Transcription de la conférence d'Yves Meyer au sujet de l'histoire de l'intelligence artificielle, mai 2025 [1]

Donc je vous remercie d'être venu si nombreux et je vais vous parler d'intelligence artificielle en montrant que c'est finalement une démarche naturelle que l'humanité est en train de prendre. Et donc je vais faire un léger recul historique pour montrer comment les méthodes d'intelligence artificielle se sont imposées dans certains domaines. Alors mon message c'est pour susciter votre attention, mais ceci est impoli parce que votre attention est pleine et vigoureuse. Je pose la question : "Et si les incroyables succès de l'intelligence artificielle venaient de ce que l'intelligence artificielle ne sait pas penser ?". Alors on va voir, on va donner un sens précis à ces mots. Donc mon plan de l'exposé, je vais commencer par le cas des voitures autonomes parce que ça va très bien expliquer le débat. Le problème de la vision humaine et de la vision de l'ordinateur (computer vision) avec les grands noms du sujet Santiago Ramón y Cajal, David Marr, David Hubel et Törsten Wiesel. Ensuite, on va aborder le deep learning, l'intelligence artificielle d'aujourd'hui, Yann Le Cun et Stéphane Mallat et ensuite les bases méthodologiques par Vladimir Vapkin. Et enfin, on terminera par l'apprentissage non supervisé qui est la voie de l'avenir. Donc voilà le plan de l'exposé en 5 points.

Alors, je commence par les voitures autonomes parce que là, d'abord le charme c'est que tout le monde va comprendre et ça va vraiment introduire le sujet. Donc on verra que dans l'exemple des voitures autonomes, les verrous, les problèmes difficiles, c'est la théorie de la vision et que ça va nous amener tout naturellement à l'intelligence artificielle, c'est que les avancées scientifiques sur le fonctionnement de la vision humaine ont inspiré la construction des réseaux neurones profonds qui sont au cœur de l'intelligence artificielle. Donc on va voir comment tout ça, bien que ça soit artificiel, est au fond très naturel et comment on ne pouvait pas faire autrement si j'ose dire. Alors donc, un robot, muni de tous les capteurs envisageables, saurait-il conduire une voiture en appliquant rigoureusement des règles soigneusement définies ? Donc ça c'est la logique et l'intelligence. Donc en 2004, c'est il y a maintenant 21 ans, la DARPA (la DARPA est une division de la défense américaine qui subventionne des programmes d'avant-garde) a lancé une compétition dont les règles étaient les suivantes. La route tracée dans le désert, longue de 200 km, comportait des obstacles. 15 véhicules autonomes sans chauffeur et sans assistance humaine (cel signifie qu'ils ne sont pas téléguidés) ont participé à la course.

L'approche utilisée en 2004 était du type système expert. Alors ça c'est très important. Une approche du type système expert, ça veut dire que l'ordinateur définissant la conduite dispose des données recueillies à chaque instant par les capteurs embarqués et leur applique une démarche strictement logique pour prendre une décision. Donc c'est la conduite intelligente en suivant des règles d'inférence. C'est l'intelligence dans son aspect logique implacable. Aucun des véhicules n'est allé plus loin que 15 km. C'est-à-dire qu'il était impossible de conduire en faisant appel à son intelligence artificielle en ce sens et à la seule logique. Ça s'est avéré impossible. Voilà les véhicules qui ont participé à cette compétition. Vous voyez, ils sont pleins de capteurs et cetera, ils sont très laids. Et là, un véhicule en marche.

1. Transcription par Downsub de la conférence visionnable à l'adresse : https://www.youtube.com/watch?v=Fs0ZeSS7eZ4 .

Donc la route traversait un désert, elle était tracée, elle traversait un désert qui allait de la Californie au Nevada. Donc ça c'est un document en anglais, c'est le témoignage de la DARPA et il dit exactement ce que je viens de vous dire. Il y avait 1 million de dollars qui était proposé en prix pour le gagnant de cette course automobile complètement autonome. Mais aucun des véhicules n'est arrivé à faire plus de 15 km pour conduire de façon autonome. Et alors, d'où venait le problème ? C'est ça le point passionnant. Le problème vient de la théorie de la vision. Les capteurs arrivaient à voir, à fabriquer des images, mais il était impossible pour le système expert d'*interpréter* ces images. C'est parce que voir, ce n'est pas seulement utiliser sa rétine, c'est interpréter l'image, voir quel est le contenu de l'image. Donc dans ce cas-là, une tache noire sur la route est-elle une ombre ou un rocher ? Donc c'est ça qui a fait que tous les véhicules se sont arrêtés et ne pouvaient pas prendre de décision. De façon logique, il était impossible de prendre une décision.

Les véhicules n'ont fait que 15 km. Donc l'année d'après, la conduite automatique a été confiée au deep learning, ou apprentissage profond, une méthode d'apprentissage artificiel, tous les véhicules ont réussi le test. C'est-à-dire que quand on conduit, en réalité, on n'utilise pas uniquement des règles logiques, on conduit par instinct en imitant des situations de conduite qu'on a déjà expérimentées. Donc ça, c'est exactement ce que fait l'intelligence artificielle. Elle procède par analogie, non pas en suivant des règles logiques, mais en comparant ce qu'elle voit à une situation déjà rencontrée qui lui ressemble. Donc il y a cette notion de proximité. Donc je résume, dans l'approche système expert, l'ordinateur déduit sa conduite en appliquant logiquement des règles fixées à l'avance. Dans l'approche deep learning, l'ordinateur compare ce qu'il voit à des situations déjà rencontrées et induit sa conduite en extrapolant à partir de ces exemples. Dans ce dernier cas, l'ordinateur, en plus, apprend et perfectionne son logiciel en temps réel. On va le voir.

Donc je résume encore. L'interprétation correcte des scènes enregistrées par les capteurs est le problème le plus difficile rencontré dans la construction des véhicules autonomes. Il faut décider si une tache noire est une ombre ou un rocher. Dans le premier cas, le véhicule continue à rouler. Dans le second cas, il ralentit. Alors donc à l'aide de cet exemple, la DARPA a conclu qu'il fallait construire des voitures autonomes, qui auraient beaucoup moins d'accidents etc. Mais il y a eu un drame. Donc l'étude de ce drame est très intéressante. Le drame s'est passé 14 ans plus tard comme vous le voyez. Donc un véhicule, une voiture Uber autonome, a tué une femme dans une rue de la ville d'Arizona qui s'appelle Tampe.

Et qu'est-ce qui s'est passé ? Ce qui s'est passé, j'ai analysé l'accident de façon détaillée, c'est que la femme a traversé la route la nuit, dans un endroit qui n'était pas un passage clouté. Voilà. La voiture était en mode autonome. Néanmoins, il y avait une conductrice pour assurer la sécurité, parce que c'était un véhicule qui expérimentait la conduite autonome. Ce n'était pas un taxi, il y avait une conductrice dans le véhicule. Mais il semblerait, d'après l'enquête policière, que la personne qui devait assurer la sécurité faisait autre chose, elle regardait son téléphone portable.

Alors maintenant, quelle est la situation au jour d'aujourd'hui ? Donc là, j'ai vérifié avec mon ami Coifman auquel j'ai téléphoné samedi dernier et qui est professeur à Yale. Donc à l'heure actuelle, vous avez des taxis qui sont autonomes de niveau 4. C'est-à-dire que ce sont des taxis sans chauffeurs mais ils se déplacent toujours dans la même ville et en suivant le même itinéraire. Donc pour

eux, ça assure une sécurité et ils opèrent à Los Angeles, Phoenix en Arizona et à San Francisco. Donc là, il n'y a plus d'accident.

Alors maintenant, nous allons faire un pas en arrière et voir comment les méthodes, les algorithmes d'intelligence artificielle ont été construits.

Donc comme vous le voyez sur cette photo (aparté : c'est Proust qui remarquait que sur les photos d'une époque donnée, tous les gens se ressemblaient). Au XIXe siècle, un savant avait à peu près cette tête. Or, il faut penser que Santiago Ramón y Cajal était beaucoup plus jeune sur cette photo que je ne le suis à l'heure actuelle. Vous voyez, on est déterminé par son époque d'une façon catastrophique. Un savant au XIXe siècle avait cette tête. Donc Santiago Ramón y Cajal était un médecin et il a consacré sa vie à l'étude du cerveau. Donc il n'a pas du tout pris des cerveaux humains, il a pris des jeunes poussins et il a fait une analyse de leur cerveau en se servant d'un révélateur de neurones et donc qui noircissait les neurones qui avait été inventés par un italien qui s'appelle Golgi. Et donc là, vous avez des dessins faits par Santiago Ramón y Cajal et c'est un excellent dessinateur, les dessins sont splendides. Et donc il a compris le premier le fonctionnement du cerveau. Il a compris que le cerveau était composé de neurones qui sont terminés par des synapses.

Ces synapes sont des portes qui peuvent être ouvertes ou fermées et elles s'ouvrent quand il y a un flux suffisant qui traverse le neurone. Alors, les neurones sont connectés à d'autres neurones lorsque les synapses sont ouvertes et un neurone est connecté à peu près à 100 000 autres neurones de telle sorte qu'un neurone reçoit aussi d'énormes connexions et quand le flux qui arrive dans le neurone est suffisamment élevé, la synapse s'ouvre et le flux passe. Donc pour chaque synapse est défini un certain seuil de transmission de l'information. Quand le seuil est atteint, l'influx nerveux traverse la synapse. Alors, ce qui est remarquable, c'est que Ramón y Cajal fit ses découvertes 50 ans avant que les synapses puissent être vu grâce au microscope électronique. Donc, en réalité, ce n'est pas de ses observations, mais c'est par son imagination qu'il a inventé les synapses, il a vu les synapses avec les yeux de l'esprit. C'est absolument extraordinaire. Donc il a eu le prix Nobel pour cette découverte en 1932 je crois. Alors ce qui est amusant pour l'histoire du prix Nobel c'est que la théorie de Ramón y Cajal était absolument contraire à la vision traditionnelle, qui disait que le cerveau représentant la personnalité, l'âme, il était forcément d'un seul tenant. Donc il ne pouvait pas être une un graphe discontinu et l'homme qui défendait la théorie contraire, c'est cet italien qui s'appelle Golgi.

Et le jury du prix Nobel n'a pas eu le courage de donner raison au seul Santiago Ramón y Cajal. il a donné le prix Nobel aux deux défendant à la fois les versions unitaire et multiple du cerveau. Alors Ramón y Cajal a compris que nos pensées ne sont pas autre chose que des arborescences de neurones, arborescences qui ensuite disparaissent, s'éteignent au profit d'autres arborescences, qui sont d'autres pensées, et par là même, il découvrit la plasticité du cerveau. Donc il savait ça à la fin du XIXe siècle. C'est ça qui est extraordinaire. Alors maintenant, j'arrive à un deuxième homme, qui est David Marr, et qui est complètement remarquable. Il a vécu 35 ans. C'était un neurobiologiste anglais intéressé d'abord par une théorie générale de la cognition. Il devint ensuite un spécialiste de la vision humaine. Alors en 77, David Mar rejoint le MIT.

Et pourquoi Marr a-t-il été invité au MIT ? Il a été invité au MIT parce que les chercheurs de

robotique du MIT n'arrivaient pas à construire un robot qui puisse utiliser ses capteurs pour se déplacer. Donc, on en revient au problème des véhicules autonomes. C'est-à-dire qu'il n'y avait aucune possibilité pour un robot d'interpréter ce qu'il voyait pour décider de son déplacement et ne pas se cogner. C'est à nouveau le problème majeur de comprendre le contenu d'une image. Alors David Marr mourut d'une leucémie 3 ans après et son livre qui est un livre sublime que je connaissais presque par cœur, c'est *Vision et computational investigation the human representation and processing of visual information.* Donc c'est un livre qui est absolument magnifique et que je vous recommande.

Alors, j'arrive maintenant à des gens que j'ai vraiment connus, Hubel et Wiesel, chercheurs sur la vision humaine. Et ensuite, on va voir comment ça nous ramène tout naturellement à l'intelligence artificielle.

Alors, qu'a fait Hubel ? Il a fait des expériences qui aujourd'hui seraient interdites parce qu'il a expérimenté sur des chats et donc mes cousines nantaises auraient protesté et poussé des cris d'horreur à l'idée qu'on tourmente les chats, sauf qu'elles ne savaient pas, mes cousines nantaises, que quand on plante une microélectrode dans le cerveau, le cerveau ne sent rien parce que l'évolution n'avait pas prévu qu'on plante une microélectrode dans le cerveau. Donc on met juste un anesthésiant au niveau de la surface du crâne. On peut planter des microélectrodes dans le cerveau sans que ça ne provoque la moindre douleur. Donc alors à l'aide d'expériences sur des chats, Hubel et Wiesel ont décodé tout le parcours du signal électrique qui part d'un cône ou d'un bâtonnet de la rétine pour rejoindre les zones cognitives du cerveau. Donc ce qu'ils ont compris, c'est que le signal est transformé par une série d'opérations algébriques qui correspondent exactement aux synapses découvertes par Santiago Ramón y Cajal et enfin visionnables. C'est-à-dire qu'il y a des centres de post-traitement, à peu près une dizaine de centres avant que le signal qui vient de la rétine n'atteigne les zones cognitives du cerveau. Et donc en réalité, ce signal est numérisé, c'est ça qui est extraordinaire. Il est transformé en une suite de pics (spikes) à 1. Et ensuite ces signaux sont traités à nouveau à peu près une dizaine de fois avant qu'ils atteignent les zones cognitives où l'information qui est digitale est recomposée.

Donc la zone cognitive recompose l'image et vous voyez une image. La preuve de cette chose-là, c'est que contrairement à ce que tout le monde croit, le câblage qui part de la rétine et qui va vers le cerveau passe devant la rétine. Vous vous reporterez au livre de Hubel qui explique ça. Si on voyait ce que voit la rétine, on verrait des barreaux de prison parce que le câblage passe devant. Alors on pourrait penser que logiquement, le câblage devrait passer par derrière, mais il passe devant la rétine pour une raison que je ne connais pas. Donc on ne voit pas ce que la rétine voit. On voit ce que les zones cognitives reconstruisent comme image, ce qui est complètement extraordinaire. Donc comme disait Léonard de Vinci, la vision est un processus mental. On ne voit pas avec ses yeux, avec la rétine, on voit avec le cerveau. Il y a un post-traitement de l'image.

Et c'est ça qui a permis à Hubel de faire la découverte majeure qui lui a permis d'obtenir le prix Nobel. Il a compris comment on traite l'amblyopie du nouveau nez. Qu'est-ce que c'est l'amblyopie ? C'est une maladie des jeunes enfants qui provient d'un strabisme. Une partie des neurones ne sont pas utilisés à cause du strabisme et comme toute partie du système nerveux qui n'est pas utilisée, cette partie dégénère et donc l'enfant devient aveugle. Et les gens, au moment du travail de

Hubel, pensaient que l'amblyopie était un problème de la rétine, de l'œil. Et il dit de façon amusée qu'aucune agence de soutien de la recherche ne lui aurait donné de l'argent s'il avait dit qu'il voulait traiter l'amblyopie. Il a compris comment traiter l'amblyopie comme conséquence de ces travaux sur le chemin de l'influx visuel. Et on le fait tout simplement de la façon suivante. On met un bandeau sur l'œil de l'enfant pendant une semaine ou deux semaines pour forcer l'autre œil à fonctionner et ensuite, on échange alternativement. Et le problème se corrige de lui-même. Lesquels d'entre vous connaissaient ce traitement de l'amblyopie... ? Ah oui, donc il y a effectivement des gens qui ont vu ce traitement à l'œuvre et ce traitement, on le doit au génie de Hubel. Donc on le doit à une étude théorique sur le fonctionnement du du système visuel humain. C'est vraiment prodigieux ! Donc vous allez voir que tout ça conduit à l'intelligence artificielle. L'intelligence artificielle est soudée à ces découvertes. Ce n'est pas quelque chose qui vient d'être inventé aujourd'hui, n'est-ce pas. C'est Hubel a eu le prix Nobel, Hubel et Wiesel en 1932, donc il y a plus de 90 ans.

Alors ça, c'est une photo de Stéphane Mallat. Donc Stéphane Mallat est une de mes connexions avec l'intelligence artificielle. Stéphane Mallat était mon élève à l'École Polytechnique. Ensuite, il a travaillé en partie sous ma direction et aujourd'hui, il est professeur au Collège de France et son sujet d'enseignement est l'intelligence artificielle.

Donc c'est pour ça que de façon indirecte, je suis lié à ces connaissances. Donc il nous parle et il est enthousiaste lui. Il dit que l'intelligence artificielle a des milliers d'applications, comme la reconnaissance de la parole, les applications biomédicales, la compréhension du langage naturel etc.

J'ai oublié de faire une remarque sur les systèmes experts. C'est une remarque au sujet du problème de la traduction. La traduction d'une langue dans une autre langue avait été abordée à l'époque de Noam Chomsky par des méthodes à base de systèmes experts. En effet, une langue a une syntaxe. Et donc l'idée naturelle, c'est qu'il suffit de suivre les règles de la syntaxe, d'avoir un dictionnaire pour le contenu sémantique des mots que l'ordinateur apprendrait et on peut traduire. Et alors là, la difficulté est identique à la difficulté de la conduite autonome. Ce système ne fonctionne pas. L'ordinateur utilisant un système expert est incapable de traduire. Il y a une complexité dans l'arborescence qui dépasse la possibilité de calcul des ordinateurs. Il faudrait des siècles pour traduire un texte. La traduction automatique est faite aujourd'hui par l'intelligence artificielle. C'est-à-dire que l'ordinateur n'apprend pas la syntaxe, il apprend des tournures de phrases. Et donc il traduit en regardant dans l'ensemble des choix possibles quelle est la tournure de phrase qui est la plus fréquemment utilisée, qui est la plus probable, ce qui conduit évidemment à des défauts terribles sur lesquels on interviendra à la fin, mais toutes les traductions, aujourd'hui, se font par intelligence artificielle. C'est à cause de l'impossibilité parce qu'il y a ce qu'on appelle une explosion combinatoire. Les méthodes logiques de traduction font dérailler la puissance de calcul de l'ordinateur à cause d'une combinatoire qui est beaucoup trop touffue et beaucoup trop complexe.

Alors les réseaux de neurones du type convolution ont été introduits par Le Cun, on va revenir sur ces méthodes, et elles sont implémentées par des convolutions au sens habituel, suivies par des non-linéarités. Alors la non-linéarité, c'est le fonctionnement des synapses. L'intelligence artificielle repose sur des algorithmes non linéaires définis par une notion de seuil. On va voir tout de suite ça sur le dessin. Donc ça, c'est exactement copié sur le fonctionnement du cerveau inventé par Santiago Ramón y Cajal. Alors donc je résume, les réseaux de neurones formels, ceux qui sont utilisés dans

l'intelligence artificielle, sont une imitation de ceux qui opèrent dans le contexte visuel primaire.

Les réseaux de neurones formels jouent un rôle essentiel dans l'apprentissage profond. Donc voilà un exemple où à gauche, vous avez une image du président américain George Washington et à droite, vous avez l'illustration représentant un réseau de neurones sur lequel s'applique un algorithme qui trouve le nom du président. Donc c'est amusant parce qu'en réalité, il y a une quinzaine de couches et on ne vous le dit pas. Alors qu'est-ce qui se passe ? Il y a des premiers capteurs qui examinent l'image à différents endroits etc., et ensuite, les résultats sont transmis à une première couche qui fait l'analogue de ce que Hubel a trouvé pour la vision humaine.

Ce sont les choses qui partent des cônes et des bâtonnets de la rétine et qui sont traités à nouveau à un niveau légèrement supérieur, et où on fait déjà des opérations algébriques et des sommes ou des différences, des additions, etc. Si ça dépasse un certain seuil, c'est transmis à la couche suivante, avec d'autres coefficients et paramètres. Ensuite, c'est transmis à la couche suivante et à la fin, si le système est bien ajusté, vous obtenez la réponse. Donc tout le problème est de savoir comment ajuster ces coefficients pour que le système trouve la bonne réponse, qui est George Washington (la sortie ou output). Alors le problème, on a l'impression que c'est un tour de magie. Le problème est de savoir comment on fabrique une telle machine. Toujours est-il que l'idée de la machine, c'est de copier le cerveau humain. Donc le prix Turing 2019 qui est l'équivalent du prix Nobel pour l'informatique et qui est l'équivalent d'autres prix, le prix Turing donc, a été attribué à Yoshua Bengio, Yann Le Cun et Geoffrey Hinton. Je fais une remarque sur ces trois chercheurs. Le travail de ces trois chercheurs a été réalisé à Montréal où ils travaillaient ensemble, sous la gouvernance de Geoffrey Hinton, et Yoshua Bengio et Yann Le Cun sont des Français. Yoshua Bengio est né au Maroc. Ensuite, il est arrivé en France et ils ont travaillé à Montréal sous la direction de Hinton. Alors, c'est très triste que ces recherches aient dû être conduites au Canada parce que ça montre le fait qu'on ne parie pas assez sur les jeunes en France. C'est un fait que l'on n'offre pas en France toutes les possibilités. C'est une remarque que le défunt Claude Allègre avait faite souvent, c'est que la structure de la recherche en France est beaucoup trop hiérarchisée. C'est-à-dire qu'il y a encore une forme de mandarinat qui fait qu'on ne donne pas tout l'argent qu'ils désirent à des jeunes pour lancer des projets un peu risqués, un peu aventureux. Et donc on loupe des occasions. Depuis Yann Le Cun est aux États-Unis, il n'est pas revenu en France. Enfin, il revient en France de temps en temps. Alors, expliquons comment le schéma précédent fonctionne. Le schéma précédent qui présente différentes couches de réseaux de neurones, il utilise des paramètres par lesquels on multiplie l'influx qui traverse un neurone entre une couche et une autre. Donc comment trouve-t-on ces paramètres ? Ces paramètres déterminent le fonctionnement. Ils sont cruciaux, ils sont déterminants, c'est-à-dire qu'on amplifie ou on atténue, on module,quoi ! Et ils ne sont pas introduits par les programmeurs mais ils sont appris à partir d'une base de données d'exemples par essai/erreur. Donc là, il y a déjà un des points faibles de l'intelligence artificielle. Ça repose sur le travail de petites mains qui sont obligés de nourrir l'algorithme avec des millions d'exemples, et où la personne qui fait ce travail donne la bonne réponse de telle sorte que l'ordinateur apprend sur ces bonnes réponses ce qu'il doit faire. Donc on utilise, pour ça, des tas de jeunes en Inde qui sont sous-payés et qui sont la base du succès commercial de l'intelligence artificielle et qui servent de professeurs pour l'apprentissage des machines. Donc des millions, voire des milliards, de paramètres sont ainsi modifiés à partir de milliers d'exemples. Alors, une fois qu'on a choisi des paramètres qui fonctionnent, on introduit un nouvel exemple, on regarde si ça fonctionne sur ce nouvel exemple,

sinon on corrige légèrement les paramètres pour que ça fonctionne. Et une fois que l'ordinateur est suffisamment nourri de ses connaissances, c'est ce qu'on appelle l'apprentissage supervisé, à ce moment-là, l'ordinateur peut se lancer en autonomie et il peut travailler sur d'autres exemples.

Donc, ce n'est pas une démarche logique, c'est une démarche qui dit que la situation nouvelle va être un compromis entre diverses situations déjà rencontrées et l'ordinateur fournit la réponse la plus vraisemblable. Ici, on a la définition mathématique de l'apprentissage supervisé. L'apprentissage supervisé, c'est qu'on a un ensemble de données ; ce sont des images, par exemple, dans le cas de la conduite autonome, et on cherche à construire une fonction, un algorithme qui modélise le jugement ou la décision prise sur la donnée X, savoir si la tache noire est une ombre ou un rocher. Quand on connaît (ça, c'est la page de la phase d'apprentissage), le résultat sur un certain nombre d'exemples, c'est-à-dire pour l'image x_1 on sait qu'on doit trouver y_1, et pour l'image x_2, on sait qu'on doit trouver y_2, et pour l'image x_n, on doit trouver y_n. C'est pendant cette phase que les petits indiens travaillent, et on cherche f, on cherche l'algorithme en privilégiant certains critères : simplicité, vraisemblance, qui sont des a priori du problème, et qui sont des contraintes injectées dans l'algorithme. Yann Le Cun explique bien mieux que moi la méthode que je viens de décrire. Voilà. Alors, Yann Le Cun, comme je'lai dit a fait son post-doctorat au sein de l'équipe de Hinton, Geoffrey Hinton, à Montréal. Alors, en 1987, il rejoint l'université de Toronto et en 1988, les laboratoires AT&T pour lesquels il développe des méthodes d'apprentissage supervisé pour la reconnaissance d'images et c'est en ce sens que les travaux de Yann Le Cun rejoignent mes travaux. Il élabore aussi l'algorithme de compression DJVU (pour déjà vu) qui repose sur l'analyse et la compression par ondelettes.

Il y a des tas de livres que vous recevez en PDF et qui sont basés sur cet algorithme DJVU. Voilà. Alors voilà Yann Le Cun qui nous parle.

> "L'explosion récente de l'intelligence artificielle et de ces applications est due en grande partie à l'utilisation de l'apprentissage machine et particulièrement des techniques d'apprentissage profond. Les idées fondatrices de l'apprentissage machine remontent aux années 1950. À cette époque, l'apprentissage machine se fait par les systèmes experts. Donc ça remonte aux années 1950 tandis que l'apprentissage profond remonte aux années 1980. Les applications pratiques ont explosé ces dernières années. Des voitures autonomes à la traduction, la recherche d'information et la médecine personnalisé."

Mais que dit la théorie ? Donc la formulation mathématique la plus générale des processus d'apprentissage supervisé, ça va vous amuser, est celle du mathématicien russe émigré aux États-Unis, Vladimir Vapnik.

Dans son petit livre publié en 1995 *La nature de la théorie statistique de l'apprentissage*, Vapnic décrit les conditions générales dans lesquelles un système informatique ou biologique peut apprendre une tâche et à quelle vitesse. Mais cette théorie générale n'explique pas pourquoi les méthodes modernes d'apprentissage profond fonctionnent si bien, bien mieux que les méthodes plus simples et mathématiquement mieux comprises. L'émergence de l'intelligence artificielle suscitera-t-elle une nouvelle théorie de l'intelligence ? Ce que je crois, c'est que la compréhension de la raison pour laquelle l'ia fonctionne si bien est un nouveau défi pour les mathématiques et cette compréhension mathématique est peut-être la clé des futurs progrès de l'intelligence artificielle.

Alors, qui était Vladimir Vapnic ? C'est intéressant de voir comment tout se rejoint dans la recherche et qui fait qu'un homme qui a vécu en Union Soviétique, finalement, arrive à travailler dans un laboratoire d'intelligence artificielle. C'est vraiment un grand paradoxe. Il est né en 1936, en Ouzbekistan, il est plus âgé que moi. Il a quitté l'Union soviétique en 1995 après la chute du mur. Il a travaillé pour Bell et AT&T, puis pour Facebook. Voilà. Alors, il reprend la phrase de Galilée : "Les mathématiques sont le langage dans lequel Dieu a écrit l'univers".

Et il dit : "Ma recherche actuelle consiste à développer des méthodes avancées de l'inférence statistique empirique.". Et alors, il dit que malgré les succès extraordinaires du deep learning, la plupart de la recherche est totalement empirique et manque complètement d'un substrat mathématique. Alors là, hier j'ai reçu mon courrier et à ma grande stupéfaction, il y avait un article magnifique de Gabriel Peyré, dont j'ai parlé (juste en disant que c'était mon petit-fils) sur les mathématiques de l'intelligence artificielle. Donc c'est dans la Gazette des mathématiciens de hier, et l'auteur de l'article est Gabriel Peyré. La Gazette est publiée par la SMF, et c'est un article absolument spectaculaire. Donc il il dit quelles sont les mathématiques qui sont derrière ces recherches de paramètres optimaux et qui font que le système fonctionne.

Donc maintenant, on revient à la DARPA, et l'apprentissage profond, aujourd'hui, est une boîte noire. Une boîte noire, ça signifie qu'on ne sait pas, en réalité, comment ça fonctionne. On a une entrée, une sortie et on ne sait pas ce qui se passe à l'intérieur. Ça fonctionne et on ne sait pas très bien pourquoi.

L'intelligence artificielle ne sait pas penser. Elle sait voir, maintenant, elle sait interpréter des images. Elle sait conduire, elle sait traduire une langue dans une autre de façon remarquable, mais elle ne sait pas penser. Et il y a maintenant un programme de la DARPA doté de 4 milliards de dollars qui est de construire une intelligence artificielle qui puisse justifier ses décisions. Alors le point est capital parce que l'intelligence artificielle est utilisée à l'heure actuelle pour faire des diagnostics médicaux parce que mon rhumatologue me dit qu'à l'heure actuelle, les examens de radiographie pour savoir si une femme a un cancer du sein ou non conduit par l'intelligence artificielle fonctionne mieux que le diagnostic du médecin, mais en même temps, de temps en temps, ce système produit des erreurs. Donc le problème auquel on est confronté, c'est le problème de la responsabilité. Est-ce que c'est l'algorithme qui est responsable ? Où se situe la responsabilité ? Donc on voudrait savoir comment l'intelligence artificielle a pris sa décision. C'est un problème qui est crucial aujourd'hui, compte tenu des applications immenses de l'intelligence artificielle, qu'elle puisse justifier ses décisions de façon logique mais il n'y a aucune logique dans le fonctionnement de la boîte noire. On en est à l'intelligence artificielle dite ia de la 3^{eme} génération. Je fais un résumé de mon exposé.

L'intelligence artificielle est née en 1970. La première vague, c'étaient les systèmes experts. On nourrissait l'ordinateur de connaissances, de l'expertise nécessitée par un domaine particulier du savoir. Une fois nourri de cette expertise, l'ordinateur répondait à une demande en appliquant des règles logiques. C'étaient les systèmes experts. Cette approche système expert est une approche très rigide. Elle échoue misérablement sur la reconnaissance vocale, la traduction automatique, les voitures autonomes etc., comme on l'a vu. Donc quel est le défaut terrible des systèmes experts ? Ils sont rigides, ils ne peuvent pas apprendre. La seconde génération de l'ia, ça a été le machine

learning, ce dont je vous ai parlé, et qui est basé sur le fonctionnement du cerveau. Le machine learning est adapté au problèmes nécessitant la perception du monde extérieur. C'est le problème de la vision, et on a vu que le problème posé par la vision est l'interprétation de ce qu'on voit. Mais le machine learning échoue sur des problèmes nécessitant l'abstraction, la modélisation et le raisonnement. C'est-à-dire qu'on est passé d'un univers à un autre complètement différent où on sait par instinct quelle est la décision à prendre mais où on est incapable de justifier cette décision. La troisième vague de l'ia qui est encore à venir devra expliquer les décisions prises par l'ia en élaborant un modèle. Alors, je vais voir. Je parle encore 5 minutes de l'approche de Coifman. Donc l'approche de Coifman consiste à remplacer l'apprentissage supervisé par l'apprentissage non supervisé. Que veut dire "apprentissage non supervisé" ? C'est que la machine par exemple observe le mouvement d'un pendule et elle découvre sans qu'on ne lui apprenne rien les lois de Galilée et les lois de la physique sur ce qu'elle voit. Cette chose-là existe déjà et a été élaboré sur des systèmes qui présentent une certaine cohérence interne.

(projetant une photo où on le voit avec Coifman) Alors là, vous me voyez en train de discuter avec Coifman au lieu d'écouter l'orateur. Voyez, les autres gens sont très sages et nous, on est en train de discuter. Alors l'approche de Coifman utilise les réseaux de diffusion (diffusion maps). Sont ce Ces réseaux permettent ce qu'on appelle aujourd'hui d'appliquer une méthode de réduction de la dimensionnalité. Ce sont des problèmes très complexes qui utilisent des millions de paramètres, et on veut simplifier ça dans un modèle qui est seulement fonction de trois ou quatre paramètres. Donc je ne peux pas entrer dans les détails, je dis juste que ça existe. Donc cette approche est particulièrement utile pour analyser des ensembles de données complexes et non linéaires en révélant les structures sous-jacentes sans nécessiter les étiquettes (les labels) ou la supervision, ou une supervision explicite (c'est-à-dire qu'on n'a pas besoin d'esclaves), on regarde ce qui se passe et l'ordinateur est capable avec ce type de nouvel algorithme de conclure sur les causes, les raisons de ce qu'il observe. Bon, je ne vais pas insister parce que c'est technique, je vais m'arrêter là. Alors, je vais faire en sorte que le pdf soit disponible à tout le monde parce qu'il contient une bibliographie qui contient énormément d'informations.

Donc, outre la conférence orale, on va mettre en ligne le pdf de mon exposé comme ça vous pourrez le relire etc. Donc je m'arrête là, le domaine est extrêmement neuf, il y a évidemment une infinité de questions à poser et comme vous le voyez c'est un domaine en pleine évolution. L'article de Peyré dans la Gazette des mathématiciens montre que finalement, jusqu'à aujourd'hui on ne savait pratiquement rien. Tout est à venir et pour les jeunes c'est vraiment le domaine où il faut travailler. Je vous remercie.