Tempted to cheat on a written exam? Artificial intelligence is 90% certain to nab you

May 30 2019

Credit: CC0 Public Domain

Combining big data with artificial intelligence has allowed University of Copenhagen researchers to determine whether you wrote your

assignment or whether a ghostwriter penned it for you—with nearly 90 percent accuracy.

Several studies have shown that cheating on assignments is widespread and becoming increasingly prevalent among high [school](#) students. At the University of Copenhagen's Department of Computer Science, efforts to detect cheating on assignments through writing analysis by way of [artificial intelligence](#) have been underway for a few years. Now, based on analyses of 130,000 written Danish assignments, scientists can, with nearly 90 percent accuracy, detect whether a [student](#) has written an assignment on their own or if a ghostwriter composed it.

Danish high schools currently use the Lectio platform to check if a student has handed in plagiarized work with passages copied directly from a previously submitted assignment. High schools have a harder time discovering if a student has enlisted someone else to write the assignment for them, something that happens to a more or less systematized degree via online services. The case of the SRP, a major written assignment in the final year of Danish high school, is particularly telling. Because the assignment counts for double, students have gone as far as tendering out their [writing assignments](#) on the Danish classified website, Den Blå Avis.

"The problem today is that if someone is hired to write an assignment, Lectio won't spot it. Our program identifies discrepancies in writing styles by comparing recently submitted writing against a student's previously submitted work. Among other variables, the program looks at: word length, sentence structure and how words are used. For instance, whether 'for example' is written as 'ex.' or 'e.g.,'" explains Ph.D. student Stephan Lorenzen of the Department of Computer Science. He, along with the rest of the DIKU-DABAI research group, recently presented their findings at a major European AI conference.

Prior to setting the trap, an ethical debate

The program, Ghostwriter, is built around [machine learning](#) and [neural networks](#)—branches of artificial intelligence that are particularly useful for recognizing patterns in images and texts. MaCom, the company that provides Lectio to Danish high schools, has made a dataset of 130,000 written assignments from 10,000 [high school students](#) available to Ghostwriter project researchers at the Department of Computer Science. For now, it is still a research project.

Stephan Lorenzen doesn't think that it is unrealistic for the program to find its way into high schools in the not too distant future, as schools must constantly stay apace with technological developments to address authorship verification. "I think that it is realistic to expect that high schools will begin using it at some point. But before they do, there needs to be an ethical discussion of how the technology ought to be applied. Any result delivered by the program should never stand on its own, but serve to support and substantiate a suspicion of cheating," believes Lorenzen.

Police and fake news

Ghostwriter's technological foundation can be applied elsewhere in society. For example, the program could be used in police work to supplement forged document analysis, a task carried out by forensic document examiners and others.

"It would be fun to collaborate with the police, who currently deploy forensic document examiners to look for qualitative similarities and differences between the texts they are comparing. We can look at large amounts of data and find patterns. I imagine that this combination would benefit police work," says Lorenzen, who emphasizes that ethical

discussions are needed here as well.

The artificial intelligence used by researchers at the Department of Computer Science to detect cheating on assignments has a wide range of applications. It has already been used to analyze Twitter tweets to determine whether they were composed by actual users or penned by paid imposters or robots.

Facts:

- The ghostwriter program uses what is known as a Siamese neural network to distinguish the writing styles of two texts. The network is trained on large amounts of data to learn from representations of writing styles, which are then compared.
- When a student submits an assignment, the network compares it against their previous assignments. For each previous assignment, the network provides a percentage score for writing style similarity against the new assignment.
- In the end, a weighted average of these scores is calculated using a calculation that also takes other factors, such as delivery time, into account. This final score is presented as a percentage and indicates the similarity between the new assignment and the student's writing style.
- The research group behind the result is the Danish Center for Big Data Analytics-driven Innovation (DIKU-DABAI). The group is headed by Professor Stephen Alstrup.

More information: Detecting Ghostwriters in High Schools. www.science.ku.dk/presse/nyhed … _in_High_Schools.pdf

Provided by University of Copenhagen

Citation: Tempted to cheat on a written exam? Artificial intelligence is 90% certain to nab you (2019, May 30) retrieved 24 May 2024 from https://phys.org/news/2019-05-written-exam-artificial-intelligence-nab.html