Exploring Artificial Intelligence Tools for Materials Science in Engineering: A Work-in-Progress in Undergraduate Classroom Integration

Rackan Sami Mansour, Texas A&M University at Qatar

Rackan Mansour is a graduating student at Texas A&M University in Qatar, where he has pursued his interests in material science, machine learning, and design independently. His work in material science has equipped him with a deep understanding of the properties and applications of various engineering materials. Rackan has also explored machine learning, focusing on its potential to influence and transform data analysis and decision-making processes. In the realm of design, he has actively engaged in product design and computer-aided design projects, including participation in the Shell Eco-Marathon. Each of these areas reflects Rackan's versatility and dedication to mastering diverse aspects of modern engineering.

Osama Desouky, Texas A&M University at Qatar

Osama Desouky is a Technical Laboratory coordinator at Texas A&M University in Qatar. Osama is currently pursuing his Ph.D. in interdisciplinary engineering from Texas A&M University at College Station. He is responsible for assisting with experimental method courses, 3D printing, mechanics of materials, material science, senior design projects, and advanced materials classes. Osama's professional interests include manufacturing technology, materials science, 3D printing, experiments, and product design, and systems engineering for development of additive manufacturing systems.

Marwa AbdelGawad, Texas A&M University at Qatar

Dr. Marwa AbdelGawad is an Instructional Assistant Professor at Texas A&M University at Qatar. She earned her Ph.D. in Mechanical Engineering from Texas A&M University (USA), where her research focused on examining the impact of microstructure on the corrosion response and mechanical integrity of magnesium alloys used in biomedical applications, specifically orthopedic implants, which resulted in the publication of several papers in prestigious journals and presentations at conferences.

Dr. AbdelGawad's interests are centered around materials and manufacturing, with a strong focus on corrosion of advanced materials, and the study of statics and mechanics. With an extensive teaching background, she has developed a keen interest in advancing innovation in engineering education. At present, she actively explores various methods to enhance student engagement and optimize their learning experiences through curriculum and course design.

Exploring Artificial Intelligence Tools for Materials Science in Engineering: A Work in Progress in Undergraduate Classroom Integration

Abstract

The onset of user-friendly Artificial Intelligence (AI) tools has significantly disrupted traditional educational methodologies within higher education. This paper explores the application of advanced AI technologies, specifically GPT-4, to enhance student assessments in material science education. It details the strategic integration of AI to develop dynamic and personalized assessments, such as multiple-choice quizzes and open-ended case studies, aiming to redefine classroom engagement by adapting to diverse student needs and learning environments.

AI is a promising transformative tool in educational assessment processes within material science. Utilizing GPT-4, the study investigates the creation of diverse assessment forms, showing AI's capability to tailor assessments to individual learning requirements and curriculum standards. This approach deepens student engagement and advances educational strategies by equipping educators with dynamic tools that respond to the evolving educational landscape. The current study particularly emphasizes prompt engineering with AI, a critical element in optimizing AI's utility for generating advanced, curriculum-aligned assessments. It assesses how effectively crafted prompts can guide AI to produce more relevant educational content, thereby enhancing learning experiences. As effective prompts are developed, GPT-4's potential to customize assessments to meet specific student needs and address the complexities of material science theories is highlighted, presenting a valuable approach to boost student engagement, and understanding.

These AI-driven methodologies aim to enhance the creative process in educational material development, offering educators an expanded array of tools for designing customized instructional materials. The role of AI in enriching educational content is expected to significantly elevate student engagement and deepen their comprehension of complex material science concepts. The study documents iterative testing and refinement of AI tools in producing and improving educational materials, providing tangible examples of AI's contributions to educational innovation. The results add important insights to the discourse on integrating AI into engineering education, underscoring its potential as a collaborative tool in a rapidly evolving academic environment.

Introduction

The field of engineering education is witnessing a transformative era with the advent of user-friendly Artificial Intelligence (AI) tools [1] [2] [3]. The introduction of this shift into mechanical engineering and material science disciplines could significantly enhance traditional instructional methods, leveraging the vast capabilities of AI [4] [5]. The integration of AI tools like GPT-4 and DALL·E 3 represents a pioneering effort in this direction [1]. This paper aims to explore and document the impact of AI on the pedagogical landscape of material science education.

At the core, this study investigates how AI tools can aid in the creation of student assessments, potentially redefining student engagement and learning processes. It focuses on two facets of student assessments: summative assessments, such as multiple-choice quizzes, and formative assessments, including open-ended case studies. By employing AI to generate these assessments, tailored to the curriculum's specific requirements and the diverse learning needs of students, this process introduces a new level of customization and complexity [2] [4]. This approach not only challenges students in more meaningful ways but also ensures alignment with learning objectives, allowing for a comparison of GPT's effectiveness across assessment types [6].

Following the initial exploration into AI's capacity for generating student assessments, this project will utilize evaluation methodologies identified from the literature. These methodologies focus on tools for assessing the effectiveness of AI-generated educational content. Drawing upon insights from an investigation into the application of AI in educational settings, a framework for evaluating the quality and pedagogical value of AI-generated assessments, like MCQs and case studies, is proposed [2]. These tools are crucial for refining AI-generated content to meet the curriculum's specific needs and the diverse learning profiles of students in material science education, validating the utility and enhancing the effectiveness of AI-generated materials.

The unveiling of AI-powered tools like ChatGPT has sparked considerable debate regarding their impact on the traditional educational landscape, particularly in engineering disciplines. While some view these advancements as potential threats to conventional teaching methods and the educator's role, others foresee a new paradigm in knowledge dissemination that could significantly enhance learning outcomes. The incorporation of AI tools in engineering education holds the promise of redefining student engagement and instructional strategies, yet it also raises concerns over academic integrity and the potential loss of jobs within the teaching profession [1].

This study aims to shed light on various aspects of AI implementation in material science education, highlighting its role in bridging gaps between theoretical concepts and practical applications, and its influence on the overall learning experience.

Current Methods in Student Assessment

The current landscape of student assessment in material science education predominantly relies on conventional methods, including standardized tests, written assignments, and practical laboratory work [6]. These assessment tools are designed to evaluate a student's understanding of theoretical concepts, practical skills, and their ability to apply knowledge in real-world scenarios [6]. Standardized tests and exams often consist of a mix of multiple-choice questions, short answer questions, and problem-solving exercises, aimed at measuring students' retention, comprehension, and analytical abilities [6].

The integration of AI into material science education provides a unique opportunity to enhance traditional assessment methods, promoting a more dynamic, interactive, and student-centered approach. Leveraging AI for personalized assessments through adaptive learning systems and simulated scenarios allows education to evolve towards fostering higher-order critical thinking skills [2] [5] [7]. Bloom's Taxonomy—encompassing knowledge, comprehension, application, analysis, synthesis, and evaluation—offers a structured method for designing educational experiences that align with these advancements [8] [9] [10]. Utilizing this model, the study will employ AI to generate MCQs for exams, testing this generation method with evaluation tools as outlined in literature [9] [10]. The capability of AI to tailor content and assessments to Bloom's various levels encourages a shift from mere knowledge recall to deeper understanding, application in novel situations, and the creative synthesis of new ideas [9] [10]. This alignment of AI-driven tools with Bloom's hierarchical model not only challenges students but also prepares them for complex real-world problems, making education more comprehensive and engaging [9] [10].

Methodology

To harness AI at its fullest potential within educational content creation, or any other field for that matter, it is imperative to master precise and effective prompting [9] [11] [12]. This approach empowers educators to guide AI towards generating material that not only aligns with specific pedagogical goals but also engages students on a deeper level [9]. Through careful crafting of prompts, faculty can ensure that AI tools like GPT-4 produce content that is directly relevant to the curriculum, enriched with the complexity and nuance required in material science education [4] [5]. By focusing on prompt engineering, educators can effectively utilize AI to create customized, high-quality educational resources that meet the rigorous demands of academic standards [9] [11] [12].

Prompt Engineering for Educational Content Creation:

Prompt engineering for large language models (LLMs) like GPT4 is a pivotal task that significantly enhances their performance across diverse natural language processing tasks [11] [4]. The essence of this process is to craft prompts that effectively direct the LLM's processing, ensuring it produces the desired output [4]. However, achieving the right prompt is not straightforward due to the LLM's sensitivity, which often leads to a labor-intensive cycle of trial-and-error to find the most effective instruction [11].

Effective prompt engineering demands detailed, carefully considered prompts that go beyond mere task descriptions. It involves a nuanced balance between providing enough context to guide the AI and avoiding the imposition of biases that might limit creative output or skew the AI's response in unintended directions [1] [5]. For instance, incorporating a step-by-step reasoning template or providing contextual background can significantly enhance the AI's performance in generating educational materials such as visualizations, physical models, and scripts [8]. This iterative refinement process, underscored by the examination of AI-generated responses and subsequent prompt adjustments, is critical in honing the AI's output to align closely with educational objectives.

Assessment and Iteration

The evaluation of AI-generated educational content is an integral component of ensuring its effectiveness and alignment with educational standards and objectives [7]. This process involves systematically assessing the quality, accuracy, and pedagogical value of the content created by AI tools, such as GPT-4 and DALL·E 3, in the context of material science education. The assessment is not a one-time task but a continuous cycle of feedback and iteration, where educators and subject matter experts review AI-produced materials and provide specific feedback for improvements.

Key to this process is the establishment of clear, measurable criteria for evaluating content [ ]. These criteria may include the scientific accuracy of the information presented, the clarity and comprehensiveness of explanations, and the engagement level of the materials. Incorporating student feedback and performance data into this assessment provides a direct measure of the content's impact on learning outcomes, further informing the iterative refinement process.

Based on the assessment findings, adjustments to the AI's prompts may be necessary to better tailor the content to students' needs and learning objectives. This iterative cycle of creation, evaluation, and refinement ensures that the AI-generated educational materials not only meet the high standards expected in educational settings but also evolve in response to feedback and emerging educational challenges. Through this process, AI-generated content would become more refined, targeted, and effective in enhancing student learning and engagement in material science.

Evolving Assessment Methods with AI Integration

Mirroring the transformative impact observed with the introduction of calculators in the assessment of mathematical skills, AI's incorporation into material science education prompts a re-examination of traditional evaluation frameworks. Our approach seeks to align assessment strategies with the enhanced capabilities introduced by AI tools.

MCQ Generation

Continuing from the strategy outlined, the process of generating MCQs for material science will involve several key steps, informed by the methodologies presented in the referenced study [9]. Initially, a comprehensive collection of learning objectives (LOs) specific to the material science curriculum will be established. These LOs will serve as the foundational guide for content generation, ensuring that each MCQ directly contributes to the educational goals.

Subsequently, leveraging the capabilities of GPT-4, prompts will be designed to encapsulate the essence of each LO, incorporating the contextual backdrop of the material science domain. This prompt engineering process is critical, as it dictates the specificity and relevance of the AI-generated content. The prompts will be formulated to encourage the generation of MCQs that not only challenge students' understanding and application of concepts but also promote analytical thinking and problem-solving skills relevant to material science.

Upon generating a preliminary set of MCQs, an iterative evaluation and refinement phase will follow. This phase involves assessing the MCQs against a set of quality criteria derived from the study's findings, such as clarity of language, alignment with LOs, and the plausibility of distractors [9]. Feedback from this evaluation will be utilized to fine-tune the prompts and generation parameters, enhancing the quality and relevance of the generated MCQs.

This adaptive approach ensures that the MCQs produced are pedagogically valuable. By integrating this AI-assisted methodology, the project aims to modify existing systems to create a robust, scalable system for developing high-quality assessments that can be customized to suit a wide range of topics within the field of material science.

To generate a set of multiple-choice questions (MCQs) on the topic of strain hardening in material science, this study employed the methodology outlined by Doughty [9]. This involved defining the specific learning objectives (LOs) within the context of material science education, focusing on the concept of strain hardening and its impact on material properties. According to the guidelines provided by [6], the process began with identifying the appropriate Bloom's Taxonomy level for the Los. For this example, the strain hardening from the 300 level Materials and Manufacturing Selection in Design course, pertained primarily to the comprehension and application levels.

Following this, a detailed prompt was crafted, designed to guide the AI in generating questions that not only cover the theoretical aspects of strain hardening but also probe students' ability to apply this knowledge in analyzing material behavior. This was done by feeding the AI with the learning objectives of the class and the paper detailing Doughty's [9] methodology to generating MCQs. The generation of MCQs was then executed using the AI tool, with the implication to ensure that each question comprised a clear stem, one correct answer, and several plausible distractors, all aimed at testing students' understanding and application of strain hardening principles. Figure 1 shows the prompt given to GPT4 to generate the MCQs, after feeding it [9]. Figures 2 and 3 display the output of the prompt.

You

i am going to give you new learning outcomes based on MEEN 360 for strain hardening. i want you to generate a new quiz:

Describe how microstructure evolves and the corresponding changes in mechanical properties due to processing and heat treatment. this will be achieved based on the following learning objective: Quantify the strain hardening of a metal and explain why strain hardening may significantly increase the force required to deform a metal

Figure 1: Prompt to Chat GPT 4 to generate an MCQ of 5 questions.

ChatGPT

Based on the provided learning outcomes and objective for the MEEN 360 course on strain hardening, here is a new set of MCQs:

1. **What is the primary effect of strain hardening on the microstructure of metals?**
   - A) Reduces the size of the grains within the metal.
   - B) Increases the number of dislocations within the metal. *(Correct Answer)*
   - C) Decreases the number of dislocations within the metal.
   - D) Has no significant effect on the metal's microstructure.
2. **Strain hardening affects the mechanical properties of metals by:**
   - A) Decreasing the metal's yield strength.
   - B) Increasing the metal's ductility.
   - C) Increasing the metal's yield strength and hardness. *(Correct Answer)*
   - D) Making the metal more brittle.
3. **During the process of strain hardening, if the force required to deform a metal increases significantly, it is primarily due to:**
   - A) A decrease in the metal's temperature.
   - B) An increase in the dislocation density within the metal. *(Correct Answer)*
   - C) A reduction in the metal's internal stress.
   - D) The metal reaching its melting point.

Figure 2: First 3 MCQs.

4. **How does strain hardening influence the force required to further deform a metal?**
   - A) It has no influence; the force remains constant.
   - B) It decreases, making deformation easier.
   - C) It increases significantly, requiring more force for further deformation. *(Correct Answer)*
   - D) It initially increases but decreases after a certain threshold.
5. **Quantify the effect of strain hardening: After undergoing strain hardening, the yield strength of a metal is observed to:**
   - A) Remain unchanged.
   - B) Double in value.
   - C) Increase, but the exact amount depends on the metal and extent of deformation. *(Correct Answer)*
   - D) Decrease by half.

Figure 3: Last 2 MCQs.

It should be noted that, in Figures 2 and 3, the AI's approach to generating MCQs was overly direct in addressing the LOs. This level of directness, while beneficial for clarity, potentially undermines the development of analytical skills by not fully challenging students to apply their understanding in varied contexts. Ideally, assessments should not only test knowledge but also encourage its application, analysis, and synthesis. Future iterations of MCQ generation will aim to properly balance addressing LOs and promoting cognitive engagement. The goal is to refine the AI's output to yield questions that compel students to think critically and apply their knowledge more broadly.

This issue could be addressed by providing the AI with a more comprehensive set of course materials, thereby enriching its understanding of the learning objectives. With a deeper reservoir of context, the AI would be better equipped to generate MCQs that are not only aligned with the learning objectives but also sufficiently sophisticated to meet the established success criteria. As it stands, the current MCQs are too rudimentary to effectively embody the nuanced success criteria desired.

Future Work

Building on the findings and methodologies implemented in this study, future research will further refine and expand the use of AI tools in material science education. The next phase will explore deeper integrations of AI for creating more dynamic and responsive learning environments. Specific areas for future exploration include:

1. **Advanced Prompt Engineering**: Continued development of more sophisticated prompt engineering techniques to improve the precision of AI-generated content. Future work will focus on minimizing the sensitivity of AI responses to prompt nuances and enhancing the AI's ability to generate content that aligns more closely with advanced educational objectives.
2. **Broader Assessment Types**: Expansion of AI-generated assessments to include more varied formats beyond multiple-choice quizzes and case studies. This could involve the development of interactive simulations and real-time problem-solving scenarios that utilize AI to adapt challenges based on student performance dynamically.

3. **Ethical and Practical Considerations**: Addressing ethical considerations and the practicality of AI in educational settings, including issues of data privacy, student autonomy, and the potential for AI to replace traditional educational roles. Future research will aim to develop guidelines and frameworks that ensure ethical AI use in education.

References

[1] K. B. Mustapha, E. H. Yap and Y. A. Abakr, "Bard, ChatGPT and 3DGPT: a scientometric analysis of generative AI tools and assessment of implications for mechanical engineering education," *Interactive Technology and Smart Education,* vol. 1, no. 1, 2024.

[2] J. Qadir, "Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education," in *IEEE Global Engineering Education Conference (EDUCON)*, 2023.

[3] J. E. Duah and P. McGivern, "How generative artificial intelligence has blurred notions of authorial identity and academic norms in higher education, necessitating clear university usage policies," *IJILT,* 2024.

[4] T. Jie, J. Hou, Z. Wu, P. Shu, Z. Liu, Y. Xiang, B. Gu, N. Filla, Y. Li, N. Liu and X. Chen, "Assessing Large Language Models in Mechanical Engineering Education: A Study on Mechanics-Focused Conceptual Understanding," *arXiv preprint,* 2024.

[5] M. Frenkel and H. Emara, "ChatGPT & Mechanical Engineering: Examining performance on the FE Mechanical Engineering and Undergraduate Exams," *Arxiv,* 2023.

[6] S. R. M. Yusop, M. S. Rasul, R. M. Yasin, H. U. Hashim and N. A. Jalaludin, "An Assessment Approaches and Learning Outcomes in Technical and Vocational Education: A Systematic Review Using PRISMA," *Sustainability,* 2022.

[7] S. Elbanna and L. Armstrong, "Exploring the integration of ChatGPT in education: adapting for the future," *Management & Sustainability: An Arab Review,* vol. 3, no. 1, 2024.

[8] M. Forehand, "Bloom's Taxonomy," *Emerging perspectives on learning, teaching, and technology,* vol. 41, no. 4, 2010.

[9] J. Doughty, "A Comparative Study of AI-Generated (GPT-4) and Human-crafted MCQs in Programming Education".

[10] C. C. L. Kwan, "Exploring ChatGPT-Generated Assessment Scripts of Probability and Engineering Statistics from Bloom's Taxonomy," in *Technology in Education. Innovative Practices for the New Normal*, 2023.

[11] Q. Ye, M. Axmed, R. Pryzant and F. Khani, "Prompt Engineering a Prompt Engineer," *arXiv,* 2024.

[12] B. Lund, "The prompt engineering librarian," 2023.