

COMMISSION DE RÉFLEXION ÉTHIQUE

PRODUCTIONS



Les enjeux de l'Intelligence Artificielle en Santé : Entre réalité, limites et perspectives

(CRE, Hôpital Saint Joseph, 4 décembre 2023)

Bonjour à tous. Le thème d'aujourd'hui se portera sur l'intelligence artificielle appliquée à la santé. Le but, c'est de comprendre un peu quelle est la réalité aujourd'hui ? Quelles sont les limites des systèmes ? Et quelles sont les perspectives pour le futur ? Je vais rentrer dans quelques points techniques, mais l'idée, c'est surtout de vous donner un goût aux enjeux éthiques qui sont soulevés par l'intelligence artificielle. Et ils sont très nombreux.

Peut-être qu'à la fin de cette présentation, vous aurez plus de clés en main pour avoir un recul critique sur ce qu'on appelle intelligence artificielle et sur les moyens de développer une intelligence artificielle qui soit efficace mais surtout responsable.

Alors, qu'est-ce que l'intelligence artificielle ?

Le terme intelligence n'est peut-être pas le plus adapté parce que, en français, intelligence est relié à une conscience et ça donne une dimension un peu effrayante.

Le terme, cependant, il vient de l'anglais et en anglais intelligence, ça témoigne plutôt d'une qualité à gérer des informations. Par exemple, CIA, c'est la Central Intelligence Agency et cela n'évoque pas une personne morale. Le principe de l'intelligence artificielle, c'est que c'est un apprentissage par l'exemple. Alors qu'une approche plutôt par compréhension, ça consiste à donner des données génétiques d'un patient et à essayer de comprendre à partir des données génétiques qu'est-ce qui va avoir un lien avec le développement de cette maladie ? En intelligence artificielle, on donne à la machine à la fois les données génétiques, mais aussi les données de qui a développé une maladie. Et ensuite on laisse l'algorithme s'optimiser tout seul. Et ce sera à lui de comprendre, sans qu'on l'ait guidé, qu'est-ce qui avait un lien entre les données génétiques d'entrée et la maladie de sortie ?

Lui, il va essayer de maximiser sa performance, c'est-à-dire la justesse de la prédiction à partir de tout ce qu'on lui donne. Le processus est itératif et il implique un très grand nombre

de patients. Une fois que l'algorithme est optimisé et qu'il est satisfait de sa prédiction, on dit qu'il est entraîné. Attention, il ne s'agit pas juste de stockage en mémoire. Le jour où l'algorithme rencontre un nouveau patient, il n'a plus accès à ces patients sur lesquels il a été entraîné. Il sait donc performer sur des situations qu'il n'a jamais vues. Pourquoi ? Parce qu'il a compris des choses dans les gènes d'entrée qu'on lui avait donnés, qui peuvent ou non avoir un lien avec le développement d'une maladie, et donc, quand il voit des nouveaux gènes, il va en déduire quelque chose sur la maladie du patient.

Cependant, ce n'est pas la même compréhension que celle d'un humain. Souvent, on a tendance à croire que l'algorithme comprend la même chose que nous, mais non, c'est rarement, voire jamais le cas. La compréhension d'algorithme est très différente de la nôtre. Alors on entend les mots « intelligence artificielle » de partout.

Pourquoi maintenant ? Parce qu'il y a certainement un phénomène de mode. Mais aussi on a plus de données qui sont disponibles, des compétences qui s'améliorent et des puissances de calcul qui deviennent de plus en plus importantes. Et donc, il y a des budgets qui sont mis en œuvre et qui permettent le développement de ce secteur.

En santé, il y a déjà eu des applications d'intelligence artificielle qui ont eu un grand succès. Je peux vous donner quelques exemples. Déjà, pour le diagnostic, on peut avoir un diagnostic rapide et précoce et automatique. Par exemple, on peut diagnostiquer des rétinopathies diabétiques à partir d'images de fonds d'oeil. On peut aussi avoir un diagnostic précis et adapté à chacun. Par exemple, on peut caractériser des tumeurs de manière très précise aux niveaux moléculaire et cellulaire, pour savoir au mieux quelle est la maladie du patient.

Ensuite, une fois qu'on a diagnostiqué, on peut traiter de manière personnalisée et adaptée. L'exemple célèbre ce sera pour les patients diabétiques qui ont des pompes à insuline. On va introduire de l'intelligence artificielle, où on aura des données nombreuses sur le patient qui permet une administration d'insuline qui est particulièrement précise et adaptée aux patients. Enfin, le dernier secteur, c'est la découverte de médicaments, où il y a eu des bonds qui ont été faits parce qu'on peut prédire mieux, grâce à l'intelligence artificielle, quel patient sera réceptif à une molécule ? Quelles molécules va interagir ? De quelle manière avec des protéines ?

Alors maintenant, si je m'adresse à vous personnel de santé, qui pouvez être confrontés à des solutions d'intelligence artificielle, et vous avez besoin d'avoir un regard un peu critique par rapport à cette solution. Qu'est-ce qu'il faut que vous regardiez en priorité ?

Il y a un proverbe en informatique qui s'appelle « garbage in garbage out ». La qualité des données est essentielle. Un algorithme qui a reçu des mauvaises données va donner des mauvais résultats, parce que l'intelligence artificielle, ce n'est rien d'autre qu'une compréhension de données d'entrée. Donc si vous donnez des mauvaises données, il ne comprendra pas bien. Cependant, les données ne sont jamais parfaites parce que c'est la réalité du monde et cela peut donner lieu à des problèmes importants.

Alors un exemple ? Il y a eu un système de montres connectées qui permettait de prédire des risques d'arrêt cardiaque. Le problème, c'est que ces montres étaient mal calibrées et ils ont commencé à prédire des arrêts cardiaques à tort et à travers. Et quand bien même les données sont ajustées, elles ne sont jamais complètes. Les données ADN par exemple, les séquences d'ADN, elles comportent toujours des lacunes. Donc il faut qu'on fasse avec. On fait avec des données qui sont peut-être moins bonnes que ce qu'elles devraient être. Enfin, le problème dont on parle souvent d'intelligence artificielle, c'est les biais. Donc les données, elles ont des biais à l'intérieur, c'est-à-dire qu'on est limité à ce que les données

contiennent. Et notamment, on va rarement avoir une représentation parfaite de la population mondiale dans les données qu'on donne à un algorithme.

Il y a eu des algorithmes qui prédisaient des mélanomes cutanés de manière très précise, mais que sur les peaux blanches. Et le jour, où on a donné une peau noire à l'algorithme, il n'a pas été capable de comprendre qu'il y avait un mélanome. Donc vraiment l'idée, c'est que les données qui sont données en entrée à l'algorithme doivent être le plus exhaustif possible.

Une fois qu'on a récolté les données, vous donc, en tant que professionnels de santé, on vous donne un algorithme qui a été optimisé de qui a, a priori, un très bon critère de performance puisqu'il a été optimisé pour maximiser ces critères de performance. Et vous, vous avez les résultats de l'algorithme. Et là vous êtes un peu coincés parce que vous ne comprenez pas quelles sont les règles de décision de l'algorithme. Pourquoi est-ce qu'il vous donne ce résultat ? On ne sait pas quel critère a été important pour lui, pour l'algorithme. Qu'est-ce qu'il a regardé en priorité parmi les données génétiques du patient pour pouvoir prédire qu'il y avait une maladie ? Vous ne savez pas quels exemples ont été utiles pour l'algorithme parmi tous les patients qu'il a vus, pour vous donner aujourd'hui le résultat qu'il vous donne.

Un exemple : c'est un algorithme qui avait été développé qui indiquait si un patient devait être hospitalisé ou rentrer chez lui. Quand le patient avait une pneumonie, il disait qu'il était hospitalisé. Quand il avait une pneumonie et de l'asthme, l'algorithme disait que le patient pouvait rentrer chez lui en toute sécurité. Pourtant, l'asthme, c'est un facteur aggravant. Donc il ne devrait pas pouvoir rentrer chez lui en sécurité. Alors, on a essayé de comprendre pourquoi il y avait ces prédictions erronées et on s'est rendu compte que dans les données patients qu'on avait données à l'algorithme au moment de son entraînement, les patients qui avaient eu de l'asthme avaient été très bien pris en charge et donc l'algorithme en avait déduit que l'asthme était un facteur qui diminuait la gravité de la pneumonie.

Donc, ce que je veux dire ici et qui rejoint ce que je vous ai dit tout à l'heure, c'est que la compréhension d'un algorithme est différente. Et s'il y a des biais dans les données d'entrée, il prend ces biais et il en fait une prédiction qui peut être erronée.

Mais alors, si la machine est en désaccord avec le médecin ou l'utilisateur de la machine, de manière générale, il y a une question du coup de responsabilité. Qui a le dernier mot ? Est-ce qu'on dit que c'est le médecin ? Pourquoi pas Mais est-ce que le but n'est pas aussi que la machine soit capable de comprendre des choses qui vont au-delà de la compréhension d'un être humain ? Donc, c'est difficile de savoir qui tranche. Enfin, du coup pour votre conduite de projet, une fois que l'algorithme a été développé, il y a un aspect qui est peu pris en compte, c'est le suivi de cet algorithme. Une fois lancée, est-ce qu'il y a une possibilité de stopper une application dangereuse ? De plus, un système informatique, il est voué à connaître des mises à jour, à être dynamique, à s'améliorer. Et dans ce cas-là, est ce qu'on va bien suivre que les mises à jour ne sont pas dangereuses ?

Aujourd'hui, la conduite d'un projet en informatique est faite un peu tête baissée. C'est-à-dire, on a en tête des opportunités, on réfléchit pas trop au risque, on développe. Et ensuite, une fois qu'on a développé, on va confronter notre algorithme pour essayer de comprendre où est-ce qu'il peut faire des erreurs. Alors qu'on pourrait, avant de concevoir l'algorithme, réfléchir à quelles sont les frontières, où je veux être sûr que l'algorithme n'ira pas explorer de ce côté-là. Par exemple, si je veux être sûr que mon algorithme ne peut pas recracher plus tard des données patients sur lesquelles il a été entraîné, et bien, je peux être spécifiquement construire un algorithme en ayant cette contrainte-là, c'est-à-dire qu'il ne puisse pas redonner des données patients telles quelles.

Alors maintenant, je vous propose d'élargir le point de vue et voir des séquences plus globales de l'intelligence artificielle sur la santé et sur la société. Je voudrais qu'on s'intéresse à des conséquences sociales, mais aussi éthiques de l'intelligence artificielle.

Alors aujourd'hui, il y a beaucoup de preuves de concept qui ont été développées. Certaines sont utilisées dans des hôpitaux un peu de manière isolée, mais il n'y a pas de consensus. Il n'y a pas de système généralisé qui soit utilisé à grande échelle et on va essayer de comprendre pourquoi. Pourquoi est-ce qu'il y a des freins qui font que, aujourd'hui, l'intelligence artificielle n'est pas développée à grande échelle dans la santé ?

Déjà, l'intelligence artificielle, elle fait intervenir des acteurs qui sont très différents et qui n'ont pas l'habitude de travailler ensemble. D'un côté, il y a des cliniciens qui sont méfiants parce qu'ils ont un système qu'ils comprennent pas forcément. De l'autre, des informaticiens qui déploient des choses, mais qui ne comprennent pas les enjeux cliniques et humains de leurs travaux parce qu'ils n'ont pas une connaissance du terrain. De plus, le système de santé connaît des réglementations très lourdes, et c'est bien, mais qui ne sont pas forcément adaptés à des systèmes informatiques.

Là, je vous donne l'exemple très connu de l'étude clinique. Donc une étude clinique, ça dure rarement moins d'un ou deux ans, mais dans le monde de l'informatique, un ou deux ans, c'est beaucoup. En un ou deux ans, vous pouvez avoir des avancées majeures qui font que les performances de votre système, elles sont obsolètes, elles sont mauvaises, elles ne sont plus du tout de la qualité de l'état de l'art.

Par exemple, si vous avez un ChatGPT ou un Deepface qui sort, et bien votre algorithme est périmé. De plus la réglementation, elle est statique. On va mettre une réglementation sur un algorithme à un instant T sans prendre en compte que, comme je vous le disais tout à l'heure, l'algorithme est voué à avoir des mises à jour, à avoir des corrections de bugs, il va se développer. Et est-ce que on refait une étude clinique pour toutes les mises à jour ?

Alors, il y a quand même des réflexions qui sont en cours par l'Union européenne notamment, qui sort des projets de loi en intelligence artificielle, notamment en santé. C'est important que on trouve un compromis qui ne freine pas la recherche mais qui en même temps, offre le cadre légal qui est quand même nécessaire.

Enfin, j'évoque un dernier frein à la généralisation de l'intelligence artificielle, c'est qu'il y a eu l'intervention d'acteurs privés très puissants qui jusque-là n'intervenaient pas en santé, notamment Google, Deep Mind, Facebook. Ces acteurs-là, on ne sait pas encore comment les faire interagir avec des systèmes publics de santé. Est-ce que on leur donne accès aux données de santé qui sont confidentielles ? D'un côté, on n'a pas envie, mais en même temps, ces acteurs-là, ils ont des puissances de calcul, des ingénieurs de grande qualité et ce serait dommage qu'ils n'aient pas le droit à développer des algorithmes performants sur des données de santé. Et donc cette mauvaise communication, elle induit des situations assez paradoxales où, par exemple, on observe des géants comme Apple ou Google qui détiennent un énorme nombre de données liées à la santé via leurs applications connectées, Apple Health, Google Fit, mais ils n'ont pas d'obligation de les transmettre à la recherche publique. Donc aujourd'hui, ils ont ces données-là et ils les gardent pour eux.

Enfin, dernière partie de cette présentation, ce sera sur les enjeux éthiques de l'intelligence artificielle qui existent déjà aujourd'hui, à la fois à l'échelle de l'individu et à l'échelle du groupe.

Je pense qu'à la sortie des premiers systèmes d'intelligence artificielle, il y avait un idéal philosophique qui était de diriger le patient vers plus d'autonomie et de liberté individuelle.

On imaginait un monde de demain où l'individu possède toutes les informations sur sa santé. Il connaît l'impact de chacune de ses décisions au quotidien. Il devient extrêmement responsabilisé, ça devient un acteur de sa propre santé et il est donc plus libre. Cependant, la réalité est tout autre. Aujourd'hui, il y a une dualité qui s'est créée entre, on appelle le moi réel, qui est donc ma personne, et le moi numérique qui est ma représentation numérique, et cette dualité, elle peut être presque schizophrène. D'un côté, on peut avoir un moi réel qui se sent en très bonne santé, mais son application connectée, ou le système de santé lui dit « Ah, mais tu n'es pas en bonne santé », alors même que lui se sent en très bonne santé. À l'inverse, on imagine un moi numérique qui dit que tout va bien alors que moi, je me sens fatigué, je me sens moins bien que d'habitude et pour autant, ce n'est pas aligné avec ce que je vois dans mon moi numérique. Et c'est étrange parce que le moi réel, c'est moi. Moi, j'ai une réalité physique, je reçois des informations de mon propre corps, je suis sûr de ce que je ressens. En revanche, le moi numérique, je ne sais pas d'où viennent les informations qu'il me donne et je ne sais pas non plus comment elles ont été calculées.

Donc si par exemple, il me dit qu'il faut que je marche pendant 10 000 pas. Qui a calculé que 10 000 pas été le meilleur nombre de pas à faire ? Et pourquoi ? Pourquoi pas 15 000 pas ? Pourquoi pas un marathon, un ultra marathon ? Et puis l'application qui me donne des conseils ne connaît pas toute ma réalité, toute ma complexité. Par exemple, si elle me dit de faire du sport et que la personne est une mère de trois enfants qui travaille, elle n'a peut-être pas le temps d'aller à la salle. Si elle me demande d'acheter des produits de bonne qualité, et bien peut-être qu'un étudiant n'aura pas le budget d'acheter des légumes tous les jours.

Donc le patient qui était censé avoir gagné en liberté, il devient au contraire soumis à une dictature numérique qui lui impose des règles qu'il ne comprend pas. Et encore plus angoissant, le patient ne sait pas quelles sont les conséquences de ne pas suivre ses règles. Et donc, il y a une grande anxiété qui se crée de ne pas savoir s'il doit ou non écouter la machine.

Et donc on se pose cette question un peu philosophique, est-ce que le fait d'être mieux décrit permet d'être plus heureux ? Mais aujourd'hui, en tout cas, on n'a pas vraiment le choix de refuser de se faire décrire, de refuser de se faire enregistrer. On n'a pas le choix de refuser de se faire diagnostiquer par une machine, par exemple.

Enfin, le dernier enjeu éthique dont je souhaite parler est très souvent négligé. Et pourtant, je pense qu'il est très important. Il s'agit de la protection des groupes.

Donc le principe même de l'intelligence artificielle, c'est à partir d'un grand nombre, de faire des groupes, de comprendre des tendances, de rassembler des gens. Aujourd'hui, on protège beaucoup l'individu et les données personnelles, mais personne ne protège les groupes. Et si les résultats d'une étude complètement publique, puisque les résultats des études sont publics, accusent toutes les personnes aux cheveux roses et aux yeux bleus d'avoir plus de chances de développer un cancer, qui peut empêcher une assurance de vous faire payer plus cher parce que vous avez les cheveux roses et les yeux bleus ?

De plus, la description des groupes est très inégale. Si vous regardez autour de vous et que vous cherchez des personnes qui ont besoin des systèmes de santé, vous imaginez des sans-abris, des personnes âgées, des personnes avec des problèmes psychiatriques. Puis ensuite, vous arrivez dans un laboratoire et vous regardez les données numériques que vous avez. Ces personnes-là ne sont pas représentées. Aujourd'hui, on améliore toujours plus la santé de populations très privilégiées, jeunes, souvent européennes ou américaines, qui ont accès au

numérique. Mais on laisse de côté des populations plus précaires et qui ont moins accès au numérique.

J'arrive donc à la fin de cette présentation. J'espère vous avoir convaincu à la fois des grandes opportunités de l'intelligence artificielle avec des magnifiques avancées, des promesses très excitantes, mais aussi des risques et enjeux de ce domaine qui est très récent.

Aujourd'hui, je pense qu'il est nécessaire qu'il y ait une réflexion groupée qui fasse intervenir à la fois des professionnels de santé et des professionnels de l'informatique, mais aussi des professionnels de la loi, peut-être des représentants des populations, afin de se tourner vers une intelligence artificielle qui soit plus responsable tout en étant toujours très efficace.

Manon CHOSSEGROS

Sorbonne Centre for Artificial Intelligence (SCIA) - APHP

