Human-Created and AI-Generated Text: What's Left to Uncover

Salter, S., Teh, P.L., Hebblewhite, R.

This is a paper presented at the Intelligent Computing, SAI 2024.

Human-Created and AI-Generated Text: What's Left to Uncover?

Steven Salter, Phoey Lee Teh(✉), and Richard Hebblewhite

Faculty of Art, Computing and Engineering, Wrexham University, Wrexham, UK
{phoey.teh,r.hebblewhite}@glyndwr.ac.uk

Abstract. The advent of generative Artificial Intelligence (AI) has brought about profound changes in society, education, and the professional realm. Machine Learning models have created remarkably sophisticated language generators, blurring the line between human-authored content and AI-generated text. This poses a challenge for educators and professionals in distinguishing authentic work from instances of plagiarism. This study investigates the fundamental distinctions between human and AI texts by analyzing perspectives from human points of view. It aims to present the outcome of those five questions that we examined about human perceptions of text composition and contrasting them with computer-generated text. By exploring this, it aids in upholding academic integrity and contributes to advancing our comprehension of Natural Language Processing. In essence, this research strives to maintain academic credibility in a landscape transformed by AI and nurtures the growth of more equitable AI technologies.

Keywords: Artificial intelligent · Chat GPT · Human generated-text · AI generated text

1 Introduction

Machine Learning (a key part of AI) has led to the creation of advanced Large Language Models (LLMs). These models can produce very realistic text that resembles human writing. It's getting harder to tell if a text was written by a person or generated by AI, and this puzzle has attracted a lot of media focus [1]. A model like ChatGPT has gained widespread attention and quickly drawn millions of users [2]. ChatGPT is like a smart tool that understands and generates human-like text. It uses a bunch of clever techniques to quickly create clear and logical sentences. It can match the way humans write, and this is the important focus of this paper. The tool uses model that learned from a lot of text. The new version, ChatGPT 4.0, released in April 2023, is even better and can handle a lot of information at once.

Generative AI has a lot of good things it can do for society, but there's a worry about fake academic work made by this kind of technology. There is a need to ensure that assignments completed by AI are not included in assessments. Additionally, it is crucial to be able to distinguish between AI-generated work and human-produced content.

ChatGPT and its associated bots have smoothly become a part of many popular websites. Chances are you've come across their capabilities without even realizing it.

K. Arai (Ed.): SAI 2024, LNNS 1017, pp. 74–80, 2024.
https://doi.org/10.1007/978-3-031-62277-9_5

Of particular importance, ChatGPT has created a versatile tool that can be utilized by a wide range of websites, software, and applications for tasks such as generating art and composing music. This expansion has considerably increased ChatGPT's popularity, drawing in users from various backgrounds, extending beyond computer scientists and engineers, who can enhance their work processes and more through its features.

Our objective is to unravel the distinctions between texts authored by humans and those crafted by machines. To achieve this, we commence with a questionnaire, which serves as our initial step. Concurrently, we gather human-generated texts for further elaboration in subsequent stages. Employing a unique approach to language comprehension, we seek to shed light on this differentiation process. It's a small study that looks into this topic, and it's important because it can help make sure that when we use AI like ChatGPT to write, we also give credit to the right sources. This research also gives us some good ideas about how tech companies, researchers, and schools can work together to stop cheating and copying. This can help students use AI in a positive way to help the world or augmenting their works.

The research question is the same as what we first asked in the title, but because the article is meant to present a smaller outcome of the study, we will only present the responses pertaining to the five questions discussed in Sect. 3.3.

2 Literature Review

Researcher has mentioned that the long-term consequences of generative AI are uncertain; it could either greatly improve our lives or lead us into a disastrous future [3]. Having spent the past year experimenting with ChatGPT, I've come to realize that allowing students to use ChatGPT for their school assignments or other tasks is actually a beneficial idea.

Generative AI can be a helpful tool for students to learn more about different theories and enhance their overall understanding. It can answer many questions without students feeling embarrassed in front of their classmates. This technology lets students explore ideas quickly and satisfies their curiosity [4]. Moreover, it allows students to ask multiple questions without burdening teachers with excessive demands. This extra help gives teachers more time to focus on important tasks. This approach can improve students' knowledge and assist with their schoolwork, broadening their understanding.

When used like this, the risk of cheating is quite low. In a paper titled "Should Using an AI Generator be Plagiarism?", Frye [5] brings up an interesting idea. He says that if students can get good answers using AI, then the questions they're asked might be too easy. Interestingly, ChatGPT is a coauthor of his paper, adding a clever twist. Though his comment might seem a bit critical, there are valid points to think about [5].

To delve deeper into how ChatGPT operates, here is a list of nine key points to consider. 1) Structured Responses: ChatGPT maintains an organized writing style, presenting ideas coherently. It typically starts by summarizing the main points of the questions and then provides a comprehensive response. This is often followed by a concluding summary [6]. 2) Thorough Explanations: ChatGPT's tendency to offer detailed responses is shaped by Reinforcement Learning with Human Feedback (RLHF). It usually provides thorough explanations unless specifically asked for brevity [7], in other words, ChatGPT tends to give long and detailed answers because of the way it learns from human

feedback. It usually provides these detailed answers unless you ask for a shorter one. Sometimes, the way it writes can affect how human perceive it [8]. 3) Neutral Handling of Sensitive Topics: When discussing sensitive subjects, ChatGPT remains neutral. This quality is especially useful for addressing politically charged issues, as it avoids bias or harmful content [9]. 4) Recognizing Knowledge Limits: ChatGPT acknowledges its limitations and refrains from answering when it lacks the necessary information [10]. 5) Occasional Fabrication of Information: At times, ChatGPT may generate incorrect details, particularly when responding to queries beyond its expertise [11]. 6) Use of Repetitive Phrases: The model might use certain terms repeatedly, resulting in lack of variety in its output [12]. 7) Following Set Patterns: ChatGPT's sentences can adhere to specific templates, leading to predictable language patterns [13]. 8) Contextual Accuracy: ChatGPT's responses might lack accuracy in the context, leading to logically or contextually irrelevant information [14]. 9) Filling Information Gaps: Even if ChatGPT doesn't fully grasp the context, it attempts to fill in gaps using patterns from its training data [15].

Text analysis researcher has demonstrated the capability of human text to convey nuanced sentiment [16], particularly in the context of text messaging and written conversations, effectively capturing the intricacies of human communication through written text. As human beings, our communication goes far beyond the literal, influenced by the nuances of language and culture. For instance, studies have explored the detection of sarcasm [17] in human language and the concealed that emotions within text could potentially reverse the intended sentiment the polarity of expression [18]. Others study also stated that the repetition of words could affect the intended sentiment [19]. As part of composing this paragraph, a straightforward query was presented to ChatGPT, prompting the generation of a sentence that incorporates both sarcasm and emotion, as demonstrated in the response: "Oh, wonderful! Another Monday morning at work, my absolute favorite time of the week." This raises the question: Can the emotions conveyed through the way words are used in writing help us tell the difference between text written by humans and text generated by ChatGPT?

3 Methodology

3.1 Data Collection

The study uses a combination of methods. It looks at both the text itself and uses numbers to analyze and model the data. The focus is on understanding how AI-generated text differs from human-written text. The study involves comparing 66 summaries of research papers abstract made by AI to those made by human. These summaries are taken from abstract from one specific author, keeping the consistency of topic/theme on check. The aim is to compare how well AI and people can summarize the same research.

To engage participants, a focus group was organized at University, where an advertisement was placed in the main foyer. This group was informed about the purpose of study, which would take around 15 min to complete. A computer lab was arranged, equipped with computer for respondents to respond. This setup allowed for in-person informed consent, immediate response to participant queries, and clear instructions. Additionally, participants were invited to a ChatGPT workshop, designed by OpenAI,

which showcased the benefits of generative AI and its practical applications. The workshop aimed to teach attendees, including teaching staff, students, and non-teaching personnel, how to utilize ChatGPT effectively. It demonstrated how to write simple prompts and showcased the system's capabilities. Participants were enthusiastic about joining the research, especially since attendance at the workshop was offered in exchange for their participation. Light refreshments were also provided as a token of appreciation for their time.

3.2 Participant

A total of 66 participants took part in the study, with 59% being male and 41% female. Among them, approximately 42% had a postgraduate background, 30% were undergraduates, and 28% had an educational background of A levels and GCSEs. Participants from various educational backgrounds were engaged in summarizing the texts. Following the summarization task, participants answered five personal perspective questions. They were given a 15-min window for summarization to encourage natural responses without overthinking. Throughout the study, participants' identities were kept confidential. Ethical approval for the study was obtained before conducting the experiment. Participants provided basic information such as gender, education, age group, and whether English was their first language.

3.3 Survey/Questionnaire Design

To gauge participants perspective towards text disguising, participant was asked to rate the score on all the following questions on a scale on 1 to 5 (with 1 being "strongly disagree", 2 being "disagree", 3 being "don't know or neutral", 4 being "agree" and 5 being "strongly agree").

1) Errors in grammar are key indicators that text has been generated by humans?
2) Repetition in text is a key indicator that text has been generated by humans?
3) Emotional meaning (underlying emotional nuances, tones, or sentiments conveyed through the choice of words, phrasing, and overall expression of the written content) is a key indicator that text has been generated by humans?
4) Mistakes around capitalisation, grammar and syntax are key indicators that text has been generated by humans?
5) Connotations (the emotional or cultural associations that are evoked by certain words, phrases, or expressions beyond their literal or dictionary definitions) are key indicators that text has been generated by humans?

4 Results

Figure 1 displaying the outcomes of the perspectives on the five questions posed earlier. Out of the total participants, 61 responded to the questionnaire, as five participants did not provide answers, rendering their responses invalid. The mean ratings for each question are as follows: Q1 - 3.39, Q2 - 2.87, Q3 - 3.80, Q4 - 3.49, Q5 - 3.70. Generally, there is not a significant difference in the ratings, except for Q2 where there seems to

be more disagreement regarding the notion that text repetition is a clear sign of AI-generated content. The responses also indicate a higher level of agreement concerning the influence of education on writing style and sentiment towards the choice of words and phrasing.

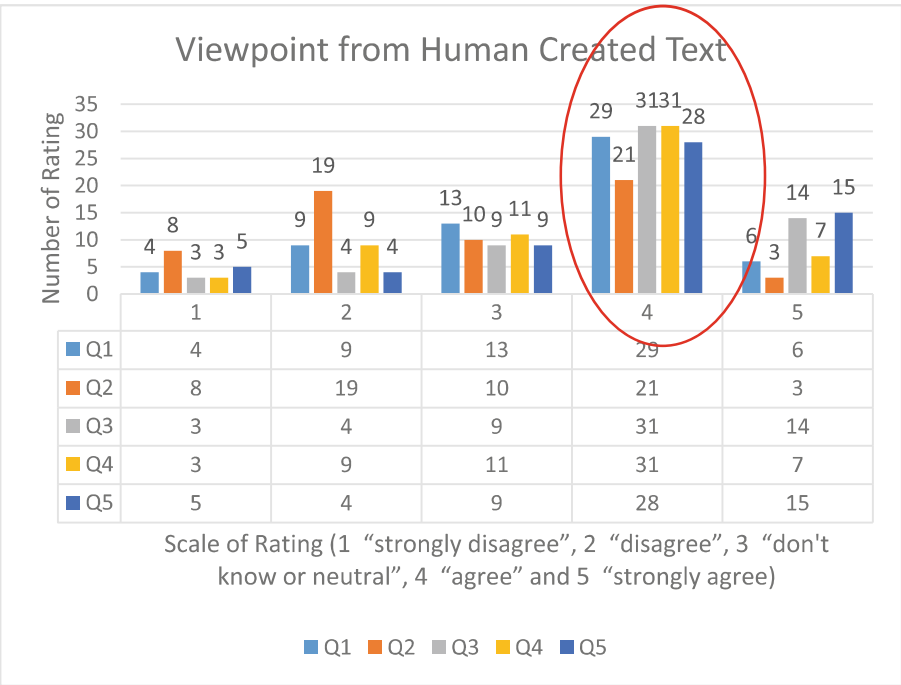

Fig. 1. Viewpoint from human created text

It’s interesting that about half of the people strongly agree that when there are mistakes in things like capital letters, grammar, and how sentences are put together, it’s a good clue that a human wrote the text. So, these little errors might be like a fingerprint that shows if a person or a machine wrote something.

Also, a lot of the participants think that some words or phrases can make you feel certain emotions or remind you of specific things from your culture. It’s like words can carry extra meanings beyond what you find in the dictionary. This kind of special feeling in the words seems to be something that humans are really good at, and it’s not easy for machines to do the same. The fact that many people agree with these ideas helps us understand that there are still parts of how humans use language that machines haven’t quite figured out yet.

5 Conclusion

In conclusion, our aim was to provide insights based on the responses to the questions, shedding light on society's viewpoint. Despite the limited responses from the UK, we managed to gain valuable insights into the aspects agreed upon by human respondents regarding their perspective on human-generated language. In the future, further studies will delve into the topic, focusing on the determination to uphold academic integrity. Furthermore, this research aims to increase awareness and organize workshops for informing different stakeholders, particularly those in academia, about the ways AI can enhance various aspects of their work. It also encourages exploration of what distinctions remain between human-generated and machine-generated text.

References

1. Zohny, H., McMillan, J., King, M.: Ethics of generative AI. J. Med. Ethics **49**, 79–80 (2023)
2. Karaarslan, E., Aydın, O.: Is ChatGPT Leading Generative AI? What is Beyond Expectations? Social Science Research Network (2023)
3. Andersen, E.: The Technological Singularity, Ubiquity Symposium (2014)
4. Crawford, J., Cowling, M., Allen, K.: Leadership is needed for ethical ChatGPT: character, assessment, and learning using artificial intelligence (AI). J. Univ. Teach. Learn. Pract. **20**(3), Article 02 (2023)
5. Frye, B.L.: Should using an AI text generator to produce academic writing be plagiarism. Fordham Intell. Prop. Media Ent. LJ **33**(947), 949–969 (2023)
6. Holden Throp, H.: ChatGPT is fun, but not an author. Science **379**(6630), 313 (2023)
7. Benzon, W.: A note about story grammars in ChatGPT, Independent, no. SSRN (2023) https://ssrn.com/abstract=4324840
8. Teh, P.L., et al.: Textual variations affect human judgements of sentiment values. Electron. Commerce Res. Appl. **53**, 101149 (2022)
9. Thunstrom, A.O., Steingrimsson, S.: Can GPT-3 write an academic paper on itself, with minimal human input. Hal Open Science, no. 03701250 (2022)
10. OpenAI: ChatGPT (2022). https://chat.openai.com/. Accessed 05 June 2023
11. Khalil, M., Er, E.: Will ChatGPT get you caught? Rethinking of Plagiarism (2023)
12. Uddin, M. S.: Stochastic Parrots: A Novel Look at Large Language Models and Their Limitations, towardsai.net (2023). https://towardsai.net/p/machine-learning/stochastic-parrots-a-novel-look-at-large-language-models-and-their-limitation. Accessed 17 June 2023
13. Lund, B.D., Wang, T.: Chatting about ChatGPT: how may AI and GPT impact academia and libraries? Library Hi Tech News (2023)
14. Rudolph, J., Tan, S., Tan, S.: ChatGPT: bullshit spewer or the end of traditional assessments in higher education? J. Appl. Learn. Teach. **6**(1), 342–363 (2023)
15. Perkins, M.: Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. J. Univ. Teach. Learn. Pract. **20**(2) (2023). L. S. G. ltd, Baldur's Gate 3, Larian Studios Games ltd, 2023. https://baldursgate3.game/. Accessed 11 Aug 2023
16. Teh, P.L., Pak, I., Rayson, P., Piao, S.: Exploring fine-grained sentiment values in online product reviews. In: 2015 IEEE Conference on Open Systems (ICOS), Melaka, Malaysia, pp. 114–118 (2015)
17. Teh, P.L., Ooi, P.B., Chan, N.N., Chuah, Y.K.: A comparative study of the effectiveness of sentiment tools and human coding in sarcasm detection. J. Syst. Inf. Technol. **20**(3), 358–374 (2018)

18. Teh, P.L., Rayson, P., Pak, I., Piao, S., Yeng, S.M.: Reversing the polarity with emoticons. In: Métais, E., Meziane, F., Saraee, M., Sugumaran, V., Vadera, S. (eds.) NLDB 2016. LNCS, vol. 9612, pp. 453–458. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-41754-7_48
19. Pak, I., Lee Teh, P., Cheah, Y.-N.: Hidden sentiment behind letter repetition in online reviews. J. Telecommun. Electron. Comput. Eng. (JTEC) **10**(3–2), 115–120 (2018)