

Le machine learning améliore le ROI pour ses pionniers

Les recherches d'IDG montrent comment les entreprises mettent en œuvre les projets d'intelligence artificielle et de machine learning, ainsi que les avantages commerciaux qu'elles en retirent.

La technologie du machine learning (ML) en est encore à ses balbutiements, mais les premières entreprises à l'avoir adoptée constatent déjà ses avantages. Planification minutieuse, préparation détaillée des données et capacité d'apprentissage au fil des tâches font partie des clés de son succès.

Ces conclusions sont tirées d'une enquête menée par IDG Research Services auprès de 200 cadres supérieurs du secteur informatique ayant participé à des projets de ML au sein d'entreprises privées comptant 1 000 employés au minimum. IDG a également effectué des entretiens approfondis avec huit cadres supérieurs de l'informatique exerçant dans de grandes entreprises, au cours desquels ils ont détaillé leur expérience du machine learning.

L'ensemble des recherches indiquent un avenir largement prometteur pour le ML et les technologies d'intelligence artificielle (IA) associées, la moitié des personnes interrogées estimant que cette technologie est d'ores et déjà incontournable pour les entreprises.

Ce livre blanc reprend les principaux résultats des recherches sur la mise en œuvre du machine learning, de ses origines à son évaluation.

Origines du projet : les projets de ML proviennent de différents secteurs d'activité

Au départ, l'intérêt pour les projets de ML émane typiquement de services de différents secteurs d'activité qui cherchent à résoudre un problème ou à saisir une nouvelle opportunité commerciale. La difficulté, ou l'idée, est généralement soumise au service informatique, qui la juge appropriée pour les technologies de ML/IA. La majorité des personnes interrogées

(67 %) déclarent que le service informatique devient alors l'initiateur des projets de ML, un nombre limité de projets continuant d'être pilotés par les services des secteurs d'activités (18 %) ou les équipes chargées de la cybersécurité (13 %).

L'aspect le plus délicat dans la mise en place des projets de ML est la capacité à concevoir un projet présentant un retour sur investissement (ROI) important. Par exemple, certaines des personnes interrogées indiquent avoir pu démontrer que cette technologie augmente la productivité des employés, grâce, notamment, à l'automatisation des réponses aux appels téléphoniques des clients et aux requêtes sur les médias sociaux, ou à la détection des faux positifs dans de gros volumes de données sur la cybersécurité.

Les responsables informatiques précisent que l'élaboration du projet nécessite de collaborer étroitement avec les responsables des secteurs d'activités afin de déterminer comment cette technologie peut répondre aux différents besoins des entreprises.

Pour un premier projet, les entreprises pionnières de l'utilisation du ML conseillent de choisir un cas relativement simple et susceptible de prouver l'intérêt de cette technologie.

Les deux tiers des personnes interrogées utilisent le ML afin de passer au crible des quantités colossales de données en vue d'analyser la sécurité, les risques et la fraude, tandis qu'elles sont 29 % à envisager de le faire dans les 12 mois qui viennent (voir la Figure 1). Sur l'ensemble des recherches, les cas d'utilisation du ML varient selon les secteurs. Les entreprises orientées consommateur ont par exemple tendance à privilégier les projets qui les aident à gérer la communication avec les clients et à stimuler les ventes. Un professionnel de l'immobilier interrogé souhaitait, quant à lui, utiliser le ML afin d'estimer les prix des biens sur le marché.



"Rien ne peut nous arrêter. Tout ce que nous faisons ou presque peut tirer avantage du machine learning et d'autres fonctionnalités intelligentes."

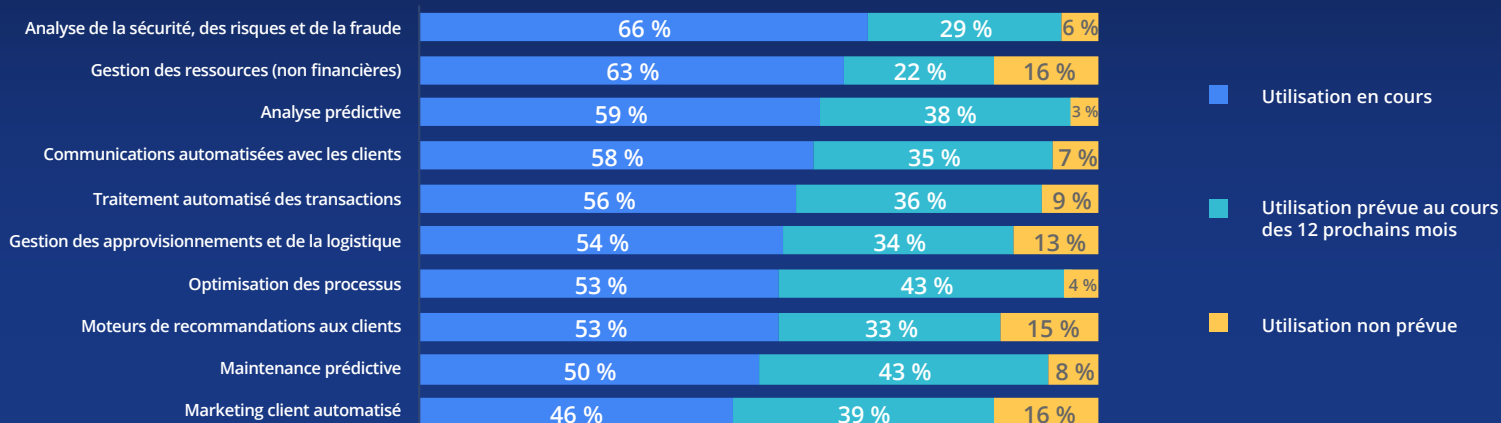
— *Le DSI d'une société immobilière au CA de 1 milliard de dollars*



"Commencez à petite échelle, puis prouvez que le concept fonctionne et qu'il génère de la valeur."

— *Le vice-président de l'informatique d'une société de services financiers au CA de 90 milliards de dollars*

Figure 1. Cas d'utilisation actuels et planifiés du machine learning





"Nous sommes noyés dans un flot d'informations organisées en silos de différentes manières. Il faut chercher à comprendre comment tirer le meilleur profit de toutes ces données."

— Le RSSI d'une société de centres d'appels au CA de 155 millions de dollars



"Nous n'avions aucun plan dont nous étions sûrs à 100 %. Nous avons appris au fur et à mesure."

— Le directeur de la technologie d'une société de services financiers au CA de 2 milliards de dollars



"Ce qui nous retient, c'est que nous n'avons pas 20 data scientists pour s'en charger."

— Le responsable des calculs d'un fabricant de pièces automobiles au CA de 330 millions de dollars

Se préparer pour le ML

Une fois le premier projet identifié, l'étape suivante consiste à se préparer à sa mise en œuvre.

Il ressort clairement des entretiens approfondis que l'une des étapes initiales les plus importantes est d'évaluer l'infrastructure en place. La plupart des entreprises interrogées possédaient déjà une stratégie d'analyse/de Big Data et percevaient les projets de ML comme le prolongement de cette stratégie. Par conséquent, elles n'ont pas dû réaliser de gros investissements dans des infrastructures au départ.

Par ailleurs, le cloud joue souvent un rôle important. Les personnes interrogées indiquent à 62 % que leur infrastructure informatique utilise des ressources cloud publiques, privées ou hybrides, avec une légère hausse prévue pour les 12 prochains mois. L'enquête indique que seul un tiers des ressources continuera d'être géré sur site dans un an.

D'après les personnes interrogées, la prééminence du cloud dans les stratégies de ML est due au fait que cette technologie est bien adaptée aux charges de travail nécessitant une forte puissance de traitement informatique sur de courtes durées, ainsi qu'au stockage des données.

Les entreprises qui comptent utiliser des ressources cloud doivent examiner les API disponibles pour déplacer les données vers le cloud. Près des deux tiers des personnes interrogées (60 %) déclarent utiliser des API pour les charges de travail analytiques, 53 % s'en servent pour les bases de données, 47 % pour les données transactionnelles et 43 % pour les charges de travail hautes performances.

La préparation des données représente un autre défi de taille. Pour 35 % des personnes interrogées, l'organisation et la préparation des données constituent le principal défi lié à la mise en œuvre du ML.

Mettre en œuvre un projet de ML

Concernant la mise en œuvre, les entretiens approfondis révèlent un consensus quant à la méthode, basée sur des tentatives par tâtonnement.

Les difficultés évoquées incluent la capture et l'analyse de données en temps réel (33 %) ainsi que la gestion de données non structurées (32 %). Seul un tiers des personnes interrogées utilise exclusivement des données structurées dans leurs projets de ML, ce qui rend d'autant plus importante l'élaboration d'un plan pour le traitement des données non structurées.

De plus, il est crucial de pouvoir optimiser les programmes de ML, les principaux enjeux étant les suivants :

- **Qualité et gestion de données complètes**
- **Une plate-forme intégrée de bout en bout pour l'automatisation de la prise de décision à partir des données**
- **Exploration interactive des données, et visualisation des résultats modélisés**

En outre, les personnes interrogées soulignent que l'environnement de ML doit être flexible et permettre de gérer plusieurs types de charges de travail. Lorsqu'il leur a été demandé de sélectionner les charges de travail qu'elles appliquent au ML, les personnes interrogées les ont toutes citées : charges de travail analytiques (77 %), avec bases de données (70 %), transactionnelles (65 %), hautes performances (56 %) et par lot (46 %).

Concernant les types de méthode de ML employés, le modèle supervisé est le plus répandu. Seuls 7 % des personnes interrogées utilisent un modèle non supervisé, selon lequel l'algorithme détermine librement les associations entre les données d'entrée et les résultats, ou trouve des structures et des associations dans un ensemble de données : les clients ayant acheté X ont aussi acheté Y, par exemple.

Près du tiers des personnes interrogées utilisent un modèle hybride ou semi-supervisé, selon lequel une partie des données est libellée, mais la majeure partie ne l'est pas. L'utilisation d'une grande quantité de données non libellées simplifie la préparation, en nécessitant moins d'expertise, et permet aux entreprises d'inclure plus de données dans le modèle, ce qui peut conduire à de meilleurs résultats. Concernant le choix du logiciel de ML, les responsables de la mise en œuvre n'indiquent pas de réelle préférence,



à une exception près : seuls 7 % utilisent une solution entièrement propriétaire. Sinon, l'utilisation se partage en trois options :



Environ un quart des personnes interrogées prennent la décision d'utiliser tel ou tel logiciel de ML au cas par cas.

De même, il n'existe pas de consensus réel sur l'utilisation de solutions Open Source, privilégiées par 35 % des personnes interrogées, par rapport aux solutions propriétaires de fournisseurs (26 %). Les deux solutions sont utilisées conjointement par 40 % des personnes au cas par cas.

Certains précurseurs jugent les offres de ML basées sur le cloud attractives : "Ce qui nous a plu [dans le cloud], c'est que nous ne savions pas quel montant nous allions investir dans cette technologie. Avec cette approche, aucun investissement matériel ni logiciel initial n'était requis," déclare le responsable de la sécurité des systèmes d'information (RSSI) d'un centre d'appels. Il ajoute qu'il a trouvé le côté "prêt à l'emploi" attractif, car cela rendait la solution plus accessible.

L'expertise, un facteur décisif

Outre les considérations liées aux infrastructures pour la mise en œuvre du ML, les difficultés liées au recrutement se sont largement imposées comme un facteur décisif. Seuls 15 % des personnes interrogées ont affirmé que leurs équipes en place possédaient toutes les compétences requises pour mener à terme un projet de ML. C'est probablement aussi la raison pour laquelle seuls 7 % des entreprises développent leur propre solution de ML.

Pour parer à ce problème, une approche fondée sur plusieurs orientations est nécessaire. Ces orientations incluent la formation du personnel, déjà assurée ou planifiée par 57 % des entités interrogées, l'inclusion de partenaires établis (48 %) et le recrutement de nouveaux employés possédant les compétences requises (39 %).

Les data scientists n'étant globalement pas assez nombreux, cela fait grimper leur salaire de base. Il s'agit même actuellement du poste le plus difficile à pourvoir, en partie parce que cette fonction est relativement nouvelle. Le salaire moyen d'un data scientist s'élève à 111 267 \$ par an.¹

Les résultats de l'enquête confirment ce chiffre. Le directeur des systèmes d'information (DSI) d'une entreprise de services de santé au chiffre d'affaires de 100 M\$ explique que ses data scientists gagnent plus que lui. Il recrute maintenant en Inde, où les employés du secteur informatique, qui possèdent un niveau de qualification standard, gagnent environ le dixième du salaire de leurs homologues américains. Cependant, il indique rémunérer ses data scientists indiens 100 000 \$ par an afin de contourner le problème de la demande.

Évaluer les résultats du ML

Si la mise en œuvre présente des défis, les personnes interrogées indiquent clairement que les résultats en valent la peine. Les entreprises attendent surtout du ML qu'il apporte un gain de valeur commerciale, fondé entre autres sur l'amélioration de la rentabilité, un avantage concurrentiel, etc. (voir la Figure 2).

Ainsi, une entreprise de services de santé est parvenue à améliorer la productivité de ses employés et à optimiser ses opérations grâce à la mise en œuvre d'un projet de ML. Cette société a appliqué la technologie de ML à son centre des opérations de sécurité afin d'augmenter le nombre de points de terminaison que chaque employé peut surveiller. Le secret pour y parvenir fut de gérer les alertes concernant des faux positifs. D'après ses recherches, ceux-ci prenaient 70 % du temps d'un technicien par alerte. La technologie de ML permet désormais d'identifier ces faux positifs à l'avance.



"Nous pensons pouvoir réduire nos faux positifs d'environ 95 %. Nous pourrions alors faire passer le temps moyen de résolution d'un problème d'environ 49 minutes à environ 5 minutes."

— Le DSI d'une entreprise de services de santé au CA de 100 millions de dollars

Figure 2. Avantages motivant la mise en œuvre du machine learning



IDG Research Services, mai 2017 ; critères identifiés comme décisifs ou très importants

Aujourd'hui, grâce à la réduction des faux positifs, chaque technicien peut surveiller 10 fois plus de points de terminaison. Au final, l'entreprise évolue plus rapidement à mesure qu'elle gagne des clients. "Si j'ajoute 50 000 licences de notre produit de sécurité, j'envisage maintenant de recruter six personnes, contre 50 auparavant," explique-t-il. De plus, ses besoins en techniciens hautement qualifiés sont réduits, étant donné que l'application de ML/IA exécute le gros du travail.

Dans certains cas, il arrive que les entreprises mettent en œuvre le machine learning afin de résoudre un problème spécifique et identifient un gain de valeur commerciale insoupçonné dans un autre domaine. Par exemple, un fabricant de pièces automobiles qui utilisait à l'origine la technologie de ML pour développer des polymères a été surpris de constater que ses clients, dont GM, appréciaient le fait de pouvoir prévoir le prix des produits. Cela a permis à l'entreprise de quasiment doubler son activité avec GM.

Un impératif commercial

Bien que le ML n'en soit qu'à ses débuts, 50 % des personnes ayant répondu à l'enquête d'IDG déclarent que cette technologie constitue déjà un impératif commercial. De même, 89 % d'entre elles s'accordent à dire que le ML peut apporter un avantage concurrentiel, tandis que 86 % indiquent y voir une possibilité de soutenir leurs initiatives en matière de cybersécurité. La technologie de ML permet de réduire les coûts d'après 81 % des sondés, et elle offre la possibilité aux services informatiques de s'adapter sans recruter (77 %).

"Je pense que mes homologues devraient s'intéresser à cette solution, étudier son fonctionnement et se livrer à quelques mesures pour parvenir à des résultats et générer davantage de revenus. Pouvoir assurer un certain retour sur investissement et faire valoir ce gain en dollars auprès de vos homologues du secteur, c'est là tout l'enjeu de cette technologie", conclut le DSI de la société immobilière.



"Il faut pouvoir prendre une autre direction lorsque les choses ne se passent pas comme prévu."

— Le directeur mondial de la technologie pour les relations avec la clientèle d'une entreprise de produits grand public au CA de 40 milliards de dollars

¹ CareerCast, www.careercast.com/jobs-rated/toughest-jobs-fill-2017