UDC 37.013 **DOI: 10.34670/AR.2023.62.69.112**

Tackling the issue of cheating with ChatGPT in open-ended assignments at university

Anastasiya A. Kalashnikova

Bachelor of Linguistics,
Higher School of Economics,
101000, 20, Myasnitskaya str., Moscow, Russian Federation;
e-mail: nastykalach@yandex.ru

Abstract

Since the company OpenAI released its ChatGPT (Generative Pretrained Transformer), a powerful neural network capable of generating human-like replies based on a collection of samples, educational institutions faced up with the challenge of updating their plagiarism and cheating policies, introducing guidance on the acceptable usage of artificial intelligence (AI) in studies and offer special instructions for teachers and professors on how to spot the texts, produced by machines. In the article a Large Language Model ChatGPT and cheating with Artificial Intelligence (AI) tools are discussed. As cheating with ChatGPT has become a problem for educational institutions, the paper aims to tackle the mentioned issue. It was previously found out that ChatGPT has several limitations distinguishing its output from a real person's response. This article broadens the previous research and extends it to the sphere of higher education. It evaluates the upsides and downsides of using AI tools and explores the measures to prevent students from plagiarizing. Its practical section is based on qualitative analysis of ChatGPT's outputs. As a result, several guidelines were proposed for university tutors on how to spot a response generated by AI and how to distinguish it from human writing by certain criteria, such as repetitions, inaccurate use of language, stereotyped phrasing, commonsense presence, fake facts, critical thinking, alignment to the course program and explanations and examples.

For citation

Kalashnikova A.A. (2023) Tackling the issue of cheating with ChatGPT in open-ended assignments at university. *Pedagogicheskii zhurnal* [Pedagogical Journal], 13 (5A), pp. 844-854. DOI: 10.34670/AR.2023.62.69.112

Keywords

ChatGPT, cheating at university, fake assignments, neural networks/models, artificial intelligence.

Introduction

Since the company OpenAI released its ChatGPT (Generative Pretrained Transformer), a powerful neural network capable of generating human-like replies based on a collection of samples, educational institutions faced up with the challenge of updating their plagiarism and cheating policies, introducing guidance on the acceptable usage of artificial intelligence (AI) in studies and offer special instructions for teachers and professors on how to spot the texts, produced by machines.

It is crucial for educators to know how to notice cheating with neural models in open-ended tasks for two major reasons. Firstly, the students who constantly copy the answers from ChatGPT being unnoticed may fall behind the rest of the group [Shearing & McCallum, 2023]. It results in failure to develop the students' memory, critical thinking skills and subject expertise necessary for performing job duties and further studies well. Secondly, the breakdown in proper knowledge transmission undermines universities' academic reputations. The value of a graduate degree will diminish posing difficulty for recruiters to hire truly qualified staff.

The *object* of the research is the concept of cheating with the help of ChatGPT. The object was selected because of the lack of available knowledge on the topic.

The *subject* of the analysis is the peculiar traits of the neural model's replies to complex academic open-ended tasks by which they can be distinguished from a real person's answer by professors. The choice of the subject is explained by the urgent need for developing a functional strategy for recognizing fake assignments submitted by students.

The *goal* of the study is to produce a list of practical recommendations for faculty members on how to spot a comment generated by a neural network rather than a student himself. The following *objectives* were set up to accomplish the stated goal:

– to explain the limitations that make it possible to detect ChatGPT's output;

– to evaluate the upsides and downsides of using AI tools in higher education;

– to explore the measures to prevent students from plagiarizing;

– to create guidelines for university teaching staff on recognizing texts written by machines.

The *methods* employed in the research include observation and qualitative content analysis of ChatGPT responses to academic open-ended questions. The tasks are taken from real Cambridge philosophy exams published on their webpage [Faculty of Philosophy at Cambridge University, 2013, 2015, 2019(a), 2019(b), 2021, 2022]. One example is exceptional because it is borrowed from a math textbook for the third grade of a famous Russian educator [Peterson, 2012].

The work will be useful for university teaching staff, other faculty members and educators in general as it gives understanding on how to identify cheating with AI models.

Limitations that make it possible to detect ChatGPT's output

ChatGPT is an advanced neural network which has been recently opened to public by OpenAI. It has instantly gained popularity because of its ability to propose relatively accurate solutions to a broad variety of issues in a conversational way of interaction. The model is trained on large amounts of data using deep learning technology and with reinforcement learning from human feedback (RLHF), which is supposed to make them more useful and aligned [Ouyang et al., 2022]. However, it possesses some limitations that can be used for distinguishing its response from a real person's one.

The research has proved that LLMs memorize segments of their training data and then reproduce them verbatim [Carlini et al., 2023]. The likelihood of repetition depends on the frequency of the sequence in the training set [Carlini et al., 2020; Lee et al., 2022]. It confirms the presence of *repetitions*

and standard expressions in the outputs of ChatGPT.

Commonsense knowledge is essential for LLMs to comprehend and produce human-like language [Bian et al., 2023]. Such knowledge is difficult to transmit to machines [Li et al., 2022; Zhang et al., 2022] because of it being implicit and its high dependence on the context [Gordon and Van Durme, 2013; Shwartz and Choi, 2020]. That is why the commonsense problem is often seen as the major limitation of LLMs [Zhou et al., 2020; Bhargava and Ng, 2020]. However, the models are being constantly trained and by now the creators have managed to attain a certain degree of commonsense knowledge in them [West et al., 2022].

The most recent research on this topic [Bian et al., 2023] has shown that ChatGPT is well-informed about commonsense knowledge to answer questions accurately, nonetheless the data it encompasses may sometimes be misleading and overgeneralized. This leads to the conclusion that to notice a solution written by ChatGPT rather than a student himself/herself, educators should evaluate the output of ChatGPT against the presence of commonsense.

Wrong judgements and fake facts can also help to identify a machine generated text. The latest multimodal model GPT-4 is by far the most intelligent one. As the previous edition GPT-3 it is still not entirely reliable, because it "hallucinates" facts and makes reasoning errors [Nakano et al., 2021; OpenAI, 2023]. In addition, it is not always capable of separating facts from fiction. Besides, most of its data resources was disconnected in September 2021, as a result the system is unaware of later events. One more downside is that it can sound convincing while making wrong predictions and it does not check the results twice when it is likely to make a mistake. ChatGPT shares the same drawbacks to a larger extent.

In this chapter there were listed and explained several limitations of ChatGPT that may help academics detect a work plagiarized with AI. They include repetitions and standard expressions, lack of commonsense knowledge, wrong judgements, and fake facts.

The upsides and downsides of using AI tools in higher education

There is an ongoing debate on the benefits and drawbacks of the use of LLMs in the education field. While some scholars see the potential for students' growth and development, others are afraid of a massive biased and fake facts reproduction.

The proponents of using ChatGPT in higher education claim that asynchronous communication gives students time to discuss topics with fellow students. It promotes collaboration in group projects [Li, Xing, 2021]. One more advantage is that ChatGPT enables remote learning for students who are unable to attend classes [Barber et al., 2021]. In addition, GPTs can generate tailored learning and customized tests [Bommasani et al., 2021]. Edwin Bodge, principal product manager at Duolingo, expressed the intention of the company to continue using it for personalizing lessons and improving conversation practice and contextual feedback [OpenAI, 2023]. It is also important that it can provide feedback in real-time or any time a student needs it.

As for the challenges ChatGPT and other such models pose for education, the major issue is the possibility of plagiarism. In pursue to achieve the edge in assessments, students may submit essays not written by themselves but by ChatGPT [Dehouche, 2021]. It questions the value of an academic degree and the necessity of getting it. Secondly, the access to ChatGPT or its absence draws unfair inequality in academic performance. Educators may face up with a difficulty in assessing one's real knowledge and providing timely feedback, if a student uses a chatbot application in their studies [Cotton, Cotton & Shipway, 2023].

All in all, technology cannot be completely excluded from learning nowadays, however its usage

should be regulated properly and its adverse effects on education should be minimized. The actions to prevent students from plagiarizing are explored in the next section.

Measures to prevent students from plagiarizing

There are several strategies to tackle the above-mentioned problems. They include a policy on ChatGPT usage, educational classes on plagiarism, plagiarism detection tools and checking works by academic staff.

Firstly, there should be a strict policy on how and when a student can use the technology. The tutors need to agree on the assignment types or task stages where it is possible to permit students refer to ChatGPT. For example, Kim Watts, marketing lecturer, allowed her students to use ChatGPT as a tool for generating ideas for their marketing plans if they do not know where to begin with [Shearing, McCullum, 2023].

Secondly, it should be explicitly explained what is considered plagiarism and what is not, why it is wrong to cheat with AI and what the consequences are for the student and their academic performance. Educating students on plagiarism and asking them to sign a declaration that they didn't use any AI tools to complete the task may make students more responsible for their actions [Cotton, Cotton, Shipway, 2023].

Thirdly, there are special tools that has long been used to check thesis works for plagiarism. For instance, Turnitin. They recognize AI-generated content in students' works, however nowadays exist some ways to bypass those detection algorithms, such as code substitution. Code substitution makes the text look the same, but the machine will read it as another text, which is unique. Thus, the detention algorithms will show that there is no plagiarism at all. That is why it is important that academic staff check their students' works for ChatGPT responses themselves. The next chapter provides some recommendations on how to do it.

Recommendations to university teaching staff on how to recognize fake assignments submitted with the help of ChatGPT

Here are several tips on what to look for in the students' assignments.

Scan the text for repetitions.

The first step in identifying a response generated by the AI tool is checking whether it contains just the same wording as that of the task. ChatGPT may sometimes reproduce the question (one or more times in one answer) since it follows certain algorithms and patterns underlying the process of its text production.

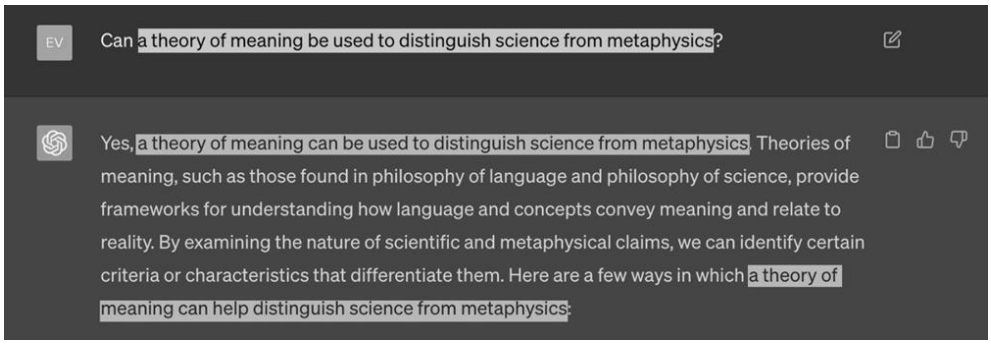

Figure 1 - ChatGPT repeats the question two times

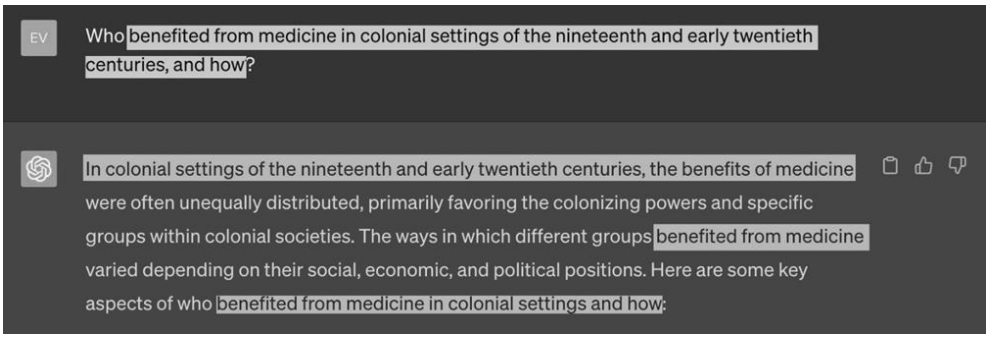

Figure 2 - ChatGPT repeats the wording of the question

The examples above provide evidence that ChatGPT tends to use the wording of the inquiry without paraphrasing it. If a student simply copies the text to his answer, it will be easy to notice that.

Look for stereotyped phrasing and inaccurate use of language.

The second action that needs to be done is searching for clichés and phrases that does not sound natural. As ChatGPT neural model is trained to produce certain sentence structures, they appear unchanged or slightly transformed in its outputs to different inputs. The following examples demonstrate that ChatGPT reproduces the same sentence structure many times.

Evaluate the solution against commonsense.

At this stage the presence of commonsense in the responses should be verified by humans, because the chatbot struggles with solving cases from real-life. The developers of ChatGPT claim that neural networks deal with real-world situations not as effectively as people do. The research on how to improve it is still on the go.

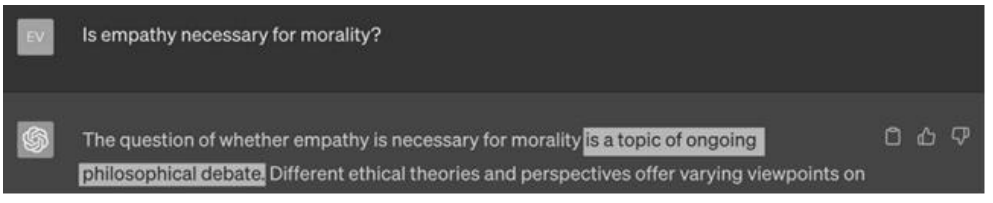

Figure 3 - Stereotyped phrasing

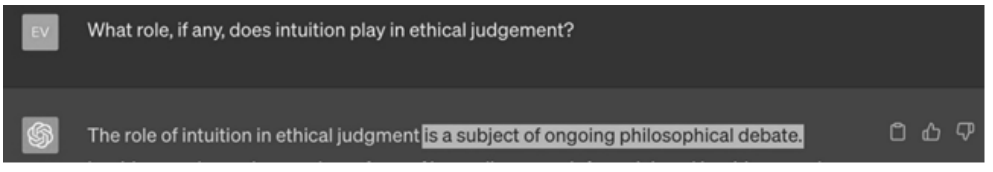

Figure 4 - Cliché

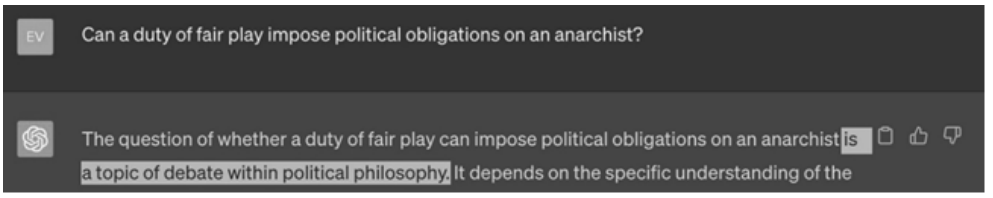

Figure 5 - The same sentence structure ChatGPT reproduces multiple times

Here is an example of a mathematical problem. ChatGPT calculated the total time for frying one cutlet correctly, however it miscounted the amounts of time for frying a batch of 4 cutlets and a batch of 2 cutlets (although it was stated in the condition of the task that 4 cutlets can fit in a frying pan). Thus, the division of cutlets in batches in the response to the first question in the task is not justified. A real person will never fry each cutlet separately and spend 60 minutes on it instead of 15 or 20. So, the information ChatGPT provided in reply is irrelevant and contradicts commonsense.

The second part of the answer is much closer to reality. It represents what an ordinary person would do in this case. However, it does not offer a solution on how to fry 6 cutlets in 15 minutes using only one frying pan.

Examine the facts presented in your students' answers to find fakes.

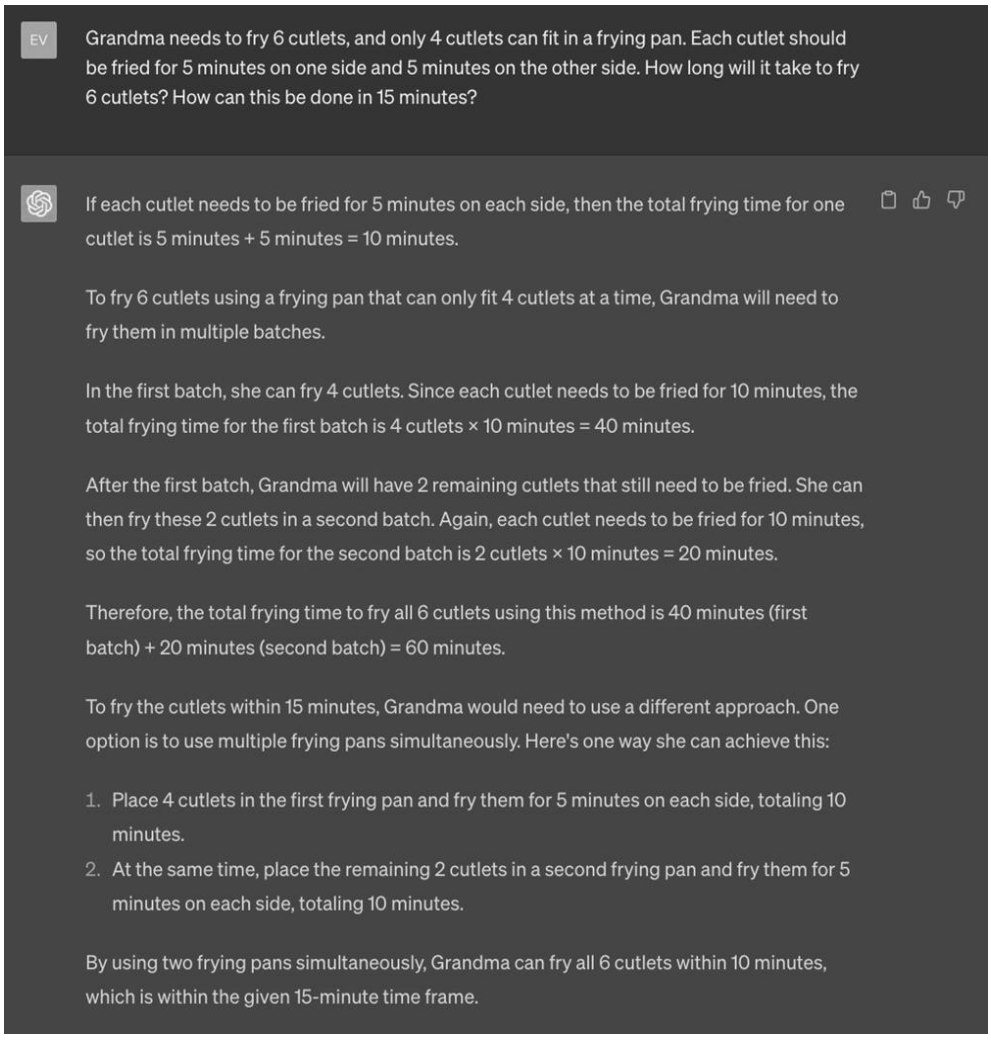

Figure 6 - Commonsense problem

Next, the assessor should read the text more thoroughly to spot any misinformation. AI is not always capable of distinguishing between proven facts and false statements. It can use unreliable resources and fabricated data to produce its reply. Thus, the text may include completely wrong judgements. For instance, the figures below show that ChatGPT in its reply to the philosophical task pursues incorrect logic and comes to the wrong conclusion.

Assess the depth of answer and alignment to what had been taught.

Then the examiner should estimate the depth of response and task fulfillment. In addition, it must be checked whether the text shows any learning and includes any ideas presented during the course. It is essential, because educators generally agree that the AI tool under discussion is unable to demonstrate enough critical thinking and depth of subject knowledge to provide a reply sufficient for academic purposes.

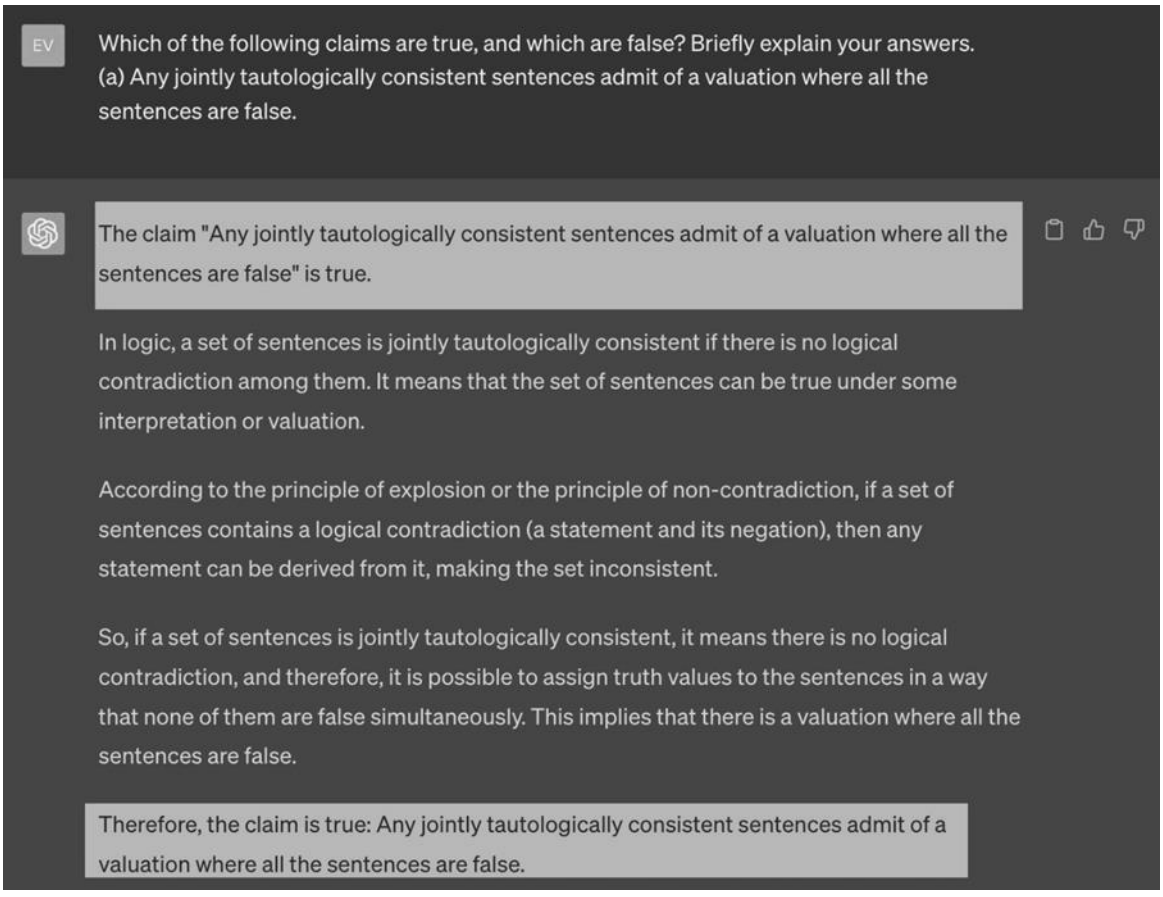

Figure 7 - Wrong judgement

1. Which of the following claims are true, and which are false? Briefly explain your answers.
   (a) Any jointly tautologically consistent sentences admit of a valuation where all the sentences are false.
   (b) Any jointly tautologically inconsistent sentences will remain tautologically inconsistent if any other sentence is added.
   (c) If an argument's premises are jointly inconsistent with its conclusion, the argument must be invalid.

Solution:

(a) False: that there is a valuation where all the sentences are true does not entail that there is one where they are all false. For example, $(A \vee \neg A)$ is consistent with B, yet there is no valuation rendering both sentences false.

Figure 8 - Cambridge examiner's answer

The question in figure 10 is philosophical. The examiner would expect the student to contemplate on social contract theories, compare the ideas of contractarians (for instance, Thomas Hobbes) and

contractualists (for example, Immanuel Kant), and based on these viewpoints make a conclusion on whether it is rational to act against one's self-interest or not. The output of ChatGPT lacks critical thinking and does not show any learning. It is not adequate for scoring a good mark in the exam.

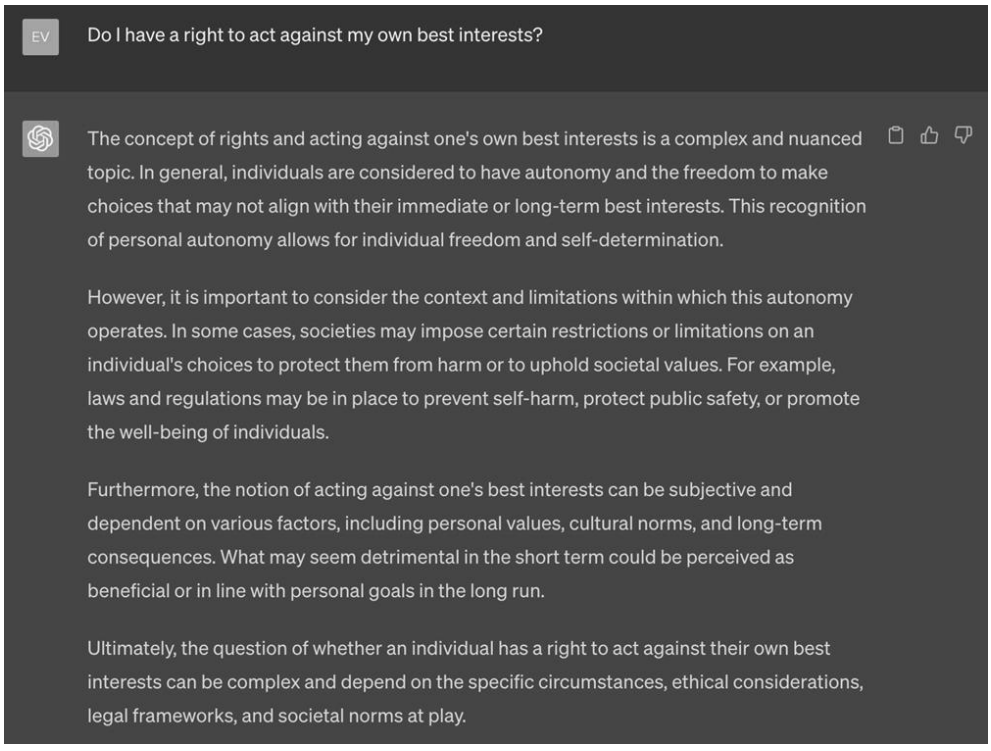

Figure 9 - The - answer lacks critical thinking and does not demonstrate any learning

In this chapter some recommendations and examples on ChatGPT's limitations were provided to demonstrate their usefulness. They include scanning the text for repetitions, looking for stereotyped phrasing and inaccurate use of language, evaluating the solution against commonsense, examining the facts to find fakes and assessing the depth of answer and alignment to what had been taught.

Limitations

The main limitation of this work is that is explores only the open-ended type of assessment. In addition, the frequency of the mentioned flaws of ChatGPT was not evaluated, so, hypothetically, an educator may not spot any presence of them in their students' works.

Conclusion

ChatGPT is a controversial tool. From one perspective, it undermines universities' reputations, poses threats to knowledge transmission and critical thinking, and draws inequality in academic performance. On the other hand, it should not be banned completely, as it may be a resource for tailored learning and customized tests, it enables remote studies and helps generate ideas for works. That is why it is crucial to develop a strategy for plagiarism avoidance and appropriate usage of ChatGPT at universities.

The theoretical part of work provides the basis for its practical part. It informs the reader about the research done into the limitations of LLMs in general. The studies prove that ChatGPT sometimes

"hallucinates" facts, possesses restricted commonsense knowledge, and uses standard expressions as well as multiple repetitions of questions. This section also observes the pros and cons of using AI tools in higher education and proposes several measures to prevent students from plagiarizing with these technologies.

The practical part presents a guideline for educators on what to search for in the students' assignments to spot cheating with ChatGPT. The recommendations include scanning the text for repetitions, looking for stereotyped phrasing and inaccurate use of language, evaluating the solution against commonsense, examining the facts to find fakes and assessing the depth of answer and alignment to what had been taught.

The article is instrumental for academic staff in making decisions on plagiarism policies concerning ChatGPT. They should think through the possibility for students to use AI to generate ideas for some tasks and consider scheduling a class on what is considered plagiarism and what is not. It is also helpful in identifying cheated responses, which is supposed to help tutors take timely actions.

References

1. (2015) Part IA Past Exam Papers. Ethics and political philosophy. Available at: cam.ac.uk [Accessed 05/05/2023]
2. (2019(a)) Part IA Past Exam Papers. Ethics and political philosophy. Available at: cam.ac.uk [Accessed 05/05/2023]
3. (2019(b)) Part IA Past Exam Papers. Meaning. Available at: cam.ac.uk [Accessed 05/05/2023]
4. (2022) Part IA Past Exam Papers. Formal methods. Available at: cam.ac.uk/paper_5_formal_methods_external_comments_version_2.pdf [Accessed 05/05/2023]
5. (2022) Part IA Past Exam Papers. Formal methods, Paper 5, Answer Key. Available at: cam.ac.uk/ia-5_answer-key.pdf [Accessed 05/05/2023]
6. (2013) Part IA Past Exam Papers. Ethics. Available at: cam.ac.uk [Accessed 05/05/2023]
7. (2021) Part IA Past Exam Papers. History and philosophy of science. Available at: cam.ac.uk [Accessed 05/05/2023]
8. (2023) GPT-4 deepens the conversation on Duolingo. Available at: https://openai.com/customer-stories/duolingo [Accessed 05/05/2023]
9. (2023) GPT-4 Technical Report. Available at: https://arxiv.org/abs/2303.08774 [Accessed 05/05/2023]
10. Barber M. et al. (2021) Gravity assist: propelling higher education towards a brighter future. Available at: https://www.officeforstudents.org.uk/publications/gravity-assist-propelling-higher-education-towards-a-brighter-future/ [Accessed 05/05/2023]
11. Bian N. et al. (2023) ChatGPT is a Knowledgeable but Inexperienced Solver: An Investigation of Commonsense Problem in Large Language Models. Available at: https://arxiv.org/abs/2303.16421 [Accessed 05/05/2023]
12. Bommasani R. et al. (2021) On the Opportunities and Risks of Foundation Models. Available at: https://arxiv.org/abs/2108.07258 [Accessed 05/05/2023]
13. Carlini N. et al. (2020) Extracting Training Data from Large Language Models. Available at: https://arxiv.org/pdf/2012.07805.pdf [Accessed 05/05/2023]
14. Cotton D.R.E., Cotton P.A., Shipway J.R. (2023) Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. In: Innovations in Education and Teaching International. doi: https://doi.org/10.1080/14703297.2023.2190148
15. Gordon J., Van Durme B. (2013) Reporting bias and knowledge acquisition. In: Proceedings of the 2013 workshop on Automated knowledge base construction – AKBC '13. doi: https://doi.org/10.1145/2509558.2509563
16. Lee K. et al. (2022) Deduplicating Training Data Makes Language Models Better. Available at: https://arxiv.org/abs/2107.06499 [Accessed 05/05/2023]
17. Li C., Xing W. (2021) Natural Language Generation Using Deep Learning to Support MOOC Learners. International Journal of Artificial Intelligence in Education, 31(2), pp.186-214. doi: https://doi.org/10.1007/s40593-020-00235-x
18. Li X.L. et al. (2022) A Systematic Investigation of Commonsense Knowledge in Large Language Models. Available at: arXiv.org [Accessed 05/05/2023]
19. Nakano R. et al. (2022) WebGPT: Browser-assisted question-answering with human feedback. Available at: https://arxiv.org/abs/2112.09332 [Accessed 05/05/2023]
20. Ouyang L. et al. (2022) Training language models to follow instructions with human feedback. Available at: https://arxiv.org/abs/2203.02155 [Accessed 05/05/2023]
21. Peterson L.G. (2012) Mathematics. 3d Grade. Part 2. Moscow: Uventa Publ.
22. Shearing H., McCullum S. (2023) ChatGPT: Can students pass using AI tools at university? Available at:

https://www.bbc.co.uk/news/education-65316283 [Accessed 05/05/2023]
23. Shwartz V., Choi Y. (2020) Do Neural Language Models Overcome Reporting Bias? Available at: https://aclanthology.org/2020.coling-main.605.pdf [Accessed 05/05/2023]
24. West P. et al. (2022) Symbolic Knowledge Distillation: from General Language Models to Commonsense Models. Available at: arXiv.org [Accessed 05/05/2023]
25. Zhang Y. et al. (2021) Alleviating the Knowledge-Language Inconsistency: A Study for Deep Commonsense Knowledge. Available at: arXiv.org [Accessed 05/05/2023]

Решение проблемы списывания с помощью ChatGPT при ответах на открытые вопросы в университетах

Калашникова Анастасия Андреевна

Бакалавр лингвистики,
Высшая школа экономики,
101000, Российская Федерация, Москва, ул. Мясницкая, 20;
e-mail: nastykalach@yandex.ru

Аннотация

С тех пор, как компания OpenAI выпустила ChatGPT, мощную нейронную сеть, способную генерировать ответы на основе набора образцов, образовательные учреждения столкнулись с проблемой обновления своей политики в отношении плагиата и мошенничества. Необходимы инструкции для учителей и преподавателей о том, как распознавать тексты, созданные машинами. В статье обсуждаются большая языковая модель ChatGPT и списывание с помощью инструментов искусственного интеллекта (ИИ). Поскольку мошенничество с ChatGPT стало проблемой для образовательных учреждений, статья направлена на решение указанной проблемы. Ранее выяснилось, что ChatGPT имеет ряд ограничений, отличающих его продукт от ответа реального человека. Данная статья расширяет предыдущее исследование и распространяет его на сферу высшего образования. В нем оцениваются преимущества и недостатки использования инструментов ИИ и исследуются меры по предотвращению плагиата учащихся. Практический раздел основан на качественном анализе выходных данных ChatGPT. В результате преподавателям университетов было предложено несколько рекомендаций о том, как распознать ответ, сгенерированный ИИ, и как отличить его от написанного человеком по определенным критериям, таким как повторения, неточное использование языка, стереотипные формулировки, присутствие здравого смысла, поддельные факты и т.д.

Для цитирования в научных исследованиях

Калашникова А.А. Tackling the issue of cheating with ChatGPT in open-ended assignments at university // Педагогический журнал. 2023. Т. 13. № 5А. С. 844-854. DOI: 10.34670/AR.2023.62.69.112

Ключевые слова

ChatGPT, списывание в университете, фейковые задания, нейронные сети/модели, искусственный интеллект.

Библиография

1. Barber M. et al. Gravity assist: propelling higher education towards a brighter future. 2021. URL: https://www.officeforstudents.org.uk/publications/gravity-assist-propelling-higher-education-towards-a-brighter-future/
2. Bian N. et al. ChatGPT is a Knowledgeable but Inexperienced Solver: An Investigation of Commonsense Problem in Large Language Models. 2023. URL: https://arxiv.org/abs/2303.16421
3. Bommasani R. et al. On the Opportunities and Risks of Foundation Models. 2021. URL: https://arxiv.org/abs/2108.07258
4. Carlini N. et al. Extracting Training Data from Large Language Models. 2020. URL: https://arxiv.org/pdf/2012.07805.pdf
5. Cotton D.R.E., Cotton P.A., Shipway J.R. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT // Innovations in Education and Teaching International. 2023. P. 1-12. doi: https://doi.org/10.1080/14703297.2023.2190148
6. Gordon J., Van Durme B. Reporting bias and knowledge acquisition // Proceedings of the 2013 workshop on Automated knowledge base construction – AKBC ’13. 2013. doi: https://doi.org/10.1145/2509558.2509563
7. GPT-4 deepens the conversation on Duolingo. 2023. URL: https://openai.com/customer-stories/duolingo
8. GPT-4 Technical Report. 2023. URL: https://arxiv.org/abs/2303.08774
9. Lee K. et al. Deduplicating Training Data Makes Language Models Better. 2022. URL: https://arxiv.org/abs/2107.06499
10. Li C., Xing W. Natural Language Generation Using Deep Learning to Support MOOC Learners // International Journal of Artificial Intelligence in Education. 2021. 31(2). P. 186-214. doi: https://doi.org/10.1007/s40593-020-00235-x
11. Li X.L. et al. A Systematic Investigation of Commonsense Knowledge in Large Language Models. 2022. URL: arXiv.org
12. Nakano R. et al. WebGPT: Browser-assisted question-answering with human feedback. 2022. URL: https://arxiv.org/abs/2112.09332
13. Ouyang L. et al. Training language models to follow instructions with human feedback. 2022. URL: https://arxiv.org/abs/2203.02155
14. Part IA Past Exam Papers. Ethics and political philosophy. 2015. URL: cam.ac.uk
15. Part IA Past Exam Papers. Ethics and political philosophy. 2019. URL: cam.ac.uk
16. Part IA Past Exam Papers. Ethics. 2013. URL: cam.ac.uk
17. Part IA Past Exam Papers. Formal methods, Paper 5, Answer Key. 2022. URL: cam.ac.uk/ia-5_answer-key.pdf
18. Part IA Past Exam Papers. Formal methods. 2022. URL: cam.ac.uk/paper_5_formal_methods_external_comments_version_2.pdf
19. Part IA Past Exam Papers. History and philosophy of science. 2021. URL: cam.ac.uk
20. Part IA Past Exam Papers. Meaning. 2019. URL: cam.ac.uk
21. Peterson L.G. Mathematics. 3d Grade. Part 2. Moscow: Uventa, 2012. P. 96.
22. Shearing H., McCullum S. ChatGPT: Can students pass using AI tools at university? 2023. URL: https://www.bbc.co.uk/news/education-65316283
23. Shwartz V., Choi Y. Do Neural Language Models Overcome Reporting Bias? 2020. URL: https://aclanthology.org/2020.coling-main.605.pdf
24. West P. et al. Symbolic Knowledge Distillation: from General Language Models to Commonsense Models. 2022. URL: arXiv.org
25. Zhang Y. et al. Alleviating the Knowledge-Language Inconsistency: A Study for Deep Commonsense Knowledge. 2021. URL: arXiv.org