How useful are AI text detection tools?

Oliver Trevisiol, Christoph Schmitt

Version: 07.05.2024

Services such as ChatGPT have created a demand for automated recognition of AI-generated texts. An article from the end of 2023 tested various products. We briefly present the results here. Overall, the detection software is only of very limited use, even if it had a very high recognition rate. This is because no verifiable evidence is provided for the result.

Note for further use

Citation:

[author/editor] Schmitt, Christoph; Trevisiol, Oliver

[title] How useful are AI text detection tools?

[date] 2024

[source] https://www.plagiatspraevention.uni-konstanz.de/lehrmaterial/leitfragen/

[access date]

On 25 December 2023, the article *Testing of detection tools for AI-generated text*[1] was published in the *International Journal for Educational Integrity*. Based on this article, we tested various detection tools using a German academic text. The services in question were:

- *https://ki.fh-wedel.de*
- *https://checkforai.com*
- *https://detectgpt.ericmitchell.ai*
- *https://contentatscale.ai/ai-content-detector*
- *https://gptzero.me*
- *https://openai-openai-detector.hf.space*
- *https://openai-openai-detector.hf.space*
- *https://platform.openai.com/ai-text-classifier*
- *https://x.writefull.com/gpt-detector*
- *https://writer.com/ai-content-detector*
- *https://www.zerogpt.com*

With the exception of the first-mentioned site, these providers are also discussed in the article.

We come to the following conclusions:

- *Some of the tools tested in the above article no longer exist. The market for detection providers is constantly and rapidly changing:*
  - *https://checkforai.com*
  - *https://detectgpt.ericmitchell.ai*
  - *https://platform.openai.com/ai-text-classifier*
- *Sometimes websites advertise dishonest writing tools with an supposedly lower probability of detection.*
- *The results differ significantly from one another. The FH Wedel website comes to different and opposite conclusions to the other providers.*
- *The tools are a black box and do not make it possible to understand the results. None of the tools provide reliable evidence for the test result. Only https://gptzero.me offers a more detailed list. The tool highlights individual sentences and sections that are likely to have been generated by AI.*
- *Some services do not accept texts in German:*
  - *https://www.compilatio.net/ia-detecteur-info*
  - *https://x.writefull.com/gpt-detector*
- *Some providers only accept a limited amount of text for checking (e.g. https://contentatscale.ai/ai-content-detector only offers 2500 characters).*
- *Apart from https://gptzero.me, the results are only addressed before the difference human/AI-generated — mixed forms are not addressed or only indirectly via the percentage of the probability of an AI origin.*

The answer to the question of whether detection tools are helpful in the assessment of suspected cases is therefore rather no. In the academic examination context, the open question remains as to what can be done with the test result of a detection tool, even if it is 100% correct? After all, it only provides an answer, but no evidence.

---

1   Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S. et al. Testing of detection tools for AI-generated text. Int J Educ Integr 19, 26 (2023). https://doi.org/10.1007/s40979-023-00146-z