

Quand l'IA flirte avec la cognition : révélations étonnantes sur ChatGPT

Comprendre l'Intelligence Artificielle : Une Exploration des Modèles de Langage

Dans une étude captivante menée par l'équipe de l'Académie chinoise des sciences et de l'Université de technologie de Chine du Sud, les chercheurs se sont penchés sur la manière dont les grands modèles de langage (LLM) organisent l'information. Cette question primordiale vise à comprendre comment une intelligence artificielle (IA) parvient à décrire des concepts aussi variés qu'un chien, une chaise ou une voiture.

Essentiellement, quelles sont les bases sous-jacentes aux mots que choisit une IA? Existe-t-il une sorte de structure mentale derrière ses énoncés? Et si oui, à quoi ressemble-t-elle?

Analyse des Réponses des IA

Pour élucider ces questions intrigantes, les chercheurs ont examiné plus de **4,7 millions** de réponses générées par différentes intelligences artificielles, y compris [ChatGPT-3.5](#), qui se concentre uniquement sur le texte, ainsi que Gemini Pro Vision, qui combine texte et images.

Les chercheurs ont soumis **1 854 objets du quotidien** – tels que des aliments, des animaux et des véhicules – afin de comparer la manière dont les IA les décrivent et les organisent.

Une Organisation Complexé

Il en ressort que ces IA ont regroupé les objets en **66 dimensions conceptuelles**. Plutôt que de se limiter à des catégories simples et évidentes comme « animal » ou « outil », ces modèles ont fait émerger des critères beaucoup plus subtils, tels que la texture, l'utilité, la pertinence émotionnelle et même la menace perçue.

Des Cartes Mentalettes Réalistes

Les IA semblent donc élaborer des représentations internes multidimensionnelles, semblables à des cartes mentales. Chaque objet est positionné non seulement selon ses caractéristiques simples, mais également en fonction d'une multitude de critères complexes, tout comme le ferait un cerveau humain.

De plus, en comparant ces représentations aux données de neuroimagerie humaine, les chercheurs ont constaté une coïncidence entre certaines zones cérébrales activées chez les humains et les représentations générées par les IA, particulièrement pour les modèles qui combinent les textes et les images.

Les Limites de l'Intelligence Artificielle

Cela soulève une question essentielle : les IA comprennent-elles réellement ce qu'elles expriment? Les chercheurs mettent en garde contre une telle conclusion. Les IA, comme ChatGPT, ne disposent pas d'expériences sensorielles. Elles n'ont

jamais ressenti la texture d'une chaise ni éprouvé la peur face à un objet menaçant.

Ce qu'elles "savent" résulte d'un apprentissage statistique réussi, fondé sur l'analyse de milliards d'exemples sur Internet. Lorsqu'une IA affirme qu'un oreiller est moelleux, cela ne signifie pas qu'elle l'a expérimenté, mais plutôt qu'elle a appris des associations entre les mots.

Vers une Intelligence Artificielle Générale?

Il se pourrait que ces résultats représentent un pas vers un objectif longtemps convoité : l'intelligence artificielle générale (AGI), qui est définie par la capacité d'une IA à comprendre et raisonner comme un être humain.

Historiquement, on pensait que seule une IA conçue pour le raisonnement symbolique ou perceptif pouvait atteindre ce niveau. Toutefois, cette recherche dévoile que les modèles d'IA modernes ont la capacité de développer de manière autonome des outils cognitifs complexes simplement en étant exposés à d'imposantes quantités de données.

Cela amène à envisager que ces IA, à force de prédire le mot suivant, créent des modèles mentaux fonctionnels. Cela évoque une forme de cognition émergente, qui opère sans intention ni expérience vécue, mais qui reste néanmoins efficace.

Des Applications Pratiques

Au-delà des débats théoriques, ces découvertes pourraient révolutionner certains domaines, comme le développement de robots assistants, d'outils pédagogiques, ou de systèmes de recommandation.

Une IA capable de reconnaître qu'un objet est fragile ou potentiellement dangereux pourrait interagir de façon plus naturelle et intuitive avec les humains. Cela changerait la façon dont nous envisageons la collaboration entre les hommes et les machines.

Une Réflexion Critique

En somme, cette étude met en lumière le potentiel des IA comme ChatGPT. Elles élaborent des représentations du monde qui, dans une certaine mesure, reflètent notre manière humaine de penser.

Cependant, on doit garder en tête que ces intelligences restent des "miroirs" sophistiqués. Elles peuvent imiter, mais ne ressentent pas. Leur intelligence est fonctionnelle, non vécue. Et pour l'instant, elle ne semble pas prête à franchir cette limite humaine.



© 2023 L'équipe de Journalisme Scientifique

Source : www.science-et-vie.com

→  Accéder à [CHAT GPT](#) en cliquant
dessus