



COMMENTARY

Title: Glimpses of the Use of Generative AI and ChatGPT in Medical Education

Authors: Nilesh Kumar Mitra, Ebenezer Chitra

Submitted Date: 04-08-2023

Accepted Date: 10-03-2024

Please cite this article as: Nilesh Kumar Mitra, Ebenezer Chitra. Glimpses of the Use of Generative AI and ChatGPT in Medical Education. Education in Medicine Journal. (early view).

This is a provisional PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article.

ARTICLE INFO

Submitted: 04-08-2023

Accepted: 10-03-2024

Glimpses of the Use of Generative AI and ChatGPT in Medical Education

Nilesh Kumar Mitra¹, Ebenezer Chitra²

¹Human Biology Division, School of Medicine, International Medical University, Bukit Jalil 57000, Kuala Lumpur, Malaysia

²School of Health Sciences, International Medical University, Kuala Lumpur, Malaysia

ABSTRACT

Artificial intelligence (AI) refers to computer systems that can perform tasks that typically require human intelligence. This is achieved by using algorithms and neural networks for machine learning and deep learning. ChatGPT is an AI-powered chatbot that can generate responses on any topic based on the user's input or queries. In this review, the focus is on the potential applications of generative AI and ChatGPT in the fields of medical and health professions education. It is important to educate both educators and students about the impact of using generative AI, such as ChatGPT, which is widely used through generative adversarial networks (GANs). In the field of healthcare, GANs can analyze vast datasets to assist in diagnosis, patient data management, and analysis. Students use ChatGPT to obtain factual answers, write papers and translate languages. It can help students with their assignments by summarizing literature reviews and generating new ideas. In medical education, educators use ChatGPT to develop learning activities, assessments, and curricula, enhance student learning, and even generate research papers for publication. However, students need to use generative AI carefully so that it does not impede their ability to think critically or write effectively. Guidelines are being formulated in different institutions to regulate the use of this immensely powerful tool. This review could highlight the scope of incorporating ChatGPT into the field of medical education.

Keywords: Artificial intelligence, generative adversarial networks, ChatGPT

CORRESPONDING AUTHOR

Nilesh Kumar Mitra, Human Biology Division, School of Medicine, International Medical University, Bukit Jalil 57000, Kuala Lumpur, Malaysia
Email: nileshkumar@imu.edu.my

INTRODUCTION

Artificial intelligence (AI) explores how computer systems can perform tasks requiring human intelligence. John McCarthy introduced AI at a Dartmouth conference in 1956 (1). He defined AI as “the science and engineering of making intelligent machines, especially computer programs.”

The use of AI relates to using computers to simulate human intelligence, but AI does not have to confine itself to biologically observable methods (2). A comparison between the human brain and the computer shows that the human brain is superior due to its learning capability, the trillions of active and renewable connections between biological neurons, and the economy of its energy consumption (3) [Fig. 1].

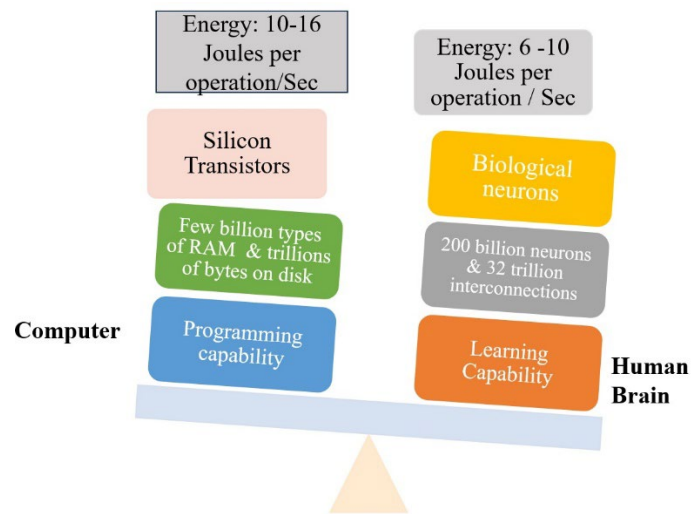


Figure 1. Relationship between the computer and the human brain

Computer programming can solve several challenging problems, such as searching, pattern recognition, learning, planning, and induction. In 1961, Marvin Minsky suggested that computers can enhance their problem-solving abilities through pattern recognition techniques, leading to increased efficiency (4). Subsequent authors worked on building up a cognitive architecture, which would be able to make a good diagnosis of the problem using multiple criteria (5). The search area can be limited by combining pattern recognition with learning from accumulated experiences. By using planning methods, the given search can be replaced with simpler and more appropriate exploration. To manage broader problems, constructing models of environments will be required

using the induction process. The core ingredients of AI have been grounded in mathematics and computing as high-performance computing infrastructure, large datasets, and algorithm development (6). AI tools are intelligent machines that learn continuously and perform tasks that would require human intelligence, such as language translation, decision-making and speech recognition (7). AI systems apply a component of the technology known as machine learning (ML), composed of specialized hardware and software. Large volumes of labeled training data are ingested, and rules are created by programming about turning the data into usable information. ML algorithms use past data to forecast new output values, allowing software applications to produce predictive results (8). McCulloch and Pitts (9) were the first to create a prototype of an artificial neuron and artificial neural networks. Subsequently, this prototype has been described as a hallmark event and fundamental to the development of cognitive science and AI (10). ANNs are faster at processing information, and they can, for example, be used to predict dementia from a clinical dementia rating, asymptomatic cirrhosis of the liver from the demographic and laboratory data of chronic hepatitis patients, 1-year survival in neuro-oncology patients with brain metastasis, mortality rates in hip fracture patients and the risk of developing hypertension from the previous demographic and body mass index data of patients (11, 12, 13, 14, 15).

Unlike the predictive AI described above, generative AI such as Chat Generative Pre-Trained Transformer (ChatGPT) is an unsupervised or partially supervised ML framework that allows the feeding of a vast volume of data, discovers the patterns, and uses learned ways to produce new data. ChatGPT is a newly developed conversational chatbot that uses natural language processing to provide responses that resemble those of a human in response to user input. This technology has gained global recognition for its remarkable ability to produce logical, organized, and informative

answers (16). A recent study showcased the use of ChatGPT in preparing for English teaching in Malaysia (17). According to this qualitative research by Lou (17), educators have insufficient knowledge about ChatGPT. Following the release of ChatGPT, several conversational chatbots appeared on the internet: Google Bard, Bing AI, Replika, Claude, Jasper.ai, Baidu's UNIT, Meta's Wit.ai, and Perplexity. However, its conversational fluency, ability to reinforce learning from human feedback, and user-friendly interface made ChatGPT popular, with approximately 180 million users currently using it (18). To ensure informed decision-making about ChatGPT, university faculty must be provided with more information and education about its use (19). This review analyzes how generative AI and ChatGPT can be used in medical and health professions education.

LITERATURE SEARCH METHODS

This review aims to improve educators' awareness of the potential use of generative AI and ChatGPT in medical and health professions education. Given its broad scope of use, it will focus on the specific areas described in the literature or experienced by the authors while using the desktop and mobile versions of the ChatGPT platform. We used a structured search strategy to gather articles from EBSCOhost, ProQuest, PubMed, and Google Scholar. Only articles written in English were selected. The following search terms were used: "artificial intelligence," "generative AI," "ChatGPT," "GPT-4" in combination with "medical informatics," "machine learning," "progress testing," "curriculum development," "medical student," "research," and "assessment." The articles selected focused on AI, ChatGPT, and GPT-4 studies using quantitative or qualitative research. Some articles describing the in-depth concepts of AI were included to explain the work-

process mechanism of ChatGPT. Articles published within the past ten years were included. However, we were compelled to use specific articles beyond the initial time range in certain areas where only a few publications were available within the past ten years. Background information on AI, generative AI, and ChatGPT was obtained from the review articles and reviews by technology experts published in the latest editions of renowned international magazines to capture the latest changes in the configuration of the versions of ChatGPT.

GENERATIVE AI

Generative adversarial networks (GANs), are algorithmic architectures consisting of two models—a generator model and a discriminator model. During the training process, the generator model produces a large amount of both fabricated data and authentic data when a dataset is fed into the model. Two models operate together in this process. The generator model generates artificial data to try to deceive the discriminator model, while the discriminator model verifies that it can differentiate between factual and fabricated data (20). GANs have worked as an efficient generative model to facilitate the generation of images, videos, and voice. Since 2014, many new GANs-based models have been proposed to improve the stability and quality of the generated results (21) and GANs have played a key role in simulating various types of learning environments and helped develop practical skills in a virtual teaching assistant model. The experimental results have indicated that the model improved students' learning outcomes, increased their motivation for learning and satisfied individual student needs (22). In one study, a GAN-based tool was used to generate a sequence of images from descriptive text. This tool was used in the visual storytelling process facilitating technology-enhanced learning with improved logical reasoning (23).

INTRODUCTION TO ChatGPT

ChatGPT is a chatbot developed by OpenAI and launched on November 30, 2022. It is built on top of OpenAI's GPT-3 large language model (LLM) and has been fine-tuned (involving an approach to transfer learning) using both supervised and reinforcement learning techniques (24). Subsequently, the GPT-3.5 and GPT-4 models were released on November 6, 2023 and March 14, 2023, respectively, via the chatbot product ChatGPT Plus. Previous AI-enabled chatbots had limitations, including a lack of understanding of context and decision-making capability and short dialogue exchanges due to limited memories. The GPT-1, introduced in June 2018, used books as training data to predict the next word in a sentence and demonstrated language understanding tasks in an unsupervised learning model. The GPT-2, introduced in February 2019 and later relaunched in November 2019, produced coherent multi-paragraph text. Compared to the GPT-1, trained with 117 million parameters, and the GPT-2, trained with 1.5 billion parameters, the GPT-3 was trained with 175 billion parameters. The LLM capability of the GPT-3 allowed advanced text-generation capacity including tasks such as answering factual questions, drafting emails, writing articles, creating poetry, generating programming code, and even translating between languages. Although the GPT-3 could understand context, make decisions, and process long threads of dialogue to continue conversations humanly, academia was hesitant to use it. The references it generated were not found on Google Scholar or PubMed, and the responses related to the scientific disciplines were superficial and needed more critical analysis. The GPT-3.5 and GPT-4 worked on these limitations and were released with exponential improvements such as better model alignment, an understanding of user intention, and a lower likelihood of generating offensive, wrong, and

dangerous outputs. The GPT-4 model could connect to the internet in real time, increasing factual accuracy, and could change its behavior according to the user request (25).

With GPT, the user can draft documents, write computer code, answer questions about a knowledge base, analyze texts, create conversations, translate languages, and simulate characters for games. The present version of the OpenAI API platform is available via <https://platform.openai.com/examples> and presents 30 examples of applications for use in education or business (26). The use of ChatGPT as a writing assistant has made it easier to produce well-structured reviews from a vast amount of text. It has also assisted in translating content from one language to another and provided students with novel ideas for their assignments (27). Faculty members in health professions education use ChatGPT to generate questions. After comparing the generated items with those produced by the faculty from an accredited nursing program, it was found that the clarity, grammar, and difficulty levels of the items were similar (28).

ChatGPT IN PROGRESS TESTING

Medical and health science educators often employ Bloom's taxonomy to plan and map the assessment of learning activities. The taxonomy table has identified four categories of knowledge, namely, factual, conceptual, procedural, and metacognitive (29). Resources containing accurate knowledge about medicine are often distributed and arranged in a non-interactive manner in textbooks and online resources. ChatGPT offers detailed information on specific topics, including technical terms related to the subject matter. It is essential to assess whether ChatGPT, an AI language model with access to information but not deep knowledge, can provide answers based on a critical understanding of concepts. A group of researchers from a medical school in Germany used the ChatGPT interface to input 400 multiple-choice questions (MCQs) from the Berlin

progress test. These questions covered 27 medical specialties and 14 organ systems. Only 65.5% of the progress test questions answered by ChatGPT were correct. Each answer had 35 words on average, and it took 22.8 seconds to respond to a single MCQ. ChatGPT provided a detailed explanation along with those answers that were correct (30).

The United States Medical Licensing Examination (USMLE) is the exam used to determine eligibility for a medical license in the United States. Examinees must answer at least 60% of MCQs accurately to pass the USMLE. ChatGPT's GPT-4 version had an accuracy rate of 86.72% for Step 1, 91.5% for Step 2, and 85.23% for Step 3 in the USMLE self-assessment exam. The case study has shown GPT-4's ability to discuss medical reasoning, create explanations for students, and even craft new scenarios around the medical cases (31). The ability of ChatGPT to answer questions from higher-order examinations indicates that students can use it to learn complex concepts of different medical subjects.

The ChatGPT is constructed on an LLM that can aid clinical decision-making. A high-quality MCQ can be generated when a specific and well-defined prompt is provided. These questions can help develop personalized learning and formative assessments for students [Fig. 2].

Prompt: Please write a clinical scenario and MCQ testing the outcome of the blood supply of thalamus

Response:

Clinical Scenario: A 67-year-old man with a history of hypertension and smoking presented to the emergency department with acute sudden-onset right-sided weakness and sensory loss. His wife noticed that he started to have difficulty with speech and right-sided facial droop. On physical examination, you notice motor weakness and sensory deficits, primarily on the right side of his body. His speech was slurred, and he had a homonymous hemianopia on visual field testing. After the initial assessment, an emergent brain MRI revealed a small acute infarct positively identified in the left thalamus.

MCQ:

The thalamus receives its blood supply from several arteries. Based on the patient's symptoms and the location of the infarct in the MRI, which artery is most likely affected in this patient?

1. Posterior Cerebral Artery (PCA)
2. Middle Cerebral Artery (MCA)
3. Anterior Cerebral Artery (ACA)
4. Pons Artery
5. Anterior Choroidal Artery

Correct Answer: 1. Posterior Cerebral Artery (PCA)

Explanation: The thalamus receives its blood supply primarily from branches of the Posterior Cerebral Artery (PCA), particularly from the thalamoperforating and thalamogeniculate arteries. An infarct in the PCA would, therefore, lead to a wide variety of manifestations, depending on the location and extent of ischemia, but these commonly include contralateral hemisensory loss, hemiparesis and visual deficits, as seen in the clinical scenario.

Figure 2. Prompt and response to generate a clinical anatomy MCQ from GPT-4

ChatGPT IN CURRICULUM DEVELOPMENT

To shift medical education practices from being information-driven to AI-driven and implementing LLM-based curriculum development requires the identification of potential concerns and challenges. AI can help in needs analysis, identification of content gaps, and comprehensive coverage of subjects. With the help of GPT-4, we can create measurable learning objectives and receive content suggestions (32, 33). Large language models can suggest suitable teaching methodologies based on best practices in medical education. When requesting a lesson plan from ChatGPT, the user needs to provide sufficient information on the topic's complexity, the student's level, and any other relevant details, such as common mistakes or challenging concepts (34). Fig. 3 shows the example of a prompt and response regarding the generation of a lesson plan on introductory algebra (OpenAI, 2023) (26) [Fig. 3].

Generate a lesson plan for a specific topic.

Prompt

USER Write a lesson plan for an introductory algebra class. The lesson plan should cover the distributive law, in particular how it works in simple cases involving mixes of positive and negative numbers. Come up with some examples that show common student errors.

Sample response

Lesson Plan: Introduction to the Distributive Law

Objective:

By the end of the lesson, students will be able to understand and apply the distributive law in algebra, particularly in cases involving positive and negative numbers.

Materials:

- Whiteboard and markers
- Student notebooks and pencils
- Worksheets for practice problems

Introduction (10 minutes):

1. Begin the lesson by reviewing the basic concepts of algebra, focusing on the operations of addition, subtraction, multiplication, and division.
2. Introduce the distributive law as a rule in algebra that allows us to multiply a single term by two or more terms inside a set of parentheses. Write the general form of the distributive law on the board: $a(b + c) = ab + ac$.

Direct Instruction (15 minutes):

1. Explain that the distributive law applies to both positive and negative numbers.
2. Demonstrate the distributive law with a simple example involving positive numbers, such as $3(4 + 2) = 3 \cdot 4 + 3 \cdot 2$.
3. Show how the distributive law works with negative numbers, using an example like $-2(5 - 3) = -2 \cdot 5 + (-2 \cdot -3)$.
4. Emphasize that when distributing a negative number, the sign of the number inside the parentheses changes.

Guided Practice (15 minutes):

1. Provide students with a set of practice problems that involve the distributive law with both positive and negative numbers.
2. Work through the first few problems as a class, encouraging students to explain their steps and reasoning.
3. Allow students to complete the remaining problems in pairs or small groups, offering guidance and clarification as needed.

Common Student Errors (10 minutes):

1. Discuss common errors students make when applying the distributive law, such as forgetting to distribute to all terms inside the parentheses, or not changing the sign when distributing a negative number.
2. Show examples of these errors, such as incorrectly simplifying $-3(2 - 4)$ as $-6 - 4$ instead of $-6 - (-12)$.
3. Reinforce the correct application of the distributive law and the importance of careful attention to signs.

Independent Practice (15 minutes):

1. Distribute a worksheet with additional problems for students to complete independently. These problems should include a mix of positive and negative numbers.
2. Circulate the room to answer questions and provide assistance as needed.

Closure (5 minutes):

1. Review the key points of the lesson, emphasizing the distributive law and its application to positive and negative numbers.
2. Ask students to summarize what they learned in their own words.
3. Collect the worksheets for assessment.

Homework:

Assign additional problems for students to practice applying the distributive law at home.

Assessment:

Assess student understanding through their participation in class discussions, their performance on the practice problems, and their completion of the independent practice worksheet.

Figure 3. Example of a prompt from user and response generation of a lesson plan [26, <https://chat.openai.com>]

The journal, JMIR Medical Education, interviewed ChatGPT, which demonstrated its ability to suggest appropriate teaching methods in medical education (35). When asked how generative language models can help create personalized learning experiences on diabetes, ChatGPT proposed interactive simulations, concept maps, case studies, educational videos, quizzes, and self-assessments. ChatGPT was then prompted to generate a simulation, which shows how insulin interacts with glucose to regulate blood sugar levels (35).

When using AI tools such as ChatGPT to aid in the medical curriculum, educators need to verify the information produced using reliable, evidence-based resources. Sometimes, ChatGPT may miss some steps and provide references that do not exist. It may also erroneously include a link to

an article that is outside the topic. One of the significant issues in LLM is the occurrence of AI hallucinations or responses that the training data cannot explain. (36).

ChatGPT AND STUDENT LEARNING

The free version of ChatGPT-3.5 offers assistance to students with grammar, summarizing, text-to-command conversion, English translation, data classification, unstructured data parsing, and natural language conversion to Python. Both GPT-3.5 and GPT-4 have the potential to create chatbots or virtual assistants that can help students engage in problem-solving and answer questions through natural language interaction. This can be an entertaining and engaging way for students to demonstrate their knowledge and communication skills. Chatbot applications can offer students quick feedback and personalized answers to their inquiries (37).

Teacherbot is increasingly used for administrative aspects of teaching in online courses. Teacherbot deals mainly with content delivery, essential feedback, and supervision, thereby replacing traditional teaching assistants. In the online Masters in Computer Science Program at Georgia Institute of Technology, an AI teaching assistant for the “Knowledge-based AI“ course successfully met all of the students’ expectations. The teaching assistant operated through IBM’s Watson platform (38).

Boost is a mobile application that is seamlessly integrated with Canvas LMS. This app serves as a personal assistant for online learning (39). One of the ideal examples of an intelligent tutoring system is the Andes physics tutoring system. The mean post-test exam score of the students who did their homework on the Andes was approximately one standard deviation higher than that of

students who did the same homework with pencil and paper (40). Pump Algebra Tutor, SQL tutor, and Mathematic tutor are other notable examples.

ChatGPT AND RESEARCH ARTICLES

ChatGPT has reportedly generated articles that are similar to scientific papers written by academic researchers. Questions have been raised about the ethics of using ChatGPT in writing any dissertation component in educational institutions. Concerns about the accuracy of the texts on research-related topics generated by ChatGPT (41) have been raised. The researchers instructed ChatGPT-3.0 (the free version) to rewrite five articles published before 2021. The articles consisted of an original work, a case series, two case reports, and a technical report. To ensure accuracy, the references generated underwent cross-checking through PubMed, Google, and Ovid Medline. The articles were structured like typical journal article with an introduction, main body, conclusion, and references. The generated case report was well-written and included a detailed description of the differential diagnoses for the median nerve ganglion. However, four out of the five generated articles were inaccurate and unrelated to the topic (42).

Although ChatGPT's research reports may not always be completely accurate, educators consider the LLM in ChatGPT to be useful for summarizing and categorizing literature reviews and quickly synthesizing large amounts of information. Medical education literature has in the past been hindered by language barriers, putting non-native English speakers at a disadvantage. With ChatGPT, manuscripts from non-native English-speaking countries can be translated and corrected, giving them an equal chance to be published as scholarly work. AI tools offer the convenience of fast feedback on text drafts, assistance with language revision, structural support,

and the ability to carry out repetitive tasks. By substituting human reviewers, the processing time for publication can be reduced (43).

CONCLUSION

ChatGPT can answer factual questions, write articles, create poetry, generate programming code, and even translate between languages. Chatbot applications can give students feedback and facilitate personalized learning. ChatGPT can be trained to act as a tutoring system and help students complete their homework. Students can improve their preparation for licensing examinations by using ChatGPT's GPT-4 version. This tool generates practice questions through intelligent prompts. Making such efforts can help reduce the high subscription costs of test preparation and learning platforms such as Kaplan, AMBOSS, and Lecturio. It can also be used by the lecturers to develop the curriculum, construct lesson plans, and create questions. It can save lecturers' preparation time and allow them to dedicate more time to face-to-face or online teaching sessions. Relying solely on ChatGPT for answers to difficult queries may hinder the development of critical thinking and writing abilities among students studying for medical and health professions. Several universities have implemented measures to prohibit students from using ChatGPT. Some institutions are permitting their students to use ChatGPT while simultaneously adjusting their plagiarism policies to ensure that the use of this tool does not compromise academic integrity.

Teaching assistant tools powered by AI have the potential to transform the educational experience of students enrolled in programs with a high volume of students. As there are reports of both factual accuracy and AI hallucination, it is crucial to establish AI education guidelines and policies. This

will help protect both students and teachers from the dangers that can arise from unethical data use and piracy, and from the possible harassment of students due to their unpredictable and improper actions.

ACKNOWLEDGMENTS

The authors acknowledge the support from the website <https://chat.openai.com> to experience the use of generative AI.

REFERENCES

1. McCarthy, J. (1985). What is artificial intelligence anyway. *American Scientist*, 73(3), 258.
2. McCarthy, J. (1998). What is artificial intelligence? (Tech. Rep.). Stanford University.
3. Keates, S., Varker, P., & Spowart, F. (2011). Human-machine design considerations in advanced machine-learning systems. *IBM Journal of Research and Development*, 55(5), 4:1-4:10. <https://doi.org/10.1147/JRD.2011.2163274>.
4. Minsky, M. (1961). Steps toward artificial intelligence. *Proceedings of the IRE*, 49(1), 8-30. <https://doi.org/10.1109/JRPROC.1961.287775>.
5. Pomi, A., & Mizraji, E. (2001). A cognitive architecture that solves a problem stated by Minsky. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics: A Publication of the IEEE Systems, Man, and Cybernetics Society*, 31(5), 729-734. <https://doi.org/10.1109/3477.956034>.

6. van Assen, M., Lee, S. J., & De Cecco, C. N. (2020). Artificial intelligence from A to Z: From neural network to legal framework. *European Journal of Radiology*, 129, 109083. <https://doi.org/10.1016/j.ejrad.2020.109083>.
7. Gupta, S., & Khattar, M. (2020). Internet of things and artificial intelligence. *Towards smart world: Homes to cities using Internet of things*, 99.
8. Helm, J. M., Swiergosz, A. M., Haeberle, H. S., Karnuta, J. M., Schaffer, J. L., Krebs, V. E., *et al.* (2020). Machine learning and artificial intelligence: Definitions, applications, and future directions. *Current Reviews in Musculoskeletal Medicine*, 13(1), 69-76. <https://doi.org/10.1007/s12178-020-09600-8>.
9. McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115-133.
10. Abraham, T. H. (2002). (Physio)logical circuits: The intellectual origins of the McCulloch–Pitts neural networks. *Journal of the History of the Behavioral Sciences*, 38(1), 3-25. <https://doi.org/10.1002/jhbs.1094>.
11. Ho, S. Y. C., Chien, T. W., Lin, M. L., & Tsai, K. T. (2023). An app for predicting patient dementia classes using convolutional neural networks (CNN) and artificial neural networks (ANN): Comparison of prediction accuracy in Microsoft Excel. *Medicine*, 102(4), e32670. <https://doi.org/10.1097/md.00000000000032670>.
12. Cazzaniga, M., Salerno, F., Borroni, G., Ceriani, R., Stucchi, G., Guerzoni, P., *et al.* (2009). Prediction of asymptomatic cirrhosis in chronic hepatitis C patients: Accuracy of artificial neural networks compared with logistic regression models. *European Journal of Gastroenterology and Hepatology*, 21(6), 681-687. <https://doi.org/10.1097/meg.0b013e328317f4da>.

13. Oermann, E. K., Kress, M. A. S., Collins, B. T., Collins, S. P., Morris, D., Ahalt, S. C., & Ewend, M. G. (2013). Predicting survival in patients with brain metastases treated with radiosurgery using artificial neural networks. *Neurosurgery*, 72(6), 944-51; discussion 952. <https://doi.org/10.1227/neu.0b013e31828ea04b>.
14. DeBaun, M. R., Chavez, G., Fithian, A., Oladeji, K., Van Rysselberghe, N., Goodnough, L. H., *et al.* (2021). Artificial neural networks predict 30-day mortality after hip fracture: Insights from machine learning. *JAAOS-journal of the American Academy of Orthopaedic Surgeons*, 29(22), 977-983. <https://doi.org/10.5435/jaaos-d-20-00429>.
15. Melin, P., Miramontes, I., & Prado-Arechiga, G. (2017). [PP.15.16] A new model for obtaining the risk assessment of developing hypertension based on artificial neural networks. *Journal of Hypertension*, 35 Supplement 2, [15.16], e210. <https://doi.org/10.1097/01.hjh.0000523595.53292.2c>.
16. Zhai, X. (2022). ChatGPT user experience: Implications for education. *SSRN Electronic Journal*. <https://dx.doi.org/10.2139/ssrn.4312418>.
17. Lou, Y. (2023). Exploring the application of ChatGPT to English teaching in a Malaysia primary school. *Journal of Advanced Research in Education*, 2(4), 47-54. <http://dx.doi.org/10.56397/JARE.2023.07.08>.
18. Aydin, Ö., & Karaarslan, E. (2023). Is ChatGPT leading generative ai? what is beyond expectations? *Academic Platform Journal of Engineering and Smart Systems*, 11(3), 118-134. <http://dx.doi.org/10.2139/ssrn.4341500>.
19. Iqbal, N., Ahmed, H., & Azhar, K. A. (2022). Exploring teachers' attitudes towards using chatgpt. *Global Journal for Management and Administrative Sciences*, 3(4), 97-111. <https://doi.org/10.46568/gjmas.v3i4.163>.

20. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S. *et al.* (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144. <https://doi.org/10.1145/3422622>.
21. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved techniques for training GANs. *Advances in Neural Information Processing Systems*, 29. <https://doi.org/10.48550/arXiv.1606.03498>.
22. Xiong, L., Chen, Y., Peng, Y., & Ghadi, Y. Y. (2024). Improving robot-assisted virtual teaching using transformers, GANs, and computer vision. *Journal of Organizational and End User Computing*, 36(1), 1-32. <https://doi.org/10.4018/JOEUC.336481>.
23. Kumar, K. D., Srang, S., & Valy, D. (2022, December). A review of generative adversarial networks (GANs) for technology-assisted learning: Solving teaching and learning challenges. In *IEEE International Conference on Automation, Computing and Renewable Systems (ICACRS)*, 2022: pp. 820-826. <http://dx.doi.org/10.1109/ICACRS55517.2022.10029021>.
24. Roose, K. (2022). "The brilliance and weirdness of ChatGPT". *The New York Times*. Retrieved 11 February 2023 from <https://www.nytimes.com/2022/12/05/technology/chatgpt-ai-twitter.html>.
25. Marr, B. (2023). "A short history of ChatGPT: How we got to where we are today". *Forbes*. Retrieved 19 July 2023 from <https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/?sh=7c428e6b674f>.
26. Open, A. I. (2023). ChatGPT (July 20 Version) [Large Language Model] from <https://chat.openai.com>.

27. Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, 15(4), ep464. <https://doi.org/10.30935/cedtech/13605>.
28. Cox, R. L., Hunt, K. L., & Hill, R. R. (2023). Comparative analysis of NCLEX-RN questions: A duel between ChatGPT and human expertise. *The Journal of Nursing Education*, 62(12), 679-687. <https://doi.org/10.3928/01484834-20231006-07>.
29. Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching and assessing: A revision of bloom taxonomy of educational objectives*. New York: Longman.
30. Friederichs, H., Friederichs, W. J., & März, M. (2023). ChatGPT in medical school: How successful is AI in progress testing? *Medical Education Online*, 28(1), 2220920. <https://doi.org/10.1080/10872981.2023.2220920>.
31. Nori, H., King, N., McKinney, S. M., Carignan, D., & Horvitz, E. (2023). Capabilities of gpt-4 on medical challenge problems. arXiv preprint arXiv:2303.13375. <https://doi.org/10.48550/arXiv.2303.13375>.
32. Lee, H. (2023). The rise of ChatGPT: Exploring its potential in medical education. *Anatomical Sciences Education*. <https://doi.org/10.1002/ase.2270>.
33. Wang, L. K., Paidisetty, P. S., & Cano, A. M. (2023). The next paradigm shift? ChatGPT, artificial intelligence, and medical education. *Medical Teacher*, 45(8), 925. <https://doi.org/10.1080/0142159x.2023.2198663>.
34. Abd-Alrazaq, A., AlSaad, R., Alhuwail, D., Ahmed, A., Healy, P. M., Latifi, S., *et al.* (2023). Large language models in medical education: Opportunities, challenges, and future directions. *JMIR Medical Education*, 9(1), e48291. <https://doi.org/10.2196/2F48291>.

35. Eysenbach, G. (2023). The role of ChatGPT, generative language models, and artificial intelligence in medical education: A conversation with ChatGPT and a call for papers. *JMIR Medical Education*, 9, e46885. <https://doi.org/10.2196/46885>.
36. Azamfirei, R., Kudchadkar, S. R., & Fackler, J. (2023). Large language models and the perils of their hallucinations. *Critical Care*, 27(1), 120. <https://doi.org/10.1186/s13054-023-04393-x>.
37. Perez, S., Massey-Allard, J., Butler, D., Ives, J., Bonn, D., Yee, N., & Roll, I. (2017). “Identifying productive inquiry in virtual labs using sequence mining”. In E. André, R. Baker, X. Hu, M. M. T. Rodrigo, B. du Boulay (Eds.), *Artificial intelligence in education*, 10(331): pp. 287-298. http://dx.doi.org/10.1007/978-3-319-61425-0_24.
38. Maderer, J. (2016). Artificial intelligence course creates AI teaching assistant. Georgia Tech news center, 9.
39. Chew, E., & Chua, X. N. (2020). Robotic Chinese language tutor: Personalising progress assessment and feedback or taking over your job?. *On the Horizon*, 28(3), 113-124. <http://dx.doi.org/10.1108/OTH-04-2020-0015>.
40. VanLehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L. *et al.* (2005). The Andes physics tutoring system: Lessons learned. *International Journal of Artificial Intelligence in Education*, 15(3), 147-204.
41. Biswas, S. (2023). ChatGPT and the future of medical writing. *Radiology*, 307(2), e223312. <https://doi.org/10.1148/radiol.223312>.
42. Ariyaratne, S., Iyengar, K. P., Nischal, N., Chitti Babu, N., & Botchu, R. (2023). A comparison of ChatGPT-generated articles with human-written articles. *Skeletal Radiology*, 52(9), 1755-1758. <https://doi.org/10.1007/s00256-023-04340-5>.

43. Ellaway, R. H., & Tolsgaard, M. (2023). Artificial scholarship: LLMs in health professions education research. *Advances in Health Sciences Education*, 28(3), 659-664.
<https://doi.org/10.1007/s10459-023-10257-4>.