MAURICE Alexandre

TD 1 TP 2

R1.11 - Bases de la communication

Exposé sur l'IA : L'IA doit-elle être *open-source* ?

B.U.T. Informatique, 1ère année, semestre 1

Table des matières

Texte :

L'intelligence artificielle est une technologie qui a commencé à se développer grandement depuis ces 20 dernières années, et elle est à présent omniprésente dans nos vies. On peut prendre l'exemple de Facebook, de Google ou encore Netflix qui utilise la puissance de ces réseaux de neurones pour nous proposer des contenus proches de ce que nous consommons habituellement, ou encore dans le développement croissant des voitures autonomes. Nous sommes tous habitués à cette idée d'être constamment surveillés par les multinationales qui emplissent le classement des sociétés les plus importantes, pour des raisons de suggestions préférentielles relatives au contenu, ou parfois, des usages bien plus controversés, comme la revente illicite de ces données (cf. le procès de Facebook). La question de la sécurité numérique est maintenant au cœur du débat public, troublé dans un monde trop surveillé mais à la fois trop peu sécurisé. Et si ces outils, à l'origine employés par des experts et les grandes multinationales étaient aussi à la portée de tous, nous sommes alors obligatoirement menés à une interrogation : l'IA doit-elle être *open-source* ? Pour répondre à cette question, nous allons d'abord discuter des problèmes que peut causer l'IA entre les mains d'une personne malintentionnée, pour ensuite conclure brièvement.

On observe depuis des années le développement d'outils très puissants permettant de réaliser des tâches demandées à une intelligence artificielle. On reste pour le moment loin d'une IA forte, mais le machine learning proposé sur des modules, tels que NumPy sur Python ou encore Tensorflow, proposé en open-source, est une extension extrêmement puissante permettant de faire des estimations de plus ou moins tout et n'importe quoi (simulation d'une balle, prédiction de cours, mesures…). Dans l'optique où quelqu'un de malintentionné voudrait se procurer de tels outils, il aurait juste à chercher sur le site, télécharger la librairie sur pip (le *package installer* de python), se renseigner sur la très vaste documentation et la pléthore de tutoriels trouvables sur le net. En l'espace de quelques jours, il peut développer un programme qui lui dispensera des moyens d'analyse de l'environnement virtuel ou réel. Dans un cas concret, on peut facilement imaginer un *hacker* qui dispose d'un micro et d'un système embarqué, se plaçant à la terrasse d'un café et essayant d'examiner la liste de touche tapées sur l'ordinateur de la personne assise à la table d'à côté. Le système embarqué, une sorte de mini-ordinateur, contient une puce qui permet d'assurer les calculs, et une banque de données correspondant à des enregistrements de milliers de pressions de touches. Grâce à cette banque

données, l'algorithme peut tenter de prédire quelle touche vous avez pressé grâce à une interpolation ou une extrapolation. Les mathématiques peuvent aussi aider : connaitre le nom de la langue que l'on est en train d'analyser permet de déterminer la fréquence d'apparition des lettres, et améliorent significativement le résultat. Ce scénario est plus que plausible, et il peut même être appliqué. Cependant, en conditions réelles, la fraude est quasi-irréalisable, il faudrait que le micro soit suffisamment sensible pour pouvoir capter tous les sons du clavier, qu'un algorithme très performant s'assure de la réduction du bruit, et que la prédiction soit juste. Mais dans le cas où le *hacker* réussit, il peut avoir accès à une information potentiellement sensible que vous avez saisie, et ce, sans que vous ne le sachiez. Ce scénario est évidemment bien alambiqué, mais d'autres cas d'utilisation bien plus probables peuvent être élaborés. Et dans un monde où l'on développe du matériel toujours plus performant, même ce script qui semble absolument improbable de nos jours pourrait être possible d'ici quelques années.

L'IA est aussi extrêmement puissante pour les attaques de force brutes, testant des mots de passe plus pertinents en fonction de l'environnement (logiciel/site web, type de ménages ciblés par la marque, informations personnelles…), mais permet aussi de trouver des cibles plus facilement attaquables, en fonction de critères similaires.

Au long de cette réflexion, nous nous sommes attardés sur des usages malveillants de l'IA. Il faut garder en tête que ceux-ci sont extrêmement marginaux, et l'IA, surtout en *open-source*, a bien plus permis des prouesses technologiques exceptionnelles qu'elle n'a causé du tort. Cependant, l'IA reste un domaine qui est très surveillé mais qui souffre encore de troubles juridiques. De ce fait, les attaques se multiplient, développant à vive allure la cybercriminalité, à coups de rançongiciels, obligeant l'industrie de la cybersécurité à se développer de plus en plus. En somme, personne n'est à l'abri d'une attaque informatique, et la responsabilité de protection revient aux utilisateurs et aux entreprises qui proposent des services numériques. Pour ce qui est de l'utilisation malveillante, dans le cas de l'*open-source*, le seul responsable est l'auteur de la fraude. Pour conclure, l'IA *open-source* n'est pas vraiment un problème en soi, de plus que quelqu'un d'extrêmement avancé pourrait très bien développer son propre réseau de neurones, mais cela facilite grandement l'accès à des méthodes puissantes surdéveloppées. Palier à ce problème est alors du ressort de l'industrie de la cybersécurité, que ce soit un organisme public ou privé.

Sources :

Fonctionnement du système de recommandations de Netflix :

https://help.netflix.com/fr/node/100639

Informations concernant le procès de Facebook :

https://fr.wikipedia.org/wiki/Critiques_de_Facebook#Vente_d'informations_personnelles

Informations concernant la désignation « *open-source* » :

https://fr.wikipedia.org/wiki/Open_source

Bibliothèque python NumPy :

https://numpy.org

Tensorflow, une plate-forme *open-source* développée par Google :

https://www.tensorflow.org/?hl=fr

Keytap et Keytap2, deux logiciels libres permettant de retrouver un texte grâce au son de pression des touches :

https://www.youtube.com/watch?v=2OjzI9m7W10

https://www.youtube.com/watch?v=Y8nWkdWl7Pg&feature=youtu.be

Démonstration du scénario évoqué dans le développement :

https://www.youtube.com/watch?v=XYZJPAzATbY

Informations concernant les attaques par la force brute :

https://fr.wikipedia.org/wiki/Attaque_par_force_brute

Informations concernant les rançongiciels :

https://fr.wikipedia.org/wiki/Rançongiciel

L'industrie de la cybersécurité, un secteur avec une croissance exponentielle :

https://www.statista.com/statistics/595182/worldwide-security-as-a-service-market-size/

Tutoriel pour développer son propre réseau de neurone :

https://fr.blog.businessdecision.com/tutoriel-machine-learning-comprendre-ce-quest-un-reseau-de-neurones-et-en-creer-un/