Dataset Creation and Analysis for Evaluating AI Text Detection in Student Essays

Teachers across the world are worried about students using Large Language Models, such as ChatGPT, to cheat on assignments. Such teachers are employing tools to detect Artificial Intelligence (AI) writing. We are collecting essays written by students to determine the accuracy of AI writing detectors, such as Turnitin AI Writing Detection. We will not ask for your name, but please remove your name from any attached essays.

Contact information: Please email nsorkin1@terpmail.umd.edu with any questions.

CONSENT TO PARTICIPATE

Project Title

Dataset Creation and Analysis for Evaluating AI Text Detection in Student Essays

Purpose of the Study

This research is being conducted by Tom Goldstein at the University of Maryland, College Park. We are inviting you to participate in this research project because you have written essays during your academic journey. The purpose of this research project is to determine the accuracy of AI writing detectors, such as Turnitin AI Writing Detection.

Procedures

The procedures involve submitting your essays, which will be used to test the effectiveness of AI writing detection tools. Please ensure you remove your name and any identifying information from the essays prior to submission. This survey should take under 15 minutes to complete. Specifically, the essay submission process is estimated to take about 10 minutes. It can be skipped if you are only giving your perspective. Following the essay submission, you will be asked to complete a survey which will take approximately 5 minutes. The survey will gauge your perspective on AI writing detection tools and false accusations in academic settings.

Potential Risks and Discomforts

There may be minimal risks from participating in this research study, mainly associated with data privacy concerns.

Potential Benefits

There are no direct benefits from participating in this research. However, we hope that, in the future, educators and institutions might benefit from this study through improved understanding of AI-written content detection.

Confidentiality

Any potential loss of confidentiality will be minimized by storing data in a secure location, such as encrypted servers and ensuring that all submitted essays are anonymized. If we write a report or article about this research project, your identity will be protected to the

maximum extent possible. Your information may be shared with representatives of the University of Maryland, College Park or governmental authorities only if you or someone else is in danger or if we are required to do so by law.

To ensure you have a record of the consent form, we advise you to download or print a copy (CTRL+P) of this page for your personal records before submitting your responses.

Right to Withdraw and Questions
Your participation in this research is completely voluntary. You may choose not to take part at all. If you decide to participate in this research, you may stop participating at any time. If you decide not to participate in this study or if you stop participating at any time, you will not be penalized or lose any benefits to which you otherwise qualify.
If you decide to stop taking part in the study, if you have questions, concerns, or complaints, or if you need to report an injury related to the research, please contact the investigator:

Tom Goldstein
8125 Paint Branch Dr IRB 4212, College Park, MD 20742
tomg@cs.umd.edu
(301) 405-0053

Participant Rights
If you have questions about your rights as a research participant or wish to report a research-related injury, please contact:

University of Maryland College Park
Institutional Review Board Office
1204 Marie Mount Hall
College Park, Maryland, 20742
E-mail: irb@umd.edu
Telephone: 301-405-0678

For more information regarding participant rights, please visit:
https://research.umd.edu/research-resources/research-compliance/institutional-review-board-irb/research-participants

This research has been reviewed according to the University of Maryland, College Park IRB procedures for research involving human subjects.

For your reference, the IRBNet package number for this study is 2105015-1.

* Indicates required question

1. **Statement of Consent**   *

By checking the box below, you acknowledge that you are at least 18 years of age; you have read this consent form or have had it read to you; your questions have been answered to your satisfaction, and you voluntarily agree to participate in this research study.

Mark only one oval.

◯ I consent   *Skip to question 2*

◯ I do not consent

Essay Release Permission

2. Essay Release Permission *

I give permission for all essays attached below to be held by Tom Goldstein's group and partner university faculty to be utilized for any research by his group. These essays will never be publicly released.

Mark only one oval.

◯ Yes, I consent   *Skip to question 3*

◯ No, I do not consent (You will jump to the AI usage punishment section)
     Skip to question 6

Essay Collection

3. Is English your first language? *

Mark only one oval.

◯ Yes

◯ No

4. Essays written **without** the use of AI. ChatGPT was publicly released November 30, 2022. (Optional)

    Upload all essays that were written without the use of AI tools like ChatGPT. Remove all personally identifiable information.

    Files submitted:

5. **Essays Assisted by AI Tools** (Optional): AI tools, like ChatGPT which was publicly released on November 30, 2022, can assist in various stages of the essay writing process. For the purposes of this study, an essay is considered "assisted by AI" if:
   Direct Integration: You've directly pasted content generated by an AI tool into your essay.
   Minimal Edits: You've made minor edits (e.g., changing a few words here and there) to content that was originally generated by an AI tool.

   Example 1 of What Doesn't Count: If you used an AI tool to summarize content (like a video script) and then, after reading that summary, wrote your essay entirely in your own words without directly including or minimally editing the AI's output, your essay would not be considered "assisted by AI."

   Example 2 of What Doesn't Count: If you used an AI tool like ChatGPT to help brainstorm topics for your essay, but the actual content and structure of the essay were created solely by you without directly integrating or minimally editing any AI-generated text, then your essay would not be considered "assisted by AI."

   If your essay meets the "assisted by AI" criteria, please upload it. Ensure that all personally identifiable information is removed from the document.

   Upload all essays that were written with the use of AI tools like ChatGPT. Remove all personally identifiable information.

Files submitted:

Survey on AI Usage Punishments

6. If a student is caught copying their answers on a writing assignment from another *
   student, what should the punishment be?

   Mark only one oval.

   ◯ Automatic zero on assignment

   ◯ Automatic XF for the course (Fail with academic dishonesty)

   ◯ Conversation with professor followed by further scrutiny

   ◯ Referral to honor council or some other university level disciplinary body

   ◯ No punishment

   ◯ Other: _____

7. Suppose a student is using a Large Language Model (LLM), such as ChatGPT, *
   when they are not allowed to use an LLM on an assignment. What should the
   consequence be when a student gets caught using an LLM?

   Mark only one oval.

   ◯ Automatic zero on assignment

   ◯ Automatic XF for the course (Fail with academic dishonesty)

   ◯ Conversation with professor followed by further scrutiny

   ◯ Referral to honor council or some other university level disciplinary body

   ◯ No punishment

   ◯ Other: _____

Demographics

8. Gender Identity (Optional)

   Mark only one oval.

   ◯ Woman

   ◯ Man

   ◯ Transgender

   ◯ Non-binary/non-conforming

   ◯ Prefer not to say

9. Choose one or more races that you consider yourself to be (Optional)

   Check all that apply.

   ☐ White or Caucasian
   ☐ Black or African American
   ☐ American Indian/Native American or Alaska Native
   ☐ Asian
   ☐ Native Hawaiian or Other Pacific Islander
   ☐ Other
   ☐ Prefer not to say

10. Role *

    Mark only one oval.

    ◯ Professor   *Skip to question 14*

    ◯ Parent   *Skip to question 14*

    ◯ Undergraduate student   *Skip to question 11*

    ◯ Graduate student   *Skip to question 13*

    ◯ Other: _____

Undergraduate Student Specific Questions

11. Class Standing *

Mark only one oval.

( ) Freshman (between 0 and 29 credits)

( ) Sophomore (between 30 and 59 credits)

( ) Junior (between 60 and 89 credits)

( ) Senior (at least 90 credits)

( ) Graduated

12. What is/was your approximate undergraduate GPA on a 4.0 scale? *

Mark only one oval.

( ) Below 2.0

( ) 2.0 - 2.5

( ) 2.6 - 3.0

( ) 3.1 - 3.5

( ) 3.6 - 4.0

( ) No GPA yet (Freshman)

( ) Prefer not to say

( ) Other: _____

Skip to question 14

Graduate Student Specific Questions

13. What is/was your approximate undergraduate GPA on a 4.0 scale? *

Mark only one oval.

◯ Below 2.0

◯ 2.0 - 2.5

◯ 2.6 - 3.0

◯ 3.1 - 3.5

◯ 3.6 - 4.0

◯ No GPA yet (Freshmen)

◯ Prefer not to say

◯ Other: _____

Skip to question 14

Comments

14. Comments (Optional)

If you have any comments to add that you would like to share, such as a personal experience with the honor council, or class policies related to AI detectors, please add it below.

_____

_____

_____

_____

Google Forms