

LEXIQUE DE L'IA

Intelligence artificielle (IA)

Domaine de la science informatique qui consiste à développer des systèmes pour imiter le comportement humain et démontrer l'intelligence des machines.

Biais

Phénomène qui se produit lorsqu'un système d'IA produit des résultats systématiquement injustes ou imprécis en raison d'hypothèses ou de présuppositions erronées dans le processus d'apprentissage automatique. Ces biais peuvent entraîner des répercussions négatives sur les individus et la société (discrimination, désinformation ou perte de confiance, par exemple). Il en existe différents types et ils ont différentes sources (biais dans les données, biais algorithmiques, partis pris humains et partis pris sociétaux).

Apprentissage profond

Sous-domaine de l'apprentissage automatique utilisant de grands réseaux profonds de neurones (artificiels) à plusieurs couches qui calculent des représentations continues (en nombres réels), à la manière des neurones organisés hiérarchiquement dans le cerveau humain. Il est très efficace lorsqu'il s'appuie sur des données déstructurées comme les images, le texte et l'audio.

Peaufinage

Procédé consistant à adapter un modèle fondateur préentraîné pour effectuer une tâche spécifique plus efficacement. La période d'entraînement est relativement courte, car elle porte sur un jeu de données annotées, bien plus réduit que celui sur lequel le modèle a été entraîné initialement. Cette formation supplémentaire permet au modèle de s'enrichir et de s'adapter aux nuances, à la terminologie et aux formes spécifiques.

IA générative

Forme d'apprentissage automatique dans laquelle une plateforme d'IA peut générer de nouveaux contenus en réponse à des requêtes basées sur des données sur lesquelles elle a été entraînée.

Hallucination

Phénomène par lequel un système d'IA produit des résultats qui ne sont pas fondés sur la réalité ou le contexte donné. Par exemple, un agent conversationnel peut inventer des faits ou des histoires, ou un système de reconnaissance d'images peut voir des objets ou des formes qui n'existent pas.

Grand modèle de langage

Réseau de neurones formé sur un large volume de texte pour imiter le langage humain. Cette catégorie de modèles fondateurs peut traiter d'énormes volumes de texte déstructuré et apprendre les relations entre les mots ou les morceaux de mots, appelés jetons. Cela lui permet de générer du texte en langage naturel pour effectuer des tâches telles que la synthèse ou l'extraction de connaissances. GPT-4 (modèle de ChatGPT) et LaMDA (modèle qui sous-tend Bard) sont des exemples de grands modèles de langage.

Apprentissage automatique

Étude de la façon dont l'IA acquiert des connaissances à partir des données d'entraînement. Dans ce sous-ensemble de l'IA, un modèle acquiert des capacités et améliore sa perception, ses connaissances, son raisonnement ou ses actions après avoir été entraîné sur de nombreux points de données ou y avoir été exposé. Les algorithmes d'apprentissage automatique détectent les formes et apprennent à formuler des prédictions ou des recommandations en traitant des données et des expériences. De cette manière, le système apprend à fournir du contenu précis au fil du temps.

Réseau de neurones artificiels

Modèle de calcul inspiré de la structure et de la fonction des neurones biologiques.

Ingénierie de requête

Technique utilisée dans l'IA pour optimiser et peaufiner des modèles de langage pour des tâches particulières et des résultats escomptés. Aussi appelée construction de requête, cette notion désigne la construction soigneuse de requêtes ou données en entrée pour les modèles d'IA afin d'en améliorer les performances sur des tâches spécifiques.

Requêtes

Instructions données à un système d'IA utilisant le langage naturel et non le langage informatique. Par exemple, la requête peut consister à demander à l'IA générative de créer du contenu d'apparence novateur ou intéressant.

Apprentissage supervisé

Type d'apprentissage automatique qui utilise des jeux de données annotées pour entraîner les algorithmes à classer les données ou à prédire les résultats. Les jeux de données annotées sont des ensembles de données auxquels des humains ont ajouté une annotation ou une catégorie.

Apprentissage non supervisé

Type d'apprentissage automatique dans lequel les algorithmes apprennent des formes à partir de données non annotées, sans orientations ou observations de l'être humain.