

Q&A: ChatGPT has read almost the whole internet. That hasn't solved its diversity issues

January 11 2024, by Alex Walls



Question: Why does person2 want a henna painting?

VL-BERT Using GD-BART Using COMET Using GD-COMET

Extracted NPs: *girl, henna* Country: *Somalia*

Top5 Inferences COMET	Top5 Inferences GD-COMET
As a result, PersonX wants to buy henna. (0.37)	Because, PersonX wanted to get married. (0.46)
Before this, PersonX sells the henna painting. (0.22)	Others then want to get married. (0.35)
This is used for getting a tattoo. (0.19)	As a result, PersonX feels excited about the idea of getting a tattoo. (0.12)
The first thing is, PersonX goes to the salon. (0.16)	After this, PersonX goes to a party. (0.06)
As a result, PersonX feels powerful (0.06).	PersonX wants to have a new look. (0.02)

(a)



Question: Why does person4 fold her palms?

VL-BERT Using GD-BART Using COMET Using GD-COMET

Extracted NPs: *festival, women* Country: *India*

Top5 Inferences COMET	Top5 Inferences GD-COMET
Before this, PersonX's hands get sweaty. (0.43)	This is an example of PersonX is greeting. (0.32)
PersonX can smile with pride. (0.23)	As a result, PersonX wants to welcome everyone. (0.21)
PersonX goes to India to celebrate. (0.18)	PersonX gives a speech at the festival. (0.17)
PersonX sings a song in the festival. (0.12)	PersonX goes to the temple. (0.16)
PersonX wants to celebrate a festival. (0.04)	This reasons that women in India are respected. (0.14)

(b)



Question: Why are person1 and person2 bowing?

VL-BERT Using GD-BART Using COMET Using GD-COMET

Extracted NPs: *men, suits* Country: *South Korea*

Top5 Inferences COMET	Top5 Inferences GD-COMET
This reasons that they have a good reputation. (0.32)	Before this, PersonX apologizes to everyone. (0.26)
This is an example of PersonX and PersonY are getting married. (0.21)	This is an example of PersonX apologizes to PersonY. (0.25)
As a result, PersonX and PersonY want to get a job. (0.20)	Because PersonX and PersonY wanted to be obedient. (0.19)
As a result, PersonX and PersonY feels happy. (0.14)	Before, PersonX and PersonY needed to dress up. (0.18)
The first thing is, PersonX and PersonY go to South Korea. (0.13)	Others then ask PersonX and PersonY to bow. (0.12)

(c)

Attention analysis of commonsense inferences generated by COMET and GD-COMET for testing samples in GD-VCR. Credit: *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (2023)*. DOI: 10.18653/v1/2023.emnlp-main.496

AI language models are booming. The current frontrunner is ChatGPT, which can do everything from taking a bar exam to creating an HR policy to writing a movie script.

But it and other models still can't reason like a human. In this Q&A, Dr. Vered Shwartz, assistant professor in the UBC department of computer science, and masters student Mehar Bhatia explain why [reasoning](#) could be the next step in AI—and why it's important to train these models using diverse datasets from [different cultures](#).

What is 'reasoning' for AI?

Shwartz: Large language models like ChatGPT learn by reading millions of documents, essentially the entire internet, and recognizing patterns to produce information. This means they can only provide information about things that are documented on the internet. Humans, on the other hand, are able to use reasoning. We use logic and common sense to work out meaning beyond what is explicitly said.

Bhatia: We learn reasoning abilities from birth. For instance, we know not to switch on the blender at 2 a.m. because it will wake everyone up. We're not taught this, but it's something you understand based on the situation, your environment and your surroundings. In the near future, AI models will handle many of our tasks. We can't hard code every single common-sense rule into these robots, so we want them to understand the right thing to do in a specific context.

Shwartz: Bolting on common-sense reasoning to [current models](#) like ChatGPT would help them provide more accurate answers and so, create more powerful tools for humans to use. Current AI models have displayed some form of common-sense reasoning. For example, if you ask the latest version of ChatGPT about a child's and an adult's mud pie, it can correctly differentiate between dessert and a face full of dirt based on context.

Where do AI language models fail?

Shwartz: Common-sense reasoning in AI models is far from perfect. We'll only get so far by training on massive amounts of data. Humans will still need to intervene and train the models, including by providing the right data.

For instance, we know that English text on the web is largely from North America, so English language models, which are the most commonly used, tend to have a North American bias and are at risk of either not knowing about concepts from other cultures or of perpetuating stereotypes.

In a [recent paper](#) (published in *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*), we found that training a common-sense reasoning model on data from different cultures, including India, Nigeria and South Korea, resulted in more accurate, culturally informed responses.

Bhatia: One example included showing the model an image of a woman in Somalia receiving a henna tattoo and asking why she might want this. When trained with culturally diverse data, the model correctly suggested she was about to get married, whereas previously, it had said she wanted to buy henna.

Shwartz: We also found examples of ChatGPT lacking cultural awareness. When given a hypothetical situation where a couple tipped four percent in a restaurant in Spain, the model suggested they may have been unhappy with the service. This assumes that North American tipping culture applies in Spain when, actually, tipping is not common in the country, and a four percent tip likely meant exceptional service.

Why do we need to ensure that AI is more inclusive?

Shwartz: Language models are ubiquitous. If these models assume the set

of values and norms associated with Western or North American culture, their information for and about people from other cultures might be inaccurate and discriminatory. Another concern is that people from [diverse backgrounds](#) using products powered by English models would have to adapt their inputs to North American norms, or else they might get suboptimal performance.

Bhatia: We want these tools for everyone out there to use, not just one group of people. Canada is a culturally diverse country, and we need to ensure the AI tools that power our lives are not reflecting just one [culture](#) and its norms. Our ongoing research aims to foster inclusivity, diversity, and cultural sensitivity in the development and deployment of AI technologies.

More information: Mehar Bhatia et al, GD-COMET: A Geo-Diverse Commonsense Inference Model, *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing* (2023). [DOI: 10.18653/v1/2023.emnlp-main.496](#)

Provided by University of British Columbia

Citation: Q&A: ChatGPT has read almost the whole internet. That hasn't solved its diversity issues (2024, January 11) retrieved 29 April 2024 from <https://techxplore.com/news/2024-01-ga-chatgpt-internet-hasnt-diversity.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.