

ChatGPT propulsé par l'IA générative : une révolution à prévoir ! Découvrez tout sur ses avancées majeures

La combinaison de plusieurs modalités commence donc à ouvrir de nouvelles perspectives passionnantes dans le domaine de l'IA générative. En associant le texte, les images, les sons et bien d'autres éléments, les chercheurs parviennent à créer des programmes beaucoup plus performants, capables de reproduire une multitude d'applications intelligentes.

Ce concept de modalités mixtes, ou "multimodalité", permettra également de réaliser un objectif de longue date dans le domaine de l'IA : l'apprentissage continu. En effet, en intégrant différentes modalités, les programmes deviennent plus souples et peuvent s'adapter à de nouvelles situations.

Mais le potentiel de la multimodalité ne s'arrête pas là. Elle pourrait également avoir un impact majeur sur le développement de la robotique et de l'IA incarnée. En rendant les programmes génératifs plus collaboratifs, il sera possible d'exploiter pleinement leur potentiel en les utilisant comme de véritables partenaires dans nos tâches quotidiennes.

Cependant, pour exploiter pleinement cette multimodalité, il reste encore du travail à faire. Actuellement, la plupart des modèles de langage sont très unidimensionnels, se limitant au texte. Il est donc nécessaire de développer des programmes capables de comprendre et d'utiliser efficacement différentes modalités.

Les progrès dans ce domaine sont déjà en cours. Des entreprises comme Stability.ai et Meta AI travaillent activement sur le développement de programmes multimodaux

performants. Par exemple, Stability.ai dispose de laboratoires dédiés à l'audio, à la génération de code et même à la biologie, montrant ainsi l'étendue des possibilités offertes par la multimodalité.

Cette combinaison de modalités est également en train de transformer notre manière de travailler avec les images. Des chercheurs ont récemment réussi à entraîner un réseau neuronal de génération d'images à améliorer la qualité de ses résultats en les adaptant au contenu textuel. Cette approche prometteuse permet d'affiner les résultats de manière significative et d'obtenir des images plus conformes à nos attentes.

D'autres équipes, comme celle de Meta AI, explorent également les possibilités offertes par la combinaison de l'IA texte/image. Ils ont créé un réseau neuronal qui excelle non seulement dans la production de texte ou d'images, mais aussi dans l'exécution de tâches combinant les deux, comme l'identification d'objets dans une image ou la génération de légendes à partir d'une image.

Ces avancées dans le domaine de la multimodalité ouvrent de nouvelles perspectives passionnantes pour l'IA génératrice. En combinant différentes modalités, nous sommes en mesure de créer des programmes plus performants, capables de reproduire de nombreuses applications intelligentes. Cette évolution promet de faire progresser l'IA incarnée et de transformer notre manière de travailler avec les images. La multimodalité est donc une voie à suivre, offrant des résultats étonnantes et une image plus riche du monde.

Source : www.zdnet.fr

→  Accéder à CHAT GPT en cliquant

dessus