

L'INTELLIGENCE ARTIFICIELLE

Une machine est capable d'exécuter un algorithme. Dans le cas de l'intelligence artificielle*, l'algorithme n'est pas explicitement programmé, mais ses paramètres sont ajustés au mieux en utilisant des données.

1 L'apprentissage automatique



Certains processus se programment bien explicitement, comme laver automatiquement du linge. D'autres vont nécessiter des calculs massifs ajustés sur beaucoup de données et des approximations : c'est ce qu'on nomme l'apprentissage automatique*. Les deux exemples les plus utilisés d'algorithmes d'apprentissage automatique permettent de dicter un message au lieu de le taper au clavier (*voice-to-text*) ou de reconnaître des objets dans une image. On les retrouve par exemple dans les enceintes connectées ou nos smartphones.

3 Ajuster les paramètres par apprentissage automatique

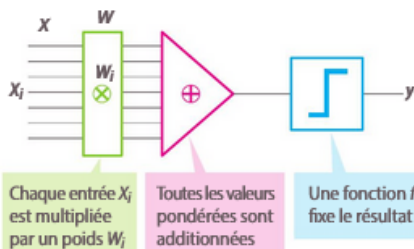
Comment ajuster les paramètres qui permettent de définir le calcul effectué par un algorithme ? On fournit un très grand nombre de données connues et on regarde le résultat du calcul. Les paramètres sont ensuite modifiés pour améliorer le résultat, un peu comme quand nous ajustons « à tâtons » les boutons d'une machine inconnue. Dès qu'on obtient une amélioration, on modifie de nouveau les paramètres. De proche en proche, on optimise ainsi le calcul pour que la transformation souhaitée soit obtenue.

2 Le fonctionnement de l'apprentissage automatique

Pour qu'une machine traite des connaissances humaines, il est nécessaire de les représenter sous une forme très simple, que la machine va pouvoir manipuler.

Un principe est de réduire toutes nos informations à des « faits » élémentaires et à des règles, qui permettent de déduire d'autres faits par des raisonnements automatiques.

Un autre principe est de réduire ces informations à des nombres sur lesquels des opérations seront éventuellement réalisées par un mécanisme comme un réseau de neurones artificiels.



Chaque entrée X_i est multipliée par un poids W_j

Toutes les valeurs pondérées sont additionnées

Une fonction f fixe le résultat

$$y = f(X_1W_1 + X_2W_2 + \dots)$$

Chaque neurone est une petite unité de calcul qui combine des entrées, notées X sur la figure ci-dessus, en les multipliant par des « poids » W plus ou moins grands. Le résultat de cette combinaison est ensuite traité pour permettre de fixer un résultat. Si l'on combine un grand nombre de ces neurones, on peut alors approximer toutes les transformations entre des entrées fournies et une sortie souhaitée.

Femme de science



Rose Dieng-Kuntz (1956-2008), scientifique sénégalaise spécialiste en intelligence artificielle, est la première femme africaine à intégrer l'École Polytechnique. À l'origine d'algorithmes « intelligents » d'acquisition et de recherche de connaissances sur le web, elle a participé à l'émergence de ce qu'on appelle maintenant le web sémantique.

*** VOCABULAIRE**
Apprentissage automatique ou apprentissage machine : ajustement des paramètres d'un algorithme en fonction de données pour prédire le résultat souhaité.
Intelligence artificielle : système ou machine comportant des algorithmes de calcul qui permettent de réaliser certaines tâches confiées auparavant à des humains.

Expliquer par quel processus une machine peut apprendre ?

NUMÉRISATION DES DONNÉES

Un ordinateur ne peut pas mémoriser une image ou analyser un son comme nous. Pour qu'il puisse exploiter ces données, elles doivent être numérisées, c'est-à-dire que l'information qu'elles contiennent doit être convertie en nombres.

Comment numériser des données ?

1 Types de fichiers : exécutables ou données

Il existe deux grandes familles de fichiers informatiques.

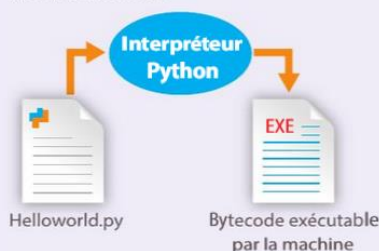
– Les fichiers « exécutables » : ils contiennent un programme informatique. Lorsque l'on charge un fichier exécutable, si le système d'exploitation le permet, cela déclenche un processus et l'exécution des instructions contenues dans le fichier. Sous Windows, par exemple, les fichiers portant l'extension « .exe » sont reconnus comme des exécutables.

– Les fichiers « de données » : ils contiennent des informations. Lorsque l'on essaie d'ouvrir un fichier de données, le système d'exploitation détermine grâce à l'extension du fichier le logiciel* capable de lire ces données (s'il est installé sur la machine) et l'exécute. Sous certains systèmes d'exploitation, un double-clic sur un fichier avec l'extension « .jpg » lance un logiciel qui permet d'afficher l'image numérique à l'écran.



Point science

Python est un langage de programmation multiplateformes. C'est l'interpréteur qui va traduire ligne par ligne le code en instructions spécifiques à la machine sur laquelle il doit s'exécuter.



– Un programme Python est une donnée.
– L'interpréteur le transforme en exécutable.

2 Le code ASCII

Le code ASCII fait son apparition dans les années 1960. L'objectif est de standardiser le codage des caractères. La table ASCII définit 128 caractères et leur fait correspondre un entier codé sur 1 octet (7 bits suffisent mais on ajoute un 0 en début pour compléter l'octet).

Caractère	Codage décimal	Codage binaire
@	64	01000000
A	65	01000001
B	66	01000010

Ainsi, le caractère « A » est associé à l'entier « 65 », c'est-à-dire « 01000001 » en binaire.

3 La taille des fichiers

Une photographie prise par un smartphone de 2019 a une définition standard d'environ 8 millions de pixels. En codant chaque pixel sur 3 octets, elle occupera 24 Mo sur l'espace mémoire. Pour limiter cet espace et faciliter les échanges, on applique un algorithme de compression qui permet de stocker cette même image en prenant moins d'espace (par exemple, 1,9 Mo pour cette image au format JPG). De la même manière, on stocke un fichier son de 4 minutes sur quelques mégaoctets au format MP3 et une vidéo de 1 h 30 sur quelques centaines de mégaoctets au format MP4. Il existe plusieurs algorithmes de compression et plusieurs formats.

- 1 Citer une extension possible pour chacun des fichiers suivants : son, image, vidéo, texte, exécutable.
- 2 Un texte codé en ASCII comporte 1 447 caractères. Calculer le nombre d'octets nécessaires à sa numérisation.
- 3 Donner un ordre de grandeur de la taille occupée en mémoire par une photographie numérique, une chanson de 3 min, une vidéo de 2 h.

* VOCABULAIRE

Logiciel : programme informatique qui apporte à l'ordinateur des fonctionnalités supplémentaires. Par exemple, le logiciel GIMP ajoute des possibilités de traitement d'image. Un logiciel est constitué de fichiers qui sont enregistrés dans la mémoire de l'ordinateur.



CORRECTION D'UN PROGRAMME

Un algorithme est considéré comme correct s'il fait ce qu'on attend de lui. Plus précisément, un algorithme prend en entrée des données et, après traitement, produit un résultat. L'algorithme sera correct s'il produit pour chaque donnée entrée le résultat escompté.

Comment détecter et corriger des *bugs* ?

Protocole

- ▶ Télécharger le programme python du doc. 1 ou le recopier dans votre éditeur de code.
- ▶ Exécuter le programme et observer le résultat dans la console Python.

L'algorithme choisi est le suivant :

- si le véhicule est un poids lourd et roule à plus de $90 \text{ km}\cdot\text{h}^{-1}$, il est en infraction ;
- sinon, si le véhicule est léger et roule à plus de $130 \text{ km}\cdot\text{h}^{-1}$, il est en infraction ;
- sinon le véhicule n'est pas en infraction.

Matériel

- Un ordinateur avec un logiciel capable d'exécuter un programme codé en Python 3.

1 Le radar automatique

Sur une autoroute, la vitesse pour les véhicules légers est limitée à $130 \text{ km}\cdot\text{h}^{-1}$ et pour les poids lourds à $90 \text{ km}\cdot\text{h}^{-1}$.



Un radar automatique possède deux capteurs :

- un premier capteur qui mesure la hauteur h du véhicule (si $h > 2 \text{ m}$, on considère que le véhicule est un poids lourd) ;
- un second capteur qui mesure la vitesse v du véhicule.

Le radar est programmé en Python.

```
def enInfraction(h, v):  
    if h > 2 and v > 90:  
        enFaute = True  
    if h <= 2 and v > 130:  
        enFaute = True  
    else:  
        enFaute = False  
    return enFaute
```

```
>>> h=1.8  
>>> v=111  
>>> enInfraction(h, v)
```

1 Exécuter le programme fourni et observer le résultat produit.

2 Proposer un jeu de données pour tester toutes les configurations possibles.

3 Indiquer si la fonction renvoie toujours la valeur attendue. Sinon, la modifier pour qu'elle soit correcte.

4 Prolongement : en consultant le site legipermis.com, améliorer le programme pour qu'il détermine automatiquement le nombre de points perdus et le montant de l'amende encourue.

Point historique

Le 4 juin 1996, la fusée européenne Ariane 5 explose en vol 36,7 secondes après son décollage.

Cet accident, qui a coûté plusieurs centaines de millions d'euros, est dû à un bogue (*bug* en anglais) dans le programme informatique d'Ariane 5.

Une partie du code a été recopié d'Ariane 4. Cependant, la nouvelle fusée est plus puissante et une variable stockant la vitesse de déplacement horizontal de la fusée n'a pas pu contenir la valeur mesurée (on parle de dépassement de capacité). Le programme a alors lancé la procédure d'autodestruction.

Cette situation a coûté beaucoup d'argent mais d'autres *bugs*, dans les domaines automobiles ou médicaux par exemple, peuvent entraîner la perte de vies humaines.



② Numérisation des données

Les savoirs du B0

Un ordinateur peut manipuler des données de natures diverses une fois qu'elles ont été numérisées : textes, images, sons. Les programmes sont également des données : 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.

Réponses aux questions

1. Extensions possibles :

- pour un fichier son : mp3, wav, ogg, etc. ;
- pour un fichier image : jpg, png, bmp, etc. ;
- pour un fichier vidéo : mp4, avi, mov, .etc. ;
- pour un fichier texte : txt, odt, etc. ;
- pour un fichier exécutable : exe.

2. (doc. 2) : En ASCII, chaque caractère est codé sur un octet, il faudra donc 1 447 octets.

3.

- Une photographie numérique occupe quelques Mo ;
- une chanson de 3 mn occupe également quelques Mo ;
- une vidéo de 2 h occupe quelques centaines de Mo.

③ Correction d'un programme

Les savoirs du B0

Un programme peut comporter jusqu'à plusieurs centaines de millions de lignes de code, ce qui rend très probable la présence d'erreurs appelées bogues (ou bugs). Ces erreurs peuvent conduire un programme à avoir un comportement inattendu et entraîner des conséquences graves.

Commentaires pédagogiques / différenciation

Cette activité requiert uniquement des connaissances de seconde : partie « Algorithmique et programmation », en particulier « Savoir utiliser les variables et les instructions élémentaires

(instruction conditionnelle) » et « notion de fonction ».

Réponses aux questions

1. L'exécution renvoie `False`.

2. Pour le couple (h, v) , on peut tester toutes les combinaisons suivantes : $(1.8, 80)$; $(1.8, 100)$; $(1.8, 140)$; $(2.1, 80)$; $(2.1, 100)$; $(2.1, 140)$

3. `enInfraction(2.1, 100)` ainsi que `enInfraction(2.1, 140)` renvoient `False` et pourtant, le véhicule est en infraction.

Fonction corrigée :

```
def enInfraction(h, v):
    if h > 2 and v > 90:
        enFaute = True
    elif h <= 2 and v > 130 :
        enFaute = True
    else :
        enFaute = False
    return enFaute
```

4. L'amélioration du programme pour qu'il détermine automatiquement le nombre de points perdus et le montant de l'amende encourue figure en haut de la page suivante. (Selon le site legipermis.com, à la date d'impression.)