

## **O Estado de São Paulo**

ESPAÇO ABERTO (13/11/2024, p. A4; copiado do site do jornal nesse dia)

### **Perigos da 'inteligência' artificial**

*Computadores só fazem escolhas lógico-matemáticas. Somente seres humanos tomam decisões*

**Por Valdemar W. Setzer**

Os leitores devem ter estranhado o emprego de *inteligência* entre aspas no título. Ocorre que a ciência não sabe o que é inteligência e como ela se processa: assim, a denominação "inteligência" artificial (IA) é errada. Um nome mais apropriado: "simulação digital de processos cognitivos". IA será aqui empregada por ser popular. Esse erro no nome já representa um perigo, pois dá a impressão de que algo artificial pode ser como a nossa inteligência e, com isso, dar a impressão de sermos máquinas, o que pode trazer terríveis consequências (por exemplo, máquinas não têm empatia nem compaixão). Por outro lado, eleva os computadores a um nível indevido, achando-se que haverá máquinas com todas as capacidades humanas. Duas destas jamais estarão em máquinas: nossas sensações e sentimentos, pois são individuais e subjetivos, ao passo que computadores são objetivos e universais – com capacidade suficiente, qualquer um pode simular exatamente ("emular") qualquer outro. Se uma pessoa come um caqui e tem a sensação do gosto dele, nenhuma outra pessoa pode ter a mesma sensação; se a pessoa gosta ou não gosta do caqui, esse sentimento também não pode ser sentido por outra pessoa.

Ultimamente, com o advento da IA generativa, os resultados da IA tornaram-se muito bons em termos linguísticos e imagéticos, tornando-se convincentes. Assim, pessoas não especializadas podem ser enganadas com resultados falsos ou viciados, um grande perigo de formar opiniões erradas ou tendenciosas. Qualquer usuário da internet já tem um perfil de seus dados armazenado em vários sistemas: duas pessoas fazendo a mesma pergunta para máquinas de busca ou para a IA generativa provavelmente terão respostas diferentes, baseadas em seus perfis. Com isso, resultados dessa IA podem ser personalizados, tornando-se ainda mais convincentes.

Um perigo muito grande é o fato de a IA generativa usar dados que foram fornecidos durante a fase de "treinamento" – novamente um nome errado, pois somente seres humanos e animais podem ser treinados. Nessa fase é dada ao sistema uma quantidade enorme de dados de entrada e os esperados de saída. Com isso, usando enormes capacidades computacionais e técnicas matemáticas complexas, são calculados muitos parâmetros (hoje em dia, milhões deles). Por exemplo, pode-se "treinar" um desses sistemas para distinguir fotos de cães e de gatos. Fornece-se como entrada milhares de fotos, cada uma com o resultado

esperado, isto é, "cão" ou "gato". Com isso, num processo de otimização (diminuição do erro), o sistema calcula milhares, milhões de parâmetros que processam uma nova foto de entrada, dando como resultado "cão" ou "gato". (Dando-se uma foto de raposa ou lobo, certamente o resultado será "cão", mostrando como o sistema é parcial e viciado.) "Treinando-se" o sistema com fotos de temas diferentes, ele poderá dar respostas a outros tópicos, cada um com um novo algoritmo. Portanto, os sistemas generativos representam uma revolução na computação: na programação tradicional de computadores, um programador sabia exatamente, matematicamente, o que cada comando de seu programa iria forçar a máquina a executar. Assim, o programa era teoricamente conhecido (mas, se tivesse erros de programação, os resultados poderiam ser imprevisíveis). A IA generativa é um sistema de gerar programas, produzindo-os com diferentes parâmetros para diferentes propósitos. O grande problema é que, como a quantidade de parâmetros calculados em cada caso é imensa, é impossível que eles sejam examinados e se compreenda como o programa funciona. Assim, os computadores adquiriram certa independência: eles mesmos geram seus programas, cujo funcionamento ninguém conhece, ninguém compreende. Temos então o maior perigo representado pela IA generativa.

Isso não é muito grave quando computadores controlam máquinas, como pilotos automáticos em aviões ou automóveis autônomos – a grande maioria dos acidentes é provocada por falhas humanas; se sistemas autônomos diminuíssem muito esses desastres, deverão ser usados. A situação torna-se muito grave quando os resultados da IA são usados para substituir decisões humanas, que afetam outras pessoas ou a sociedade. Nesse caso, deveria haver leis que proibissem o uso de resultados da IA sem que seres humanos examinem esses resultados e aprovem os resultados. Em outras palavras, um resultado que influencia diretamente pessoas deveria ser usado exclusivamente como sugestão. Um exemplo disso é o diagnóstico médico. Aliás, aqui há um outro grande perigo: os médicos confiarem nesses diagnósticos e perderem a capacidade de fazer os seus. Computadores só fazem escolhas lógico-matemáticas. Somente seres humanos tomam decisões.

Finalmente, outro grande perigo é o congelamento da evolução cultural humana. Os dados usados pela IA para chegar aos seus resultados são sempre dados do passado. Obviamente computadores não têm a intuição humana, que pode e deve trazer inovações para o futuro, que não são simplesmente combinação de coisas passadas.

É PROFESSOR APOSENTADO DO DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO DA USP. SITE: [WWW.IME.USP.BR/~VWSETZER](http://WWW.IME.USP.BR/~VWSETZER)

Em tempo (VWS): sobre sensações e sentimentos, e muito mais, ver meu artigo [www.ime.usp.br/~vwsetzer/conceito-cerebro.pdf](http://www.ime.usp.br/~vwsetzer/conceito-cerebro.pdf)