect their grade.
12. After the course, evaluate the success of integrating LLMs through:
    a) Conversations (semi-structured interviews) with +/- 5 students (with more and less programming experience) to evaluate their experiences.
    b) A survey to ask all students about their findings, learning experiences, their currently perceived programming skills level (scale 1-10), and the experienced added value of LLMs for programming.
    c) Evaluate the impact of LLMs on the scores and programming skills of students based on the answers provided by students in the final assignment and on the answers provided during steps 2, 5, and 12b.

Steps 2 to 5 will also be implemented in the "economic evaluation in healthcare" course (bachelor Gezondheidswetenschappen) since students in this course have to perform practical assignments in R and encounter the same issues as the master students. This implementation will also be evaluated as in step 12, which will provide more insights concerning how to effectively implement this technology.

A dissemination plan within and beyond the UT

We aim to disseminate the results of our project through the following activities:

1. Presentation within the section Health Technology & Services Research to create awareness among other teachers within the Gezondheidswetenschappen and Health Sciences programs. We will invite the Health Sciences program management team to attend this presentation.
2. Presentation during a BMS teaching academy activity.
3. Writing an article about the (findings of) the project that will be published within the Digital Competence Centre newsletter of the University of Twente.
4. Writing a scientific journal article describing the results of the project and our experiences.
5. Archiving the developed teaching materials on an open repository.
6. Presentations at conferences such as the R for HTA conference, the European ISPOR conference, and the Lowlands Health Economics Study Group.