

ressources en ligne

Machines, données et apprentissage : relations et enjeux

Federica Minichiello



Édition électronique

URL : <http://ries.revues.org/5755>

ISSN : 2261-4265

Éditeur

Centre international d'études
pédagogiques

Édition imprimée

Date de publication : 1 avril 2017

Pagination : 12-15

ISBN : 978-2-85420-614-2

ISSN : 1254-4590

Référence électronique

Federica Minichiello, « Machines, données et apprentissage : relations et enjeux », *Revue internationale d'éducation de Sèvres* [En ligne], 74 | 2017, mis en ligne le 01 avril 2017, consulté le 19 mai 2017. URL : <http://ries.revues.org/5755>

Ce document a été généré automatiquement le 19 mai 2017.

© Tous droits réservés

ressources en ligne

Machines, données et apprentissage : relations et enjeux

Federica Minichiello

- 1 Apprentissage par la machine, apprentissage profond, intelligence artificielle : des sujets récurrents dans les réflexions actuelles sur l'évolution de la société, du marché de travail et de l'éducation.
- 2 L'association entre machines, données et apprentissage laisse présager une réponse adaptée aux besoins de chaque élève, une personnalisation et une optimisation possible des parcours de formation ainsi qu'une évolution de certains métiers, exposés aux risques d'automatisation.
- 3 Dans cet article, nous proposons des ressources pour mieux appréhender ces évolutions : dans la première partie, nous tentons de clarifier les différents concepts et les relations sous-jacentes ; la deuxième partie traite plus en particulier de l'utilisation d'informations liées à l'apprentissage (learning analytics).
- 4 En 2015, la conférence internationale OCDE/CERI « Governing Complex Education Systems project » résumait ainsi la situation : « we know so much, but we understand so little¹ » : malgré une quantité exponentielle de données, il demeure compliqué d'extrapoler des éléments d'aide à la décision et d'identifier les leviers d'action pertinents.
- 5 Sitographie arrêtée le 15 février 2017.

Des concepts proches, mais distincts

« Intelligence artificielle, apprentissage machine et apprentissage profond »

- 6 Cet article d'Y. Le Cun, chercheur en intelligence artificielle (et l'un des pères de l'apprentissage profond) propose une série de définitions. L'intelligence artificielle est décrite comme un ensemble de techniques permettant aux machines de réaliser des actions normalement réservées aux hommes et par conséquent, de les imiter : dans leur capacité de reconnaître des images ou de gérer une conversation, par exemple. Par l'apprentissage, un système intelligent affine ses performances sur des tâches données ou apprend à en exécuter de nouvelles. Dans l'apprentissage par la machine (ou automatique), l'unité d'information est la donnée (« une image est un tableau de nombres indiquant la luminosité ou la couleur de chaque pixel »). Lorsque le nombre des données est suffisamment élevé, la machine acquiert une capacité de généralisation et saura, par exemple, classifier correctement des images jamais vues auparavant. L'apprentissage profond s'inspire du cerveau humain « comme l'avion est inspiré de l'oiseau » : un réseau artificiel de neurones qui hiérarchise les unités de l'information, les ordonne des plus simples aux plus complexes en toute autonomie et qui peut ainsi « apprendre ». La difficulté principale pour la machine réside dans l'apprentissage non supervisé, la connaissance que seul un être vivant sait traduire de l'observation du monde ou, tout simplement, de l'imprévu.
- [<http://www.college-de-france.fr/>]

« AI vs Deep Learning vs Machine Learning »

- 7 Le site Data Science Central propose de nombreuses ressources dans le domaine du « Big data ». Selon cet article, l'intelligence artificielle, l'apprentissage profond et par la machine sont des sujets liés qui présentent des zones d'intersection, mais aucun n'est un substitut parfait de l'autre. Si une machine est encore loin de fonctionner comme un être humain dans un environnement désordonné et chaotique, l'intelligence artificielle « faible » lui permet de reproduire un comportement humain dans un périmètre délimité. « Le traitement de l'image, du texte et du dialogue sont les yeux, les oreilles, les mains et la bouche de l'intelligence artificielle ». La majorité des avancées dans ces traitements automatiques seraient liées à l'utilisation d'une catégorie d'algorithmes « d'apprentissage profond », qui peuvent par exemple reconnaître des images après en avoir vu des millions similaires, sans avoir reçu en entrée préalable toutes les clés de lecture. L'apprentissage par la machine sous-entend également une utilisation d'algorithmes pour extrapoler des tendances à partir des données de façon autonome. Cette absence de supervision le distingue cependant de l'apprentissage profond, dans lequel il faut toujours confirmer la validité des résultats, étape après étape.
- [<http://www.datasciencecentral.com>]

Artificial intelligence and life in 2030

- 8 L'Université de Stanford a lancé en 2014 un programme d'études sur cent ans de l'intelligence artificielle dans toutes ses possibilités d'impact sur nos vies ; l'enjeu principal pour des biens communs comme l'éducation et la santé est de réussir à reproduire une fluidité d'échange aujourd'hui réservée aux relations entre êtres humains : traitement du langage naturel, gestion de dialogue avec des agents virtuels, traduction automatique. Le dernier rapport prospectif (2016) prévoit dans les quinze prochaines années une généralisation de l'assistance aux enseignants par des tuteurs intelligents et une démocratisation d'applications d'apprentissage basées sur la réalité virtuelle.
[<https://ai100.stanford.edu/>]

Deep Learning.net

- 9 Ce site propose une large palette de ressources sur l'apprentissage profond : des ouvrages de référence, des groupes de recherche comme les universités de Toronto, Montréal, New York (ainsi que des laboratoires de sociétés comme Google et Facebook), des tutoriels et une sélection de prototypes qui expliquent, par exemple, comment un réseau artificiel de neurones apprend à reproduire un texte dans plusieurs versions d'écriture manuelle.
[<http://deeplearning.net>]

« *What if ... computers were trillions time faster?* »

- 10 Les évolutions sont souvent liées à des avancées techniques : par exemple, le fait de disposer de processeurs puissants et à faible coût a rendu financièrement possible de tester la reconnaissance automatique sur des millions d'images, réduisant ainsi les taux d'erreur des machines. Cette note récente du service de recherche du Parlement européen (2017) évoque la prochaine révolution annoncée : l'informatique quantique. Si les ordinateurs actuels stockent et traitent de l'information selon une logique binaire (0 et 1), l'informatique quantique permettrait une superposition d'états ou « la possibilité d'être 0 et 1 à la fois ». Ce qui se traduirait par une puissance encore plus importante pour les machines en termes de calcul, de collecte et d'analyse de données et qui bouleverserait de nombreux fonctionnements actuels, comme la sécurité informatique (cryptage des données).
[<http://www.europarl.europa.eu/stoa>]

L'analyse des données d'apprentissage

« *Les learning analytics : quand le big data s'intéresse à l'éducation* »

- 11 UTICE, le centre de formation à la pédagogie et au numérique de l'Université Bretagne Loire, a organisé en janvier 2017 un séminaire autour des learning analytics. En France, ces données seraient encore peu exploitées à des fins de pilotage. Aux États-Unis, quatre typologies d'utilisation se diffusent dans l'enseignement supérieur : des systèmes d'alerte, pour identifier les étudiants à risque ; des visualisations des données, comme les tableaux

de bord proposés aux étudiants ou aux enseignants ; l'orientation, par des suggestions de cours à suivre en fonction du profil académique ; des systèmes d'apprentissage adapté, comme le tutorat intelligent. Les conclusions de cette journée ont relevé plusieurs enjeux, comme la continuité de l'analyse des données dans une perspective de formation tout au long de la vie et les mécanismes d'appropriation des données personnelles par chaque individu.

[<https://utice.u-bretagne-normandie.fr>]

Hubble

- 12 Le projet HUBBLE associe neuf laboratoires de recherche français pour la création d'un observatoire national des processus d'analyse de données d'apprentissage. L'objectif est de modéliser les traces laissées dans des environnements de formation en ligne, proposer des outils pour agréger et visualiser les données et décrire les processus d'analyse pour les différents décisionnaires (enseignants, concepteurs de plateforme en ligne, administrateurs de l'éducation, etc.).

[<http://hubblelearn.imag.fr>]

Learning analytics and student success – assessing the evidence

- 13 Le dernier rapport (2017) de l'organisation à but non lucratif anglaise JISC liste les deux principales sources de données liées à l'apprentissage : 1) les plateformes, où les étudiants accèdent aux emplois du temps, aux ressources pédagogiques, interagissent avec leurs pairs dans des forums (environnements virtuels, MOOCs) ; 2) les systèmes d'information de gestion de l'éducation, qui regroupent des données comme les qualifications, les situations socio-économiques, les notes. S'y ajoutent d'autres informations, comme les traces laissées sur des plateformes de livres électroniques, les statistiques de fréquentation d'une bibliothèque, etc. Le JISC travaille actuellement pour mettre à disposition d'établissements d'enseignement supérieur une architecture ouverte de collecte et de gestion de ces données, avec des bénéfices espérés en termes de taux de réussite (et de baisse du décrochage). Selon une étude de la Nottingham Trent University, 27 % des étudiants témoigneraient d'un changement de comportement provoqué par la prise de connaissance de données et la comparaison avec leurs pairs.

[<http://www.jisc.ac.uk>]

Society for Learning Analytics Research (SoLAR)

- 14 SoLAR est un réseau interdisciplinaire et international de chercheurs qui étudient l'impact des données dans l'enseignement, la formation et l'apprentissage. Parmi ses initiatives, on peut citer la conférence internationale sur les learning analytics (en mars 2017 à l'université canadienne Simon Fraser) et la revue en accès libre Journal of Learning Analytics. Parmi les sujets traités : l'apprentissage autodirigé par l'apprenant, la visualisation des émotions pendant le processus d'apprentissage, l'analyse des données produites par les enseignants, etc.

[<https://solaresearch.org/>]

Research Evidence on the Use of Learning Analytics: Implications for Education Policy

- 15 Ce rapport récent (2016) du Joint Research Center (JRC) illustre différentes initiatives dans le domaine, comme en Norvège, avec le centre de recherche dédié SLATE [<http://www.slate.uib.no/>] ou au Danemark, avec un système généralisé d'inscription en ligne dans l'enseignement secondaire et un entrepôt de données public qui permet de comparer les données des étudiants entre institutions, districts ou régions. Le JRC dresse plusieurs constats : premièrement, une attention excessivement tournée vers l'offre, au détriment de la demande. Peu d'efforts seraient faits pour mettre en relation les données avec les changements que les éducateurs voudraient impulser à partir de ces informations. Deuxièmement, il n'y aurait pas d'évidence à ce jour de corrélations entre l'analyse des données et de meilleurs processus d'apprentissage. [<http://publications.jrc.ec.europa.eu>]

Learning analytics report card (LARC)

- 16 Ce projet pilote de l'Université d'Édimbourg vise l'implication des étudiants dans la collecte et l'analyse de leurs informations, issues de l'activité individuelle dans un cours. Au lieu de rendre les résultats sous une forme institutionnelle agrégée et souvent peu lisible, la *Report Card* permet aux étudiants de manipuler leurs propres données, de choisir la façon de les trier ou les croiser, la fréquence à laquelle produire des rapports, etc. [<https://larcproject.wordpress.com>]

Learning analytics community exchange (LACE)

- 17 Le Learning Analytics Community Exchange est un projet financé par le septième programme-cadre de l'Union européenne pour la recherche et le développement qui aspire à réunir autour des learning analytics une communauté élargie associant à la fois les écoles, les universités et les lieux de travail. Terminé en juin 2016, le projet a notamment soulevé des questions d'éthique et mis en évidence la nécessité d'un cadre juridique ; une alerte a été aussi lancée au sujet des outils « dans le nuage » qui se répandent dans l'éducation à tout niveau et posent de problèmes complexes liés à la propriété des données, à l'identité virtuelle, des défis encore plus accentués au sein de l'école, avec la nécessité d'inclure les familles et les élèves en tant que parties prenantes. [<http://www.laceproject.eu/>]

NOTES

1. . Littéralement : « Nous en savons tellement, mais nous comprenons si peu ».

INDEX

Mots-clés : données massives en ligne, apprentissage par la machine, intelligence artificielle

Palabras claves : datos masivos en línea, aprendizaje automático, inteligencia artificial

Keywords : massive online data, machine learning, artificial intelligence

AUTEUR

FEDERICA MINICHELLO

Federica Minichiello est chargée de veille au centre de ressources et d'ingénierie documentaires (CRID) du Centre international d'études pédagogiques (CIEP). Courriel : minichiello@ciep.fr