

4. Las Guerras de Tonner

ChatGPT: Oportunidad y reto para la docencia.

Vídeo 5

Marc Alier

@granludo / ICE - UPC



Escribe una biografía de isaac asimov de 5 párrafos



Isaac Asimov fue un escritor y profesor estadounidense conocido por sus obras de ciencia ficción y divulgación científica. Nació en Petrovichi, Rusia, en 1920 y emigró a Estados Unidos con su familia en 1923. Comenzó a escribir ciencia ficción a una edad temprana y publicó su primer relato en 1939 en la revista "Amazing Stories". A lo largo de su carrera, escribió más de 500 libros, incluyendo la serie de la Fundación, la serie de los Robots y varios libros de divulgación científica. También fue profesor de biología en el Boston University durante más de 30 años. Asimov murió en Nueva York en 1992, pero su legado literario y científico continúa inspirando a lectores de todo el mundo.



hazla más larga



Isaac Asimov fue un escritor y profesor estadounidense conocido por sus obras de ciencia ficción y divulgación científica. Nació en Petrovichi, Rusia, en 1920 y emigró a Estados Unidos con su familia en 1923. Comenzó a escribir ciencia ficción a una edad temprana y publicó su primer relato en 1939 en la revista

emigró a Estados Unidos con su familia en 1923. Comenzó a escribir ciencia ficción a una edad temprana y publicó su primer relato en 1939 en la revista "Amazing Stories". A lo largo de su carrera, escribió más de 500 libros, incluyendo la serie de la Fundación, la serie de los Robots y varios libros de divulgación científica. También fue profesor de biología en el Boston University durante más de 30 años.

Asimov es considerado uno de los escritores de ciencia ficción más importantes de la historia. Su serie de la Fundación, compuesta por siete libros, es considerada una de las mejores obras de ciencia ficción de todos los tiempos. En ella, Asimov imagina un futuro en el que un grupo de científicos y académicos se esfuerzan por preservar la cultura humana en el espacio. También escribió la serie de los Robots, en la que se desarrolla una historia en la que los robots están evolucionando y tienen conciencia, y la relación entre humanos y robots.

Asimov también es conocido por sus libros de divulgación científica, en los que explica temas complejos de manera accesible para el público en general. Sus libros abarcan una variedad de temas, desde la historia de la ciencia hasta la física y la biología.

AI Content Detector

Some search engines **penalize** your page **ranking** if they detect content that **reads like it was entirely produced by AI**. Evaluate your text with this detector

and decide if you want to make adjustments before you publish your content.

Want to use this feature in the Writer application? It's **coming soon**.

Looking to detect AI-generated content at scale? Writer [Team](#) customers get **500k words per month** through the [API](#). [Get in touch](#) to find out about higher limits.

Add a URL

https://

Add some text

Paste text or write here

%

DETECTION SCORE

Analyze text

AI Content Detector

Some search engines **penalize** your page **ranking** if they detect content that reads like it was entirely produced by AI. Evaluate your text with this detector and decide if you want to make adjustments before you publish your content.

Want to use this feature in the Writer application? It's **coming soon**.

Looking to detect AI-generated content at scale? Writer [Team](#) customers get **500k words per month** through the [API](#). [Get in touch](#) to find out about higher limits.

Add a URL

<https://>

Add some text

Ciencia ficción es una forma temprana y popular su primer relato en 1939 en la revista "Amazing Stories". A lo largo de su carrera, escribió más de 500 libros, incluyendo la serie de la Fundación, la serie de los Robots y varios libros de divulgación científica. También fue profesor de biología en el Boston University durante más de 30 años.

>

Asimov es considerado uno de los escritores de ciencia ficción más importantes de la historia. Su serie de la Fundación, compuesta por siete libros, es considerada una de las mejores obras de ciencia ficción de todos los tiempos. En ella, Asimov imagina un futuro en el que un grupo de científicos y académicos se esfuerzan por preservar la cultura humana en el espacio. También escribió la serie de los Robots, en la que se desarrolla una historia en la que los robots están evolucionando y tienen conciencia, y la relación entre humanos y robots.

>

Asimov también es conocido por sus libros de divulgación científica, en los que explica temas complejos de manera accesible para el público en general. Sus libros abarcan una variedad de temas, desde la historia de la ciencia hasta la física y la biología.

>

100%
HUMAN-GENERATED CONTENT

Fantastic!

>"Lost in the Flood" es una canción escrita por Bruce Springsteen que apareció en su álbum debut "Greetings from Asbury Park, N.J." de 1973. La canción narra la historia de un joven llamado Billy, que vive en una ciudad industrial en ruinas y se enfrenta a la desesperación y la soledad. El tema central de la canción es la pérdida y la desesperación, tanto en términos personales como en términos de la sociedad en general.

En la canción, Springsteen describe a Billy como un joven que se siente atrapado en una vida que no deseaba y que está luchando por encontrar su lugar en el mundo. El "flood" (inundación) se refiere tanto a las condiciones físicas de la ciudad en ruinas como a las emociones de Billy, que se siente abrumado y perdido.

La canción también contiene críticas sociales y políticas, como la descripción de la ciudad como un "infierno industrial" y las referencias a la guerra de Vietnam y la opresión del gobierno. Estas críticas sugieren que Billy no es el único que se siente perdido y oprimido, sino que esta sensación es común en la sociedad en general.

En general, "Lost in the Flood" es una canción que refleja la desesperación y la soledad del individuo en una sociedad en crisis, y cómo esta sensación se ve reflejada en la vida cotidiana de las personas. La canción es considerada una de las primeras obras maestras de Springsteen y sigue siendo una de sus canciones más aclamadas por la crítica.

File submissions

1 task submitted



lost .docx Ouriginal:0%

24 January 2023, 4:35 PM



API RESEARCH BLOG ABOUT

New AI classifier for indicating AI-written text

We're launching a classifier trained to distinguish between
AI-written and human-written text.

January 31, 2023



We've trained a classifier to distinguish between text written by a human and text written by AIs from a variety of providers. While it is impossible to reliably detect all AI-written text, we believe good classifiers can inform mitigations for false claims that AI-generated text was written by a human: for example, running [automated misinformation campaigns](#), using AI tools for academic dishonesty, and positioning an AI chatbot as a human.

Our classifier is not fully reliable. In our evaluations on a “challenge set” of English texts, our classifier correctly identifies 26% of AI-written text (true positives) as “likely AI-written,” while incorrectly labeling human-written text as AI-written 9% of the time (false positives). Our classifier’s reliability typically improves as the length of the input text increases. Compared to our previously released classifier, this new classifier is significantly more reliable on text from more recent AI systems.

We’re making this classifier publicly available to get feedback on whether imperfect tools like this one are useful. Our work on the detection of AI-generated text will continue, and we hope to share improved methods in the future.

Try our free work-in-progress classifier yourself:

[TRY THE CLASSIFIER ↗](#)

Limitations

Our classifier has a number of important limitations. **It should not be used as a primary decision-making tool**, but instead as a complement to other methods of determining the source of a piece of text.

1. The classifier is very unreliable on short texts (below 1,000 characters). Even longer texts are sometimes incorrectly labeled by the classifier.
2. Sometimes human-written text will be incorrectly but confidently labeled as AI-written by our classifier.
3. We recommend using the classifier only for English text. It performs significantly worse in other languages and it is unreliable on code.
4. Text that is very predictable cannot be reliably identified. For example, it is impossible to predict whether a list of the first 1,000 prime numbers was written by AI or humans, because the correct answer is always the same.
5. AI-written text can be edited to evade the classifier. Classifiers like ours can be updated and retrained based on successful attacks, but it is unclear whether detection has an advantage in the long-term.
6. Classifiers based on neural networks are known to be poorly calibrated outside of their training data. For inputs that are very different from text in our training set, the classifier is sometimes extremely confident in a wrong prediction.

<https://arxiv.org/pdf/2303.11156.pdf>

Can AI-Generated Text be Reliably Detected?

Vinu Sankar Sadasivan
vinu@umd.edu

Aounon Kumar
aounon@umd.edu

Sriram Balasubramanian
sriramb@umd.edu

Wenxiao Wang
wxw@umd.edu

Soheil Feizi
sfeizi@umd.edu

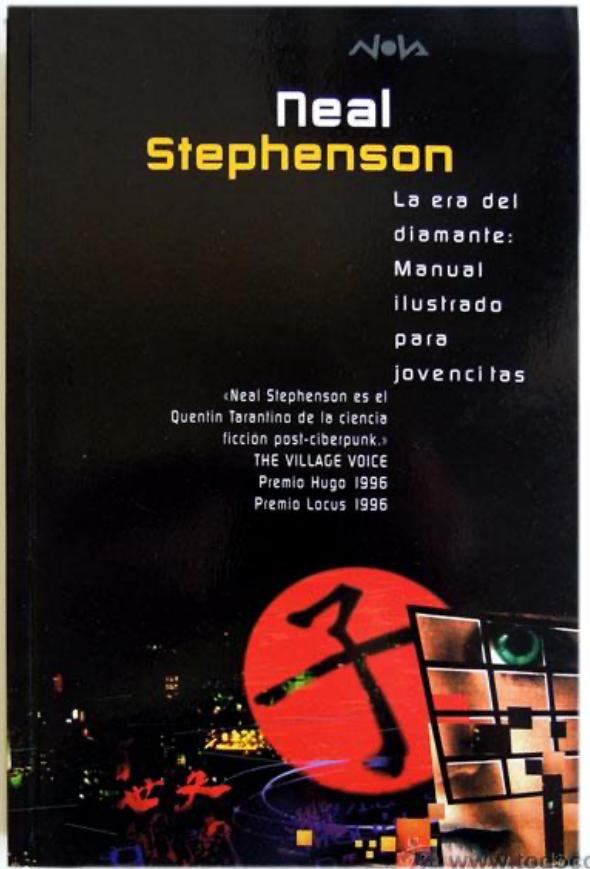
Department of Computer Science
University of Maryland

Abstract

The rapid progress of Large Language Models (LLMs) has made them capable of performing astonishingly well on various tasks including document completion and question answering. The unregulated use of these models, however, can potentially lead to malicious consequences such as plagiarism, generating fake news, spamming, etc. Therefore, reliable detection of AI-generated text can be critical to ensure the responsible use of LLMs. Recent works attempt to tackle this problem either using certain model signatures present in the generated text outputs or by applying watermarking techniques that imprint specific patterns onto them. In this paper, both empirically and theoretically, we show that these detectors are not reliable in practical scenarios. Empirically, we show that *paraphrasing attacks*, where a light paraphraser is applied on top of the generative text model, can break a whole range of detectors, including the ones using the watermarking schemes as well as neural network-based detectors and zero-shot classifiers. We then provide a theoretical *impossibility result* indicating that for a sufficiently good language model, even the best-possible detector can only perform marginally better than a random classifier. Finally, we show that even LLMs protected by watermarking schemes can be vulnerable against spoofing attacks where *adversarial humans* can infer hidden watermarking signatures and add them to their generated text to be detected as text generated by the LLMs, potentially causing reputational damages to their developers. We believe these results can open an honest conversation in the community regarding the ethical and reliable use of AI-generated text.

Abstract

The rapid progress of Large Language Models (LLMs) has made them capable of performing astonishingly well on various tasks including document completion and question answering. The unregulated use of these models, however, can potentially lead to malicious consequences such as plagiarism, generating fake news, spamming, etc. Therefore, reliable detection of AI-generated text can be critical to ensure the responsible use of LLMs. Recent works attempt to tackle this problem either using certain model signatures present in the generated text outputs or by applying watermarking techniques that imprint specific patterns onto them. In this paper, both empirically and theoretically, we show that these detectors are not reliable in practical scenarios. Empirically, we show that *paraphrasing attacks*, where a light paraphraser is applied on top of the generative text model, can break a whole range of detectors, including the ones using the watermarking schemes as well as neural network-based detectors and zero-shot classifiers. We then provide a theoretical *impossibility result* indicating that for a sufficiently good language model, even the best-possible detector can only perform marginally better than a random classifier. Finally, we show that even LLMs protected by watermarking schemes can be vulnerable against spoofing attacks where *adversarial humans* can infer hidden watermarking signatures and add them to their generated text to be detected as text generated by the LLMs, potentially causing reputational damages to their developers. We believe these results can open an honest conversation in the community regarding the ethical and reliable use of AI-generated text.



www.todocolección.net



[Home](#) / [2023](#) / [January](#) / [31](#) /

arXiv announces new policy on ChatGPT and similar tools

By [ame5](#) · January 31, 2023 · [arXiv updates](#)

The recent release of AI technology that generates new text has raised serious questions among the research community. For one, “Can ChatGPT be named an author of a research paper?”

The resounding answer from arXiv leaders and advisors is, “No.” A computer program cannot, for example, take responsibility for the contents of a paper. Nor can it agree to arXiv’s terms and conditions. [Other organizations agree.](#)

To address this issue, arXiv has adopted a [new policy](#) for authors regarding the use of generative AI language tools.

The official policy is:

Subscribe

[Log in](#)
[Entries RSS](#)
[Comments RSS](#)
[CU Blog Service](#)

arXiv policy for authors' use of generative AI language tools

January, 31 2023

arXiv recognizes that authors of scientific works use a variety of tools to do the science on which they report, and to prepare the report itself, from simple ones to very sophisticated ones. Community opinion on the appropriateness of such tools may be varied and evolving; AI powered language tools have in particular led to significant debate. We note that tools may generate useful and helpful results, but also errors or misleading results; therefore, knowing which tools were used is relevant to evaluating and interpreting scientific works.

In view of this, we

1. continue to require authors to report in their work any significant use of sophisticated tools, such as instruments and software; we now include in particular text-to-text generative AI among those that should be reported consistent with subject standards for methodology.
2. remind all colleagues that by signing their name as an author of a paper, they each individually take full responsibility for all its contents, irrespective of how the contents were generated. If generative AI language tools generate inappropriate language, plagiarized content, errors, mistakes, incorrect references, or misleading content, and that output is included in scientific works, it is the responsibility of the author(s).
3. generative AI language tools should not be listed as an author; instead authors should refer to (1).

