

L'Intelligence Artificielle : quelles opportunités pour les industries agroalimentaires

Prof. Antoine CORNUEJOLS,

Dpt Modélisation Mathématiques Informatique Physique, UMR AgroParisTech Inrae _ Paris Saclay

Au cours des dernières années, le terme « intelligence artificielle » est devenu omniprésent dans les médias. Dans cet article, nous expliquons d'où vient cette révolution et ce que font les algorithmes existants en apprentissage automatique. Nous examinons ensuite l'impact que nous pouvons en attendre à court terme dans les industries agroalimentaires. De la production agricole aux modes de consommation en passant par la transformation et les chaînes logistiques, tous les maillons peuvent profiter d'une optimisation accrue et de nouvelles possibilités de gestion intelligente et d'aide à la décision. Nous concluons en soulignant les défis à relever pour tirer pleinement parti du potentiel de l'intelligence artificielle, c'est-à-dire d'une intelligence augmentée.

ABSTRACT

In the last few years, the term «artificial intelligence» has become ubiquitous in the media. In this paper, we explain where this revolution comes from and what existing algorithms in machine learning do. Then we examine the impact that we can expect in the short term in the agri-food industries. From agricultural production to consumption patterns, processing and supply chains, all links can gain from increased optimization and new possibilities for intelligent management and decision support. We conclude by highlighting the challenges that need to be addressed in order to benefit from the full potential of artificial intelligence, i.e. augmented intelligence.

1. INTRODUCTION

Depuis un peu plus d'une décennie, une révolution est en cours qui va affecter profondément nos manières de vivre, de produire et de consommer. Plusieurs évolutions technologiques ont convergé en effet pour ouvrir tout un champ d'opportunités, dont nous ne faisons encore qu'entrevoir les potentialités.

La première évolution concerne **les données**. En 1982, seuls 2% des informations glanées par l'humanité au cours de son histoire étaient disponibles sous forme numérisée. Il y a trente ans environ, les entreprises commençaient à rassembler leurs données dans des systèmes de gestion de bases de données, et les chercheurs en intelligence artificielle et en apprentissage automatique étaient heureux quand ils pouvaient tester leurs algorithmes sur quelques centaines de données. Il y a vingt ans, l'émergence d'Internet a permis l'échange à grande échelle d'informations et s'est accompagnée de la numérisation massive des informations : images, vidéos, textes, musique, données génomiques, etc. au point qu'aujourd'hui, il est estimé que c'est plus de 98% des données produites qui sont numérisées. La dernière décennie a vu se multiplier les moyens de récolte de données : smartphones, plates-formes numériques, réseaux sociaux, capteurs en tous genres, et maintenant l'internet des objets, et l'on a pu parler d'« avalanche de données ».

Dans le même temps, les ressources en **puissance machine** et en espace mémoire se sont énormément accrues. Un ordinateur portable d'aujourd'hui a la même puissance qu'un super-ordinateur des années 1980. Et si cela ne suffit pas, le cloud computing donne accès potentiellement à des milliers de machines.

Finalement, la recherche en intelligence artificielle et en apprentissage automatique (ce que les médias popularisent sous le terme « machine learning ») a inventé, mis au point et fourni en accès libre (une exception très notable par rapport à d'autres

disciplines) de **nouveaux algorithmes** d'apprentissage automatique et d'aide à la décision qui ne sont plus des jouets mais bien des outils très puissants dès lors que suffisamment de données sont disponibles pour les entraîner.

2. L'INTELLIGENCE ARTIFICIELLE : QU'EST-CE QUE C'EST ?

Mais d'abord finalement, que font ces algorithmes ? Comment transforment-ils des masses de données en connaissances potentiellement exploitables ?

Ici, il est utile de distinguer plusieurs types d'algorithmes et d'apprentissages :

- l'apprentissage descriptif
- l'apprentissage prédictif
- l'apprentissage prescriptif

L'**apprentissage descriptif** vise à re-décrire les données fournies pour aider l'expert à identifier des régularités. En un sens, le but est de compresser l'information. Par exemple, partant de données récoltées sur des milliers ou des millions de clients, un algorithme d'apprentissage descriptif pourra fournir un résumé de ces données en proposant une catégorisation de ces clients, c'est-à-dire une partition en sous-groupes distincts qu'un expert pourrait ensuite interpréter (e.g. les retraités aisés, les étudiants dépendant de leurs parents, ...). C'est ce que l'on appelle du « clustering » parce que l'on recherche des groupements d'exemples, ou nuages de points. Un deuxième type d'apprentissage descriptif cherche des « motifs fréquents » dans les données. Par exemple, il se pourrait qu'à partir de tickets de caisse et d'un croisement avec des données météo, un algorithme note qu'il est fréquent qu'un week-end estival s'accompagne d'achats important en merguez, en moutarde et en assiettes en carton.

L'**apprentissage prédictif** cherche des dépendances fonctionnelles entre la description d'un exemple, e.g. un client, un produit alimentaire, et une prédiction relative à cet exemple, e.g. le client remboursera son crédit à tel terme, le produit alimentaire sera apprécié ou pas par telle catégorie de consommateurs. L'apprentissage prédictif a connu récemment un développement spectaculaire dans la reconnaissance d'images. Les réseaux de neurones profonds excellent en effet à associer à une image une « étiquette », par exemple reconnaître ou pas la présence d'une tumeur à partir d'imagerie cérébrale. L'apprentissage prédictif est également employé pour filtrer la littérature scientifique et ne présenter à un spécialiste que les articles portant sur un certain

sujet. Il faut pour cela qu'une base d'entraînement comportant suffisamment d'exemples d'articles pertinents et non pertinents ait été disponible pour que l'algorithme apprenne à faire cette distinction. Une contrainte importante de l'apprentissage prédictif est qu'il nécessite des bases comportant de nombreux exemples « étiquetés », ce qui nécessite souvent un expert, et donc a un coût.

L'apprentissage prédictif peut être très performant en identifiant seulement des corrélations entre une description et une étiquette. Souvent, cependant, le but recherché n'est pas seulement de prédire, mais est aussi de savoir sur quels facteurs agir pour modifier un résultat. Or, corrélation n'est pas causalité. Observer un baromètre peut être suffisant, jusqu'à un certain point, pour prédire le temps du lendemain. Changer la position de son aiguille ne permettra pas pour autant de le modifier. De même, identifier des corrélations ne suffit pas à répondre à des questions sur ce qui pourrait arriver si ... Par exemple, quel serait l'impact sur mes ventes si je doublais le prix du produit ? La découverte de liens de causalité est l'objectif de l'**apprentissage prescriptif**. C'est beaucoup plus exigeant que la découverte de corrélations et est encore du domaine de la recherche.

UN CHANGEMENT DE PARADIGME

Le foisonnement des sources de données accompagné de la mise au point de nouvelles techniques d'apprentissage automatique et de la disponibilité désormais de ressources en calcul et en mémoire considérables conduit également à un changement de paradigme dans l'étude du monde. Dans le paradigme classique, on commençait par construire une hypothèse, par exemple le capital culturel est le principal facteur de reproduction des élites. On concevait alors un plan d'expérience susceptible de tester la validité de l'hypothèse. Ce plan d'expérience permettait de récolter un volume limité de mesures qui ne servaient qu'à tester l'hypothèse émise. Dans le nouveau paradigme, la démarche est de demander aux machines de découvrir toutes les corrélations multi-variables existantes dans un énorme volume de données souvent bruitées, puis seulement ensuite d'examiner ce qui peut présenter un intérêt dans cette masse de liens potentiels. De manière alternative, on peut demander aux machines de détecter ce qui émerge comme étant la norme et, à partir de là, d'identifier des « signaux faibles », c'est-à-dire des phénomènes étranges, hors norme, qu'il peut être intéressant d'examiner. De plus, les corrélations ainsi découvertes peuvent à leur tour servir d'entrées pour d'autres mécanismes de « data mining », participant ainsi à un processus d'enrichissement (ou de pollution) cumulatif et potentiellement exponentiel.

Il est en tous les cas essentiel de bien réaliser que d'un questionnement orienté et raisonné, on passe avec le « Big data » à une exploration tous azimuts de corrélations ou de signaux faibles ou de tendances, pour ensuite les filtrer, les recouper, et alimenter l'univers numérique. Il n'y a plus en pratique de question de taille d'échantillon, et, de plus, les données ne servent plus seulement à répondre à une question pour laquelle elles ont été récoltées, mais elles sont ré-utilisables à l'infini en fonction de nouveaux traitements que n'importe quel « data scientist » qui en dispose peut imaginer. Pour reprendre cet exemple, les sciences humaines et la sociologie deviennent des sciences quantitatives grâce à l'analyse des réseaux sociaux et de l'usage massif des Smartphones et autres objets connectés. On parle désormais de « physique sociale », de « fouloscopie », etc. avec des questionnements et des découvertes inenvisageables avant l'avènement du big data.

3. QUELLES CONSÉQUENCES POUR LES IAA

Le foisonnement incroyable des applications faisant appel à de l'intelligence artificielle dans tous les domaines : diagnostic, recommandations, prévisions, maintenance prédictive, etc., les annonces pour demain ou après-demain à propos des véhicules autonomes, des assistants personnels encore plus intelligents, et autres smart cities, donnent le tournis et ne laissent place qu'à une certitude : l'avenir sera encore plus surprenant que ce que nous attendons. L'imagination humaine, bientôt assistée par l'intelligence artificielle, est sans limite discernable.

Pour autant, il est intéressant de chercher à identifier des grands domaines des IAA dans lesquels l'intelligence artificielle va jouer un rôle dans un avenir proche.

En partant de l'amont, l'apprentissage automatique joue désormais une place importante dans *l'étude du monde*. L'analyse des relations entre plantes, sol et climat par exemple reposera de plus en plus sur une analyse assistée par l'intelligence artificielle. De même, l'exploitation des images satellitaires avec recoupement avec des données terrain fait maintenant intervenir de manière massive des techniques telles que les réseaux de neurones profonds.

Les *exploitations agricoles* connaissent une mutation rapide avec le recours croissant à des ressources numériques et à la robotisation. Les étables entièrement automatisées font maintenant partie du paysage. Bientôt, des capteurs placés sur les animaux, mesurant l'infrarouge et les mouvements, permettront de diagnostiquer leur état en temps réel et d'adapter et d'optimiser leur alimentation et leur traitement par un vétérinaire. En Chine, des capteurs sonores et de la reconnaissance vocale sont en test dans les élevages de cochons pour identifier le cri des porcelets menacés d'écrasement par leur mère. Le machinisme agricole couplé avec des capteurs dans les champs, sur les tracteurs et par drones et/ou satellites, permet d'optimiser l'import des intrants et l'utilisation de pesticides. La reconnaissance automatique des insectes ravageurs et la prédiction de leur arrivée permettront également une meilleure protection des productions agricoles et des arbres tout en étant plus respectueux de l'environnement. Mais l'intelligence artificielle ira au-delà de systèmes de reconnaissance et de prédiction. Elle apportera aussi des outils d'aide à la décision combinant plusieurs sources de données et d'expertise, l'analyse d'historiques et des prévisions sur les marchés, la météo et des informations sur les besoins et goûts des consommateurs pour éclairer les choix possibles en terme de plantations et de calendrier de traitements et de récolte. Déjà en Inde, des assistants numériques de partage d'information et de conseil permettent aux petits agriculteurs d'orienter leur choix de production et de date de récolte en fonction des prix des marchés locaux et de la météo.

La conception de *nouveaux aliments* par des algorithmes explorant des combinaisons originales d'ingrédients et de saveurs va certainement se développer avec la recherche par les consommateurs de sensations nouvelles et la nécessité aussi de valoriser de nouvelles ressources, comme les insectes pour fournir des protéines. Le *génie des procédés* est également concerné. Les laboratoires de recherche étudient de nouvelles méthodes d'intelligence artificielle pour optimiser les ressources et la durée de transformation de produits tels les fromages. Comme dans d'autres domaines, la maintenance prédictive des appareils et machines réduira les coûts et les délais de maintenance.

Les *chaînes logistiques* seront aussi impactées par la disponibilité de capteurs à très bas coût. La traçabilité des produits va s'en trouver transformée et il va devenir possible, par suivi en temps réel de l'état des

produits frais, d'adapter par exemple leur destination et/ou leur date limite de consommation en fonction des circonstances rencontrées pendant le transport. En croisant l'historique des capacités et d'utilisation des chaînes de transport et des lieux de stockage avec des données de saisonnalité et des prévisions sur la demande, les moyens de transport, la main d'œuvre et l'emploi des entrepôts pourront être optimisés et les goulots d'étranglement évités.

Finale­ment, l'intelligence artificielle pourrait aider à réduire le gaspillage alimentaire en prévoyant la demande des consommateurs en fonction des circonstances : fêtes, saisons, météo, etc. À tous les niveaux, de la production à la chaîne de transport et de distribution, les stocks pourront être mieux ajustés. La restauration elle-même pourrait, par l'utilisation de méthodes prédictives, réduire le gaspillage en anticipant les achats et la préparation des mets en fonction de prévision sur les attentes des clients.

Ce panorama d'applications de l'intelligence artificielle existantes ou en développement ne donne qu'un aperçu de ce qui va être inventé et utilisé à large échelle dans les années à venir. La numérisation des informations, leur partage possible et les traitements par l'intelligence artificielle pour en exploiter les potentialités offrent des possibilités immenses à qui sait imaginer les services non encore attendus mais qui deviendront évidents une fois proposés.

4. QUELS DÉFIS DEVANT NOUS

Toutes les promesses décrites ci-dessus s'accompagnent également de défis pour les concrétiser.

D'abord, ce qui est clair c'est que l'on est en train d'assister à un transfert massif de pouvoir des acteurs économiques qui maîtrisent les techniques et les procédés de fabrication ou de services vers ceux qui maîtrisent l'information, c'est-à-dire qui détiennent les données et savent en tirer des régularités exploitables et des prédictions. Des entreprises très jeunes, dont particulièrement les GAFA (Google, Amazon, FaceBook et Apple) ont détrônées les acteurs traditionnels grâce au fait qu'elles récoltent énormément de données sur les utilisateurs, leurs machines, leurs comportements, et peuvent ainsi devenir les vrais donneurs d'ordre reléguant les autres entreprises à de la sous-traitance.

La souveraineté des données et des algorithmes est un enjeu majeur. En particulier, l'adoption par les agriculteurs des nouvelles technologies de récolte et de traitement des données dépendra certainement beaucoup des cadres réglementaires et des solutions proposées pour que les données concernant leur exploitation restent soumises à leur contrôle. Mais ceci sera également vrai des autres acteurs des IAA.

Cependant, ce sont les défis en termes de *compétences* qui sont prééminents et vont conditionner l'avenir du « Big data ».

La connaissance des nouveaux outils de stockage et de traitement des données est nécessaire, mais c'est surtout la compréhension des problèmes liés à l'exploitation de données multi-sources, très hétérogènes dans leurs formats et leur fiabilité, et la maîtrise des techniques d'intégration de données et de ré-expression qui vont être déterminantes. De même, explorer massivement les corrélations potentielles et les signaux faibles demande des experts qui savent utiliser les méthodes adéquates, mais aussi, de manière tout aussi cruciale, sont conscients des risques de découverte de régularités fortuites sans signification réelle et savent comment s'en prémunir. La recherche de « causalités » au milieu des corrélations est encore un problème de recherche, mais devra également faire partie bientôt des compétences à maîtriser par un « data scientist ». Pour finir, ces spécialistes devront aussi être informés de tous les aspects juridiques et éthiques, ainsi que des problèmes de sécurité, liés à l'exploitation de données massives comprenant souvent des données individuelles.

Les estimations sur les besoins en spécialistes de ce type sont pharamineuses, se chiffrant par exemple à plus de 100 000 en France dans les 6 prochaines années et à plus d'un million aux Etats-Unis. Même si ces chiffres sont peut-être surestimés, ils donnent une idée de l'urgence de former des ingénieurs en nombre très significatif sur ces métiers, ainsi sans doute que d'organiser une formation continue adaptée.

5. CONCLUSION

L'intelligence artificielle (IA) est partout dans les médias. Peut-être trop, et peut-être pas assez, du moins pas là où il faudrait. La victoire d'AlphaGo sur le champion du monde de go, la voiture autonome promise, le mythe de « terminator » en ligne de mire, dévient le regard vers une intelligence artificielle qui pourrait nous remplacer, et vers des questions : la nature de la conscience, le libre-arbitre, la possibilité d'une super-intelligence, qui, aussi intéressantes soient-elles, nous éloignent d'une perspective juste sur la place de l'IA dans la société de demain.

L'IA est d'abord un ensemble de méthodes et d'outils pour une intelligence augmentée. Elle est en train de trouver une place dans toutes nos activités, pour recommander, prédire, conseiller, optimiser des processus complexes. Bien conçue, bien utilisée, l'IA offre des opportunités inégalables pour rendre la production agricole plus respectueuse de l'environnement, améliorer les processus de production et inventer de nouveaux produits, rendre plus efficaces les chaînes logistiques et aider les consommateurs à consommer mieux.