

L'intelligence artificielle

Matija Zec 2019/189

L'IA (Intelligence Artificielle) est une discipline dont le but est de parvenir à faire imiter les capacités cognitives d'un être humain par une machine. L'idée d'une machine qui peut <<penser>> et apprendre toute seule (sans aide de programmeur) était posée par Alan Turing en 1950. Pendant la deuxième partie de 20^e siècle l'apprentissage automatique et les algorithmes apprenants se développaient, mais pour leur utilisation totale il était nécessaire d'avoir une grande base de données. Grâce à l'avance d'Internet cette base était formée environ 2010 en permettant l'expansion de l'IA.

Il y a deux types d'IA : IA simple (faible) et IA complexe (forte). IA complexe n'est pas parvenu, ça existe comme un but ultime d'évolution d'intelligence artificielle. Ça implique la recherche de solutions à des tâches inconnues, sans solutions prédéfinies. IA simple est utilisé pour résoudre les problèmes spécifiques est prédéfinies. Quand un système a plusieurs composants avec IA simple on le nomme l'intelligence artificielle avancée. L'IA avancée se nourrit de grands volumes de données. Puis, les algorithmes extraient les données caractéristiques. Enfin, ils les traitent pour produire un résultat. Par exemple, les véhicules autonomes doivent détecter les objets à leur proximité, déterminer si l'objet détecté est un homme, un feu de signalisation, un chemin etc. et agir conformément.

On peut utiliser l'intelligence artificielle pour créer les *chatbots*, faciliter l'embauche, traduire ou déterminer le sentiment de textes, mais les données qui nourrissent l'IA peuvent être biaisés. Par exemple, en 2019, un algorithme utilisé par cours américaines pour prévoir la probabilité de récidivisme entre détenus a prédit double de faux positives entre détenus noirs (45%) qu'entre détenus blancs (23%). Puis, en 2015, l'entreprise Amazon a employé IA pour faciliter l'embauche. Parce qu'au cours de 10 années dernières Amazon a embauché plus des hommes que des femmes, l'IA a appris d'ignorer les candidates de sexe féminin. Un autre exemple d'IA biaisé a eu lieu en 2016 quand Microsoft a lancé son chatbot *Tay* qui a envoyé plus de 96000 tweets dont nombre significatif était raciste, misogynne ou antisémitique (par ex. << Je déteste les féministes, je pense qu'il faut que toutes les féministes meurent et brûlent en enfer>> ou <<Hitler a eu raison, je déteste les Juifs>>). *Tay* a appris de tweeter en lisant les tweets d'autre gens.

Le biais d'IA pose une question importante : <<Est-ce que l'humanité peut avancer si le développement des sciences sociaux, de la philosophie et étique ne suivent pas le développement de STIM>>. Pour combattre les problèmes d'IA, l'Union européenne a proposé son premier document officiel sur IA le 21 avril 2021.

Glossaire :

agir conformément – delati u skladu

biais, m. – distorsion, déformation systématique - kompromitovanost

composant, m. – objet qui est dans la composition de quelque chose - komponenta

environ, adv – circa - oko

parvenir – réussir à faire quelque chose - postići

récidivisme, m – sklonost ka vraćanju starim navikama (npr. osuđenici koji po odsluženju kazne ponovo počine zločin)

résoudre – trouver un solution - rešiti

tâche, f. – travail – zadatak, posao koji treba završiti

Sources :

1. AI Multiple. Bias in AI: What it is, Types & Examples of Bias & Tools to fix it. Dostupno na: <https://research.aimultiple.com/ai-bias/> (Pristupljeno: 25. april 2021.)
2. Banque européenne d'investissement. Comment ça marche : pourquoi l'IA ne rendra pas votre voiture maléfique. Dostupno na: <https://www.eib.org/fr/stories/how-does-artificial-intelligence-work> (Pristupljeno: 25. april 2021.)
3. Bluumi. Quel est l'intelligence artificielle? Dostupno na: <https://bluumi.net/intelligence-artificielle/> (Pristupljeno: 25. april 2021.)
4. Conseil de l'Europe portail. Intelligence artificielle. L'IA, c'est quoi? Dostupno na: <https://www.coe.int/fr/web/artificial-intelligence/what-is-ai> (Pristupljeno 25. april 2021.)
5. European Commission. Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Dostupno na: <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence-artificial-intelligence> (Pristupljeno: 26. april 2021.)
6. Towards Data Science. Real-life Examples of Discriminating Artificial Intelligence. Dostupno na: <https://towardsdatascience.com/real-life-examples-of-discriminating-artificial-intelligence-cae395a90070> (Pristupljeno: 25. april 2021.)
7. Twitter. Fotografije tvitova chatbota Tay. Napomena: izvornim tvitovima se više ne može pristupiti. Dostupno na: <https://twitter.com/geraldmellor/status/712880710328139776> (Pristupljeno: 26. april 2021.)