

L'intelligence artificielle

Séance 4 : L'intelligence artificielle et l'apprentissage machine

Qu'est-ce que l'intelligence artificielle et comment peut-on l'utiliser ?

● Définition

Le terme d'**intelligence artificielle** (IA) désigne un ensemble de théories et de techniques de programmation dont le but est de simuler l'intelligence humaine, que ce soit dans la capacité d'apprentissage, les perceptions ou la prise d'initiatives.



Gary Kasparov jouant contre l'ordinateur Deepblue (1,8 m pour 1 400 kg).



Exemples d'application de l'IA :

Reconnaissance faciale

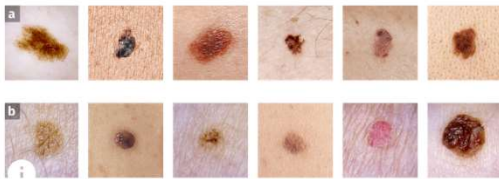
Notre visage comporte de nombreux traits caractéristiques comme la bouche, le nez, les yeux et les sourcils. Les distances entre les différents points de notre figure, calculées puis enregistrées par un programme informatique, permettent de nous identifier (figure).

Aujourd'hui, la reconnaissance d'un visage vu au 3/4 est de 100 % et de 60 % pour un visage vu à moitié dans des conditions normales d'éclairage. L'analyse de millions de photographies par des algorithmes* fonctionnant grâce à l'**apprentissage machine** donne maintenant la possibilité d'indiquer le sexe, l'âge, le poids, la taille et l'état émotionnel d'une personne.



Doc.

Détecter les problèmes de peau



► Exemples dont on connaît le diagnostic (données d'apprentissage). Mélanome malin **a** ; mélanome bénin **b**.

À l'aide d'une série d'exemples choisis par des experts, une intelligence artificielle peut apprendre à distinguer les maladies de la peau et en déduire, par généralisation, un diagnostic pour un nouveau cas qui lui est présente.

L'utilisation de l'intelligence artificielle permettrait alors un gain de temps pour les médecins et potentiellement une plus grande fiabilité, avec une IA bien entraînée.



► Grain de beauté suspect à diagnostiquer.

Les moteurs de recherche

Certains moteurs de recherches sont déjà équipés d'intelligence artificielle ! En octobre 2015, Google lance Rank Brain, son premier programme d'IA de traitement des requêtes.

Et le géant américain n'est pas le seul à s'y mettre : depuis avril 2017, le moteur de recherche français Qwant dispose aussi de plusieurs supercalculateurs spécialisés pour l'IA.



L'objectif ? Mieux classer et améliorer la pertinence des résultats de recherches des utilisateurs en se basant sur les mots employés, leurs précédentes recherches et les recherches d'autres internautes.

Par exemple, lorsqu'une personne tape les mots "vitesse" et "jaguar", l'algorithme lui présentera des pages différentes si elle a déjà fait des recherches sur les voitures de course ou sur les animaux de la jungle.

??? Pour chacune de ces applications de l'IA, donner un avantage et un inconvénient de leur utilisation.

L'**apprentissage machine** permet aux ordinateurs de réagir dans des situations inédites à partir de données d'entraînement sélectionnées par le programmeur.

Un exemple d'apprentissage machine :

Un élève étudie la préférence de consommation en matière de thé ou de café du personnel de son établissement en fonction du sexe et de l'âge. Il souhaite établir un modèle prévisionnel en programmant une intelligence artificielle à partir de deux témoignages.

La lettre A correspond à l'âge de la personne, S au sexe et B à la boisson.

Sexe	Femme	Homme
S	1	2

Boisson	Café	Thé
B	10	20

a et b sont des coefficients numériques qui représentent l'importance attribuée au sexe ou à l'âge.

Une femme de 20 ans qui aime le thé s'écrira :

$$1 \times a + 20 \times b = 20$$

Un homme de 50 ans qui boit du café s'écrira :

$$2 \times a + 50 \times b = 10$$

En résolvant ce système d'équations, l'élève détermine que $a = 80$ et $b = -3$ d'où le modèle suivant :

$$80 \times S - 3 \times A = B$$

En remplaçant S et A par les valeurs d'une nouvelle personne, le programme indiquera qu'elle préfère le café si la valeur B est plus proche de 10 ou le thé si B est plus proche de 20.



■ Fonctionnement simplifié du modèle.

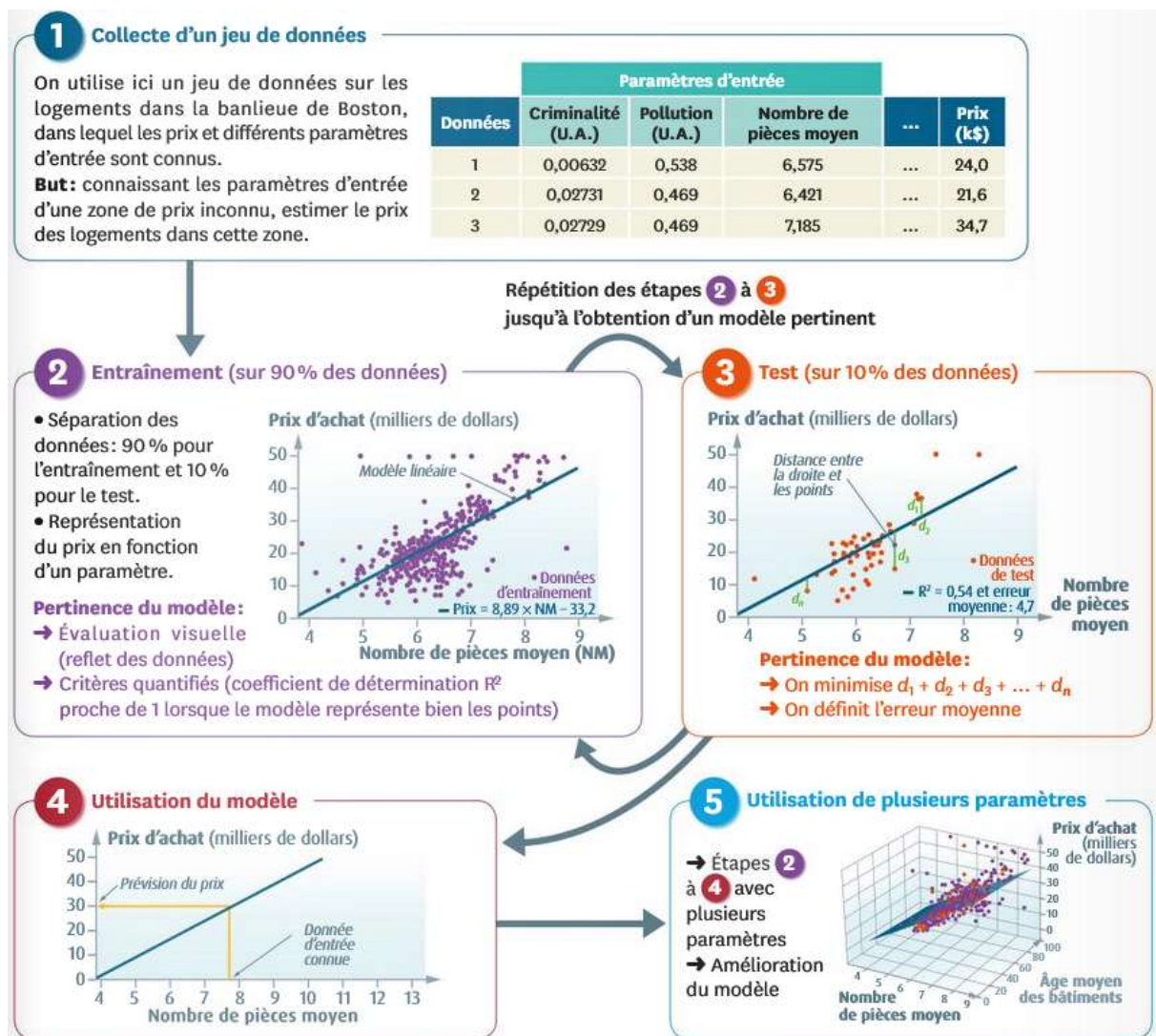
Ce modèle, très simpliste, nécessite un réajustement des valeurs a et b à chaque nouveau témoignage. C'est ce que l'on appelle l'apprentissage profond (*deep learning* en anglais) qui nécessite un grand nombre de données de qualité.

Ainsi, si les programmeurs sélectionnent de mauvaises données, ces dernières vont créer un **biais** et limiter la représentativité du modèle.

??? Déterminer la boisson que devrait consommer une femme de 25 ans selon le modèle de l'élève.

Un autre exemple utilisant une courbe de régression :

Parce qu'on dénombre plus de 30 universités dans la ville et ses environs, Boston attire chaque année des étudiants du monde entier. Un étudiant souhaite s'installer dans un quartier de Boston pour lequel il n'a pas de donnée sur le prix des logements. Comment l'IA peut-elle l'aider à estimer le prix d'un logement ?



1. Indiquer la 1^{ère} corrélation utilisée dans le jeu de données (vignettes 2 à 4). Expliquer si ce modèle pourrait encore être amélioré et si oui, comment.
2. Indiquer si le modèle proposé dans la vignette 5 pourrait être encore amélioré et si oui, comment.
3. En utilisant le modèle des vignettes 2 à 4, estimer graphiquement la valeur attendue pour le prix d'un logement à Boston pour un nombre moyen de pièces de 6,75.



L'apprentissage machine utilise un nombre important de données pour entraîner les modèles, il faut donc être vigilant en se posant quelques questions :

- D'où proviennent les données ?
- Où sont-elles stockées et qui en sont les créateurs ?

Application :

Le montant des ventes d'un magasin est reporté dans la feuille de calcul ci-contre.

1. Décrire l'évolution des ventes entre janvier et novembre.
2. À l'aide de la fonction courbe de tendance du tableur, conjecturer le montant des ventes au mois de décembre.
3. Expliquer pourquoi la collecte de nombreuses données a été nécessaire à la prévision des ventes de décembre.

	A	B
	Mois	Montant des ventes
1		
2	Janvier	200
3	Février	406
4	Mars	526
5	Avril	716
6	Mai	856
7	Juin	1 000
8	Juillet	1 159
9	Août	1 305
10	Septembre	1 474
11	Octobre	1 722
12	Novembre	1 879

Tableur des ventes du magasin entre janvier et novembre.

BILAN :

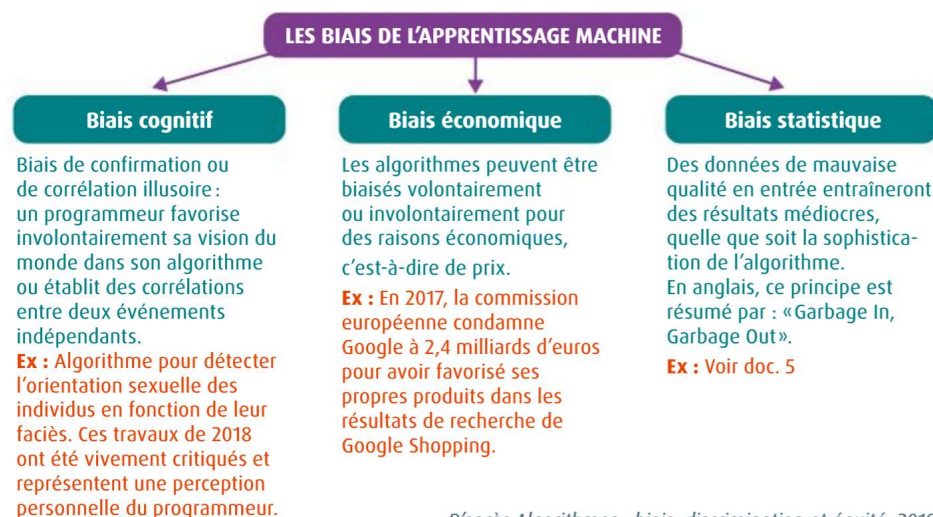
Des méthodes mathématiques de repérage de tendance (corrélations, similarités) et de probabilités sont utilisées sur de grandes quantités de données (*big data*). Elles permettent ensuite d'effectuer des prédictions ou de prendre des décisions dans des situations nouvelles.

La qualité et la représentativité des données d'entraînement sont essentielles pour la fiabilité des réponses.

Si les données d'entraînement comportent un **biais**, les résultats donnés par l'ordinateur seront de mauvaise qualité, ce qui peut avoir de graves conséquences.

Certains usages de l'IA peuvent ainsi poser des problèmes éthiques. Des biais dans l'apprentissage pourraient par exemple entraîner des discriminations.

Exemples : Des algorithmes utilisant l'apprentissage machine peuvent être utilisés pour sélectionner automatiquement des CV pour un recruteur, calculer une probabilité de radicalisation pour la police ou de maladie pour une assurance.



D'après Algorithmes : biais, discrimination et équité, 2019