



Turnitin AI: beware of its false positive in using for similarity and AI source checking for an article

Rujittika Mungmunpantipantip¹, Viroj Wiwanitkit²

¹Private Academic Consultant, Bangkok Thailand.

²Dr DY Patil Vidhyapeeth, Pune, India.

Abstract

The emerging use of computational AI tool leads to convenience and advantage. In academic, the use can be in several ways. Turnitin is an AI tool for similarity checking and tracing for the AI generated text. This tool is widely used AI application in academic and it is usually for checking of similarity of student's work or submitted article to the journal. The result from Turnitin check is usually referred information for accusation and management of unethical academic problem. However, as an AI tool, it can have false positive and hallucination. If user does not recognize or has no critical appraisal on the result, there might be an unfair unwanted complication. For example, it might lead to unfair decision on a submitted article or student's work. In the worst case, it can lead to an unfair and unethical management on an innocent author due to the accused plagiarism based on false positive result given by Turnitin.

Keywords: AI, Turnitin, false, positive, similarity, source, checking, article.

Introduction

The growing application of computational AI tools brings benefits and convenience. Academics can employ it in a variety of ways. Turnitin is an AI-generated text similarity checking and tracing tool. This commonly used AI tool in academia is used to compare student work or articles that have been submitted to journals for similarity [1]. The results of a Turnitin check typically lead to information that can be used to regulate and file charges of unethical academic behavior. As an AI tool, it is susceptible to false positives and hallucinations, though [2 - 3]. If the user is unaware of the outcome or does not critically evaluate it, an unjust and unwelcome difficulty may arise. It could result in an unjust decision on an article that a student submits, for instance. Here, a case study is demonstrated and this case can urge our attention to have a critical appraisal when the interpretation of the results given by Turnitin check is done.

Case study

A journal editor notifies to a submitting author of an article that there is high similarity score, up to 50 %, and high AI-generated content in the submitted article and point out that this is plagiarism problem and reject the article. The Turnitin checking result is used as supporting data for accusation (Figure 1). The submitting author replies to the journal and the false positive results from Turnitin AI checking is highlighted.

A. Similarity index

The similarity index should highlight common separated words rather than phrases and counts, which is a concern that could lead to false positive similarity check. Also, it indicates the source of the plagiarism as "student paper" "Submitted to Institute of International Studies," so what is it? that is the student that was plagiarized in this computational AI plagiarism checking report? How can we get to it (there is no specific website, journal name, book name, or thesis database for this quote)? or is it merely a "submitted work"? If it is a submitted work that has not been published, how can the AI plagiarism/similarity detection program identify it as similar? There is no institute titled "Institute of International Study".

Also, the author of the article presents the results derived from a different similarity detection tool (Figure 2) and it can show that there is no plagiarism. This might highlight the fact that Turnitin might generate a hallucinated result and false positive can lead to the problem in unfair decision making of the editor of the journal.

B. AI generated content

For AI application, the author of the article claims that he uses a language modification computational tool to improve language, therefore this is conceivable because there are AI-derived words and sentences. There is no need for generative language models. And the final work is devoid of plagiarism (Figure 2). It is critical to discern between AI use and plagiarism. There are various types of AI that can be used in academic work, including generative models, language edition models, and similarity check models. Any model can be used, as long as it is done ethically. Using a generative model, such as a chatbot, may be a major worry if the author simply copies and pastes the text created by the model. This does, in fact, indicate an explanation of the results of Turnitin's AI-generated content checks.

21 Dear Editor

22 The COVID-19 immunization is a necessity globally. A main area of consideration is

23 coverage in remote and hard-to-reach zone. A recent report from an Indian distant site

24 published in the journal is a good illustration [1]. But if the setting is in a distant zone with

25 local conflict, it might get trickier. There are various of explanations why many people in

26 conflict areas, like Yemen and Syria, choose not to get the SARS-Co-V2 vaccine. These

27 include the area's deteriorating living conditions and health as a problem due to local war,

28 which has led a problem of ignorance and backwardness that has brought the need of

29 awareness and education campaign about the importance, effectiveness, and safety of getting

30 the immunization [2-4]. The continuous violence and devastation in conflict zones like Syria

31 and Yemen have caused the locals with a lot of difficulties, such as limited access to basic

32 necessities and a low education.

33 Many people in conflict areas do not accept to receive the SARS-Co-V2 vaccine for various

34 reasons [2-4]. There is a lot of ignorance and disinformation out there about the safety and

35 significance of receiving the SARS-Co-V2 vaccine due to the unstable climate. Due to this, a

36 lot of people in these areas are reluctant to get the immunization, which superimposes the

37 already serious humanitarian and health problems. It is typically challenging to perform a

38 study in conflict areas, and data from both vaccinated and unvaccinated individuals might be

39 collected to determine the mild and serious side effects of the SARS-Co-V2 vaccine.

40 Findings raised that concerns about vaccine adverse effects were the most common reason for

41 not receiving the immunization. The study's dependence on self-reported data is a common

42 weakness. Furthermore, the study do usually not specify which social media sites were

43 applied to disseminate the survey, which might have affected the work's generalizability. To

44 improve the statistical power of the study, a larger sample size can also be advantageous.

45 However, if the conflict, such as a local war, still persists, it may overly speculative.

mj_28_24_2

ORIGINALITY REPORT

60%
SIMILARITY INDEX

0%
INTERNET SOURCES

0%
PUBLICATIONS

60%
STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to Institute of International Studies
Student Paper

60%

Exclude quotes On
Exclude bibliography On

Exclude matches < 1%

A. Similarity check



Page 2 of 6 - AI Writing Overview

Submission ID trmid:1:295034025

How much of this submission has been generated by AI?

50%

of qualifying text in this submission has been determined to be generated by AI.

Caution: Percentage may not indicate academic misconduct. Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Frequently Asked Questions

What does the percentage mean?

The percentage shown in the AI writing detection indicator and in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was generated by AI.

Our testing has found that there is a higher incidence of false positives when the percentage is less than 20. In order to reduce the likelihood of misinterpretation, the AI indicator will display an asterisk for percentages less than 20 to call attention to the fact that the score is less reliable.

However, the final decision on whether any misconduct has occurred rests with the reviewer/instructor. They should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in greater detail according to their school's policies.

How does Turnitin's indicator address false positives?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be AI-generated will be highlighted blue on the submission text.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.

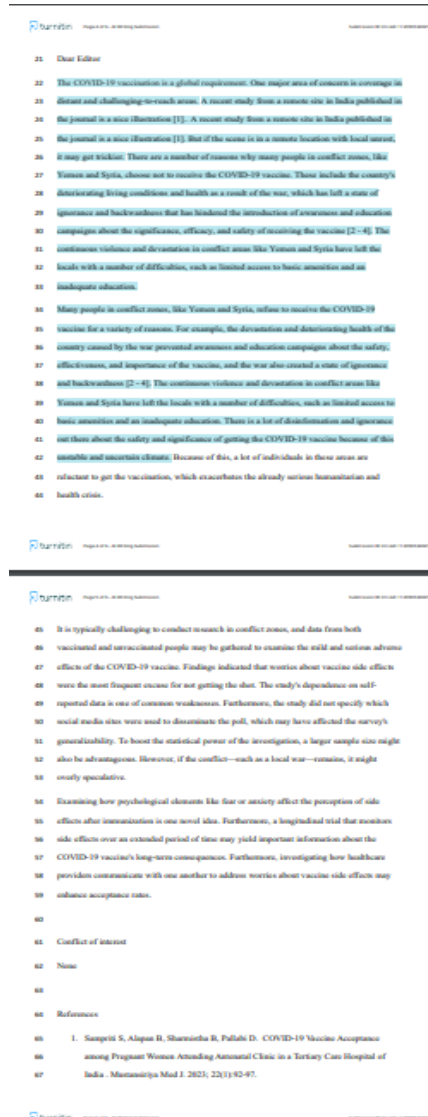
What does 'qualifying text' mean?

Sometimes false positives (incorrectly flagging human-written text as AI-generated), can include lists without a lot of structural variation, text that literally repeats itself, or text that has been paraphrased without developing new ideas. If our indicator shows a higher amount of AI writing in such text, we advise you to take that into consideration when looking at the percentage indicated.

In a longer document with a mix of authentic writing and AI generated text, it can be difficult to exactly determine where the AI writing begins and original writing ends, but our model should give you a reliable guide to start conversations with the submitting student.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (it may misidentify both human and AI-generated text) so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.



B. AI generated text check
Figure 1: Turnitin checking results

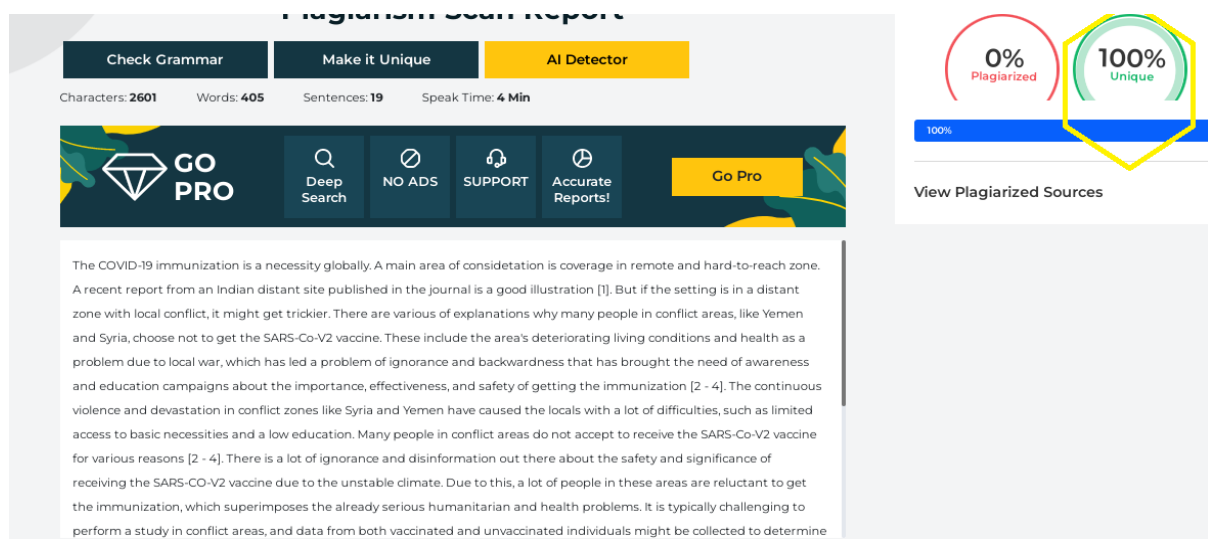


Figure 2: The similarity check by other computational tool

Discussion

The increasing use of computational AI techniques is advantageous and practical. It can be used in many different ways by academics. Turnitin is a text similarity checking and tracing program created by AI. This widely used AI tool in academia compares student work or submitted articles to journals to see if they are comparable [1]. Typically, the findings of a Turnitin check provide information that can be utilized to control and report instances of unethical academic behavior. However, it is prone to false positives and hallucinations as an AI tool [2–3]. An unfair and unwanted problem could occur if the user is ignorant of the result or does not assess it critically.

Turnitin is an AI-based tool, so it may experience accuracy and reliability issues similar to those of other AI applications. When using AI source checking or comparable checks, the results may be hallucinatory, and if a user with little experience accepts Turnitin results without question, there may be issues. These days, a lot of journals employ Turnitin as a tool, which may lead to false positive results from Turnitin, which can lead to accusations of plagiarism and harsh treatment of innocent authors of submitted work. This reality needs to be acknowledged, and like with other AI tools, attention and critical evaluation must be used when interpreting the data that Turnitin provides.

Conclusion

As an AI based tool, Turnitin can have similar problem on reliability and accuracy similar to other AI tools. Using in similar check or AI source checking might give hallucinated result and if the low experience user believes the result from Turnitin without critical appraisal it can cause problem. Many journals now use Turnitin as tool and might be the cause of accusation of plagiarism and unfair management to innocent authors of a submitted article if there is a false positive result caused by Turnitin. This fact must be recognized and it has to be careful and have critical appraisal on interpretation of results provided by Turnitin, similar to other AI tool.

References

- Meo SA, Talha M. Share Turnitin: Is it a text matching or plagiarism detection tool? Saudi J Anaesth. 2019 Apr;13(Suppl 1):S48-S51.
- What to do when you're accused of AI cheating. Available online at <https://www.washingtonpost.com/technology/2023/08/14/prove-false-positive-ai-detection-turnitin-gptzero/>
- False Positives for AI on Turnitin: Help for Teachers & Accused Students. Available online at <https://www.youtube.com/watch?v=a-Xj5vGmkis>