

Chapitre 11 – L'intelligence artificielle

L'histoire du traitement de l'information – Activité 1

- Jusqu'au début du xx^e siècle, les machines traitant l'information sont limitées à une ou quelques tâches prédéterminées : tisser grâce à un ruban ou des cartes perforées, trier un jeu de cartes perforées, séparer des cartes selon un critère, sommer des valeurs indiquées sur ces cartes, etc. Comme le montre l'exemple de la machine analytique imaginée par Babbage (1834) et celle du métier à tisser configurable de Jacquard (1801), les progrès techniques du traitement de l'information se sont construits en utilisant et en adaptant les techniques antérieures.
- Au début du xx^e siècle, des questions mathématiques nouvelles émergent : peut-on tout calculer ? Pour toute propriété mathématique, peut-on décider si elle est vraie ou fausse ? Pour répondre à ces questions, Turing (1936) est le premier à proposer le concept de machine universelle. En s'inspirant des travaux de Babbage et de Turing, Von Neumann décrit l'architecture d'une telle machine dont la matérialisation est effective au début des années 1950 avec les premiers ordinateurs.
- Conformément à l'architecture de Von Neumann, les premiers ordinateurs sont constitués à minima d'un processeur et d'une mémoire vive.

La numérisation de l'information – Activité 2

- La première machine analytique de Babbage utilisait des cartes perforées. Pour effectuer et programmer (Lovelace, 1843) que des calculs numériques, un ordinateur actuel manipule des données de natures diverses une fois qu'elles ont été numérisées : textes, images, sons. Pour contrôler les actions qu'un ordinateur exécute, on utilise des programmes écrits avec des langages de programmation comme Python ou Scratch.
- Les programmes sont donc également des données : comme les données numérisées, ils peuvent être stockés, transportés, et traités par des ordinateurs. En particulier, un programme écrit dans un langage de programmation de haut niveau (Python, Scratch) peut être traduit en instructions spécifiques à chaque type de processeur.

Les bugs – Activité 3

- Comme c'est le cas par exemple pour une voiture autonome, un programme peut comporter jusqu'à plusieurs centaines de millions de lignes de code. La présence d'erreurs, de défauts de conception, appelées bogues (ou bugs) est donc très probable.
- Ces erreurs peuvent conduire un programme à avoir un comportement inattendu et gênant et, dans certains cas, entraîner des conséquences graves.

L'utilisation de jeux de données est un des outils qui peut permettre d'éviter l'apparition de bugs.

Savoir-faire

- Analyser des documents historiques relatifs au traitement de l'information et à son automatisation.
- Recenser des outils numériques utilisés dans la vie courante, identifier ceux qui sont programmables, et par qui (ordinateur, thermostat d'ambiance, téléphone, intelligent, boîte Internet, ordinateur de bord d'une voiture, etc.).
- Savoir distinguer les fichiers exécutables des autres fichiers sous un système d'exploitation donné.
- Connaître l'ordre de grandeur de la taille d'un fichier image, son, vidéo. Savoir calculer la taille en octets d'une page de texte (en ASCII et non compressé).
- Étant donné un programme très simple, proposer des jeux de données d'entrée permettant d'en tester toutes les lignes.
- Corriger un algorithme ou un programme bogué simple.

Mots clés

Bit : contraction de Binary digit, qui signifie chiffre binaire et apparaît en 1948.

Le bit représente la plus petite unité de mémoire utilisable sur un ordinateur.

Cette mémoire ne peut prendre que deux valeurs, symbolisées le plus souvent par 0 ou 1.

Fichier exécutable : un **fichier exécutable** est un fichier contenant un programme et identifié comme tel par le système d'exploitation. Le chargement d'un fichier exécutable entraîne la création d'un processus et l'exécution du programme.

Octet : un **octet** est composé de huit bits.

L'intelligence artificielle et l'apprentissage

automatique – Activités 4 et 5

- L'intelligence artificielle (IA) est née en 1956. Elle visait alors à simuler sur ordinateur les facultés cognitives humaines et recouvrait des approches relevant de l'informatique, des mathématiques et des sciences cognitives.
- Aujourd'hui, le terme d'IA est plutôt attribué à l'un de ses sous-domaines, celui de l'apprentissage automatique. Il s'agit d'un processus par lequel un algorithme évalue et améliore ses propres performances, non pas sous l'intervention d'un humain chargé de programmer la machine, mais en répétant son exécution sur des jeux de données de natures variées (mesures de capteurs pour des prévisions, textes pour la traduction, sons pour la reconnaissance vocale, images pour la reconnaissance visuelle, etc.).
- L'apprentissage machine exploite des méthodes mathématiques qui, à partir du repérage de tendances (par exemple, des corrélations ou des similarités) sur de très grandes quantités de données (données massives, ou big data), permettent de faire des prédictions ou de prendre des décisions sur d'autres données.
- La qualité et la représentativité des données fournies sont essentielles pour la qualité des résultats. En effet, l'un des risques de

l'apprentissage automatique réside dans l'amplification des biais des données. Par ailleurs, une interprétation trop rapide des données et un amalgame entre corrélation et causalité peuvent aboutir à des résultats erronés.

Inférence bayésienne, apprentissage automatique et intelligence artificielle – Activités 6 et 7

- L'inférence bayésienne est une méthode de calcul de probabilités de causes à partir des probabilités de leurs effets.
- Elle est utilisée en apprentissage automatique pour modéliser des relations au sein de systèmes complexes, notamment en vue de prononcer un diagnostic.

Savoir-faire

- Analyser des documents relatifs à une application de l'intelligence artificielle.
- Analyser un exemple d'utilisation de l'intelligence artificielle : identifier la source des données utilisées et les corrélations exploitées.
- Sur des exemples réels, reconnaître les possibles biais dans les données, les limites de la représentativité.

- Sur des exemples simples, montrer qu'une corrélation ne correspond pas toujours à une relation de causalité.
- Expliquer pourquoi certains usages de l'IA peuvent poser des problèmes éthiques.
- Dans le contexte d'un diagnostic médical fondé sur un test, produire un tableau de contingence afin de calculer des fréquences de faux positifs, faux négatifs, vrais positifs, vrais négatifs et en déduire le nombre de personnes malades suivant leur résultat au test.
- Utiliser cette démarche dans d'autres contextes de classification (reconnaissance d'images, détection de messages non sollicités envoyés en masse, etc.).
- Expliquer pourquoi certains usages de l'IA peuvent poser des problèmes éthiques.

Mots clés

Algorithme : vient du nom du grand mathématicien persan Al Khwarizmi (vers l'an 820) et désigne une méthode ou « procédure mécanique » pour résoudre un type de problèmes.

Apprentissage automatique : technologie d'intelligence artificielle permettant aux machines d'apprendre à résoudre certaines tâches sans avoir été au préalable programmées spécifiquement à cet effet.

Big data : l'ensemble des données numériques produites par l'utilisation des nouvelles technologies. Ces données présentent une grande variété, arrivent en volumes croissants et à grande vitesse.

C'est ce que l'on appelle les trois « V ».

Courbe de tendance : représentation graphique d'une fonction permettant d'illustrer les tendances des données existantes et de fournir des prévisions pour des données futures.

Diagnostic : raisonnement menant à l'identification de la cause (l'origine) d'un problème observé. Dans le cas médical, raisonnement permettant d'identifier une probable maladie à partir de symptômes.