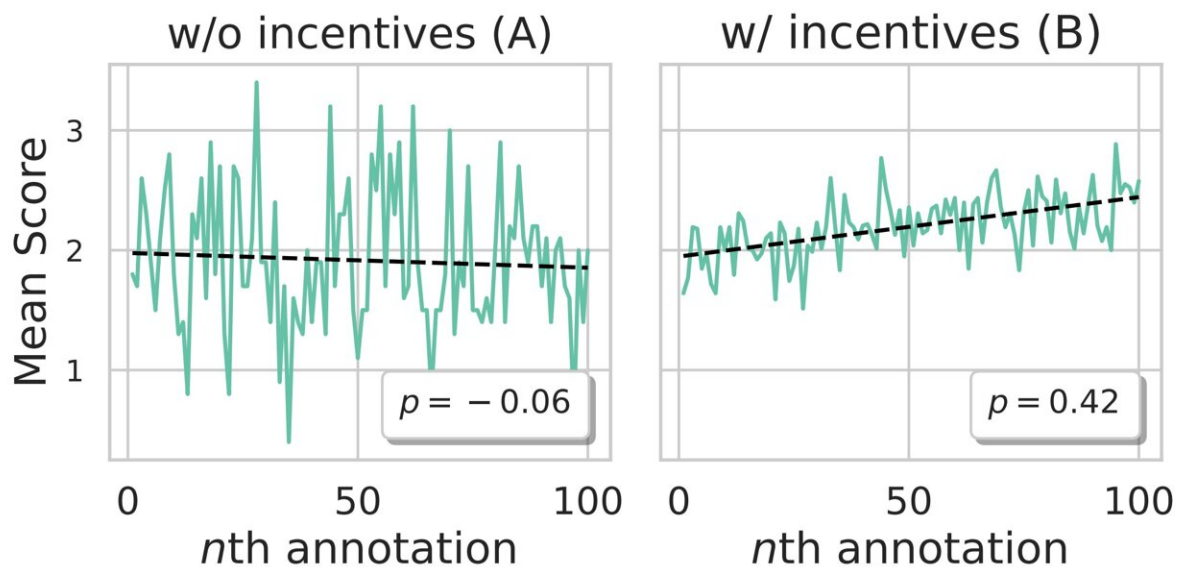# Is the text real or fake? Tell the difference with science's help

March 17 2023

Performance over time for the two player groups (§3.4). Players in Group B, who were given extra instruction and incentives, improved over time (? = 0.42) while those in Group A did not (? = −0.06). Credit: *arXiv* (2022). DOI: 10.48550/arxiv.2212.12672

AI tech tools such as ChatGPT are stealing the headlines. From writing poetry in a style reminiscent of 16th-century England to authoring academic research papers, chatbots designed for real-world services aren't going anywhere.

In fact, these not-so-bad actors will only keep getting better and better at their craft. Will we be able to keep from being fooled by these natural language processing models that are becoming a mainstay in our society?

Fool me once

A research team at the University of Pennsylvania School of Engineering and Applied Science in the United States carried out the largest-ever human study on AI detection to provide some help. They gathered data from the web-based training game [Real or Fake Text?](#) created by the university itself.

The findings are presented in a paper at an Association for the Advancement of Artificial Intelligence meeting in February. The study demonstrated that we can learn to detect human-written and machine-generated text.

"We've shown that people can train themselves to recognize machine-generated texts," stated Chris Callison-Burch, associate professor at the Department of Computer and Information Science (CIS) in a news item. "People start with a certain set of assumptions about what sort of errors a machine would make, but these assumptions aren't necessarily correct. Over time, given enough examples and explicit instruction, we can learn to pick up on the types of errors that machines are currently making."

A little training goes a long way

"AI today is surprisingly good at producing very fluent, very grammatical text," explained study co-author Liam Dugan, a Ph.D. student at CIS. "But it does make mistakes. We prove that machines make distinctive types of errors—common-sense errors, relevance errors, reasoning errors and logical errors, for example—that we can

learn how to spot."

"People are anxious about AI for valid reasons," added Prof. Callison-Burch, who led the research. "Our study gives points of evidence to allay these anxieties. Once we can harness our optimism about AI text generators, we will be able to devote attention to these tools' capacity for helping us write more imaginative, more interesting texts."

"My feeling at the moment is that these technologies are best suited for creative writing," he continued. "News stories, term papers, or legal advice are bad use cases because there's no guarantee of factuality."

Dugan sees the positive in all this: "There are exciting positive directions that you can push this technology in. People are fixated on the worrisome examples, like plagiarism and fake news, but we know now that we can be training ourselves to be better readers and writers."

So how good are you? To find out, play any of the game's four categories (short stories, news articles, recipes, presidential speeches) containing thousands of texts! And remember, while honing your recognition skills, you're also contributing to academic research.

The study is published on the *arXiv* preprint server.

 More information: Liam Dugan et al, Real or Fake Text?: Investigating Human Ability to Detect Boundaries Between Human-Written and Machine-Generated Text, *arXiv* (2022). DOI: 10.48550/arxiv.2212.12672

Provided by CORDIS