Have We Passed Peak AI?

Evie Cooper

I have actually wondered if this truly is possible, or if we have passed “peak AI” already and have unwillingly artificially capped our ability to produce machine-learning models from quality data that are meant to emulate as many features of the human mind as possible! Part of the joy of being alive as a human is that we have access to plentiful “training data” that is summed up from thousands of years of us walking around: tons of ancient stories, and tons of different people with subtly-different thought patterns; AI, on the other hand, is totally hopeless in this regard: not only has the well been poisoned, but the entire ocean of training data has been poisoned. Why? Well, two reasons:

1. The current generation of large-language models and image diffuser models are trained directly from content scraped from the Internet. Once LLMs and image diffusers were let loose on the Internet back in 2023, the quality of training data fell because those third-generation LLMs (really first generation, but the original ChatGPT is technically the third iteration of a GPT transformer model) weren't very effective, accurate, or worth using by any means. Sadly, their sloppy output made its way onto the internet... ruining the quality of future LLMs as they trained. Likewise, a similar effect occurred with image diffuser/image transformer models: everyone knows how the ChatGPT art style looks like sloppy cartoon graphics loaded with noise and is so yellow you'd think the picture was left out in the sun for a decade).
2. LLMs can't actually do any form of reasoning. The “thinking” part of modern LLMs is actually just a pseudo-RNN stage, wherein generated text output is put back into the model. What does this accomplish? Technically, nothing. If you examine the length of the intermediate token generations, they vary wildly in length and is not proportional to the difficulty of the problem contained in the prompt (for instance, Deepseek might generate 100+ tokens for a prompt consisting of “Hi”).
3. Because the video-generation models were just transfer-learning-products of an image diffuser and an LLM, they (almost straight away) were “doomed from the start” to be very inefficient, sloppy, and ultimately are a commercial failure (see the recent shutdown of Sora 2).
4. Some (Anthropic) have sought to acquire entire libraries of books and machine-scan/OCR them. Sadly, they also made the OCR model... and that just so happens to inject some “text recognition bias” into something (since their model is tuned to return a word for every chunk of writing/text in a book, even if what it returns is wrong) as it “reads” this also results in training data pollution,Ì but not as bad as “just scrape the internet for everything, even though upwards of 70% of content on the internet is AI-generated now.”

The other thing that I don't think people realize adequately is that the activation function of neurons in a neural network is an emulation of that of a real neuron, not a stand-in for an actual neuron. The current consensus is that a linear activation function is used on the output layer, and a ReLU function (itself is a max() function) for the hidden layers/input layers. The model architecture of an LLM is totally different from that of the left-brain interpreter module; even though both kinda do the same thing, the left-brain interpreter is capable of driving substantially larger downstream loads than just an 8-bit ASCII lookup table to put text on a screen. Because of the pyramidal shape of these models (which can be seen on high-level architectural diagrams of them), substantial bias comes from the hyperparameters applied to the normalization layers at the end of the model.

This actually proves to be a pretty interesting point, because the purpose of those normalization layers is to remove any noise that may accumulate from the feed-forward layers. Unfortunately, the removal of noise is almost definitely something that you can train your brain to readily realize, and it seems that people in my generation are shockingly good at doing so: for instance, I can “clock” something as being AI-generated pretty much instantly upon looking at it, if it's text or image images are easier to clock, but text requires some small rhetorical analysis skills and typographic analysis on top of that. My friends can do the same, somehow... and none of them put in any effort to learn “oh, this is the tell for whether or not something is AI-generated;” they just can somehow realize. I tried to put some method behind the madness, and this is what I've come up with:

1. AI-generated images have a very odd noise distribution. If you take a DFT of an image that's from a real camera, there will be a roughly uniform Gaussian noise distribution. Since AI images are actually generated by feeding a colored “TV static” image through a bunch of feed-through RNN layers, the resultant image has a very unusual noise distribution that does not match a real image. This is why AI-generated images

look so glossy in portions the model is extremely confident in generating, and why the less-confident portions will have a very specific noise signature that shows up in the DFT.

2. AI-generated text writes with multiple voices at the same time. This makes sense because of how an LLM works (it just predicts the next word from millions of different authors all competing to see who can do the same thing the most times, such that it makes it into the training dataset), and this actually experiences an emergent-property effect in other words, this occurs at not only a low level, but a high level too. Note very carefully the sentence I just wrote there (with the whole "em-dash delimited dependent clause that contains a ´not x, but y' phrasing style"), because that is a very common "tell" for a lot of AI-generated text now.[1]
3. Some other AI tells:
    - Excessive generation of bulleted lists that contain approximately 30 tokens after a "Some entry:" colon-delimited heading
    - Because the output is generated linearly with no backtracking (sorry, C language fans), whatever the "title" is will substantially alter the "body" that follows it. It will look like the title was written retroactively to fit the content in the body, but this is an illusion (a casual reader can get fooled on this, but the reality is that the text is already there in the training data and someone that actually is a writer already did the do-diligence of going back and placing a title)
    - Noun/preposition colocation always preludes the subject of a sentence. For example: "The advantage of AI is its efficacy in solving certain tasks." (these also often show up in independent clauses)

I would also like to challenge the notion that an LLM can do math: it can't. It can definitely generate tokens that some other program can detect, perform computation on, then slipstream the result back into the output data (in fact, ChatGPT uses Python scripts that run on a virtual machine to do this any time it needs to compute a hard math problem or produce a diagram). A flat LLM, unaided by external tooling, cannot actually do math at all... instead, it just memorizes the answer for the same reason the language-synthesis portions of our brains are not really helpful when it comes to solving math problems. This is actually an ongoing indication that LLMs have hit a scaling brick wall: no matter how many neurons seem to be added to the models, they aren't developing any emergent properties outside of language translation.

The final "nail in the coffin" in my morning-breakfast essay here will be challenging the notion that an LLM will make work more productive. Actual real-world studies have revealed that while an LLM can type quite fast (especially writing programs), they do not ever "hit the nail on the head" for what the user wants. The sad effect of this is that LLMs (especially when used for coding tasks) actually amplify the amount of work that the human must do. This is simply because if the human operator cannot go tweak out errors in a program, it must prompt the LLM over and over... and then wait for it to regenerate the entire program with the small changes applied.

Last, but not least: a common trope amongst people in my generation is that a lot of us are diehard believers in "the Dead Internet Theory," wherein most of the content on the Web is no longer real. Evidence for this exists in many empirical ways:

- Sites like Linkedin are estimated to be approximately 85% AI-generated content, and have singlehandedly resulted in the "fixing" of the LLM writing style to the above-discussed writing style.[2]
- A large amount of short-form vertical-format social media videos are AI-generated, and this "comes and goes" with huge spikes in generation once new models are produced.
- A LOT of voiceovers on YouTube videos now are AI-generated, and it's always this one monotone-sounding man that talks with the mannerisms of the biology teacher in WarGames. The female

[1] Side note: the reason for this is something I spent about a month figuring out, and it's because when the LLMs were trained, they were configured to favor the phrase that produced the most tokens; that sample text blurb is absolutely packed with short words and punctuation that cranks the token count up

[2] This is of particular interest because of what the purpose of Linkedin actually is, but that is an email for another morning as I eat breakfast!

voice counterpart sounds like a depressed bank teller.

- The r/changemyview Subreddit famously ran an “experiment” wherein they let a bunch of LLMs loose on the subreddit, and their goal was to change the viewpoints of posters. Users were not made aware of this, but those like myself that have the Gift of AI Clocking knew rather quickly what was going on.
- Pretty much all random articles on various topics on the Web are AI-generated now. If I perform a Google search for some random topic, like what the best dog bowl is for my girlfriend's sister's pet cat's brother's mom's aunt's dog, I'm very likely to find AI-generated articles.