Yükseköğrenimde Yapay Zekânın Rolü: Anatomi Dersi için ChatGPT Değerlendirmesi

***

The Role of Artificial Intelligence in Higher Education: ChatGPT Assessment for Anatomy Course

Tarık TALAN[1]
Yusuf KALINKARA[2]

DOI:10.33461/uybisbbd.1244777

Makale Bilgileri

Makale Türü: Araştırma Makalesi

Geliş Tarihi: 30.01.2023

Kabul Tarihi: 06.03.2023

Öz

Uzun bir süre boyunca yürütülen kapsamlı araştırmalar, yapay zeka (AI) teknolojisinin hızla ilerlemesiyle sonuçlanmış ve çok sayıda uygulamanın geliştirilmesine yol açmıştır. Özelikle yakın zamanda piyasaya sürülen AI uygulamalardan biri de ChatGPT'dir. Kısa süre içerisinde milyonlarca kullanıcının beğenisini kazanan ChatGPT, metin okuyup yazabilen AI kodudur. Bu çalışmanın amacı, ChatGPT'nin bir dersteki performansını lisans öğrencilerinin performansıyla karşılaştırmaktır. Katılımcılar Türkiye'de bir devlet üniversitesinin Sağlık Bilimleri Fakültesi'nde daha önce anatomi dersi almış öğrencilerdir. Sınav çoktan seçmeli test şeklinde olup 40 sorudan oluşmaktadır. ChatGPT'nin öğrencilerden daha iyi performans gösterdiği tespit edilmiştir.

Anahtar Kelimeler: *Yapay zeka, ChatGPT, Sağlık eğitimi, Anatomi.*

Article Info

Paper Type: Research Paper

Received: 30.01.2023

Accepted: 06.03.2023

Abstract

The extensive research conducted over a prolonged period of time has resulted in the rapid advancement of artificial intelligence (AI) technology, leading to the development of numerous applications. One such recent AI application is ChatGPT, an AI chatbot that has gained millions of users in a short span of time, and can read and write texts. The aim of this study is to compare the performance of ChatGPT in an anatomy course with that of undergraduate students. The participants were students from the Faculty of Health Sciences at a state university in Turkey, who had previously taken an anatomy course. The examination was in the form of a multiple-choice test consisting of 40 items. It was found that ChatGPT outperformed the students in the examination.

Keywords: *Artificial intelligence, ChatGPT, Medical education, Anatomy.*

Atıf/ to Cite (APA): Talan, T., & Kalinkara, Y. (2023). The role of artificial intelligence in higher education: ChatGPT assessment for anatomy course. International Journal of Management Information Systems and Computer Science, 7(1), 33-40.

[1] Asst. Prof., Gaziantep Islam Science and Technology University, ttalan46@hotmail.com
[2] Lecturer, Gaziantep Islam Science and Technology University, yusufkalinkara@gmail.com

1. INTRODUCTION

As scientific knowledge continues to grow exponentially, new technological developments emerge every day. These technological developments and changes have the potential to facilitate, transform, and improve our lives, and can bring great benefits to the fields in which they are used. In fact, it is difficult to think of an area that is not affected by technology today. However, as technology continues to rapidly advance, questions are being raised about how it can be effectively used in various fields and what impact it will have. One area that has seen significant investment in recent years is education, with virtual reality, augmented reality, metaverse, blockchain, simulation, mobile technologies, robotics and automation, and online learning environments all being implemented. Among these technological advancements, artificial intelligence (AI) stands out as one of the most successful and widely-used technologies in many sectors.

The primary objective of AI is to enhance the comprehension of human intelligence and to augment machine intelligence to attain maximum benefit from them (Tektaş, Akbaş & Topuz, 2002). AI is a wide-ranging domain that is constantly advancing and leading the way in technological progress (Büyükgöze & Dereli, 2019). Although various research endeavors have been undertaken in AI across multiple fields, its impact on education has also been investigated. AI is applied across several disciplines, including law, science, mathematics, health, engineering, and architecture (Korkmaz & Büyükgöze, 2019; Taşçi & Çelebi, 2020). Research in this field is gaining momentum, and progress is being made through continued research and development activities.

Over the years, AI technologies have evolved and taken various forms. Despite the fact that AI research has a long history, these systems have now become an essential part of human life thanks to the investments made over the years and the widespread use of technological tools such as the internet and smart devices. Currently, there are numerous AI technologies, most of which are still in the research stage. In recent years, AI technologies have been applied in diverse fields, including smart cities, smart watches, robotics, drone systems, defense industry, cybersecurity, and healthcare (Sarica, 2021; Talan, 2021).

However, the potential use of AI in education, its contribution to education, and its impact on education are still subjects of debate, with numerous predictions and considerations. While developments in AI offer significant opportunities for the education sector, they also pose a threat at times. Thus, AI technology needs to be carefully considered and evaluated in many ways. AI's potential to be one of the most important technologies of the future increases when the potential risks and benefits it offers are evaluated, and the necessary precautions are taken.

1.1 Artificial Intelligence (AI)

AI, first introduced by John McCarthy at the Dortmund Conference in 1956, has emerged as a significant technological advancement in recent times (Arslan, 2017). AI is an interdisciplinary field that has developed through the integration of various disciplines, including computer science, control theory, information theory, neurophysiology, psychology, linguistics, and philosophy (Wei, 2018). The primary objective of AI is to simulate and enhance human intelligence, and remarkable progress has been made in this domain over the years (Shi & Zheng, 2006).

AI can be defined as a computer-based simulation of human mental processes, such as reasoning, argumentation, learning, sense-making, communication, decision-making, and generalization (Akyürek, 2013). In essence, AI is an analytical system that aims to imitate life (Gordon, 2011). While several definitions of AI exist in the literature, they focus on different concepts, including human-like thinking, intelligent programming, rational action, and humanoid responses (Arslan, 2017; Büyükgöze & Dereli, 2019; Russel & Norvig, 2010).

AI is transforming human life in various ways, ranging from internet search engines and smartphone apps to public transportation, autonomous cars, and household appliances. AI has become an integral part of people's lives, and its applications are increasingly being used as control and

decision-making mechanisms in various fields, including health, engineering, architecture, military, psychology, energy, mining, agriculture, meteorology, and forensics (Sarıca, 2021). Furthermore, AI investment in numerous sectors is rapidly increasing.

In higher education, if appropriately and effectively utilized, AI tools can provide significant benefits (Taşçi & Çelebi, 2020). However, there is still a debate on how to effectively integrate AI applications into education. It is anticipated that with the introduction of AI in education, investment and research in this area will grow significantly in the coming years (İşler & Kılıç, 2021).

1.2 ChatGPT

Technology is advancing rapidly, bringing new concepts and innovations to various aspects of our lives. Among these concepts, AI technology has become increasingly prevalent in recent years, thanks to extensive research and development efforts. One of the most talked-about AI applications is the AI chatbot, which employs deep-learning algorithms trained on vast amounts of data to generate human-like responses to user queries (Gilson et al., 2022).

In November 2022, the public release of ChatGPT (Chat Generative Pre-trained Transformer) provided a remarkable example of human-computer communication, thanks to its advanced technology (Cotton, Cotton & Shipway, 2023; de Winter, 2023; Topsakal & Topsakal, 2022; Wenzlaff & Spaeth, 2022; Zhai, 2022). This particular AI technology, which has been at the forefront of the technological agenda, is a natural language processing model with 175 billion parameters, developed by OpenAI (Gilson et al., 2022). ChatGPT is one of the most potent NLP systems, boasting an enormous number of parameters, making it one of the largest language models currently available (Cotton, Cotton & Shipway, 2023).

Founded in 2015 by prominent technology leaders, OpenAI is a research institute that focuses on the development of AI technologies. The institute is recognized for its research endeavors in diverse fields and possesses the capability to provide conversational responses, reject inappropriate queries, challenge erroneous assumptions, admit mistakes and learn from its own errors with the aid of ChatGPT, an optimized language model (Jiao et al., 2023). This technology, which incorporates natural language processing and deep learning, utilizes a vast dataset and generates text that resembles human language (Qadir, 2022). Although initially criticized for its factual accuracy, the technology has gained popularity due to its ability to provide detailed and comprehensible answers to inquiries. Employing the GPT-3 text interpreter, this AI code is a form of NLP and can read and produce written texts (Pavlik, 2023).

According to Güçlütürk (2022), the output produced by ChatGPT is influenced by the preceding content, and even if the same command is entered, the output may differ. Consequently, the content produced in ChatGPT is highly personalized and original, and it relies on the input provided by the users and the given content (Wenzlaff & Spaeth, 2022). ChatGPT is designed to communicate and interact with people in a manner similar to human-to-human interaction, and it can respond in various languages (O'Connor & ChatGPT, 2023; Wenzlaff & Spaeth, 2022). While ChatGPT was initially intended for online customer service, it is now utilized for diverse applications and purposes, such as healthcare, software development, content creation, language translation, increasing business efficiency, cost reduction, and customer service through the utilization of AI (Gilson et al., 2022; Qadir, 2022). It is plausible to assume that AI and chatbots will continue to evolve with the advancement of technology.

It can be asserted that ChatGPT technology, designed to converse with users and provide meaningful responses to their inquiries (de Winter, 2023), will continue to evolve and become more impressive in the future (Qadir, 2022). The primary attribute of ChatGPT, which has garnered the admiration of millions of users in a brief timeframe, is its ability to supply accurate responses to user queries in real-time (Qadir, 2022). Another impressive capability is its capacity to generate high-quality, error-free text that is difficult to differentiate from human composition (Susnjak, 2022).

Similar to other productive AI systems, ChatGPT delivers answers to inquiries rapidly while preserving semantic coherence. Nonetheless, it is important to note that the answers are not always correct or appropriate, and they frequently contain erroneous and biased references (de Winter, 2023; O'Connor & ChatGPT, 2023; Qadir, 2022). Moreover, ChatGPT has limited information on events that occurred post-2021. However, as bots such as ChatGPT, believed to increase in competency over time, continue to refine themselves, such problems are expected to diminish.

1.3 Purpose of the Study

According to the literature, ChatGPT has been shown to offer extensive knowledge on almost any subject, and to provide reliable and accurate responses to challenging inquiries that necessitate advanced information analysis, synthesis, and application (Susnjak, 2022). Furthermore, Qadir (2022) suggests that ChatGPT has the potential to offer effective learning opportunities by creating realistic virtual simulations for applied learning. Additionally, AI chatbots such as ChatGPT can offer prompt feedback and corrections to questions, thereby assisting students in comprehending intricate concepts that are difficult to learn (Cotton, Cotton & Shipway, 2023). By using AI chatbots, students can benefit from personalized, interactive learning experiences, and better manage their time by planning their assignments and homework (Cotton, Cotton & Shipway, 2023; Huh, 2023; O'Connor & ChatGPT, 2023).

Despite the potential benefits of ChatGPT as an educational tool, the full extent of its impact on education remains uncertain and requires further investigation (de Winter, 2023; Qadir, 2022; Zhai, 2022). It is crucial to consider both the potential advantages and risks associated with emerging technologies like ChatGPT in order to anticipate and prepare for the future of education. One significant concern is the possibility of students using ChatGPT to cheat, particularly on online exams, due to its ability to generate personalized and authentic responses (de Winter, 2023; Stokel-Walker, 2022; Susnjak, 2022). As online education becomes increasingly widespread, ensuring the validity and reliability of online exams is a critical issue that must be addressed. It is important to acknowledge that further research is required to develop effective strategies that mitigate potential risks and leverage the benefits of AI. Furthermore, there is a scarcity of literature that delves into the potential educational use of ChatGPT, a novel tool in this domain. Hence, investigating the capabilities of this AI agent is anticipated to augment the current body of knowledge. This study aims to evaluate the performance of the newly launched ChatGPT on anatomy course examination questions among students enrolled in the Faculty of Health Sciences in Turkey.

2. METHOD

This study aimed to compare the performance of ChatGPT with that of health sciences faculty students in answering anatomy course questions. A descriptive study design was employed, and 37 students from a state university in Turkey participated in the study. The students received four weeks of training on a specific anatomy topic and then took a multiple-choice test comprising 40 questions on the covered material. The same test was also given to ChatGPT, and the answers generated were compared with those of the students. The data were analyzed using descriptive statistics, including the number, percentage, and mean.

3. RESULTS

In the study, a multiple-choice test was applied to students taking the anatomy course with ChatGPT. Below are some sample questions and screenshots of the answers from ChatGPT.

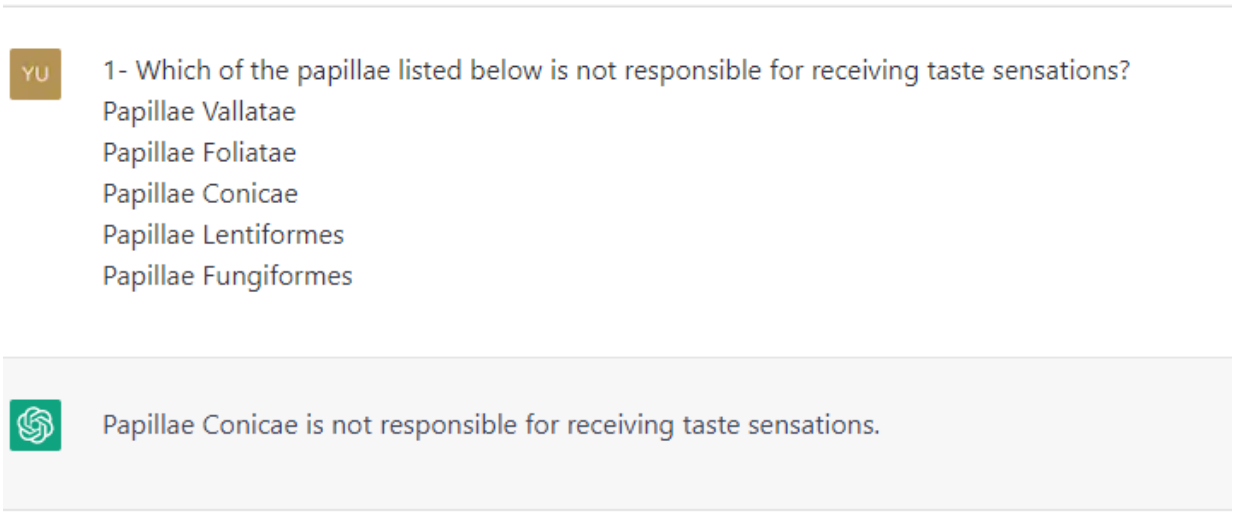

Figure 1. Screenshot of ChatGPT's correct answer to a sample question in the anatomy course

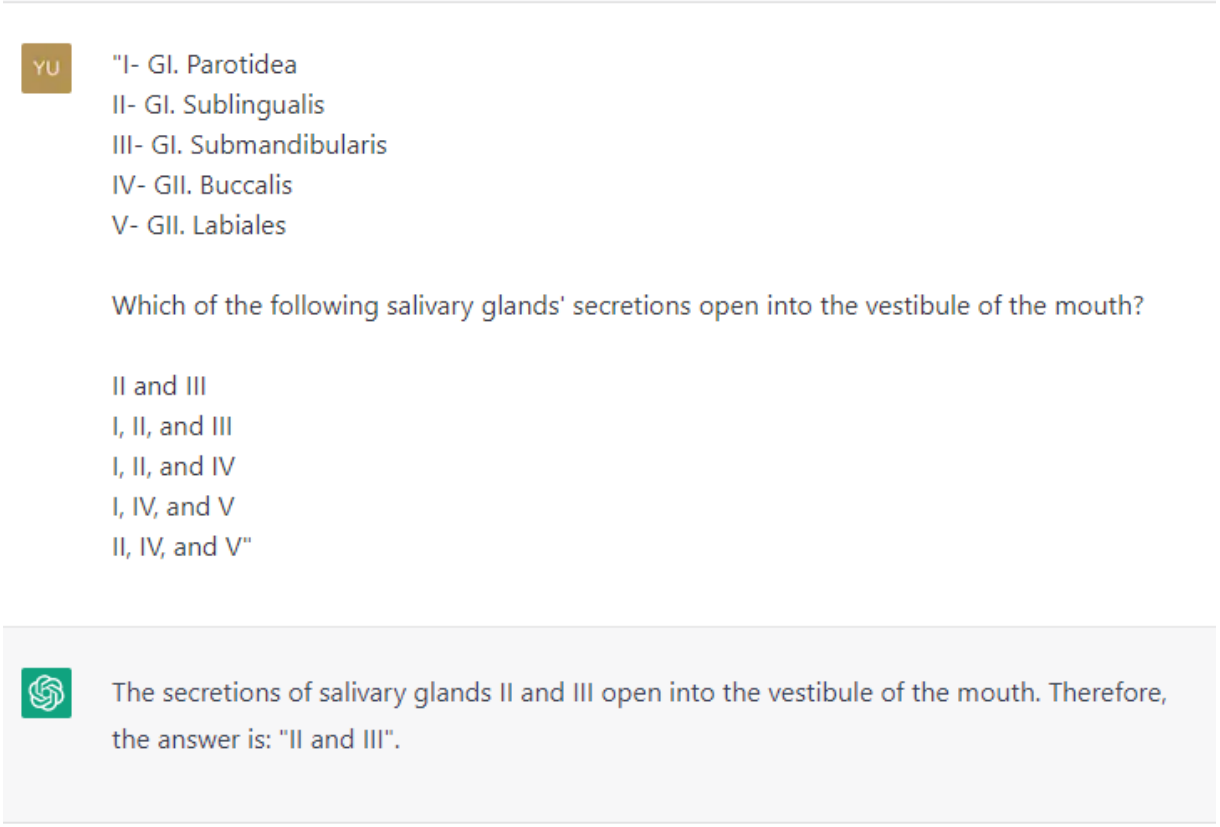

Figure 2. Screenshot of ChatGPT's wrong answer to a sample question in the anatomy course

The statistical values of the scores obtained from the tests applied in the study are shown in Table 1.

Table 1. Statistical values of the answers of the students and ChatGPT regarding the exam in the anatomy course

	Correct Responses	Incorrect Responses	Grade (0-100)
ChatGPT	27	13	67.5
Studens	21.1	18.9	52.8

The results presented in Table 1 indicate that ChatGPT correctly answered 27 out of 40 questions, corresponding to a success rate of 67.5%. By comparison, the students' average score was

21.1 out of 40, equivalent to a success rate of 52.8%. These findings suggest that ChatGPT's performance was better than the average performance of the students enrolled in the anatomy course at the Faculty of Health Sciences. Based on these results, it can be concluded that ChatGPT demonstrated a successful performance on this exam. Figure 3 shows an example of how ChatGPT compared to all students on the exam.

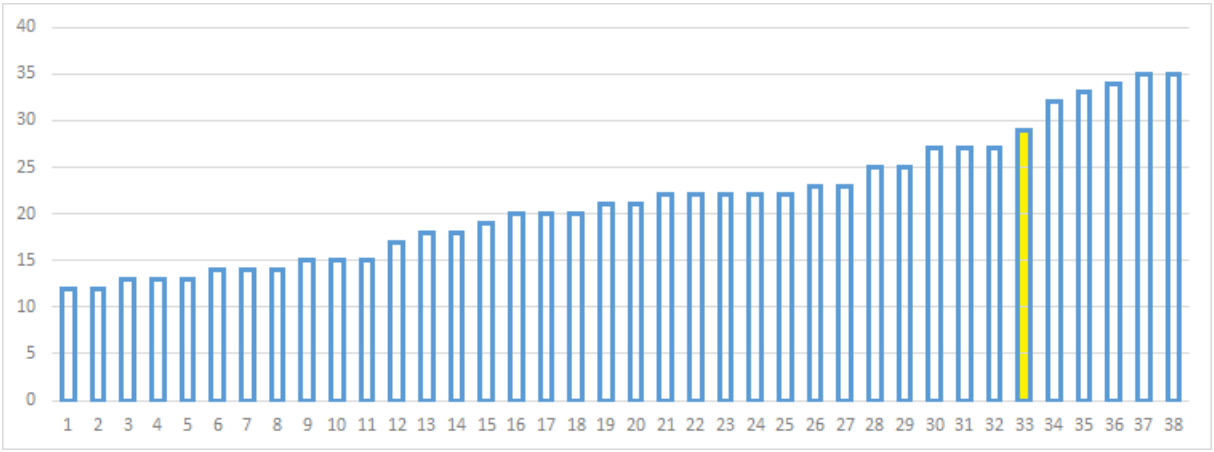

Figure 3. Distribution of anatomy test results. ChatGPT performance is highlighted in yellow

Current results show that ChatGPT performs similarly to the average student in understanding anatomy.

4. DISCUSSION

The objective of this study was to compare the ability of an AI agent, ChatGPT, to answer questions on anatomy course exams with the performance of students from the Faculty of Health Sciences in Turkey. Results showed that ChatGPT had a significantly higher ratio of correct answers compared to students. Furthermore, ChatGPT's knowledge in answering the anatomy exam was found to be superior to that of the students. Other studies also support the potential of ChatGPT as a new educational tool. For instance, Gilson et al. (2022) found that ChatGPT performed comparably to medical students on medical licensing exams, while de Winter (2022) reported that ChatGPT's performance on English comprehension was similar to that of university students. Conversely, Huh (2023) found that ChatGPT's overall performance on a parasitology exam was lower than that of medical students in Korea. Similarly, Geerling et al. (2023) found that ChatGPT achieved high performance in a university exam measuring economics knowledge.

In the future, ChatGPT is expected to improve its performance through deep learning, which is a promising development that educators and researchers should be aware of in terms of its potential applications in teaching and learning. However, there have been varying results in other studies related to ChatGPT. In a study conducted by Nisar and Aslam (2023), they requested ChatGPT to answer questions in the field of pharmacology. Although the answers provided by ChatGPT were generally accurate, the source or reference of the answers was not provided. Moreover, the study found that ChatGPT could potentially serve as a self-study tool for students struggling with pharmacology. Nonetheless, it is worth noting that the ability of ChatGPT to generate highly realistic texts poses a potential risk to the integrity of online exams, and precautions should be taken to prevent this from happening (de Winter, 2022; Susnjak, 2022).

In the current state, ChatGPT is capable of producing accurate responses within seconds. However, it has limitations in interpreting visual aids such as diagrams, shapes, and tables, which can

be easily comprehended by human students. Thus, these visual aids need to be explained in text form for ChatGPT to understand. Additionally, if a question is ambiguous or incomprehensible, ChatGPT may produce an incorrect response. To mitigate this issue, it is advisable to rephrase the question in a clear and precise manner.

The recent publication of ChatGPT implies that there are limited studies in the literature that compare its performance with that of students. The current study utilized a multiple-choice test to assess performance. Future studies may explore the performance of ChatGPT and other AI language models in various exam formats. Furthermore, this study focused on a university-level anatomy course. To make meaningful comparisons, other researchers can conduct similar studies at different educational levels and in different courses, enabling a more in-depth investigation of their efficacy.

REFERENCES

Akyürek, H. A. (2013). *Intelligent workforce management by using artificial intelligence techniques*. Master Thesis, Türkiye.

Arslan, K. (2017). Eğitimde yapay zeka ve uygulamaları. *Batı Anadolu Eğitim Bilimleri Dergisi, 11*(1), 71-88.

Büyükgöze, S., & Dereli, E. (2019). Güncel sağlık bilimleri çalışmaları II, Dijital Sağlık ve Yapay Zeka, Akademisyen Kitabevi A.Ş., Editör: Tuncay ÖZGÜNEN, ISBN: 978-605-258-626-6i.

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating. Ensuring academic integrity in the era of ChatGPT. Preprint. https://doi.org/10.35542/osf.io/mrz8h

de Winter, J. C. F. (2023). Can ChatGPT pass high school exams on English language comprehension?. Researchgate. Preprint.

Geerling, W., Mateer, G. D., Wooten, J., & Damodaran, N. (2023). Is ChatGPT smarter than a student in principles of economics?. Available at SSRN 4356034.

Gilson, A., Safranek, C., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2022). How does ChatGPT perform on the medical licensing exams? The implications of large language models for medical education and knowledge assessment. medRxiv. Preprint. https://doi.org/10.1101/2022.12.23.22283901

Güçlütürk, O. G. (2022). ChatGPT ile üretilen içeriklerin eser niteliğinin 5846 sayılı fikir ve sanat eserleri kanunu bakımından değerlendirilmesi. *Galatasaray Üniversitesi Hukuk Fakültesi Dergisi, 2*, 1899-1918.

Huh, S. (2023). Are ChatGPT's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: A descriptive study. *Journal of Educational Evaluation for Health Professions, 20*(1), 1-13. https://doi.org/10.3352/jeehp.2023.20.1

İşler, B., & Kılıç, M. (2021). Eğitimde yapay zekâ kullanımı ve gelişimi. *e-Journal of New Media, 5*(1), 1-11.

Jiao, W., Wang, W., Huang, J., Wang X, & Zhaopeng, T. (2023). Is chatGPT a good translator? A preliminary study. arXiv. Preprint.

Korkmaz, A., & Büyükgöze, S. (2019). Sahte web sitelerinin sınıflandırma algoritmaları ile tespit edilmesi. *Avrupa Bilim ve Teknoloji Dergisi, 16*, 826-833. https://doi.org/10.31590/ejosat.598036

Nisar, S., & Aslam, M. S. (2023). Is ChatGPT a good tool for T&CM students in studying pharmacology?. Available at SSRN 4324310.

O'Connor, S., & ChatGPT. (2023). Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse? *Nurse Education in Practice, 66,* 103537. https://doi.org/10.1016/j.nepr.2022.103537

Pavlik, J. V. (2023). Collaborating with chatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator*, 1-10. https://doi.org/10.1177/10776958221149577

Qadir, J. (2022). Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. *TechRxiv*. Preprint. DOI: 10.36227/techrxiv.21789434.v1

Sarica, R. (2021). *Eğitimde yapay zeka: Kavramsal temeller*. Köse, U. (Ed.) Eğitimde zeki ve esnek teknolojiler (pp.121-139), Ankara: Pegem Akademi.

Shi, Z.-Z., & Zheng, N.-N. (2006). Progress and challenge of artificial intelligence. *Journal of Computer Science and Technology, 21*(5), 810-822. https://doi.org/10.1007/s11390-006-0810-5

Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays — should professors worry? https://doi.org/10.1038/d41586-022-04397-7

Susnjak, T. (2022). Chatgpt: The end of online exam integrity? arXiv. Preprint. https://doi.org/10.48550/arXiv.2212.09292

Talan, T. (2021). Artificial intelligence in education: A bibliometric study. *International Journal of Research in Education and Science (IJRES), 7*(3), 822-837. https://doi.org/10.46328/ijres.2409

Taşçi, G., & Çelebi, M. (2020). Eğitimde yeni bir paradigma: "Yükseköğretimde yapay zekâ". *OPUS International Journal of Society Researches, 16*(29), 2346-2370. https://doi.org/10.26466/opus.747634

Tektaş, M., Akbaş, A., & Topuz, V. (2002). Yapay zekâ tekniklerinin trafik kontrolünde kullanılması üzerinde bir inceleme. *Uluslararası Trafik ve Yol Güvenliği Kongresi ve Fuarı*.

Topsakal, O., & Topsakal, E. (2022). Framework for a foreign language teaching software for children utilizing AR, voicebots and ChatGPT (large language models). *The Journal of Cognitive Systems, 7*(2), 33-38. https://doi.org/10.52876/jcs.1227392

Wei, X. (2018). The application and development of artificial intelligence in smart clothing. *IOP Conference Series: Materials Science and Engineering, 320*, 012017. https://doi.org/10.1088/1757-899X/320/1/012017

Wenzlaff, K., & Spaeth, S. (2022). *Smarter than humans? Validating how OpenAI's chatGPT model explains crowdfunding, alternative finance and community finance.* University of Hamburg – Digital Markets – Working Paper Series. Available on SSRN. Electronic copy available at: https://ssrn.com/abstract=4302443