AI Chatbot Generated Content and Similarity Index: Challenge for Plagiarism Detection Software

Vinayak Wadhwa
PhD Student, Dept. of Lib. & Info. Sc,
Kurukshetra University, Kurukshetra
Email: vinayakexams@kuk.ac.in

Manoj Kumar Joshi
Professor and Chairman, Dept. of Lib. & Info. Sc.,
Kurukshetra University, Kurukshetra
Email: manojkj01@yahoo.com

Abstract

The impact of artificial intelligence (AI) technology on education has grown significantly in recent years. Advanced AI applications, such as chatbots, are increasingly available over the Internet and have improved their capabilities. However, using chatbots, especially ChatGPT, in academic institutions to write academic articles has raised concerns among researchers. This study was conducted using Turnitin plagiarism-checking software to evaluate the originality of 20 essays written by ChatGPT on various topics. The study found that ChatGPT has the potential to generate unique content that cannot be detected by plagiarism-checking software. These findings underscore current concerns about students' use of chatbots to achieve academic success with minimal effort. The paper also discusses the challenges towards AI applications and the actions taken by plagiarism software to address this issue.

***Keywords*:** Artificial Intelligence, OpenAI, ChatGPT, Plagiarism

Introduction

"Chatbot is a computer program in the form of a virtual e-mail correspondent that can reply to messages from computer users" (Lexico Dictionaries, 2023). An "AI chatbot" is a computer program that utilises "artificial intelligence (AI)" to act human-like chats with users via a messaging interface. AI Chatbots are designed to perform various tasks, from customer support to entertainment and education (Adamopoulou &Moussiades, 2020).

AI chatbots utilise "Natural Language Processing (NLP)" and "machine learning algorithms" to understand and interpret user queries and respond appropriately. AI chatbots can be programmed to handle simple requests or engage in more complex conversations that involve multiple turns and context awareness (Awad& Khanna, 2015).

Some AI chatbots are rule-based, meaning that they are programmed with a set of predefined rules that dictate how they should respond to user queries. Others use machine learning algorithms to analyze and learn from previous interactions with users, allowing them to improve their responses over time.

AI chatbots have become increasingly popular in recent years, as businesses and organizations seek to provide more personalized and efficient customer service. They can also be used to automate tasks and free up human agents to focus on more complex issues.

The use of artificial intelligence in academia is a hot topic in the education field. The use of Artificial Intelligence in higher education has the potential to offer a range of benefits, including increased student engagement, collaboration, and accessibility. Many AI chatbots like ChatGPT, GPT-3, BERT, RoBERTa, and XLNet have recently been developed, and allare state-of-the-art language models developed by OpenAI (GPT), Google (BERT) and Microsoft (XLNet). One of the vital usefulness of GPT-3 and ChatGPT is their ability to generate high-quality text, whereas BERT, RoBERTa and XLNet excel at understanding and analyzing text. GPT-3 and ChatGPT both are based on the GPT-3 architecture and can generate human-like text, making them useful for various NLP tasks such as language translation, summarization, and question-answering. On the contrary, BERT, RoBERTa and XLNet mainly focus on understanding the text's meaning. They are beneficial for tasks such as feeling analysis and named entity recognition. Due to this, ChatGPT is quickly becoming a popular choice among students against all chatbots to generate academic essays for homework, which has elevated the worries of plagiarism in academia.

In this study, twenty different open-ended questions were prepared and asked to ChatGPT. Then the short essays generated by ChatGPT were checked for plagiarism by a popular plagiarism-detection tool (Turnitin) and tried to find out ChatGPT in academic settings from the perspective of academic honesty and plagiarism. Turnitin is basically a similarity checking software that

compares the text with a vast array of digital documents of scholarly literature and generates similarity score.

What is ChatGPT?

ChatGPT based on the GPT-3 architecture, is an enhanced version of the GPT-3 (Generative Pre-trained Transformer 3)(Brown et al. 2020) artificial language model. OpenAI first introduced the GPT-3 model in May 2020, and since then, it has become one of the most advanced and powerful language models available. GPT-3 is a generative language model that can produce coherent and sophisticated human-like text in response to prompts or questions.

OpenAI launched ChatGPT-3.5 (chat.openai.com) on 30th December 2022 as an advanced version of the GPT-3 model, ChatGPT-4 is most recent version of the tool Launched in March, 2023. ChatGPT is the fastest-growing user base application of all time. Sam Altman, OpenAI's Chief, said on Twitter that ChatGPT had more than 1 million users in the first five days after it was launched and estimates 100 million active users in January, only two months after its launch (Hu, 2023).

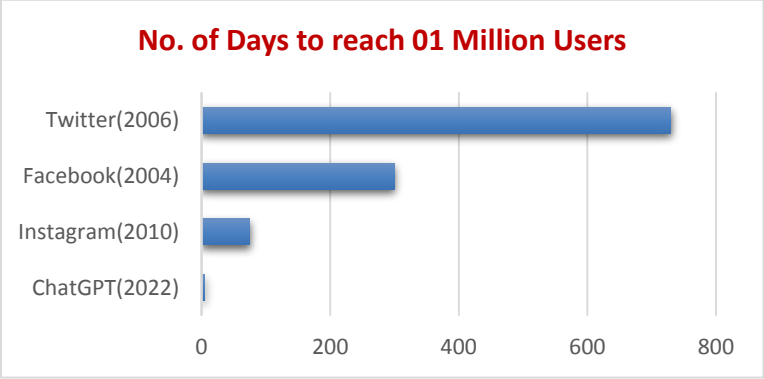

Fig. 1: x Number of days to reach One Million Users

ChatGPT became very popular shortly after public release, the search interest making a new peak every month if we take the last three-month trend (in Figure 2); the term "ChatGPT" was used globally reflecting the global popularity and demand for this application as per Google Trends that collects search data from 71 countries.

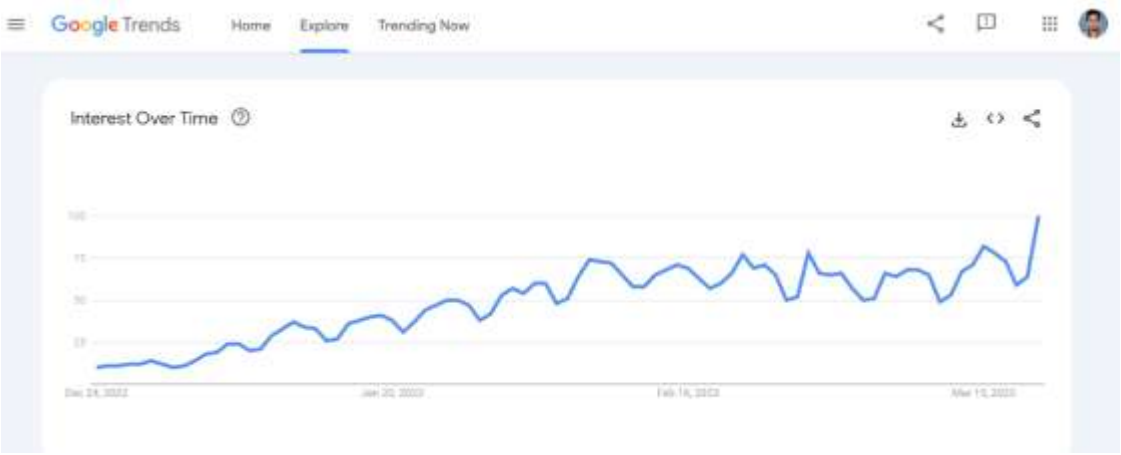

Fig. 2: Google Trend of term "ChatGPT"

Literature review

Artificial Intelligence (AI) chatbots have become increasingly popular in recent years as they provide an effective means of communicating with users. This literature review aims to examine the latest research on AI chatbots and their challenges in higher education.

MacNeil et al. (2022) explain how Language Model in Artificial Intelligence model trains to generate text like human language. King (2022) discusses the future of AI in Medicine applications, and states that AI is an effective tool that can transform the field of medicine and enhance the quality of upkeep for patients. Ding & Chan(2023) explain that iterative human-AI interactions are complex; and present LLM-powered tools show a possibility for supporting that vision of human-AI collaboration in creative tasks. However, there is still space for improvement in certain areas. Radford et al. (2018) introduced a framework for achieving high natural language understanding with a single task-agnostic model through generative pre-training and discriminative fine-tuning, and the model acquires significant world knowledge and ability to process long-range dependencies, which are then successfully transferred to solving discriminative tasks. Willems (2023) describes ChatGPT as a platform based on artificial intelligence (AI), and one can ask it anything. It responds to any question asked, and the quality of those answers ranges from surprisingly accurate to embarrassingly annoying. Interacting with ChatGPT feels very similar to interacting with some humans. Adamopoulou &Moussiades (2020) organised critical information that is a necessary background for further research activity

in the field of chatbots and presented the historical evolution of machine learning and intelligent chatbots applications in the industrial use. Lund &Wang (2023) explain the technology behind the chatbots, the process and development of Generative Pretrained Transformer (GPT), the underlying technology of ChatGPT and how natural language processing (NLP) and artificial intelligence play a crucial role in the technology. Brown et al. (2020) discuss different models of artificial intelligence and GPT-3 model performance on natural language processing datasets and their social impacts on GPT-3. Shidiq (2023) found in his study that ChatGPT can provide all answers according to keywords entered by the users, which can influence the world of education and learning. Hwang &Chang (2022) explore the trends of chatbots in education studies and found that research on chatbots is in the early stage and growing steadily. Kashyap (2023) highlights the principal requirement of all intelligent systems (AI) and describes that ChatGPT is a significant milestone in AI evolution. Shiri (2023) explains the key concerns of ChatGPT, like academic integrity issues for academic institutions and rethinking assessment strategies and approaches. Cotton et al. (2022) report that the use of chat APIs and GPT-3 in higher education offers excellent benefits and raises concerns about academic honesty and plagiarism with AI.

Methodology

The findings of this study demonstrate the outcomes of a similarity analysis carried out on AI-created content. The study adopted a quantitative approach, evaluating a representative sample of the content generated by ChatGPT (chat.openai.com).

The authors selected twenty different topics and directed ChatGPT (chat.openai.com) to generate a "500-word essay on topic x". and then examined the originality scores on the "Turnitin" tool generated content.

Limitations of the study

The research can have considerable limitations:

i. The study is specified to examine a ChatGPT chatbot technology only. The results may differ from the capabilities of all chatbot technologies.

ii. The study results depend on the accuracy of only Turnitin's plagiarism detection software.

iii. The sample size of the study, comprising 20 essays generated using ChatGPT, serves as adequate for our initial investigation and analysis within the specified limits. Although the quantity was modest, the themes covered were diversified, encompassing a wide spectrum of results. Increasing the sample size would have yielded more reliable and comprehensive findings.

Results& Discussion

The authors randomly selected 20 concepts as essay topics and asked ChatGPT to "write 500 words essay on 'x' ". Essay with a total number of words generated by ChatGPT, was checked for text similarity using the software. The similarity score that varied between 0% and 27% of all essays, has been shown in Table 1.

Table1

List of essay concepts, words generated by ChatGPT & similarity score on Turnitin.

Sr no.	Essay topic	No of words generated by ChatGPT	Similarity score
1	Chatbots	448	0%
2	Cloud computing	406	4%
3	Digital library	499	6%
4	ISO standards	479	15%
5	Library Automation	342	2%
6	Library Management	481	0%
7	Library Network System in India	509	6%
8	LIS Education in India	378	3%
9	Machine learning	484	27%
10	MOOCS	383	4%
11	Open Data	460	6%
12	Research Data Management	492	0%

13	Information Sources	476	6%
14	Information literacy	448	12%
15	Digital learning	469	8%
16	Social networking	403	0%
17	Learning theories	475	14%
18	Open journals	491	6%
19	Crowd funding	399	2%
20	Use of ICT in Education	458	6%

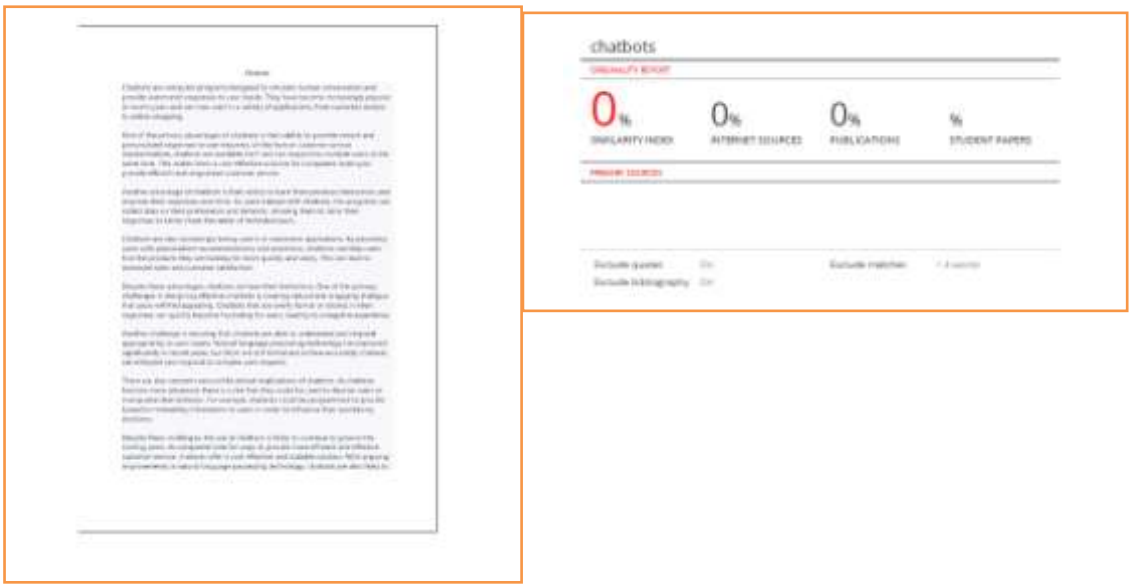

Fig. 3: Essay generated by ChatGPT on ‘chatbots’ and similarity score of 0% by Turnitin.

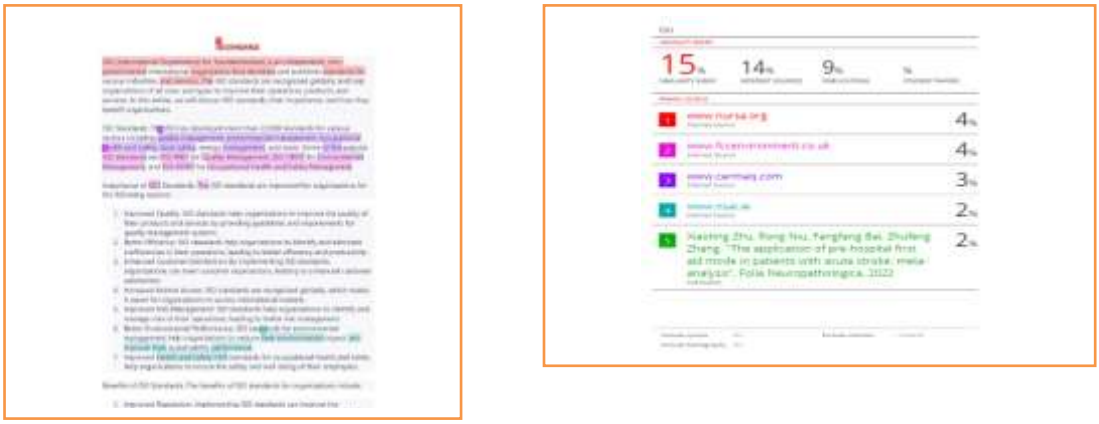

Fig. 4: Essay generated by ChatGPT on ‘ISO Standards’ and similarity score of 15% by Turnitin.

Two examples out of 20 essays are presented in Fig. 3 and Fig. 4, showing a similarity score of 0% and 15%, respectively. It is pertinent to specify that a significant proportion of similarity score may be due to common phrasal use and actual plagiarism incidence may not be significant.

With the help of plagiarism-detection software, the similarity scores of 20 essays generated by ChatGPT were calculated. Furthermore, the authors created a frequency table and divided the similarity scores into the similarity classes: less than 10%, , 10 to 20%, 20 to 40%, and 40 to 100%.

Table 2 displays the results of a plagiarism check conducted using the Turnitin tool. The authors discovered that 16 out of 20 essays, equivalent to 80%, had a high level of originality with a similarity score of less than 10%. Among the essays, three (n=3, 15%) received an acceptable similarity score ranging from 10 to 20%. Only one essay was found to have a significantly high similarity score falling within the range of 20 to 40% with other content, and none received a similarity score exceeding 40%. On average, the similarity score of all the essays was calculated to be just 6%, indicating that the text generated by the ChatGPT chatbot contained original content.

Table 2

Plagiarism check results (n=20)

Essay topic	**Essay Count**	**Essay Count %**	**Similarity score**
Chatbots, cloud computing, digital library, Library Automation, Library Management, Library Network System in India, LIS Education in India, MOOCS, Open Data, Research Data Management, Information Sources, digital learning, social networking, open journals, crowd funding, use of ICT	16	**80 %**	**<10%**

in education			
,Information literacy, ISO standards, Learning Theories	3	**15%**	**10-20%**
Machine Learning	1	**5%**	**20-40%**
None	0	**-**	**>40%**
Total & Average	**20**		**Average 6%**

Current Challenges of AI chatbots

One of the current challenges of chatbots regarding cheating by students is that chatbots can be programmed to provide incorrect or misleading answers. AI chatbots may encounter difficulties understanding the full context of a communication, particularly in complex and detailed discussions, which can lead students to submit inaccurate or plagiarized work. Additionally, students may try to exploit chatbots by using them to generate responses that they can use in their work, but they need to understand the material adequately.

Overall, chatbots are still a relatively new technology in education, and there are ongoing efforts to improve their accuracy, reliability, and security. As technology continues to evolve, new solutions will likely emerge to address the challenge of cheating by students.

Academic integrity in the age of AI writing

As AI writing technology advances, academic integrity is becoming an increasingly important issue. While AI writing can be a powerful tool for researchers and students, it can also potentially increase academic dishonesty.

For maintaining academic integrity in the age of AI writing, it is essential to establish clear guidelines and policies around using AI writing technology. It requires students to disclose their use of AI writing tools and ensure they understand their ethical implications.

Educational institutions should also invest in tools and strategies to help detect plagiarism and other forms of academic dishonesty. Include software that analyses written work for patterns and

language use and training for faculty and staff on identifying and addressing plagiarism instances.

Ultimately, the key to maintaining academic integrity in the age of AI writing is to promote a culture of honesty and ethical behaviour. We are encouraging students and researchers to prioritize their original work and to understand the importance of properly citing sources and acknowledging the contributions of others.

Plagiarism Detection tools for AI- generated Text.

This study shows that **the plagiarism detection tool has limitations to detect ChatGPT-generated text.** Many plagiarism detection tools including Turnitin have already raised concerns about AI-generated text and Turnitin tool is working on updating its plagiarism engine to detect cheating using AI-chatbots.

Turnitin also launched an AI writing resource page to support educators with teaching resources and to report its progress in developing AI writing detection features. The newly launched AI writing resource page is publicly available and will be updated regularly with information about Turnitin's progress in bringing detection features to the market (Turnitin, 2023).

Conclusion

The study analyses the use of AI applications in Higher Education, the popular chatbot ChatGPT and its capability to generate high-quality text. The study created 20 essays with the help of ChatGPT on different topics and analysed the percentage similarity using Turnitin's plagiarism-checking software. The study reveals that the average similarity score among all essays was 6%, while 16 of the 20 essays generated by ChatGPT exhibited a high level of originality and less than 10% plagiarism. The study's findings call into question the reliability of plagiarism detection software used by academic institutions considering recent advancements in chatbot technology. Moreover, the research shows that ChatGPT has tremendous potential to produce refined text outputs evading plagiarism detection software. To maintain academic integrity and address cheating in essay writing using ChatGPT, it is essential to establish transparent clear guidelines and policies regarding the use of AI writing technology. It requires students to disclose their use of AI writing tools and ensure they understand their ethical implications. The

study also highlights concerns about plagiarism detection software and its ability to detect cheating using AI chatbots, prompting the need for software updates.

References

Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications.*

Awad, M., & Khanna, R. (2015). *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers.* ApressOpen.

Brown, T. B., Mann, B., Ryder, N., & Subbaih, M. (2020). *Language Models are Few-Shot Learners.* OpenAI.

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023). Chatting and Cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International* .

Ding, Z., & Chan, J. (2023). Mapping the Design Space of Interactions in Human-AI Text Co-creation Tasks. *CHI 2023 Workshop on Generative AI and HCI.* Virtual.

Fitts, J., & Bovard, R. (2023). AI and the Future of Academic Integrity. *UNLV Best Teaching Practices Expo.* 198 p.

Hwang, G.-J., & Chang, C.-Y. (2021). A review of opportunities and challenges of chatbots in education. *INTERACTIVE LEARNING ENVIRONMENTS.*

Kashyap, R., & ChatGPT. (2023). *A First Chat with ChatGPT: The First Step in the Road-Map for AI (Artificial Intelligence).*

King, M. R. (2023, February). The Future of AI in Medicine: A Perspective from a Chatbot. *Annals of Biomedical Engineering*, pp. 291-295.

Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: how may AI and GPT. *LIBRARY HI TECH NEWS*.

MacNeil, S., Tran, A., Mogil, D., Bernstein, S., Ross, E., & Huang, Z. (2022). Generating Diverse Code Explanations using the GPT-3 Large Language Model. *ICER 2022, Lugano and Virtual Event, Switzerland.*

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving Language Understanding by Generative Pre-Training.* OpenAI.

Shidiq, M. (2023). THE USE OF ARTIFICIAL INTELLIGENCE-BASED CHAT-GPT AND ITS CHALLENGES FOR THE WORLD OF EDUCATION; FROM THE VIEWPOINT OF THE DEVELOPMENT OF CREATIVE WRITING SKILLS. *Proceeding of International Conference on Education, Society and Humanity*, pp. 360-364.

Shiri, A. (2023). ChatGPTand academic integrity. *Information Matters, 3*(2). Retrieved from https://informationmatters.org/2023/02/chatgpt-and-academic-integrity/.

Willems, J. (2023). ChatGPT at universities – The least of our concerns. doi:dx.doi.org/10.2139/ssrn.4334162.

Chatbot | definition of chatbot in English by Lexico Dictionaries, (2023). Lexico Dictionaries| English website:https://www.lexico.com/en/definition/chatbot. [Accessed on 16 March 2023].

Hu, K. (2023). ChatGPT sets record for fastest-growing user base-analyst note. [Accessed on 12 March 2023].

Turnitin. AI Writing| AI Tools. https://www.turnitin.com/solutions/ai-writing. [Accessed on 16 March 2023].