



ILLUSTRATION: ZARIF FAIAZ

How to detect AI-generated text

JINAT JAHAN KHAN

With the rapid advancement of AI content generation tools such as ChatGPT, it has become increasingly tricky to determine whether a document is AI-generated or human-written. Many companies and individuals are already using AI text creation platforms to churn out relevant works of texts for a variety of professional tasks.

While the use of AI in text generations have many potential benefits across industries, like any other tool, they too may be misused in a variety of ways. To counter the misuse of AI-generated texts, there are some tools that may come handy. Keep in mind that with the rapid growth of AI systems, it is best not to expect 100% fool-proof results all the time.

Repetitive patterns and phrases

In AI-generated content, it may use the same word or the same phrase for two or three consecutive lines. If you notice any particular line or the same structure repetitively used in the particular content, it may be a sign that it is generated by AI. Even general plagiarism detection tools identify such repetitive cases, as these texts are directly copied or slightly paraphrased from different internet resources. Common words such as 'it', 'the', and 'is' tend to pop up more than often in AI-generated text.

Unusual transitions and inaccuracies

Artificial intelligence cannot think as humans do. Thus, at times, the coherence between sentences and the logical flow of human-written texts tends to be absent. AI is not able to comprehend the exact context of a creative write-up, especially

when opinions or experiences are involved. Errors in punctuation and grammar are common as well, which eventually makes the content unclear and incomprehensible. Moreover, paragraphs seem to have a disjointed structure in AI texts.

Lack of human emotions and personal experiences

One of the common traits of AI-generated texts is that it often misses the natural flow of human language. A large piece of content with a minimum level of variation can be a trait of texts written by AI. Moreover, AI writing tools are not able to display any emotion or subjective experience. They are unable to express their perspective as there is no set of rules to explain how one should feel about a thing or how to explain themselves. As such, having a flat tone or forced emotion may be a sign that it is crafted by AI. Since AI-generated text lacks such emotions, these AI generators may cover them up by using heavy and unnecessary heavy facts.

Using AI content detector tools

If the above steps don't help you as much, you can try the more 'automatic' method of detecting AI text - using AI text detector tools. There are several you can try out, from free to paid versions, all providing degrees of successful results in detecting AI-generated text in a given content.

AI text classifier

The AI classifier is released by OpenAI themselves to determine whether any text is written with AI or not. Although it does not claim to be accurate all the

time, this detector can be simple and free to check texts for a quick review. To use the AI text classifier, there is a requirement of a minimum of 1000 characters or 150-250 words. On their website, each content is labelled as 'very unlikely', 'unlikely', 'unclear', 'possibly', and 'likely' AI-generated according to the situation.

Writer.com

This AI detector is extremely simple and completely free to use. Just go to their website, paste the writing into their text box or add the URL link to analyse the whole content. It shows how much of the content is human-generated, along with a line of suggestions. For instance, if only 16% of the content is human-generated, it suggests editing the text until there is less detectable AI content. Writer.com can check up to 1500 characters at a time.

GPTZero

GPTZero, an AI-generated text analysing tool, shows the 'perplexity' and 'burstiness' scores of written content to identify whether the text is AI-generated or human-written. Here, 'perplexity' measures text randomness and 'burstiness' implies the appearance of non-common terms in any content. Human-written texts show higher perplexity scores as these are more unpredictable than AI texts. On the contrary, AI texts show lower scores. The AI-generated text also has lower burstiness scores as they have a more consistent structure. The tool is free for up to 5000 characters per document on their website. But for further access,

subscription plans need to be bought.

Copyleaks

Copyleaks can identify AI contents by GPT-2, GPT-3, GPT-4, ChatGPT, T5, Google Bard and Jasper. It is also able to detect mixed texts where human-written contents have been amended with AI texts. One can analyse content for free on their website or use their free Chrome extension to check content directly. Copyleaks is free but it will require a paid plan if one wants to use an API for scanning many documents in a short time.

Originality.ai

Originality.ai is, so far, one of the most accurate tools to detect AI-generated text. It can identify texts written by GPT-2, GPT-3, GPT-4 and ChatGPT. This is also the only paraphrase (Quillbot) detection tool in the market, with an accuracy of about 96% in most cases. Unfortunately, this tool does not have a free version. One needs to sign up on their website to access services, and the price is measured in terms of credit. One credit goes for \$0.01 per 100 words to detect AI-generated texts.

With an increase in AI text generation tools, there is a rise in AI detector tools. But they are very new in the market, and the results from such tools are still not 100% accurate. However, it is expected that the quality and accuracy will rise soon as the demand for such tools is at its peak. As long as you're aware of the latest AI updates and keep your eyes open for inconsistencies, you should be able to detect AI-generated text with some time.