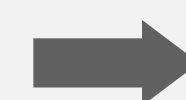




CREIEU QUE ESTEM EN UN PUNT D'INFLEXIÓ EN ELS USOS I ABUSOS DE LA IA GENERATIVA DESPRÉS DELS LLANÇAMENTS DE #CHATGPT I #MIDJOURNEY?

- Una majoria de vosaltres (46%) creieu que estem en un punt d'inflexió ens els usos i abusos de la IA generativa després dels llançaments de ChatGPT i Midjourney, mentre que la resta esteu en desacord (27%), no ho sabeu (17%) o us mostreu ambivalents (10%).
- Hi ha un cert consens que la IA generativa amb models comercials com, per exemple, BERT, DALL-E, LaMDA, GPT-3 o ChatGPT, que s'entrenen amb grans volums de dades i que són adaptables a una àmplia gamma de tasques posteriors podrien representar un punt d'inflexió.
- Molt probablement, aquest canvi de paradigma (pels més agosarats) té a veure amb una característica clau que dècades de ciència-ficció ens ha ensenyat, i és un futur d'alta tecnologia on la IA ens dona accés instantani a col·leccions de coneixement aparentment il·limitades.
- Tot i que (generalment) els models d'IA generativa es basen en l'aprenentatge profund estàndard i l'aprenentatge de transferència, la seva gran escala dona lloc a noves capacitats emergents constantment i la seva eficàcia en tantes tasques incentiva l'homogeneïtzació.
- És per aquest motiu, que es considera que representen un punt d'inflexió ja que poden convertir-se en la base de gairebé tots els sistemes basats en IA, el que significa que poden acabar alimentant un gran nombre d'aplicacions posteriors en diversos camps.
- Evidentment, aquesta homogeneïtzació exigeix molta precaució ja que els possibles defectes d'aquests models es transmetran a tots els models posteriors, la qual cosa podria provocar problemes generalitzats si no es governen deliberadament.
- Però juntament amb els problemes sociotècnics i de governança, també hi ha aquells d'impacte social relacionats amb la desigualtat, el mal ús, l'impacte econòmic i ambiental, i en consideracions legals i ètiques.





- Per exemple, aquests poden esdevenir una manera molt potent de difondre informació errònia o falsa de manera convicent, de generar continguts estereotipats o amb prejudicis, i d'escriure programari maliciós i plagiar contingut, obres i invents amb drets d'autor.
- Sens dubte, també hi ha moltes oportunitats (per exemple, en llenguatge, visió, robòtica, raonament, interacció humana) i aplicacions (per exemple, en sanitat o educació) però és crític tenir una comprensió clara de com funcionen, quan fracassen i de què són capaços.
- Segurament, la primera consideració és que la IA generativa, que s'entrena amb grans volums d'informació d'Internet i, per tant, amb entrades i sortides impredecibles, té una combinació paradoxal de pèrdues previsibles tal com es plasma en les anomenades "lleis d'escala".
- Un bon exemple és la darrera joia de la corona d'OpenAI, el ChatGPT. Si bé aquest és capaç de crear un text que sembla coherent, cal remarcar que el sistema no té cap comprensió del que està produint, tampoc cap intenció comunicativa i molt menys és responsable del que diu.
- Però el principal problema és que, malgrat l'imminent desplegament generalitzat d'aquests models, actualment encara no tenim una comprensió clara de com funcionen, de quan fracassen i fins i tot de què són capaços a causa de les seves propietats emergents.
- I fins aquí per avui. Això si, amb dues lectures sobre el tema.
- [Predictability and Surprise in Large Generative Models](#)
- [3 Ways to Tame ChatGPT](#)