Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
Paper ID #46122

WIP: AI-Driven Personalized Learning for an Introductory Computing Course

Dr. Nina Kamath Telang, University of Texas at Austin

Nina Telang is a Professor of Instruction in the Department of Electrical and Computer Engineering at the University of Texas at Austin. She received the B.Tech degree in Engineering Physics from the Indian Institute of Technology, Mumbai in 1989, and the M.S. and Ph.D. from the University of Notre Dame in 1992 and 1995.

Sneha Ballabh, University of Texas at Austin
Bishweshwor Rimal, University of Texas at Austin

Work in Progress: AI-Driven Personalized Learning for an Introductory Computing Course

Abstract:
In this work in progress paper, we present the design and implementation of AI-driven applications that create personalized learning experiences for first-year engineering students in an introductory computing course. This foundational course is crucial for Computer Engineering students, covering essential topics from computer system fundamentals to assembly language programming. While there are no prerequisites, students with some prior computer programming experience tend to be better prepared. To address this disparity, we have developed a range of comprehensive resources, including notes, solved problems, videos, and sample programs, which are beneficial for all students. However, to further enhance student mastery, we propose the use of self-assessment and personalized tutoring.

Our project leverages generative AI to develop customized question banks based on course materials within the Canvas learning management system. This approach provides a personalized learning experience tailored to each student's needs. The key innovation is to ensure that the generated questions are strictly aligned with the course content, avoiding the use of external internet sources to maintain relevance and accuracy.

Genux, an innovative startup founded by one of our co-authors, has developed 'agentic apps,' which are AI-driven applications that generate context-aware interactive graphical user interfaces (GUIs) for user interaction. These apps use artificial intelligence to create user-friendly interfaces, featuring dynamic question formats, tables of contents, and suggestions for additional questions.

Our project has yielded two advanced tools: **Textbook** and **Modules**. **Textbook** serves as an AI-powered book expert, allowing students to ask questions and receive detailed explanations about the course's textbook content, making the learning process more interactive and personalized. **Modules**, which are provided to students weekly through our learning management system, offer interactive questions and exercises based on textbook chapters, reinforcing understanding and retention. Both tools are built on the *Genux* platform, enabling the creation of dynamic user interfaces generated by AI agents.

In this paper, we discuss the structure of this AI-driven tool in a required first-year computing course, the level of usage and reported usefulness by students, and the impact of this tool on student performance in this course.

Motivation for Study:
Educators across the globe are investigating the impacts of generative AI tools on student learning outcomes in various fields of study such as computer programming [1], sciences [2], [3], economics [4], and medicine [5]. The reports suggest that these tools have an overall positive

effect for reasons such as getting instant feedback from chatbots resulting in student questions getting resolved immediately, thereby promoting student engagement [6]. In a recent article summarizing the impact of ChatGPT on a variety of engineering assessments, the authors concluded that introductory-level programming assessments can be very accurately solved by ChatGPT, and that instructors must add complex features to their assessments to deter student cheating [7].

However, there is no published research on the usage of generative AI to offer customized question banks and explanations of course concepts based only on course materials, for an introductory computing course, within the learning management system, thereby providing a personalized learning experience tailored to each student's needs. The primary innovation in our study lies in ensuring that the generated questions remain entirely focused on the course content, strictly avoiding the use of external internet sources to preserve relevance and accuracy.

Research Questions:
To assess the extent of engagement with this new personalized generative AI tool on first year computer engineering students, and the impact on their learning, this report addresses the following questions:

1. What percentage of students repeatedly used the personalized generative AI tool offered to them, and what was their reported experience?
2. What were some of the reported reasons for not using this personalized generative AI tool?
3. Was there any relationship between the level of usage and the student's final course GPA?

Design and Implementation:
The pipeline for our AI-driven personalized learning tools is designed to ensure that all content is strictly aligned with the course materials, providing accurate and context-specific support for students. The process begins with the OpenAI Assistants API, which is used to generate questions based on the course's foundational topics. These questions are constructed by creating context-aware prompts, ensuring that they align with the specific chapters, lecture notes, and learning objectives provided within the course material by providing the course material and strictly informing the assistant to stay within these boundaries.

Once generated, the prompt, including the question bank, is processed by *Genux*, an advanced AI platform for creating Agentic Apps. Agentic apps leverage *Genux's* AI Agents to dynamically generate interactive user interfaces in real time, delivering a uniquely personalized experience for each user. *Genux* stands out by combining adaptive content delivery with real-time interface generation, ensuring that students not only receive accurate and relevant answers but also engage with content in a way that feels intuitive and responsive to their needs. Its robust capabilities include generating diverse question formats, providing detailed answers, crafting interactive explanations, and formulating follow-up questions to reinforce understanding. To maintain strict adherence to course material, the prompt construction phase explicitly excludes external internet sources, ensuring all outputs are relevant, accurate, and aligned with the course content.

The dynamic question-and-answer content produced by the pipeline is integrated into two primary tools: **Textbook** and **Modules**. The **Textbook** tool acts as an AI-powered book expert, allowing students to ask questions about the course material and receive detailed, interactive explanations. The **Modules** tool provides tailored exercises and quizzes based on specific textbook chapters, enabling students to test their knowledge and reinforce learning through practice. Usage of these tools is based on student need and is student-driven, resulting in a personalized learning experience, which is described in the section below.

Interactive Explanations of Course Concepts:
To provide students with targeted assistance, the AI-driven **Textbook** and **Modules** tools offer detailed explanations of course concepts, tailored to the specific needs of each student. One key feature of these tools is their ability to generate interactive, context-aware responses, allowing students to engage in multi-step clarifications and follow-up questions. This interactive dialogue mimics a tutoring session, helping students reinforce their understanding through iterative questioning resulting in a personalized learning experience.

A prominent example of this process involves the concept of stack vs. heap memory, a fundamental topic in computer systems. As students interact with the AI tool, they often pose initial questions and refine their understanding through additional clarifications. The following figures presents examples of AI-generated explanations in response to student query and follow-ups, demonstrating how the tool adapts to different learning needs.

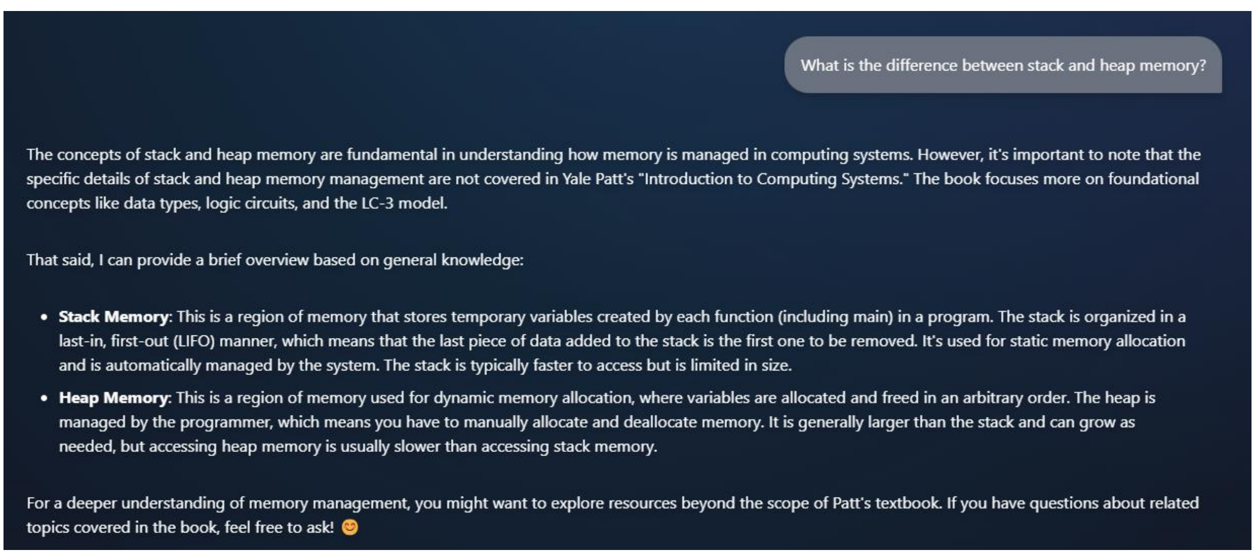


Figure 1: The student's question to the textbook along with the response

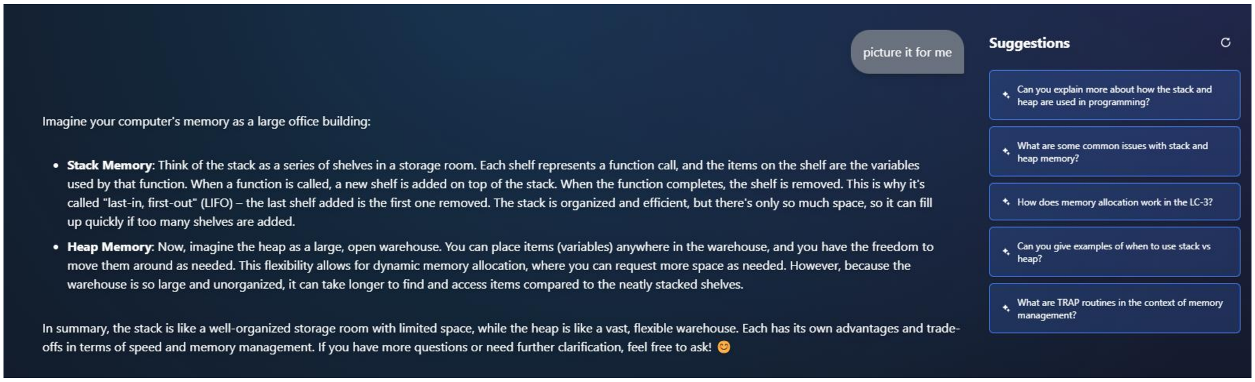


Figure 2: The student's follow-up query to Figure 1.

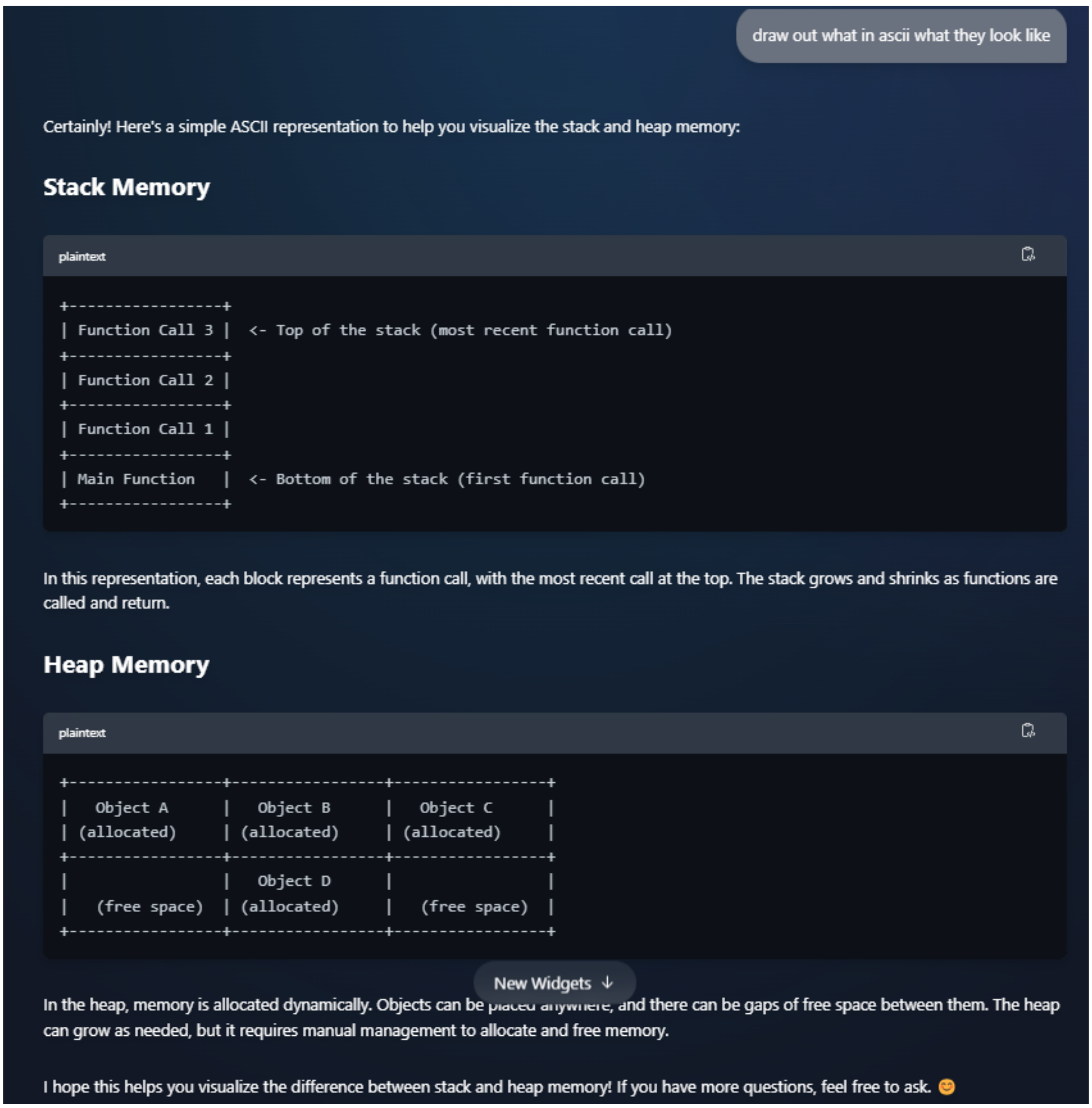


Figure 3: The student's follow-up query to Figure 2.

In Figure 4 we show an example of the **Modules** tool being used to generate questions on topics covered in a specific module in the Canvas (the learning management system) course. The tool also gives suggestions for more challenging questions. The student can choose not to answer this question and instead ask for further explanation of the concept. This is illustrated in Figure 5.

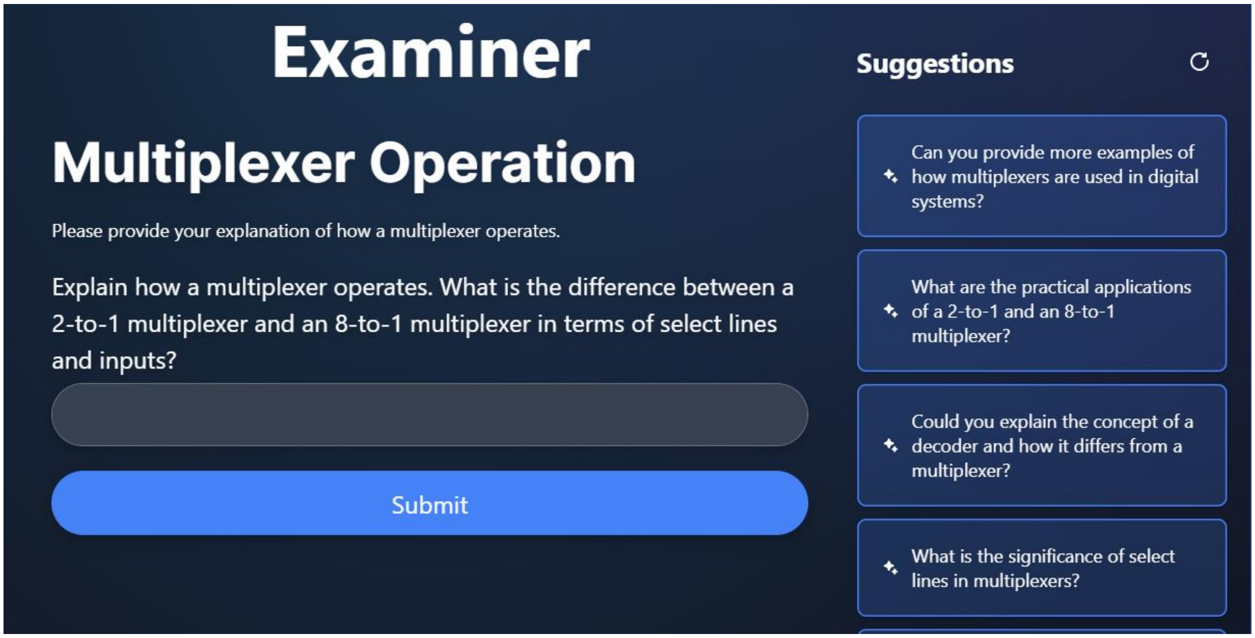


Figure 4: An example of the Modules tool for a specific course topic.

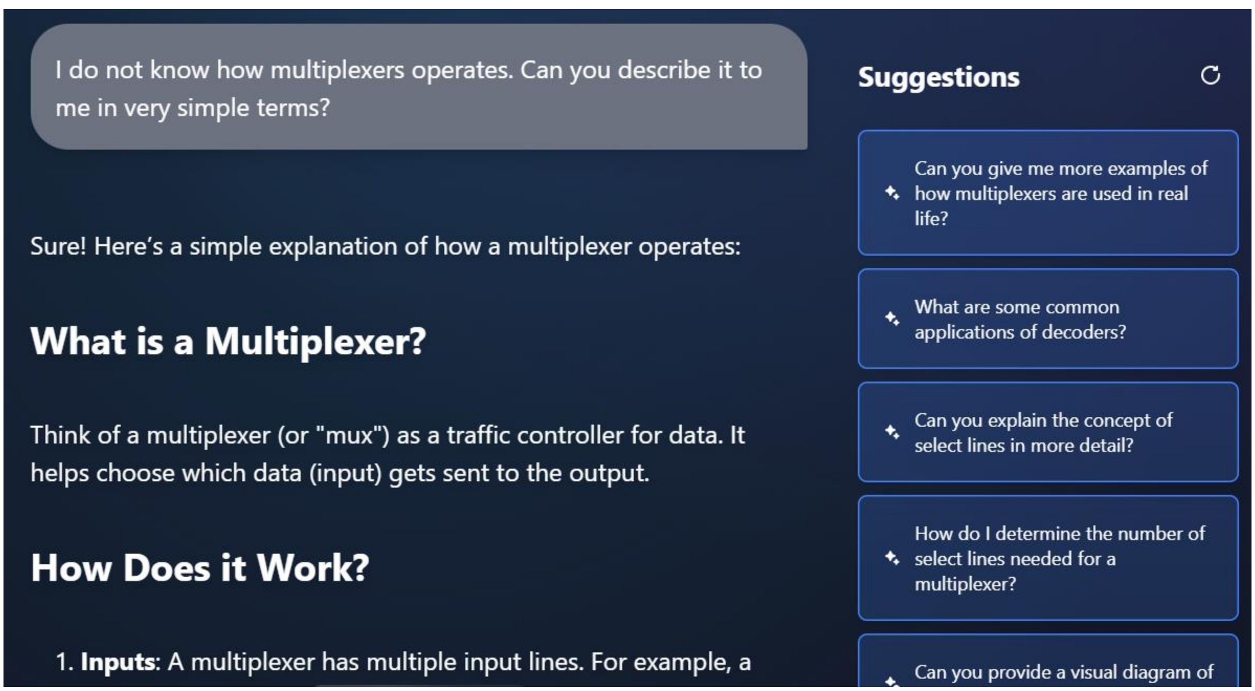


Figure 5: The student's follow-up query to Figure 4.

These new AI-driven **Textbook** and **Modules** tools served to complement the traditional study resources such as the textbook, lectures, notes, homework problems, and videos that were also available to the students. These two generative AI tools offered to the students in the Fall 2024 semester were the only resources that were not offered to students taking this introductory computing course in previous semesters.

There was a total of three surveys conducted to collect data to better understand the students' learning experiences associated with these two new AI-driven tools. These surveys were conducted every trimester of the course, including one optional survey in the first trimester (Week 3), a mandatory survey in the second trimester (Week 7), and a mandatory survey in the last trimester (Week 12). Table 1 below summarizes the dates of the surveys.

Survey #	Survey Date
Survey 1: Optional initial survey	9/13/24
Survey 2: Mandatory survey in second trimester	10/11/24
Survey 3: Mandatory survey in last trimester	11/15/24

Table 1: Survey dates

Each survey included the following three questions:

1. Did you use any of the course Chatbots when studying for homework or the midterm exam?
2. If you used the Chatbots, then did you find them useful?
3. What has been your experience so far with the Chatbots? If you used them for exam preparation, what did you find the most useful? If you did not use them for exam preparation, what was your reason for not utilizing this resource?

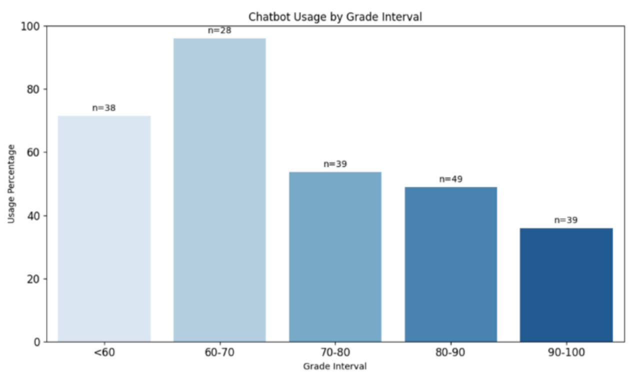
Findings:

Our analysis of student interaction with the personalized generative AI tool revealed interesting trends regarding usage patterns, reported experiences, and the relationship between usage levels and academic performance. Based on data collected and analyzed through the surveys and usage statistics, we observed the following findings. Usage is based on students reporting whether they used the chatbot or not, and usefulness is based on students reporting whether they found the chatbot useful for their studying habits.

The first research question explored the percentage of students who repeatedly used the personalized AI tool and their reported experiences. In Week 3 of the semester, an optional survey taken by 41/197 students in the class revealed that nearly 100% of the respondents had utilized the chatbot, with all these students finding it helpful for understanding conceptual material. In Week 7, a mandatory (graded for completion) survey was administered. Responses from 191/197 students showed that 53.4% of the class had used the chatbot. This percentage rose to 60.2% in Week 12 (when a similar mandatory survey was conducted with 191/197 respondents), after the first midterm exam and right before the second midterm. The rise in adoption suggests growing familiarity and trust in the tool's capabilities. Additionally, the usefulness of the chatbot remained remarkably high, with approximately 96% of students consistently reporting that the tool was beneficial for understanding course concepts, completing homework, and preparing for practice exams.

The following graphs show the trends of chatbot usage and usefulness as rated by students before the first midterm exam (Figure 6a, 6b), before the second midterm exam (Figure 7a,7b), and before the final exam (Figure 8a,8b). Surveys 2 and 3 were used as they included questions specifically about chatbot usage and usefulness together and had a much higher response rate compared to survey 1.

For the usage graph, the n- values are the total number of students in the class who fell into that grade bucket. For the usefulness graphs, the n- values are the total number of students who used the chatbot in that grade bucket. This methodology is used so that we can gather the conditional data of chatbot usefulness given chatbot usage.



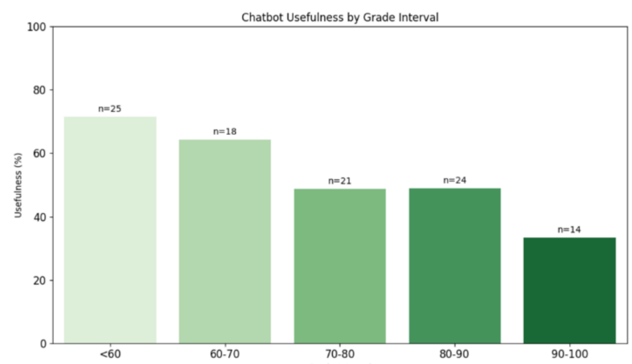


Figure 6a, 6b: Survey 2- Before Midterm 1, with buckets based on midterm 1 scores

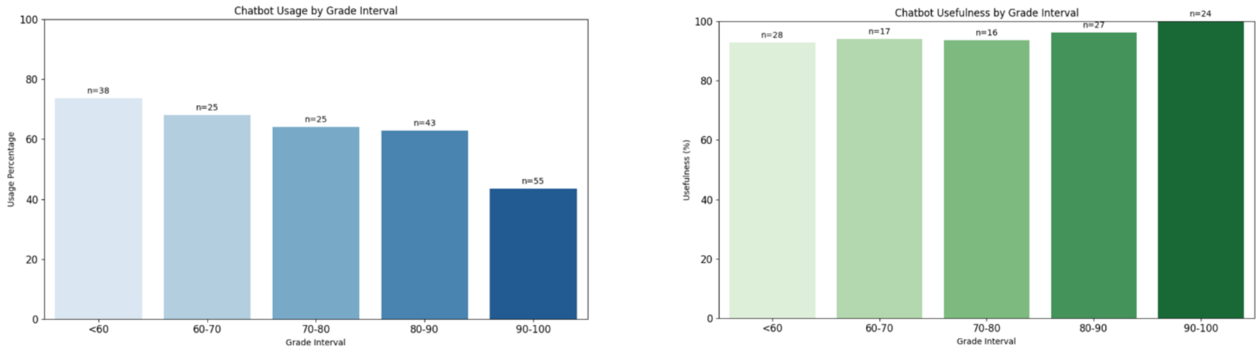


Figure 7a, 7b: Survey 3- Before Midterm 2, with buckets based on midterm 2 scores:

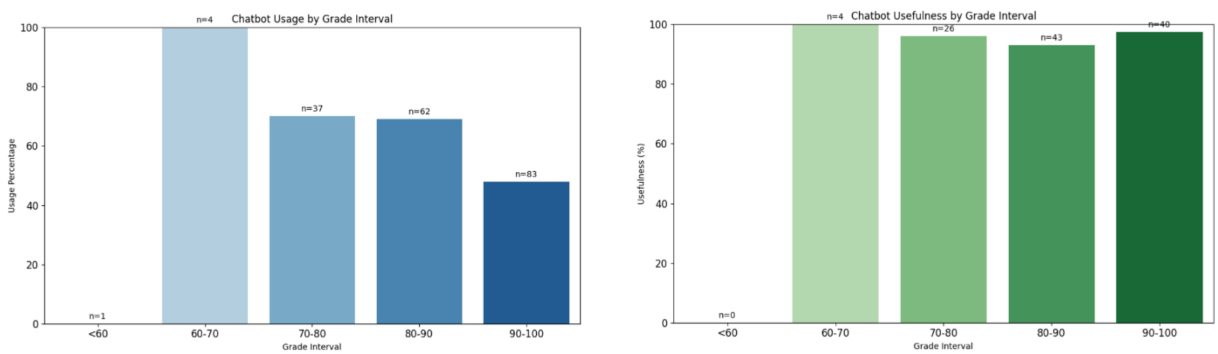


Figure 8a, 8b: Survey 3- Overall usage and usefulness, with buckets based on Final Course GPAs

As seen, there is a clear trend of increased chatbot usage as the semester progressed, and high levels of usefulness across all grade intervals, pointing at the fact that this chatbot is a versatile tool that can help all students alike. From midterm 1 to midterm 2, the average chatbot usage rose by 7%, suggesting that students realized the usefulness of this tool for reinforcing concepts.

One interesting finding was that students who used the chatbot were more likely to see positive improvements in their grade in future exams compared to students who did not use the chatbot. Analyzing the improvement data from midterm 2 scores and final scores (Figure 9), we can see that students who used the chatbot earned an increase of 1.13 points on average on their subsequent exam, while students not using the chatbot demonstrated a decrease of 1.33 points on average on their next exam.

While not statistically significant, there is a small effect of chatbot use on student scores as exemplified by Cohen's d effect size, which came out to 0.2225. T-test results yielded a p-value of 0.1393.

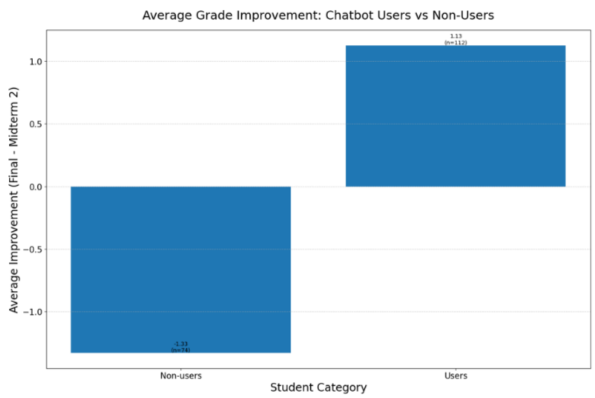


Figure 9: Improvement in score from midterm 2 to final exam for non-users vs. users.

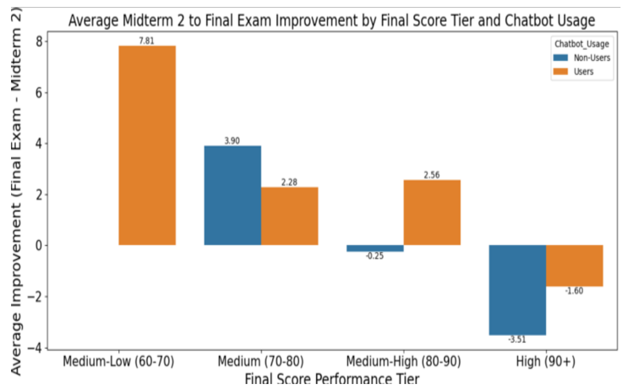


Figure 10: Midterm 2 to Final Exam Grade improvement from data processed from surveys 2 and 3

The second research question focused on reasons why some students did not use the chatbot. One commonly reported issue was a lack of trust due to occasional numerical inaccuracies, as evidenced by the following student survey comments: "*I didn't use them because I wasn't confident in them being correct. I'm hesitant to use any AI*", "*The chatbots seemed pretty useful for generating practice problems. I did not use the Chatbot for exam preparation because I wasn't too sure about accuracy.*", "*So far I have found the Chatbots to be very helpful in answering any questions I may have. They are not very good at computing answers however*". This limitation made students hesitant to rely on the chatbot to solve computational problems. Furthermore, several students indicated that they felt confident with their existing knowledge or preferred using other class resources, such as notes or past exam papers or study groups, to address specific learning needs such as practicing problems that targeted more computational skills. This is evidenced by the following student survey comments: "*I was using practice questions from the past exams, so I didn't find myself needing more practice.*", "*When studying for the exam, I focused completely on practice problems and past exams to study, and if I needed help then I'd ask my peers to teach me. I feel like that in person interaction helps me more than reading words online.*".

The third research question investigated whether there was a relationship between chatbot usage levels and students' final course GPAs. Analysis of Week 7 data indicated an inverse relationship between GPA and chatbot usage. Students with GPAs below 60% showed the highest usage rate at 71.4%, while students with GPAs above 90% had the lowest usage rate at 36.0%. An interpretation of this analysis is that the chatbot served as a valuable tool for students with lower academic performance, potentially helping to bridge gaps in understanding and mitigate learning challenges. The consistently high usefulness metric across all GPA groups (ranging from 90.5% to 100%) further highlights the tool's effectiveness in supporting diverse student needs, regardless of academic standing.

The usage of the chatbot went up as the semester progressed, indicating that the tool was well-received. It can also be seen that exam scores throughout the semester had an upwards trend for chatbot users in most final GPA buckets compared to their non chatbot user counterparts, with a detailed analysis below shown in Figure 10. While still not statistically significant, the graphs do show a trend that chatbot use positively affects improvement potential.

Overall, the findings demonstrate that personalized generative AI tools had a positive impact on learning outcomes for most users. Students particularly appreciated their ability to explain concepts, which proved to be a valuable resource for conceptual learning and exam preparation. However, addressing numerical inaccuracies and enhancing the chatbot's computational reliability could further increase its adoption and effectiveness among students, which we discuss in the next section.

To assess the overall impact of this generative AI tool on overall student performance, we compared the average course GPA and percentage of students earning D's and F's from the past 4 semesters that the instructor taught this course. All other aspects of this course offered in previous semesters were identical (with similar resources, course structure, exam difficulty, and teaching assistant support). Hence, we can consider the students who took this course in previous semesters (Spring 2024, Fall 2023, Fall 2022) comprising the control group, and students who took this course in Fall 2024 as the treatment group. The following table (Table 2) summarizes these results.

Semester (n=Class size)	Mean Course GPA	Percentage of D's and F's
Fall 2024 (n=197)	3.14	2.03
Spring 2024 (n=90)	2.83	10
Fall 2023 (n=91)	2.94	9.9
Fall 2022 (n=168)	3.22	5.95

Table 2: Mean course GPA and %DF grades

The Fall 2024 semester was the only semester when this generative AI tool was offered as a resource to students. The comparison of the mean course GPA across the four semesters does not indicate any significant trend. While not statistically significant, the percentage of D's and F's in Fall 2024 is about ⅓ to ⅕ of the corresponding value in previous semesters.

Future Work:
One key area for future work is improving the formatting and accuracy of numerical calculations. Feedback from students highlighted occasional inaccuracies in numerical problem-solving, which reduced trust in the chatbot for computational tasks. A critical improvement will involve updating the AI's prompts to require that it recomputes its answers multiple times for error-checking purposes. This iterative verification process will help ensure consistency and accuracy in the generated solutions. Additionally, context-aware prompts specifying the desired level of precision and rounding rules will help minimize ambiguity and errors. Formatting improvements, such as incorporating the ability for the tool to process LaTeX instead of just Markdown, will further enhance clarity and usability. A feedback loop will also be included in the chatbot, enabling students to flag errors or ambiguities directly within the system through a comment box per module bot. These flagged issues can then be used as training data to refine and improve the AI further. Finally, the system will undergo rigorous stress testing with edge cases, such as complex equations and extreme values, to improve accuracy.

A potential AI-driven test maker could build on the foundation of the **Textbook** and **Modules** tools, leveraging their structured alignment with course materials and interactive capabilities. Implementation would involve utilizing prompt manipulation techniques to guide the *Genux* AI Agent in generating personalized assessments. By analyzing data from student interactions with these tools, such as performance on module exercises, frequently asked textbook questions, and areas where students seek repeated clarification, the system could identify knowledge gaps and generate tailored practice exams. Prompt templates would be designed to specify the desired topic, difficulty level, and question format, to make sure the exams are aligned with course content.

The test maker would utilize adaptive testing techniques, where the difficulty of questions dynamically adjusts based on the student's responses. For instance, prompts could instruct the **Modules** chatbot to generate progressively challenging problems for topics where the student demonstrates proficiency or to create detailed, step-by-step questions for areas requiring improvement. Additionally, the test maker would provide detailed feedback for each question, linking back to relevant textbook sections or module exercises for further study. This feedback loop would be implemented using specialized prompts that instruct the agent to generate correct answers and explanations. This feature would likely be an addition to the **Modules** chatbot.

Appendix:

Data Analysis on Google Collab:
https://colab.research.google.com/drive/1THSLRLCfCnxlX0nquINq72qrwuKJMwCC?usp=sharing

References:

1. Yilmaz R., and Yilmaz, F., "The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation", Computers and Education: Artificial Intelligence, Volume 4, 2023. https://doi.org/10.1016/j.caeai.2023.100147.

2. Moore, S., Nguyen, H. A., Bier, N., Domadia, T., and Stamper, J., "Assessing the quality of student-generated short answer questions using GPT-3", EC-TEL 2022: Educating for a new future: Making sense of Technology-Enhanced Learning adoption, pp 243-257.

3. Lee, Y., Hwang, G., and Chen P., "Impacts of an AI-based *chat*bot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation", Educational technology research and development, 70, 2022, pp 1843-1865.

4. Gerling, W., Mateer, G. D., Wooten, J., and Damodaran, N., "ChatGPT has mastered the principles of Economics: Now what?", 2023, Available at SSRN 4356034, 2023.

5. Gilson, A., Safranek, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., and Chartash, D., "How does ChatGPT perform on the United States medical licensing examination? The Implications of large language models for medical education and knowledge assessment", JMIR Medical Education, 2023, 9(1), e45312.

6. Lo, C. K., "What is the impact of ChatGPT on Education? A Rapid Review of the Literature", Educational Sciences, 2023, 13(4), 410. https://doi.org/10.3390/educsci13040410.

7. Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., Lyden, S., Neal, P., and Sandison, C., "ChatGPT versus engineering education assessment: a multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity", European Journal of Engineering Education, 48:4, 559-614, https://doi.org/10.1080/03043797.2023.2213169.